FQS FORUM: QUALITATIVE SOCIAL RESEARCH SOZIALFORSCHUNG

Volume 12, No. 3, Art. 18
September 2011

Archiving Qualitative Data: Infrastructure, Acquisition, Documentation, Distribution. Experiences from WISDOM, the Austrian Data Archive

Andrea Smioski

Key words: qualitative data, archiving, data documentation, data acquisition, secondary analysis, establishment of an archive for qualitative data

Abstract: The debate about archiving and reusing not only quantitative data but also the rich resources generated through qualitative enquiry has reached a broader audience since the UK's Economic and Social Research Council (ESRC) first launched its qualitative service Qualidata in 1994. Almost seventeen years later, a look at the European situation reveals that very few countries have been able to accomplish the same culture of sharing qualitative research data. Nevertheless, the archiving and sharing of qualitative research data is gaining momentum. An increasing number of countries—including Austria—launched qualitative archives, which to date are at various stages of development.

This paper aims at exploring some of the most essential requirements for successfully establishing a national qualitative archive. Technical issues, tools and standards necessary for proper data documentation, archiving and distribution of high quality datasets will be explained, as well as the underlying legal conditions. I will illustrate the separate steps of the archiving process. Data acquisition, data processing and documentation, and the promotion of secondary usage will be discussed against the background of structural conditions, such as archiving policies by research councils, favoring or impeding the development of a culture of data sharing.

Table of Contents

1. Archiving Qualitative Data—State of the Art

Archiving and secondary analysis of quantitative data are both firmly established in social science research culture. It is hard to imagine that the reanalysis of major national or international survey data, as for example the ISSP (International Social Survey Programme)[1], the Eurobarometer[2] or the European Social Survey[3] would be called into question. To a certain degree, such programmes are even created with their secondary usage in mind. They allow researchers from different backgrounds to explore various kinds of subjects and research questions, and thereby to gain new knowledge, fuel a scientific dialogue or provide relevant guidelines for social policy. [1]

The idea of sharing research data has something genuinely egalitarian to it. Research outputs should be openly available for the scientific community, more so if the research has been financed with public funds (BERLIN DECLARATION). For some researchers, it would never be possible to create datasets large enough to produce results of any relevance to the scientific community. This might have different reasons. For one, data generation is resource intensive, and due to the difficult economic situation a lot of research institutions and funding organizations had to cut back on their research grants. Openly available datasets would offer students and researchers at the start of their career the possibility to draw on existing resources and data collections, instead of creating very small data collections for their own theses. Another reason might be that it can be difficult and time-consuming to gain access to some populations, for instance if the data collected is sensitive or if the populations are exclusive, or even facing extinction. An example for such a case is the data collected for the recently archived witnesses-project *Lebensspuren. Erlebte Zeitgeschichte im Land der 1000 Hügel* (DRESSEL & HAGENHOFER, 2007). This study conducted oral history interviews with elderly men and women from a rural area in Austria. The interviewees talked about their experiences in war- and post-war time. Some of the interviewees had been 90 years of age and older at the time of the interview, and their memories would have been lost if the interviews had not been conducted and then archived at WISDOM, the Wiener Institute for Social Science Data Documentation and Methods. [2]

Besides enabling all researchers to make use of already available material, there are other, more substantial arguments for data sharing. To "stand on the shoulders of giants," as MERTON (1965) put it, has the advantage of seeing farther. Even though analyzing someone else's data was not really what MERTON referred to, the statement nevertheless fits. In many cases, more profound insights can be gained if research material is further analyzed. This

1 The ISSP is a continuing annual programme of cross-national collaboration on surveys covering topics important for social science research.

2 Since 1973, the European Commission has been monitoring the evolution of public opinion in the Member States.

3 The ESS is designed to chart and explain the interaction between Europe's changing institutions and the attitudes, beliefs and behavior patterns of its diverse populations. Now preparing for its fifth round, the survey covers more than 30 nations.

offers the chance of multiple perspectives on the data, even enabling interdisciplinary exchange. The possibility of using existing resources as basis for new data generation facilitates cumulative and comparative studies. Thus duplication can be prevented, and in the long term research will be more effectively based on previous studies. Transparency of what has been done in primary investigation can improve methodical approaches and thereby the quality of research data. [3]

The benefits of sharing primary research data are obvious, as shown above. Moreover, all the arguments listed, and they are only the most prominent ones, are valid for quantitative as well as qualitative data. Nevertheless, the idea of archiving and sharing qualitative research data was not put into practice formally until the mid 1990s. Before that, most qualitative research data remained in the hands of individual researchers or sitting in little-known local archives, and was therefore lost to empirically based comparison, extensive use and as base for further research. [4]

