PROVIDING ACCESS TO SURVEY DATA FOR SOCIAL RESEARCH: HISTORICAL BACKGROUND AND CURRENT PRACTICES AT THE FRENCH NATIONAL INSTITUTE FOR DEMOGRAPHIC STUDIES (INED)

Arianna Caporali[1], Amandine Morisset[2], Stéphane Legleye[3]

Institut national d'études démographiques, 133 boulevard Davout, 75980 Paris Cedex 20

[1] arianna.caporali@ined.fr
[2] amandine.morisset@ined.fr
[3] stephane.legleye@ined.fr

Résumé. Permettant le partage de données, la mise à disposition des enquêtes représente un atout pour les sciences sociales. En effet, elle permet d'accroitre l'utilisation des données, de justifier les coûts élevés des enquêtes, d'éviter la répétition d'études similaires, et de re-tester les travaux scientifiques. Néanmoins, les chercheurs sous-estiment parfois le partage de données, consacrent peu de temps à leur documentation, et attachent peu d'importance aux activités de mise à disposition des enquêtes. Ce papier retrace le développement des archives de données en sciences sociales dans le contexte international et français. Il décrit ensuite les activités de mise à disposition des enquêtes du Service des Enquêtes et Sondage (SES) de l'Institut national d'études démographiques (Ined). Les premières archives furent créées dans les années 1950 pour encourager la recherche comparative en sciences politiques. Elles organisèrent des réseaux internationaux et fixèrent des normes pour la documentation des données. Les principaux avancements furent réalisés grâce aux progrès des technologies de l'information. En France, la mise à disposition des données commença plus tard que dans d'autres pays, et à l'Ined elles commencèrent au début des années 2000. Le SES fut co-fondateur du Réseau Quetelet, qui gère aujourd'hui l'accès aux enquêtes en sciences sociales. L'expérience du SES suggère que la préparation des données nécessite beaucoup de temps. Pour réduire ce temps, les chercheurs devraient préparer les métadonnées à l'avance. La mise à disposition des enquêtes devient de plus en plus importante avec la demande croissante d'accès aux données publique en « Open Data ».

Mots-clés. Archives de données en sciences sociales; Le partage des données; Normes internationales de documentation des données; Données quantitatives d'enquête; Métadonnées.

Abstract. Providing access to surveys is paramount to social research, because it enables data sharing. Data sharing increases data use by the scientific community and, therefore, justifies the high costs of surveys; it avoids duplication of studies and allows re-testing scientific results. Researchers often underestimate data sharing and devote little time to documenting data before depositing them to archives. They often disregard the activities that archives perform to provide access to deposited data. This paper reviews the development of social science data archives in the international and the French contexts. It then illustrates the activities to provide access to surveys at the Survey Department (SES) of the French National Institute for Demographic Studies (INED). Fist archives were created in the 1950s to encourage comparative research in political science. They organised international networks and fixed standard for documenting data. Major developments were taken thanks to innovations in information technology. In France, activities to provide access to data were established later than other countries. INED began these activities in the early 2000s.

Its SES was a co-founder of the *Réseau Quetelet*, which today manages the access to social science surveys. The experience at SES suggests that data preparation requires a lot of time. To reduce this time and increase the quality of metadata, researchers need to plan metadata ahead. Activities to provide access to surveys are becoming more and more important due to the trend toward open data.

Keywords: Social Science Data Archives; Data Sharing; Data Documentation Initiative; Quantitative Survey Data; Metadata.

1 Introduction: the importance of data sharing from a scientific point of view

Activities which allow providing access to survey data are of key importance for social sciences because they allow for data sharing. Data sharing permits replication of studies through which researchers evaluate existing studies, build on them and, therefore, contribute to the progress of their disciplines (King 1995; ICPSR 2012; Silberman 1999). Moreover, data sharing increases acknowledgement of researchers' work through citation, it justifies expensive costs of surveys through data re-use, and it allows for testing of methodologies of data collection (Silberman 1999).

