The Concept and Architecture of the Bremen Life Course Archive

Diane Opitz & Andreas Witzel

Abstract: The Bremen Life Course Archive was built up based on the recommendation from the German Research Council (DFG) in 2000 to make the data material of the Special Collaborative Centre 186 "Status Passages and the Risk of Life Course" available to prospective users. Approximately 700 interview texts have been made anonymous, documented and deposited into a digital archive.

Due to the sensitive nature of qualitative data, we prepared a detailed concept of anonymity and data protection. Special analyses of interviews are hardly possible without an awareness of biographical details, and as a rule it is necessary to know the whole context since the structure of the narration forms a basic requirement for text interpretation. There remains the question of clarifying under which conditions qualitative data can be given to other researchers for the purpose of secondary use.

At the moment we realize a nationwide feasibility study to examine whether and to what extent social scientists are eligible as potential data providers or as future users of secondary qualitative data for research and theory. The question of feasibility should not only make us look more closely at the development of technical solutions for professional processing of the data material or the development of digitized administration and archives, but it should also take the conceptions of the secondary users seriously.

Table of Contents

1. The Starting Point
   1.1 The inventory of partial archives
   1.2 The Bremen Life Course Archive
   1.3 Data protection and the ethics of research
   1.4 The strategies of anonymization
2. Empirical Research and Perspectives
   2.1 The pilot study regarding the reuse of data material
   2.2 The feasibility study

References
Authors
Citation

1. The Starting Point

In contrast to the growing significance of qualitative research methods in all areas of research (including those areas outside scientific institutions) and in contrast to an ever-increasing development of the infrastructure for archiving qualitative data, for example in England (FIELDING 2000), there is hardly any qualitative data in Germany which has been archived. This is remarkable considering the increasing
development of qualitative analysis programs which utilize digitally archived data in research and teaching made much easier due to the widespread use of computers in science. [1]

Therefore, like in England (CORTI & THOMPSON 1998), a lot of qualitative data in Germany has already been lost. In recent years the importance of creating archives of empirical data and documentation of the processes of research have been emphasized continually and by different fields. [2]

The use of archived data for re-analysis and secondary analysis has numerous methodological and methodic advantages which have partially been put forward in connection with the tradition of quantitative methods (MOCHMANN 1968; FRIEDRICH 1983; THORNE 1994; HAMMRSLEY 1997; SZABO & STRANG 1997; HEATON 1998; MAUTHNER, PARRY & BACKET-MILBURN 1998; WILD & BECK 1998; KROMREY 1990; THOMPSON 2000):

- By using replica studies it is possible to examine empirically the validity of theories and to determine the usefulness of methods of investigation and evaluation.
- Archived data material can serve as a "pathfinder," or as a "store of ideas" in preparation for newly-planned investigations.
- The analysis of data which has already been collected may be carried out in the form of historical and culturally-related comparative studies.
- The possibilities to analyze qualitative data (and this depends on the data's specific character) have, as a rule and in primary evaluation, not been fully utilized. There is the example that narrative parts of interviews lead to an abundance of useful information for secondary analyses. This information arises as a result of the subjective context of events and the complexity of the relationship between orientation, actions and experiences and therefore it often goes further than the respective questions of research within projects with special time limits.
- New findings may arise in secondary analyses by applying new theoretical features and changed research perspectives and may appear within the framework of specific or originally minor themes compared with the questions of the primary study.
- A further important aspect is the use of archived data in academic theory: it may be used for comparisons of different techniques of interviews, courses for interviewers or for the practical application of different processes of text analysis for methodological training within empirical social research. Archived data may also be seen as the basis for qualification work (GLÄSER & LAUDEL 2000) which has so far had to be restricted to a small number of case studies for lack of research infrastructure. By economizing extensive and complex qualitative studies and transcriptions of their verbal data, it is possible for the interviewers to answer questions empirically by applying a theoretical spot-check on a greater amount of data which is already available.

© 2005 FQS http://www.qualitative-research.net/fqs/
The interviewers may also concentrate on evaluating the material and in the long run may gain results which are more expressive.