Over the past two decades qualitative data archiving and sharing has become more established in social science research culture. From a European perspective, the launch of the Qualidata archive[4] in 1994, the first national archive for qualitative resources in the UK, was the date of founding of qualitative data archiving. Since then, increasingly heated discussions have taken place. Special issues of several journals addressed topics such as data archiving and the secondary use of qualitative research material, for example the *FQS* issues "Text. Archive. Re-Analysis" (CORTI, KLUGE, MRUCK & OPITZ, 2000), "Qualitative Inquiry: Research, Archiving, and Reuse" (BERGMAN & EBERLE, 2005) and "Secondary Analysis of Qualitative Data" (CORTI, WITZEL & BISHOP, 2005); the *Methodological Innovations Online* issue "Making Qualitative Data More Re-Usable: Issues of Context and Representation" (CORTI, 2006); or the *Sociological Research Online* issue "Reusing Qualitative Data" (BARBOUR & ELEY, 2007). [5]

In the UK and elsewhere, various models for facilitating qualitative data sharing have been developed. To date, there exist small scale qualitative archives in most European countries. However, these developments are occurring in piecemeal fashion with little provision for sharing skills, good practice or for enhancing collaboration between researchers and archivists across the European community. Data resources, standards and tools for processing data as well as infrastructure and funding situations are diverse. In April 2009, an international workshop in Bremen, Germany, brought together researchers and archivists with an interest in the reuse of qualitative data, and in the development of qualitative

4 ESDS Qualidata is a specialist service of the Economic and Social Data Service led by the UK Data Archive (UKDA) at the University of Essex. The service provides access and support for social science qualitative datasets, promoting and facilitating increased and more effective use of data in research, learning and teaching.

research and data resources across Europe.[5] [6] The main output of the workshop was a report which maps out existing resources and infrastructure across Europe. Plans for a European network of qualitative data collections, researchers and projects, all co-ordinated through databases and mailing-lists, were developed and are still pursued. A major task for the next stage of development will be the implementation of common standards and tools to facilitate the shareability of the data, not only in a national but in an international context as well. [6]

The Council of European Social Science Data Archives (CESSDA), in its function as an important umbrella organization, facilitates the development of social science data infrastructure in Europe; although heretofore, much of its work has been oriented toward quantitative data and resources. Part of the challenge of broadening the scope of CESSDA is that the ethos of sharing qualitative data has been slow to develop. This indicates one aspect of a general problem, namely the reluctance of researchers to support qualitative archives. This hesitancy is rooted mainly in ethical and legal concerns due to the sensitive content of qualitative material. The other aspect becomes visible in the fact that many funding institutions do not provide the financial means for the collation, organization and reuse of data. In addition, they hesitate to introduce archiving policies which oblige researchers to deposit their data in a repository. [7]

Nonetheless, the wheels are turning, and gradually things are changing. There is a growing awareness of the necessity of preserving qualitative resources[7] as well as an increasing interest in using existing data for research and teaching. Step by step, the commitment of funding bodies is strengthening proactive initiatives to build up repositories and archives for social science data. [8]

Using the Austrian qualitative data archive WISDOM as a model, the following points outline some of the requirements necessary for acquiring, documenting, processing and distributing qualitative social science research data. [9]

2. WISDOM

Documentation and archiving of social science research data in Austria started in the 1980s. WISDOM was founded in 1985. The Ministry of Science and Research supported these efforts. WISDOM started to acquire, process and distribute social science survey data, in close cooperation with the University of Vienna and Austrian commercial institutes for market and social research, such as the

5 The Bremen Workshop: Qualitative and Qualitative Longitudinal Data Resources in Europe: Mapping the Field and Exploring Strategies for Development. The workshop was organized as a collaborative venture between the UK Data Archive, the Timescapes Qualitative Longitudinal Study and Archive, and the Archive for Life Course Research (ALLF) in Bremen, with support from the Council for European Social Science Data Archives (CESSDA).

6 IASSIST, the International Association for Social Science Information Services and Technology, supports similar initiatives: In 2000, a first network of qualitative archives was brought together; a follow-up meeting of the Bremen participants was organized in October 2010.

7 Although it is a general assumption that researchers at the start of their career are more open to the idea of sharing their data, it is in fact researchers who plan their retirement in the near future who are the most important group of data depositors. The potential loss of data is more obvious to them than to those who are only starting to collect their own datasets.