Data documentation is paramount to data sharing, because "without adequate documentation, scholars often have troubles replicating their own results months later" (King 1995, p. 444). However, since the preparation of metadata usually has not specific budget and it does not end with scientific publications, there is little incentive for researchers to dedicate some time to it. As a consequence, archivists who assemble the documentation are often confronted with issues of quality. Researchers often attach little importance to data sharing and sometimes they are unaware about the activities carried out by archivists to provide access to data with good documentation.

This paper acknowledges the importance of data sharing. After an historical account of social science data archives at the international level and in France, it illustrates the development and the current organization of the activities that provide access to survey data at the Survey Department of the French National Institute for Demographic Studies (*Institut national d'études démographiques*, INED). It ends with a summary and an outlook over possible developments about data access in social science.

2 The development of social science data archives at the international level

First social science data archives were created after the Second World War thanks to the initiatives of researchers in political science. The post-war geo-political context encouraged international comparative studies which motivated the need for sharing survey data (Bisco 1966; Silberman 1999). In the USA, the Roper Public Opinion Research Centre opened in 1957 at the Williams College (Williamstown, Massachusetts) (Bisco 1966; Doorn and Tjalsma 2007; Hastings 1964; Silberman 1999). In Europe, Erwin K. Scheuch and Günther Schmölders founded at the University of Cologne the *Zentralarchiv* in 1960 which gathered data supplied by commercial as well as academic research institutes in the Federal Republic of Germany (Bisco 1966; Scheuch 2004). Several other European and American archives were founded in the 1960s and in the 1970s. The majority originated in the academic world with the purpose of sharing surveys produced by researchers. They developed in contexts were researchers preferred to carry out their own studies, rather than to use official statistics (Silberman 1999).

Through its International Social Science Council (ISSC), UNESCO (United Nations Educational, Scientific and Cultural Organization) promoted a debate about the costs and benefits of archiving data. It organized conferences in the 1960s and in the 1970s which encouraged the creation of social science data archives and played a role towards the development of international networks in Europe (Bisco 1966; Rokkan and Scheuch 1963; Rokkan 1966; Scheuch 2004). For example, the

Council of European Social Science Data Archives (CESSDA), still active today, was created in 1976. During the same decade, the International Association for Social Science Information Services and Technology (IASSIST) was created as an organization of individual data archivists.

The conferences organised by IASSIST and CESSDA experts' seminars encouraged the definition of a common international standard for documenting metadata (Doorn and Tjalsma 2007; Scheuch 2004; Silbermann 1999). Major developments towards this direction were especially taken thanks to innovations in information technology. Today the standard used by main social science data archives was developed since 1995 and it is called Data Documentation Initiative (DDI). DDI replaces paper codebooks with metadata on electronic human-readable format[1]. An important driver of its use has been the creation of the software Nesstar (Network Social Science Tools and Resources)[2] (Vardigan, Heus, Thomas 2008).

Today, data sharing and secondary analysis have become more and more important. Some journals have adopted a "data availability policy" (DAP), which requires authors to deposit the data used in articles. Some archives have started to acquire and provide access to qualitative data (Corti 2000).

3 The French context: catching up on a delay

Although French researchers have participated to early discussions about sharing data, social science data archives in France were not created before the mid-1980. French social scientists could rely for a long time on the aggregated data published by the National Institute of Statistics and Economic Studies (INSEE) which was rather reluctant to share the surveys it produced. INSEE surveys covered a large spectrum of topics while there was in France a lack of institutional support for the realization of large scale university surveys that were, as discussed above, at the basis of the development of social science data archives in other countries (Silberman 1999).