- Archived data makes it possible to verify research results by means of re-analysis (DFG 1998; WILD & BECK 1998; WAGNER 1999).
- The progress made as a result of computer technology and the development of text analysis systems represents a significant step towards the optimal use of the advantages of re-analysis and secondary analysis:
  - Data which is stored in a computerized way reduces expenditure for the infrastructure of an archive by lowering its storage capacity and costs as well as costs for photocopying and transport.
  - As a result of the process of digitalization, the availability of archived data for potential users within research and theory is made considerably easier. [3]

Text analysis systems like winMAX, ATLAS.ti, TEXTPACK or NUD.IST may be used to ensure systematic access to the data (KELLE 1995; WEITZMAN & MILES 1995; FIELDING & LEE 1998; ALEXA & ZUELL 1999; KUCKARTZ 1999). Such systems ensure in particular that the classification of previously-coded text passages or the search for individual terms proceeds much more easily and quickly. The systems facilitate the further administration and evaluation of numerous data as well as documentation of the process of research. The need for the lengthy process of fully transcribing qualitative data may be (partly) reduced by coding primary raw audio material directly (HAUPTMANN 2005). This means that the use of these techniques in the primary study also considerably alleviates the difficulty of secondary analysis. Code schemes used in the computer-assisted analysis of the primary data do not necessarily prejudice or contradict fresh coding of the data (MEDVEDOVIC & WITZEL 2005). [4]

Whilst it is taken for granted that quantitative data results must be handed in to the Central Archive for Empirical Social Research in Cologne (ZA) after the research study has been completed—and this has been the case since the establishment of the ZA in 1960—the archiving of qualitative data usually remains the responsibility of the individual researcher as there is no centralized organization which uses the data in research and theory in spite of the great amount of qualitative social-scientific data. [5]

1.1 The inventory of partial archives

Apart from the Life Course Archive there are only a few existing de-centralized archives in Germany for qualitative data, and these are in some instances limited to very specialized themes of research. Up to now only a small amount of this data has been archived using computers. In most of the German archives there is the possibility of gaining information about the inventory by looking through catalogues, but the data cannot be found or cannot yet be found in a computerized format (KLUGE & OPITZ 2000). The exception to this is the number of archived texts on open questions from different surveys which are available at the ZA in Cologne. [6]
1.2 The Bremen Life Course Archive

Based on the recommendation from the German Research Council (DFG) for the last phase of research (1999-2000) for the Special Collaborative Centre 186 "Status Passages and the Risk of Life Course," University of Bremen, to make the data material of individual projects available to prospective users, a large number of qualitative interviews have been made anonymous, documented and deposited into a digital archive. At the end of 2001 the Life Course Archive in Bremen was integrated into the newly-founded Graduate School of Social Sciences (GSSS) at the University of Bremen. [7]

In addition to dealing with the complex processing of extensive qualitative data, the archive includes approximately 700 interview texts from four different projects and a free QDA-program which is appropriate to interview texts (http://www.lebenslaufarchiv.uni-bremen.de/). [8]

In this regard we must emphasize the connections between qualitative and quantitative investigations within the process of research which have arisen during most of the projects, such that there is an extensive amount of standardized data which is available for re-analysis or secondary analysis, as well as interviews, which can be found in the archives. [9]

1.3 Data protection and the ethics of research

Due to the sensitive nature of qualitative data, it was necessary to prepare a detailed concept of anonymity and data protection before creating the archives. In Germany the legally guaranteed freedom of research within science must appreciate the informational self-determination right of the individual. In 1983 the German Federal Constitutional Court pointed specifically to the ruling of the law of the population census. This law ensures that the individual is protected against unlimited inquiry, storage, use and passing on of his/her personal data, since modern data processing has not produced clear possibilities of inspection and influence (BverfGE 65, 1 [42f]; SIMITIS 1984). The judgment stated that the basic right of informational self-determination (Grundgesetz, Art. 2 Abs. 1) guarantees that every individual may "decide for him/herself about the disclosure and use of his personal data" (BverfGE 65, 1 [43]). [10]

However, since the laws of data protection only apply to personal details (§1 Abs. 1 and 2 BDSG), scientific data can be passed on to other research projects if they are made anonymous such that "the individual information about personal or essential relations can no longer be allocated to a definite or definable person or when it is only possible to make an allocation after huge expenditure of time, costs and effort" (§3 Abs. 7 BDSG). Therefore (as is the case for the Central Archive in Cologne) it is possible to pass on quantitative data to other research projects if they have been made effectively anonymous beforehand (MÜLLER et al. 1991). [11]
The situation for qualitative data is, however, more difficult, because the interviewees in biographical interviews mostly give very detailed reports about their lives. As a rule it is not sufficient to erase names, places or years: the data material often contains differentiated descriptions of life circumstances which, with knowledge of additional material, could lead to the re-identification of the respondent. Longer texts in which the interviewee explicitly describes different experiences cannot be simply erased or changed like the variables of a quantitative data study. Special analyses of interviews are hardly possible without an awareness of biographical details, and as a rule it is necessary to know the whole context since the structure of the narration forms a basic requirement for text interpretation (KLUGE & OPITZ 1999). [12]