Institute for Empirical Social Studies (IFES) and the Fessel Institute for Market Research. In the 1980s, it was relatively easy to acquire data. Bureaucratic and administrative barriers were low and data protection and data abuse were not the sensitive subjects they are today. Since then, WISDOM has acquired 868 quantitative datasets; more than 480 are processed and documented. They are easy to access in an online catalog and available for secondary users in digital format. In addition to smaller studies, WISDOM holds major national and international public opinion surveys such as the Austrian Social Survey and the Microcensus, which are repeated at regular intervals. Other data-sets, such as the ISSP and the Eurobarometer, are also available at WISDOM. In the early 1990s, WISDOM joined CESSDA, the Council of European Social Science Data Archives and thereby became the acknowledged national social science data repository in Austria. [10]

Finally in 2007, WISDOM began to extend the scope of acquired data to also include qualitative datasets and datasets with mixed method designs. This required different expertise and standards. Tools for data documentation and data management had to be adapted or newly designed. Data protection as well as confidentiality issues had to be addressed in response to researcher's worries and fears about qualitative data archiving. In 2008 and 2009, the qualitative archive was financed only by one-year third-party project funds. A feasibility study was conducted, followed by a phase of extensive data acquisition. In 2010, the Austrian Ministry for Science and Research signed the commitment to support the quantitative and qualitative archive on a long-term basis.[8] This provides the archive with the necessary means to build up and promote an infrastructure that guarantees the long-term preservation of research data. Initiatives, for example the negotiation with funding agencies for data policies, the development of tools for qualitative data documentation or the in-depth work on specific topics such as research ethics, can now be planned over a longer period of time and are thus more sustainable than solely result-focused, short-term activities. Consequently, a culture of data sharing in Austria can begin to take root. [11]

3. Requirements for Qualitative Data Archiving

In the case of WISDOM, the foundation of the qualitative archive was facilitated since the necessary infrastructure already existed. Tools for data documentation and processing, software for the display of data and a legal framework for the conditions of data access and usage were already in place and had only to be adapted to meet the requirements of qualitative data. Notwithstanding the above, I will elaborate on the requirements necessary to set up a qualitative archive. [12]

8 This was done within the framework of the CESSDA project (CESSDA ERIC): In 2006, the European Strategy Forum on Research Infrastructures (ESFRI) developed a roadmap for the establishment of new pan-European research infrastructures (ESFRI, 2006). From a total of 35 projects, there are six projects from the social sciences and humanities. CESSDA, the Council of European Social Science Data Archives, is one of them. Besides a small umbrella organization, there are several national centers across Europe, which have to be supported by national research councils. Since 2010 the Ministry of Science and Research and the Ministry of Labor, Social Affairs and Consumer Protection support the long-term establishment of a national social science data archive for both quantitative and qualitative data, which is part of the CESSDA consortium.

3.1 The archive and the archivist

Following the intention to establish a qualitative data archive, an appropriate local facility must be found as well as a researcher or an archivist who will be responsible for the concerns of the archive. This person must be capable both of preparing and documenting the data for preservation as well as of providing full methodological knowledge and support. It is advisable to choose someone with a social science background who has already done empirical qualitative research and who has therefore a profound understanding of the data that is to be dealt with. Depending on the specific national situation, for instance varying degrees of support from funding agencies or diverse infrastructural preconditions, this phase is more or less resource intensive. Thus, depending on the situation, one half-time position, better yet, a full-time position or more, should be provided to effectively run the archive. [13]

3.2 Data formats and space

How to deal with the data is the next concern. Data generated through qualitative research can be textual (e.g. interview transcripts, written diaries, documents), audio, visual or audio-visual (e.g. recordings, video tapes, photographs). Since in most cases not only recent data collections are of interest, the archive has to deal with digital as well as non-digital data, which needs to be stored in a secure location. When the decision to archive is made, it should be considered that there might be large amounts of data that have to be stored. The same holds true for digital data. It might be easy to archive digital interview transcripts and to have several back-up copies of the data in different locations. Archiving digital video material is a completely different matter however, since it requires a lot of digital storage space. Thus, when starting an archive, the question should be which kind of data shall be archived and what data formats can be handled reasonably. Experiences in Austria showed that anthropologists and visual sociologists have a great interest in archiving video material. However, the necessary infrastructure to manage the amount of data and—if the material is non-digital—an easy way to make it accessible should be kept in mind before accepting the data. Cooperation with specialized repositories and archives might help to find solutions for this situation so that the offered data does not have to be rejected. [14]

Since a lot of older and non-digital formats run the risk of becoming unusable over time[9], WISDOM is working on the digitization of archived non-digital resources. This endeavor requires a variety of technical equipment which sometimes can be quite expensive. Again, cooperation with other repositories can help a lot. Besides the technical requirements, digitization is very time-consuming and therefore not something that should be promoted right from the beginning—unless there are enough resources available. Scanning hundreds of pages of interview transcripts, putting them through an OCR process, listening to and editing audio files is a workload not to be underestimated. Sometimes, data