---

[1] The DDI (http://www.ddialliance.org/) is the follow up of the Standard Study Description (SSD) agreed at Copenhagen in 1980 by CESSDA members and of the OSIRIS standard, developed at the Inter-University Consortium for Political and Social Research (ICPSR) at University of Michigan during the 1970s (Marker 2013). Based on XML schemas, the DDI consists of a group of items that enables the general description of empirical studies down to the level of each variable. The XML format presents the advantage of being robust, uniform and extensible and of being supported by many applications. DDI is also compatible with other metadata standards, such as Dublin Core (i.e., the bibliographic standards) (Mochmann and Vardigan 2011; Leighton 2002; Vardigan, Heus, Thomas 2008). DDI version 1.0 was launched in 2000. Today, there are two lines of developments of the DDI specification. The DDI-Codebook (DDI-C or DDI 2), currently at version 2.5., was introduced in 2002 and meant for documenting simple survey data. It is document-centric and focused on the elements of a traditional social science codebook. Today, this specification is widely used. The other specification, called DDI-Lifecycle (DDI-L or DDI 3), currently at version 3.1., was introduced in 2008 and designed to document surveys across their entire life cycle. It can be used since the beginning of a survey project to document any of its phases. Among its innovative aspects, it is particularly suited for longitudinal studies, because it contains features that allow explicit comparisons between items of different waves (Hansen *et al.* 2011; Kramer *et al.* 2011). Furthermore, it is compatible with the SDMX standards (Statistical Data and Metadata Exchange, Gregory and Heus 2007), which are used to document aggregated data by official statistics, such as Eurostat, Organization for Economic Cooperation and Development (OECD), and the United Nations Statistical Division (UNSD), etc.

[2] Developed in the framework CESSDA expert seminars and of EU founded projects in the second half of the 1990s (Marker 2013), Nesstar is today managed by Norwegian Social Science Data Services (NSD). It allows managing documentation in DDI-C format, without the archivists having to know the XML language. It also allows publishing data and metadata on internet. The online window of the software provides a convenient and intuitive way for the users to search, browse and explore metadata, and visualize and analyze the corresponding data online. In particular, it is possible to access variable's documentation (including a frequency table), create personalized crossed tabulations, graphs and even produce simple analyses (correlations, regressions). Survey documentations as well as results of analysis can be exported in various format. Despite other software have been developed that support DDI (for a complete list, see DDI webpage), Nesstar is the most widely used among European social science data archives. This is mainly thanks to its online functionalities (further information on Nesstar can be found on its website at http://www.nesstar.com/).

First social science data archives were created in 1986: the Bank of Socio-political Data (BDSP)[3], and the Laboratory of Secondary Analysis and Methods Applied to Sociology (LASMAS)[4]. These archives were established thanks to the initiative of single engineers rather than as a common institutional effort[5] (Silberman 1999). Access to data was regulated through special agreements between the archives and research institutions. Access to personal and confidential data was regulated within the 1978 Data Protection Act[6]; the National Commission for Informatics and Freedom (CNIL) was responsible for its enforcement. However, the law did not provide for the use of personal and confidential data with research purposes[7]. In 1994, the so-called "Balladur Circular" stated that, while access to data produced with public funding was free of charges, institutions wanting to acquire it had to pay the costs of the activities that provide access to data.

In this context, Claude Allègre minister of National Education, Research and Technology commissioned to Silberman, the director of LASMAS, a study with the aim to adequate French social science to the greater circulation of survey data at the national and international levels. Based on a survey among CNRS and University laboratories, Silberman (1999) pointed out problems related to data access, data use and data production. She called for a "real archival structure" (p.47) and for a major reform of French policy regarding access to social science data. Possibility for access to data for research purposes needed to be implemented in the law and any survey produced with public funds needed to be made available for re-use. The experience in other countries showed that archives could encourage quantitative sociology and a greater involvement of researchers in the production of surveys. A larger archival structure was necessary for harmonizing the ways in which surveys were documented and for allowing France to play a role at CESSDA.

The Silberman report laid the foundations for the creation of the *Centre Quetelet* (then called *Réseau*) which today manages in France the access to social science surveys produced by public statistical offices and researchers, and to some international surveys[8]. Established in December 2001 as a CNRS institution, it had three founding members: 1) CIDSP, 2) LASMAS, and 3) INED through its SES[9]. The *Réseau* works along three main axes: data documentation, data access, data valorization (Arduin 2004). First, its members gather all the documentation about each survey and restructure it according to international standard. Following the recommendations by CESSDA, they have adopted DDI and implemented it through the software Nesstar. Second, they distribute the datasets to applicants in the respect of deontological rules and give advice as to the use of datasets. Third, they inform users about data availability and latest innovations about software and methods.