There remains the question (based on the existing legal regulations) of clarifying under which conditions qualitative data can be given to other researchers for the purpose of secondary and re-analysis. The preliminary concept of the Life Course Archive consists of binding the data users strictly to the legal regulations and committing them to comply with further conditions of the project. [13]

The following points should be agreed upon in a written contract:

- data may only be used for research purposes or for a specially determined research project,
- data must not be passed on to a third party and must be stored in such a way that a third party cannot gain access,
- data must remain anonymous,
- person-related details may not be quoted or published, and
- notification of the end of the research project must be given and the data which has been made available must be erased. [14]

1.4 The strategies of anonymization

As a rule, person-related details like names, addresses or places are erased when the interviews are transcribed in order to prevent the re-identification of the interviewee concerned. This procedure conforms to the legal regulations which demand anonymity "as soon as it is possible for the research purposes" (§40 Abs. 3 BDSG).

However, the archive must check the process of making data anonymous by the researchers before the data is put into archives and before it is passed on. This is for two reasons: [15]

First, person-related details which are relevant for the examination of the research question are not erased or changed at the time of transcription so that a suitable interpretation and evaluation of the data material is not endangered. Secondly, while making biographical interviews anonymous there may arise frequent problems when the life course of the interviewee shows outstanding experiences and/or his place of residence is well-known. Such details can, however, be re-written at a more abstract level so that information which is
relevant for the project and necessary for an adequate interpretation of the cases does not go missing. There is here the example of the company Mercedes which was not described as a "large industrial company" but as a "large car manufacturer." In this case it is possible to evaluate the interviews in a meaningful way in spite of making the data material anonymous. As a rule it is of importance for the interpretation to know the branch of industry (e.g. car manufacturer) and the size of the company that is involved. The specification of the company name is seldom of significance and may only be required when a number of car manufacturers are to be compared with each other. [16]

For every individual study the decision must be taken as to which details should be made anonymous in order to prevent a re-identification of the interviewed person. In exactly the same way it is necessary to look at every individual case by considering the whole context and possible future research questions in order to decide how the details should be made anonymous. Regrettably, we cannot formulate any general solutions for making qualitative data anonymous because such data are extremely heterogeneous in terms of the themes and areas of life already mentioned. Further solutions concerning anonymization of data will rely on the legal recommendations carried out in the framework of our feasibility study (see 2.2). [17]

2. Empirical Research and Perspectives

As a result of the growing significance of qualitative methods since the 1970s and of the progress of computer-supported data registration and processing, our concept of an extension of the Life Course Archive to a nation-wide central archive is under consideration. This is also based on the development of those data analysis programs in collaboration with social-scientific data storage. Our feasibility study will determine the extent to which there is a need for the reuse and secondary use of digitized qualitative data. [18]

A preliminary examination of the archive highlights the importance of storing qualitative data in archives. [19]

2.1 The pilot study regarding the reuse of data material

Within the framework of the final phase of the Sfb 186, a pilot study regarding the use of qualitative data (n=77) which was carried out by Diane OPITZ evaluates the reuse and secondary use of data in a positive way. The results show a similar picture to those of the studies of the "Qualidata" archive in England at the beginning of the 1990s (CORTI & THOMPSON 1998): indeed over one third of the interviewees continue to use the material but in most cases the data material is still stored either in their office (n=42) or at home (n=27). It has already been destroyed in 14 cases. Nevertheless 15 people told researchers that they had put their data in archives. [20]

More than half of the researchers expressed their willingness (under certain conditions) to make data material available so that it could be put into archives for
the purposes of secondary analyses and comparative studies (60 entries), for re-
analyses (n=47) and for use in theory (n=46). Moreover, the majority of those asked (n=69) would use data material already found in archives for their own research projects. Only nine of those asked refused to make their data available. In this regard it is interesting to note that 34 of those questioned have already made their material available for purposes other than research, for example as teaching material (n=11) or for qualification work (n=18). [21]

2.2 The feasibility study

On the basis of these considerations and preparatory work, the German Research Foundation (DFG) financed a cooperation project of the Life Course Archive at the Graduate School of Social Sciences (University Bremen) and the Central Archive for Empirical Social Research (ZA), University of Cologne, to explore the possibilities of providing an infrastructure for archiving qualitative data and to identify the needs of the scientific community for such an infrastructure. Based on a stock of qualitative social-scientific research projects in Germany since 1984 and limited to interview data, the project examines whether and to what extent social scientists are eligible as potential data providers or as future users of secondary qualitative data for research and theory. [22]