9 The life expectancy of an audio cassette is, for instance, about 10 to 15 years.

depositors will take over some of this work, but most often this has to be done by the archive. [15]

3.3 Data protection and research ethics

When WISDOM initiated its qualitative archive in 2007, the first endeavor was to conduct a feasibility study (SMIOSKI, MÜLLER, CASADO ASENSIO & KRITZINGER, 2008). One of the results was that legal and ethical questions regarding qualitative data archiving had to be addressed by the archive in order to overcome researcher's concerns.[10] [16]

Copyright is a central issue for archiving research materials because it confers privileges of ownership and control of the dataset (PARRY & MAUTHNER, 2004, p.141). In qualitative research two separate sets of copyright are involved. The first one is the copyright of the spoken word which is owned by research participants, and the second one is the copyright of the recording owned by the research institution which conducted, or the funding agency which commissioned the study. Copyright owners can assign their copyright elsewhere or license others to use their work. Therefore, archives urge researchers to gain written assignment of copyright from their participants. Furthermore, WISDOM developed a license agreement for data depositors which clearly states that the copyright remains with the data depositor. [17]

Independent of copyright issues, research participants always retain the right that their personal information remains confidential. In Austria, this is regulated by the Data Protection Act 2000 (§46), which states that research data may only be processed for purposes other than those for which they were originally obtained, if any personal information is removed or research participants gave their agreement to it being used further. Archives recommend that researchers include archiving and scientific reuse as an option in confidentiality agreements and copyright declarations, which are used in the project. Nevertheless, it is often difficult to protect the interests of research participants once the data has been made available to secondary users. This is especially the case with qualitative data which is rich in detail and often focused on small or vulnerable populations. WISDOM as well as other archives[11] additionally anonymizes sensitive data, carefully restricts its use and obliges secondary users to sign end license agreements, which ensure contractually that users will respect the rights of participants.[12] This further protects research participants from any harm that could result from the use of the data. [18]

For users, there are different levels of access to data. These access conditions are defined in consultation with data depositors and encompass open access for research and teaching, access after consultation of the data depositor, access

10 See BODDY (2001) for similar experiences in the UK.

11 See CORTI, DAY and BACKHOUSE (2000) for the measures taken at ESDS Qualidata in the UK.

12 All WISDOM forms mentioned in this paper can be downloaded from the WISDOM homepage: http://www.wisdom.at/Datendokumentation/Da_Datendoku.aspx.

after a period of closure and closed access (preservation only). In addition to these access restrictions, users must always register officially with the archive and submit a request in which they state the intended use of the data. Primary researchers have to protect the confidentiality of research subjects. Therefore, it is necessary for an archive to tie the usage of archived datasets to clear and standardized formal procedures with legal obligation. [19]

The lively discussion about ethics and politics of qualitative data archiving and reusing both in literature (PARRY & MAUTHNER, 2004; HEATON, 2004; RICHARDSON & GODFREY, 2003; CORTI, DAY & BACKHOUSE, 2000; KLUGE & OPITZ, 1999) as well as amongst a broader scientific community[13] illustrates that these issues still have not been addressed satisfactorily. However, in light of the current trend towards archiving, it is necessary that national repositories and archives, funding agencies and disciplinary bodies debate the practical, legal and ethical challenges and limitations of archiving and reusing qualitative data more profoundly. [20]

3.4 Data documentation

Data documentation is essential for enabling sound secondary analyses of archived studies. It establishes the bridge between primary researcher and secondary user, providing the latter with 1. the meta-information necessary to understand the content and structure of the dataset, and 2. the context information necessary for the user to comprehend the institutional, theoretical and methodical background of the study, to trace the process of data generation and processing throughout the research process and to learn about primary researcher's reflections on the research project. While the first task is mainly administrative (and focused on describing data with concise bibliographic information as well as technical issues, such as file naming, labeling or relating files to each other), the second task is centered on the documentation of the research process. Primary researchers should, for example, describe the conditions under which a study was funded and the resulting influences on the research questions. They should provide details about the specifics of the data collection situation and the interaction with the interviewees. In addition, a review of the data collection methods or sources of error should be laid open. Furthermore, valuable additional information on the research process should be provided, for instance the conditions under which the interview was conducted. Often, this information is implicit knowledge which researchers are not used to making explicit. However, this information may be indispensable for further analyses of the data, depending on the methodological approach of secondary users. [21]

A common project of the CESSDA archives is the development of a shared data portal. Users will be able to search and download data from all the member archives via one online catalog. The archives need a common documentation

13 Since 2008, WISDOM has organized a series of workshops and panel discussions in which the topics of data protection and research ethics have been debated with great vehemence.