The *Réseau* maintains a website (http://www.reseau-quetelet.cnrs.fr/spip/) where all the data provided by its members are catalogued and documented. The general principle is that the access is free of charges and only for research purposes. Data requests can be sent through the website. They are then forwarded to the interested partners, who review them. The decisions as to whether access

[3]The BDSP was established at the Institute of Political Studies in Grenoble and then integrated to the Centre of Computerization of Socio-Political Data (CIDSP). It has then moved to Paris at SciencePo and it is called Center for Socio-Political Data (CDSP, more information is available here: http://cdsp.sciences-po.fr/).

[4] LASMAS was created in Paris and then transformed in the *Centre Maurice-Halbwachs* (CMH, more information is available here: http://www.cmh.ens.fr/greco/adisp.php). One of its missions was to provide access to data produced by official statistical offices (INSEE, ministries, governmental agencies and administrations).

[5]The creation of the BDSP was encouraged by Frédéric Bon. LASMAS was created thanks to the initiative of Alain Degenne and the engineers who had previously worked with Jacqueline Frisch at the Centre for Sociological Studies directed at that time by Raymond Boudon.

[6] Law No. 78-17 of 6 January 1978 relating to computers, files and freedom.

[7]See Silberman (2011) for a detailed account of the French regulations about access to personal and confidential data.

[8] The *Réseau Quetelet* was provided for in the article no. 10 of the Decree no. 2001-139 of 12 February 2001. It implements the policy of the Committee for the Consultation about Social and Human Sciences Data (CCDSHS), together with the Academic Data Platforms (PUD). This Committee is in charge of laying down the national public policy about social and human science data.

[9] Today also the Secure Remote Access to Confidential Data Centre (CASD) is part of the *Réseau Quetelet*.

can be granted or not are taken based on a description of applicants' research project (for more information on data access conditions, visit the website). These rules are in line with the 1978 Data Protection Act, which was reformed in 2004, following the European Parliament and Council Directive 95/46/EC of 24 October 1995 - effective in the member states since 1998 - on the protection of individuals with regard to the processing of personal data and on the free movement of such data. Although there was no obligation in the directive to regulate data access for research purposes, French researchers obtained that the law also included this possibility (Silberman 2011).

4 Providing access to surveys at INED

4.1. The organizations of specific activities

Since its creation in 1945, INED has conducted surveys dealing with socio-economic and demographic topics such as family, employment, fertility, sexuality, migrations, health and life conditions, etc. By the 1970s, the Institute has focused on its obligation to archive data, in collaboration with the National Archives (*Comité d'archivage de l'Ined* January 2001). First reflexions as to whether and how INED should provide access to its surveys started during the first half of the 1990s. In 1995, Michel Bozon, then head of the SES, was commissioned a report to be discussed at the Scientific Council, on how to regulate the access to INED surveys. Bozon (1995) consulted various INED researchers. Access to INED surveys had so far been based on an implicit consensus and any potential problem had been solved in an informal way. Surveys were becoming more costly and some were co-funded by external agencies. Requests to access the data for secondary analyses were increasing. Hence, the need that secondary analysts, including researchers from outside INED, could access INED surveys and re-use them while having a complete documentation (Bozon 1995).