In this nationwide quantitative survey of researchers (n=430), project leaders working with qualitative interviews were asked about the methods of and possibilities for the use of secondary data, about the worthiness or willingness of depositing data to an archive and about the demands of technical and organizational arrangements securing a trouble-free data dissemination. [23]

Based on an examination of the whole stock of quantitative data, the analysis of the qualitative survey of approximately 45 expert interviews will give answers to the following questions which are the focus of interest:

a. The willingness to deposit qualitative data into archives and to transfer such data for use. The central question of this part of the examination is where qualitative data from completed research projects should be stored, as well as conditions for transferring such data for further use (e.g. transfer of data to an archive or the placement and consultation of the archive).

b. The willingness to deposit qualitative data into archives and to transfer such data for use. The central question of this part of the examination is where qualitative data from completed research projects should be stored, as well as conditions for transferring such data for further use (e.g. transfer of data to an archive or the placement and consultation of the archive). [24]

First, cautious conclusions from the yet unfinished data analysis suggest that, on the one hand, an archive has to make an effort to preserve threatened data material and to develop concepts and guidelines for preparation and documentation of this data. On the other hand, an archive also has to deal with the reasons for the—especially in Germany—still not widely accomplished re-
usage of qualitative data (OPITZ & MAUER 2005). Therefore we wish to establish a concrete interest in the secondary use of verbal data material which is already available. It is important to examine the advantages and disadvantages which may arise when carrying out secondary and re-analyses (e.g. CORTI, WITZEL & BISHOP 2005). If we take one look at the fears of social scientists from England, we see how important it is to examine such questions empirically. Social scientists in England point to a lack of putting the available data about orientation and action in proper context. They also point to a lack of research documentation (HAMMERSLEY 1997; HEATON 1998; THORNE 1994; SZABO & STRANG 1997) and they are skeptical as to the feasibility of outside data for their own research projects. They fear methodological criticism of their own studies, checks on the validity of their findings by others as well as a general increase in competition amongst colleagues (MAUTHNER, PARRY & BACKET-MILBURN 1998). [25]

By extending upon these studies we will develop an innovative model for depositing qualitative data to archives and hope to estimate the effort needed to create expert and user-friendly documentation and data processing of material which can be regarded in Germany as worthy of being deposited to archives. [26]

The question of feasibility should not only make us look more closely at the development of technical solutions for professional processing of the data material and the development of digitized administration and archives, but it should also take the conceptions of the secondary users seriously. The planned extension of a national data archive includes international co-operation to integrate experiences from other institutions, like ESDS Qualidata in Essex, and to bring together existing resources in Germany. But it is not meaningful to create an archive which does not offer the services of data consultation, schooling, network and information provision for qualitative social research and qualitative management. By bringing together the existing resources in Germany it may be possible to make these tasks clearer. In close collaboration with the Society for Social Scientific Infra-structure Institutions (GESIS) and other existing or developing international centers, an innovative concept should be developed in order to ensure a reliable and permanent basic provision of information and communication for qualitative social research (WITZEL 2004). This serves the development of integrative structures for a Centre for Qualitative Social Research and Qualitative Management within the social and technical sciences (c.f. the contribution by Katja MRUCK in this issue). [27]

References


Authors

**Diane OPITZ**

Present position: Researcher ("Archiving and Secondary Analysis of Qualitative Interviews in Germany") at the Life-Course-Archive, Graduate School of Social Sciences (GSSS), University of Bremen.

Major research areas: archiving qualitative data; concepts of anonymization; software for qualitative data analysis.

Contact:
Diane Opitz
Bremen Life Course Archive
Graduate School of Social Sciences (GSSS)
University of Bremen
Wiener Strasse/ FVG Mitte
28334 Bremen, Germany
E-mail: dopitz@gsss.uni-bremen.de
URL: http://www.lebenslaufarchiv.uni-bremen.de/

**Andreas WITZEL**

Present position: Director of the Life-Course-Archive, project management ("Archiving and Secondary Analysis of Qualitative Interviews in Germany") at the Graduate School of Social Sciences (GSSS), University of Bremen.

Major research areas: archiving qualitative data; interpretative social research; occupational socialization; biography and life course.

Contact:
Andreas Witzel
Bremen Life Course Archive
Graduate School of Social Sciences (GSSS)
University of Bremen
Wiener Strasse/ FVG Mitte
28334 Bremen, Germany
E-mail: awitzel@gsss.uni-bremen.de
URL: http://www.lebenslaufarchiv.uni-bremen.de/

Citation