model or standard, which allows capturing meta-data (data about data), for such a project to be feasible. [22]

The Data Documentation Initiative (DDI) is an effort to create such an international standard for describing social science data. Expressed in the open source language XML, the DDI meta-data specification offers a comprehensive set of elements to describe social science datasets as completely and as thoroughly as possible. Thereby it provides the potential data analyst with broader knowledge about a given collection. This is necessary for an effective, efficient, and accurate use of datasets, more so if the data is used many years after its collection and primary researchers can no longer provide information. Relevant elements are for instance, title, ID number, authoring entity, copyright, depositor, deposit date, bibliographic citation, keywords, topic classification, abstract, time period covered, date of collection, and the like. [23]

The DDI meta-data model was planned and is mainly used for documenting quantitative data, but it can be used and elaborated for the special needs of qualitative data as well. For in-depth discussion on the use of DDI for qualitative data refer to KUULA (2000) as well as CORTI and GREGORY (2011), who describe how the elements of DDI can be adapted to the needs of qualitative data description. Moreover, the DDI Alliance set up a special working group, which is engaged in the amplification of existing DDI elements to better deal with the complex information necessary to understand the content and procedure of qualitative studies. The working group also develops a standardized vocabulary for the description of qualitative data, which guarantees the coherence and thereby an easy attainability of recorded information across the archives. [24]

Another helpful tool that advances international cooperation is the ELSST (European Language Social Science Thesaurus), a multilingual thesaurus for the social sciences, which has been developed over the years by the members of CESSDA. Its main aim is to facilitate access to data resources across Europe, independent of domain, resource, language and vocabulary.[14] Additionally to the thesaurus, CESSDA developed a multilingual thematic classification system which allows allocating a data collection to general categories as well as specific sub-categories. This simplifies the search insofar as users can browse topics of interest directly. [25]

3.5 Displaying the data

The main purpose of an archive is making research data more easily accessible and thus heightening the output of scientific research. One essential prerequisite to reach this goal is the visibility of information on datasets on the web, as created with the DDI standard. A data catalog offers the possibility of electronic search and display of meta-data. The Nesstar (Networked Social Science Tools and Resources) software, developed as a CESSDA project by the University of

14 See http://www.cessda.org/accessing/ [Date of Access: September 1, 2010].

Essex and the Norwegian Social Science Data Service (NSD), is a software system for data publishing and online analysis. [26]

The software consists of tools which enable data providers to disseminate their meta-data and data on the web. Nesstar handles survey data and multidimensional tables but also text resources. Users can search for, browse and analyze the data online. Nesstar is used by most of the CESSDA data archives to publish meta-data and to allow users to browse the data catalog, thus increasing the accessibility of data resources. [27]

3.6 Structural prerequisites of data acquisition

To this point, I have dealt mostly with physical requirements: personnel, location, and technical equipment. What is most important for successfully establishing a social science data archive for qualitative (as well as quantitative) data is, however, the existence of structures which favor the preservation of research data. In the UK, the Economic and Social Research Council requires all award-holders to offer their data for archiving. Funds for the preparation of the data for archiving are available so the data can be prepared to a standard which may be used by a third party (ESRC Data Policy). The final payment of an award is withheld until the data has been deposited. The ESRC also collaborates with other national and international agencies in funding and promoting access to datasets. [28]

The level of support for data sharing is unevenly distributed among countries (LAAKSONEN, BORG & STEBE, 2006, p.5). In Austria, as well as in many other European countries, there is no national policy for data archiving and sharing. However, the Austrian Ministry of Science and Research enables Austrian participation in CESSDA and will probably develop an archiving policy in the near future. The Austrian Science Fund (FWF) launched an archiving policy in January 2010, obliging funded project leaders to make their research data accessible. However, it does not collaborate with the national data repository for social science research data WISDOM. In addition, the Anniversary Fund for the Promotion of Scientific Research and Teaching of the Austrian National Bank encourages archiving of empirically generated research data in its policies. Though, since it does neither provide the means for data preparation nor enforce the deposition of data in any way, these recommendations are of limited efficiency. [29]

Nevertheless, in a piecemeal fashion the research councils are beginning to recognize the importance of preserving research data for future use. The archiving policy of the Austrian Science Fund marked this change, as well as the long-term commitment of the Ministry of Science and Research which will sustain the quantitative and qualitative archive at WISDOM for the next two years at least. Slowly, but surely, a culture of archiving and sharing research data for secondary use is developing. Top-down strategies, as induced through archiving policies by the research councils, are an important lever to speed up this process.