Clear intentions to provide access to INED surveys were expressed in the strategic counseling document for the years 2002-2005 (*Orientations stratégiques de l'Ined* 2002-2005). They were confirmed through the participation as co-founder partner to the establishment of the *Réseau Quetelet*. A charter setting out the principles of providing access to INED surveys had to be discussed and drafted by INED Scientific Council (*Orientations stratégiques de l'Ined* 2002-2005). It was necessary to preserve priority access for INED researchers and partners participating in the funding and design of the surveys. The Institute would provide access only to data produced by INED in whole or in partnership with other institutions. Specifically, the SES[10] was given the task to enforce the institute's policy about providing access to its surveys. The implementation of this policy started to be effective in 2004 (*Archives de l'équipe mise à disposition de l'Ined* 2004). In 2006, about 30 surveys were available for access by the research community (*Orientations stratégiques de l'Ined* 2006-2009). Cécile Lefèvre, SES's head at that time, pointed out that the activities aimed at providing access to surveys had become urgent (*Comité d'archivage de l'Ined* 2006). Providing access to surveys implied meticulous documentation and re-formatting of data files, activities that went beyond those concerned by data archiving.

A new policy was aimed at promoting access to INED surveys for the whole scientific community (*Comité d'archivage de l'Ined* 2009). In 2008, some surveys began to be made available even before being archived. In 2009 SES was given the task of providing access to surveys produced within the European funded research project *Generations and Gender Programme* (GGP)[11]. As recommended by CESSDA and the *Réseau Quetelet*, the entire catalogue of INED surveys was organized on the

[10] INED Surveys Department has three missions: 1) it offers expert know-how at all stages of the data collection process for projects conducted by INED researchers and in partnership with other institutions (INSEE, INSERM, etc.); 2) it conducts research about innovative survey techniques and protocols; 3) it manages access to INED surveys.

[11] For information on this project, please visit http://www.ggp-i.org/.

Nesstar software[12] and documented according to the DDI standard. The rules set by DDI had been screened and tested by Arnaud Bringé (*Comité d'archivage de l'Ined* 2003).

The catalogue INED-Nesstar (*Catalogue Nesstar des enquêtes de l'Ined*), was officially launched in June 2012. It replaces the dissemination tool that was previously used and it is available at: http://nesstar.ined.fr/webview/. Access to the data can be asked through the portal of *Réseau Quetelet*. It is granted according to the principles established by the *Réseau*. In the Nesstar catalogue, surveys are divided into two categories (each organized by decades):

1) *Surveys that are accessible*, i.e., for which it is possible to order the data file(s);
2) *Surveys that are not accessible*, i.e., for which it is not possible (at this time) to order the data file(s) (either because it does not exist, or because the access to it is not given yet). These surveys are described through a fact sheet.

4.2. The work of data preparation and data documentation

In order to provide access to a survey, a meticulous work of preparation and documentation must be done. It is indeed necessary to gather all the documentation about the survey itself and its data file(s). This documentation is scattered: some documents can be stored in electronic files, some others in boxes of INED archives, some other documents consist in articles published thanks to survey data. As to the data files (if available[13]), two scenarios are possible. For the more recent surveys, files are in "current" formats, such as SAS, SPSS or Stata that don't require to be transformed. In case of old surveys, it's not rare to have only paper documents, hand-written or typed with writing machine, and data in text format for example. Documents are sometimes incomplete, variable names can be missing. There can be a lack of documents, or on the contrary, some information can be duplicated. In case of anonymous files, additional work is required to detect potentially identifying variables. These variables are removed from the files made accessible.

Once all the data files and related documentation are ready, they are uploaded in Nesstar. Metadata are documented according to the DDI fields chosen by the SES and classified in 3 groups:

1) Document description: information on the Nesstar file (which survey it concerns, the person who compiled it and when);
2) Study description: information on the survey (abstract, keywords, researchers involved, producers, financiers, dates of the data collection, data collection methodology, sampling procedure, weighting, description of questionnaire, response rates), links to the questionnaire(s), interviewers' instructions, associated surveys, and the related bibliography[14];
3) Data files description: information about the dataset(s), such as how they are constructed, how the missing data are coded, which variables are replaced with derived ones, etc.

Additionally, each variable is examined in great detail. As to its label and modalities, the possibility to import already documented databases represents significant saving of time. Additional information regards e.g., the question tests (text before the question, literal question, text after the question, instructions to interviewers), the universe (i.e., the persons to whom the question was asked), the questionnaire from which the variable comes from (when the survey involves several different questionnaires) and, in the case of derivate variables, the variables and the software programs used in the calculation. To provide all this documentation, each dataset has to be entirely reviewed, and variables are reorganized so as to follow questionnaires' sections.

---

[12]License for the Nesstar software is acquired by the *Réseau Quetelet* for all its members.

[13]In the past, data were recorded on perforated cards. They were then transferred on computer files, but these operations led to losses of information. As a consequence, certain surveys will never see their data accessible.

[14]The online bibliography is a part of the survey's metadata. It consists of a bibliographic reference list (INEDOC), updated in connection with INED's Documentation department.

This work of documentation requires a great amount of time that can be reduced by the research teams who conducted the surveys. Exhaustive, clean and anonymous data files, good codebook and extensive documentation on the study methodology can accelerate the work required to provide access to data. A greater collaboration from survey producers would optimize the whole procedure

5. Conclusions

This paper reviewed the development of the activities to provide access to quantitative social science data at INED. It did so after a brief account of the development of social science data archives at the international level and in France. The need to organise the access to survey data started in the 1950s within the field of political science. The willingness to develop European and international comparative research encouraged the establishment of national data archives both in the USA and Europe. International networks fostered the creation of harmonised standards for documenting metadata. In France, due to the lack of a strong tradition in quantitative sociology (Silberman 1999), the establishment of a national data archival structure is more recent. The *Réseau Quetelet* was created in 2001 as a structured organisation of until then scattered data archives. In 2004, the reform of the Data Protection Act regulated the use of personal and confidential data for research purposes. The development of the activities to provide access to surveys at INED follows the same path. Despite first reflexions in the 1990s, it was not before 2001 that the Institute began these activities. The SES, co-founder of *Réseau Quetelet*, today offers its expertise in these activities to international projects, such as the GGP.

Because survey metadata are often not structured to the DDI standard, the work of data preparation often requires a lot of time. To reduce the time needed for data documentation and to increase its quality, it is necessary a greater awareness of DDI standards and metadata should be planned ahead. The trend toward open data and data deluge in social science (Silberman 2013) permitted by IT progresses is increasing the need for providing rapid and high quality access to social science data. Data access has become a transnational issue and actions are being taken at the EU level to integrate and strengthen research infrastructures which provide access to data. With CESSDA being identified as an ERIC candidate (Marker 2013), greater and greater resources will be devoted to providing access to social science data.

References

[1] Arduin, P. (2004), Présentation du Centre Quetelet : description des missions, activités, partenaires et projets de mise à disposition des données, Presentation to the *Journée d'information sur l'archivage historique et la mise à disposition des données d'enquêtes en SHS*, Institut national d'études démographiques (INED), January 29, 2004.
[2] Archives de l'équipe mise à disposition (2004), Internal documents of INED Surveys Department.
[3] Balladur, E. (1994), Circulaire du 14 février 1994 relative à la diffusion des données publiques, *JORF n°42 du 19 février 1994*.
[4] Bisco, R. L. (1966), Social Science Data Archives: A Review of Developments, *The American Political Science Review*, 60 (1), 93−109.
[5] Comité d'archivage de l'Ined (2001-2009), Comptes-rendus et relevés de décisions, INED internal documents.
[6] Corti, L. (2000), Progress and Problems of Preserving and Providing Access to Qualitative Data for Social Research−The International Picture of an Emerging Culture, *Forum: Qualitative Social Research*, 1(3), Art. 2, http://www.qualitative-research.net/index.php/fqs/article/view/1019/2197.
[7] Bozon, M. (1995), *L'accès aux enquêtes de l'INED. Principes généraux, règles pratiques, problèmes en suspens*, INED Internal report.