Still, the participation of the research community is necessary to establish a culture of data archiving. [30]

To involve the research community, WISDOM arranged workshops all over Austria to inform researchers about the services available at the archive, to promote archiving and secondary use and to discuss issues that are still under debate such as the handling of sensitive data or access conditions for qualitative data. WISDOM, furthermore, organized a panel discussion, bringing together funders, representatives of university departments, archivists and researchers to widen the debate on issues of data archiving and sharing in the social sciences. The feedback from these events was thoroughly positive and helped to stimulate a debate within the scientific community which is still carried on. Additionally, a number of contacts and later collaborations with researchers and institutes were established in the course of these events which fostered the acquisition of datasets. Therefore, it is advisable to address the scientific community directly when establishing a qualitative data archive. Workshops and panel discussions can be one alternative. A regular newsletter and well prepared user guides, giving information on sensitive issues in relation to data archiving supplement these events. Moreover, the new WISDOM homepage, which will be launched in 2011, will include an online-forum where debates can be carried out and thus reach a broader audience. [31]

3.7 Data acquisition

When it comes to actually deciding if a dataset can be archived, it is advisable to provide researchers with information about the requirements their data has to fulfill. Some archives confine their data collections to certain thematic fields[15]. Others set disciplinary limits or restrict their collections to certain types of data, for instance interview data or audio-visual data. WISDOM accepts qualitative data collections from social science disciplines without making restrictions to certain thematic areas or types of data; however, we have to cooperate with other repositories to deal with large amounts of video material. More important for the decision to archive are the following criteria which pertain to content and are applied to data offered for deposition. These are: 1. that the data has to be primary data, 2. that adequate description of methodological and methodical procedures has to be made available, 3. that documentation has to provide transparency of methodical and theoretical foundations of the project, 4. and that comprehensive documentation of the data itself (e.g. information on data collection situation, sample design, documentation of key decisions, etc.) is necessary. [32]

Since policies for data archiving are not yet taking effect, it turned out to be difficult to obtain data from recent projects. WISDOM puts a lot of effort into negotiations with funding agencies and generally advocates the idea of a top-down approach to data archiving via contractual guidelines. Nevertheless, we still try to acquire data from researchers themselves and support unsolicited data

15 E.g. the Archive for Life Course Research in Bremen, Germany, which focuses on transitions and status passages in the life course.

deposition. One effective strategy proved to be a project adopted from *ESDS Qualidata*, which is named Pioneers of Qualitative Research. With slight deviation from the British project, which actually seeks out surviving research data from pioneering examples of social research, WISDOM—with input from the research community—identifies prominent Austrian researchers renowned in their subject area. Data collections of these researchers, which are on average 20 to 30 years old, are deposited at the archive. Deposition often includes extensive digitization of valuable material which would otherwise be lost. As these researchers mostly studied in the 1970s and 1980s, they by now possess considerable amounts of data and are often more willing to share their data, and thereby have it processed and managed for the future, than researchers who still partake in more result-oriented contract research and are therefore exposed to scientific competition. In addition to the archived studies, WISDOM conducts interviews with the pioneering researchers. These are qualitative interviews with a life history approach, centering on the research biography of the pioneers. [33]

The Austrian pioneers-project pursues two aims: 1. to generate a pool of high-quality data collections available for secondary research; and 2. to stimulate the reuse of data, since these studies were carried out by well-known scientists and are thus of special interest to the research community. However, WISDOM also archives datasets from independent researchers and promotes these studies via leaflets and a periodical newsletter. [34]

3.8 Data processing

When progressing to processing acquired datasets, first experiences revealed that it is not common practice to fully methodologically and theoretically document a research project so that it can be properly archived and easily comprehended by other researchers (SMIOSKI, MÜLLER & BISCHOF, 2009, p.28). Often, there is too little information about concrete research methods, sample design, or theoretical considerations and their consequences. Additionally, due to dispersed project members or technical problems, not all data is available. Hardly anyone generates data with archiving or secondary usage in mind, since there are no regulations from research councils as to what should happen with research data after a project is finished. This implies that a lot of additional work remains to be done if datasets are to be archived. The archive can take over some of this work (e.g. anonymization or digitization), but it is the original researchers who need to take responsibility for managing and documenting their data comprehensively, especially in times when an archive starts to ingest a steady flow of data. [35]

To improve this situation, WISDOM offers training courses on data documentation and data management, also inviting experts from other archives as lecturers. Furthermore, we collaborate with the Faculty of Social Sciences at the University of Vienna to integrate courses on data documentation and data management into the curriculum of social science studies—both on a master and a PhD level. Moreover, WISDOM offers methodical consultancy for data depositors, advising researchers to seek counsel in the early stages of the research process, so data management plans can still be adapted to include archiving at the end of

the research project. A lot of information on data management and data documentation is available on the WISDOM homepage as well. [36]