[8] Doorn, P. and Tjalsma, H. (2007), Introduction: Archiving research data, *Archival science*, 7, 1−20.
[9] Gregory, A. and Heus, P. (2007), DDI and SDMX: Complementary, not Competing, Standards, *Open Data Foundation Papers*, http://www.opendatafoundation.org/papers/DDI_and_SDMX.pdf.
[10] Hansen, S.H., Iverson, J., Jensen, U., Orten H. and Vompras J. (2011), Enabling Longitudinal Data Comparison Using DDI, *DDI Working Paper Series – Longitudinal Best Practice*, No. 2, http://www.ddialliance.org/resources/publications/working/BestPractices/LongitudinalData.
[11] Hastings, P.K. (1964), Le Roper Public Opinion Research Center, centre international d'archive de données d'enquêtes par sondage, *Revue internationale de sciences sociales*, XVI, 1, 102−108.
[12] Inter-university Consortium for Political and Social Research (ICPSR) (2012), *Guide to Social Science Data Preparation and Archiving: Best Practice Throughout the Data Life Cycle* (5th ed.), Ann Arbor, MI, http://www.icpsr.umich.edu/files/ICPSR/access/dataprep.pdf.
[13] King, G. (1995), Replication, Replication. *PS: Political Science & Politics*, 28, 3, 443−499.
[14] Kramer, S., Banks, R., Chang, V., Sieber I., Vardigan M. and Zenk-Möltgen, W. (2011), Presenting Longitudinal Studies to end Users effectively Using DDI Metadata. *DDI Working Paper Series – Longitudinal Best Practice*, No. 4, http://www.ddialliance.org/resources/publications/working/BestPractices/LongitudinalData.
[15] Leighton, H.V. (2002), Developing a new Data Archive in a Time of Maturing Standards, *IASSIST Quarterly*, 26, 5−9, http://www.iassistdata.org/downloads/iqvol261leighton.pdf.
[16] Marker, H.J. (2013), "Strengthening cooperation between European social science data archives: The evolving role of CESSDA", in Kleiner B., Renschler I., Wernli B., Farago P., Joye D. (eds), *Understanding Research Infrastructures in the Social Sciences*, Zurich: Seismo Press, 39−46.
[17] Mochmann E. and Vardigan M. (2011), "L'analyse secondaire de données et la documentation pour leur diffusion", in Chenu A., Lesnard L. (eds.), *La France dans les comparaisons internationales*. *Guide d'access aux grandes enquêtes statistiques en sciences sociales,* Paris, Presses de Science Po, 173−181.
[18] Orientations stratégiques de l'Institut national d'études démographiques (2002-2005) and (2006-2009), INED internal reports.
[19] Rokkan, S. and Scheuch, E. K. (1963), Conference on Data Archives in the Social Sciences, *Social Science Information*, 2, 4, 109.
[20] Rokkan, S. (1966), *Data Archives for the Social Sciences*, Mouton Co. and Ecole Pratique des Hautes Etudes, Paris – The Haugue.
[21] Scheuch, E. K. (2004), Les services de données en sciences sociales : historique et perspectives. *Revue internationale des sciences sociales*, 3, 177, 433−449.
[22] Silberman, R. (1999), Les sciences sociales et leurs données, Report for the Ministère de l'éducation nationale, de la recherche et de la technologie, Juin 1999, http://www.ladocumentationfrancaise.fr/var/storage/rapports-publics/004000935/0000.pdf.
[23] Silberman, R. (2011), La protection des données individuelles en France et la recherche en Sciences Sociales, in Chenu A., Lesnard L. (eds.), *La France dans les comparaisons internationales*. *Guide d'access aux grandes enquêtes statistiques en sciences sociales,* Paris, Presses de Science Po, 183−204.
[24] Silberman, R. (2013), Transnational access to official micro-data: The *Data without Boundaries* European network, in Kleiner B., Renschler I., Wernli B., Farago P., Joye D. (eds), *Understanding Research Infrastructures in the Social Sciences*, Zurich: Seismo Press, 47−66.
[25] Vardigan, M., Heus, P. and Thomas, W. (2008), Data Documentation Initiative: Toward a Standard for the Social Sciences, *The International Journal of Digital Curation,* 3, 1, 107−113, http://www.ijdc.net/index.php/ijdc/article/viewFile/66/45.