3.9 Promoting secondary use

While data archiving is slowly starting to gain momentum, promoting the secondary use of archived resources is still a different matter. Whereas it is common practice to re-analyze and re-consult quantitative research data for different purposes than those of the original research context, this tradition is practically non-existent for qualitative research. To foster a culture of data sharing implies creating living archives that are not only storing data but also redistributing it to the research community. In the words of JOYE (2005, p.11): "For an archive, the issue is not only storage and conservation of the data themselves, but its capacity to be used rather than merely stored, in other words, the aim is publication and not only conservation." [37]

WISDOM published a number of high-quality data-sets in the catalog only recently, but it needs time to draw researcher's attention to the services of a hitherto unknown infrastructure. We promote our studies via a newsletter, but again: since we are at the beginning of the archiving project, we have so far only had a small number of requests for archived data, and in these cases the data were mainly used for teaching purposes. [38]

In the UK, the ESRC promotes not only the archiving of research data but, additionally, also the reuse of existing material. Applicants for ESRC research funds are required to demonstrate in their application that no data is already available and that there is a need for new data collection. Furthermore, the ESRC facilitates secondary analyses of datasets and encourages the citation of datasets in published materials (ESRC Data Policy) as an alternative means to acknowledge research performance. [39]

In Austria, there is still some way to go to reach this level of support on the part of the research councils. Meanwhile, WISDOM is preparing case studies of secondary analyses done with archived projects. Thereby, users can get an overview of the added value and possible benefits of secondary analysis so that they might be inspired to reuse archived data themselves. [40]

4. Conclusion

In this article, I have tried to give a comprehensive overview of the state of the art of qualitative data archiving focusing on the situation of WISDOM, the Austrian qualitative and quantitative data archive. The archiving and sharing of qualitative research data is gaining momentum. Since the establishment of Qualidata in 1994, more and more countries launched qualitative archives which to date are at various stages of development. Through joint projects and working groups, which are strongly supported by IASSIST, the International Association for Social Science Information Services and Technology, it is now possible to use common standards and tools for data documentation and data processing, providing the

user with easily accessible research data for secondary usage. The broader scientific community is taking up the discussion on open issues regarding the sharing of qualitative material and funding bodies are beginning to introduce guidelines as to what should happen with data from projects financed by their research grants. [41]

Yet, many archives are still at an early stage or even in an initial phase of development. Taking WISDOM as a good example, I have intended to give some recommendations for successfully establishing an archive for qualitative social science research data. There are numerous crucial factors which need to be considered to make data available for a wider scientific community, and they must thus be processed and documented accordingly. Meta-data standards and appropriate technical solutions have already been developed, in addition to possible solutions for ethical questions. Apart from the infrastructural context, I have also elaborated on the structural conditions; for instance, how policies of research councils and funding agencies can cultivate or constrain the establishment of a culture of data sharing significantly. WISDOM tries to encourage the willingness of researchers to deposit and reuse data via diverse activities. Nevertheless, much remains to be done. Data documentation and management strategies generally need to be improved, ethical questions must be discussed and pioneering work on the secondary analysis of qualitative research data has to be done. However, the current trend towards data archiving and sharing suggests that in this respect a lot will be in motion in the near future. [42]

References

Barbour, Rosaline S. & Eley, Susan (Eds.) (2007). Refereed special section: Reusing qualitative data. *Sociological Research Online*, *12*(3), http://www.socresonline.org.uk/12/3/contents.html [Date of Access: August 30, 2011].

Bergman, Manfred Max & Eberle, Thomas S. (Eds.) (2005). Qualitative inquiry: Research, archiving, and reuse. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, *6*(2), http://www.qualitative-research.net/index.php/fqs/issue/view/12 [Date of Access: August 30, 2011].

Boddy, Martin (2001). *Data policy and data archiving: Report on consultation for the ESRC Research Resources Board.* Bristol: University of Bristol.

Corti, Louise (Ed.) (2006). Making qualitative data more re-usable: Issues of context and representation. *Methodological Innovations Online*, *1*(2), http://erdt.plymouth.ac.uk/mionline/public_html/viewissue.php?id=2 [Date of Access: August 30, 2011].

Corti, Louise & Gregory, Arofan (2011), CAQDAS comparability. What about CAQDAS data exchange? *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, *12*(1), Art. 35, http://nbn-resolving.de/urn:nbn:de:0114-fqs1101352 [Date of Access: August 30, 2011].

Corti, Louise; Day, Annette & Backhouse, Gill (2000). Confidentiality and informed consent: Issues for consideration in the preservation of and provision of access to qualitative data archives. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, *1*(3), Art. 7, http://nbn-resolving.de/urn:nbn:de:0114-fqs000372 [Date of Access: August 30, 2011].

Corti, Louise; Witzel, Andreas & Bishop, Libby (Eds.) (2005). Secondary analysis of qualitative data. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, *6*(1), http://www.qualitative-research.net/index.php/fqs/issue/view/13 [Date of Access: August 30, 2011].

Corti, Louise; Kluge, Susann; Mruck, Katja & Opitz Diane (Eds.) (2000). Text. Archive. Re-Analysis. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, *1*(3), http://www.qualitative-research.net/index.php/fqs/issue/view/27 [Date of Access: August 30, 2011].

Dressel, Gert & Hagenhofer, Johann (2007). Lebensspuren. Erlebte Zeitgeschichte im Land der 1000 Hügel. *Study P200701q, WISDOM*.

ESFRI (European Strategy Forum on Research Infrastructures) (2006). *European roadmap for research infrastructures. Report 2006*, http://ec.europa.eu/research/infrastructures/pdf/esfri/esfri_roadmap/roadmap_2006/esfri_roadmap_2006_en.pdf#view=fit&pagemode=none [Date of Access: August 30, 2011].

Heaton, Janet (2004). *Reworking qualitative data.* London: Sage.

Joye, Dominique (2005). Qualitative or quantitative? Data archiving in documentation, research and teaching. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, *6*(2), Art. 39, http://nbn-resolving.de/urn:nbn:de:0114-fqs0502394 [Date of Access: August 30, 2011].

Kluge, Susann & Opitz, Diane (1999). Die Archivierung qualitativer Interviewdaten. Forschungsethik und Datenschutz als Barrieren für Sekundäranalysen. *Soziologie*, *4*, 84-63.

Kuula, Arja (2000). Making qualitative data fit the "Data Documentation Initiative" or vice versa? *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, *1*(3), Art. 19, http://nbn-resolving.de/urn:nbn:de:0114-fqs0003194 [Date of Access: August 30, 2011].

Laaksonen, Helena; Borg, Sami & Stebe, Janez (2006). Setting up acquisition policies for a new data archive. *IASSIST Quarterly*, *30*(1), http://www.iassistdata.org/downloads/iqvol301laaksonen_0.pdf [Date of Access: August 30, 2011].

Merton, Robert K. (1965). On the shoulders of giants. The Post-Italianate edition. New York: Free Press.

Smioski, Andrea; Müller, Karl H. & Bischof, Christian (2009). Archiving qualitative and quantitative research data at the Faculty of Social Sciences of the University of Vienna. *Research Report*, *14*, http://www.wisdom.at/Publikation/pdf/FoBerichte/2009_01_fobe_14_Archivierungqualitativerquantitativer Daten.pdf [Date of Access: August 30, 2011].

Smioski, Andrea; Müller, Karl H.; Casado Asensio, Juan & Kritzinger, Sylvia (2008). Feasibility study for the preservation and dissemination of qualitative research data in Austria. *Research Report*, *12*, http://www.wisdom.at/Publikation/pdf/FoBerichte/2008_07_fobe_12_Feasibilitystudie.pdf [Date of Access: August 30, 2011].

Parry, Odette & Mauthner, Natasha S. (2004). Whose data are they anyway? Practical, legal and ethical issues in archiving qualitative research data. *Sociology*, *38*(1), 139-152.

Richardson, Jane C. & Godfrey, Barry S. (2003). Towards ethical practice in the use of archived transcripted interviews. *International Journal of Social Research Methodology*, *6*, 347-355.

Author

Andrea SMIOSKI is Qualitative Data and Support Services Manager at the Austrian data archive WISDOM, based in Vienna. Her responsibilities include data acquisition, data processing and user support. She is teaching qualitative research methods at the University of Vienna. In the past she has worked as a researcher in the field of youth research. Her areas of interest are youth culture, transitions in the life course, qualitative research methods, and the reuse and documentation of qualitative data.

Contact:

Andrea Smioski

WISDOM
Wiener Institute for Social Science Data Documentation and Methods
Liechtensteinstraße 22a/2/17, 1090 Vienna, Austria

Tel.: +43 1 319 41 317
Fax: +43 1 319 41 318

E-mail: andrea.smioski@wisdom.at
URL: http://www.wisdom.at/

Citation

Smioski, Andrea (2011). Archiving Qualitative Data: Infrastructure, Acquisition, Documentation, Distribution. Experiences from WISDOM, the Austrian Data Archive [42 paragraphs]. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, *12*(3), Art. 18, http://nbn-resolving.de/urn:nbn:de:0114-fqs1103181.