

One Bird in the Hand ...: The Local Organization of Surveys and Qualitative Data

Anne Margarian

Key words:

documentation;
primary data
archive; work-flow
organization;
meta-database;
secondary
analysis

Abstract: In many organizations, and among many researchers, there is considerable resistance to the systematic large scale storage of survey data. Thereby, not only data storage but also work-flow organization and documentation are often left to the individual researcher. Such practice leads to inefficiencies due to a lack of positive specialization effects and to problems in traceability due to the existence of small-scale ad hoc solutions. I propose small scale solutions with survey data archives for individual institutes, projects or researchers that not only support comprehensive documentation but also raise awareness on survey-based projects' complexity and the resulting high demands for information. The proposed small scale solutions must adhere to common standards and need to be attractive for the user in order to motivate application. Attractiveness increases with elements like master-detail interfaces and tree-views that provide the user with immediate overviews and by a structure that allows for simultaneous work-flow organization and meta-data entry. The meta-database SuPER is described as a paradigmatic example of how a database could reflect the complexity of a survey data-based research project.

Table of Contents

- [1. Introduction](#)
- [2. Saving Data Along the Work-Flow](#)
- [3. Defining a Survey](#)
- [4. Planning and Organizing a Survey](#)
- [5. Data Processing and Analysis](#)
- [6. Concluding Remarks](#)

[References](#)

[Author](#)

[Citation](#)

1. Introduction

Regular large scale surveys usually deliver quantitative data that are not only used by those who collect them but also by secondary users. These surveys are designed with the purpose of delivering reliable information about a population as a whole or about clearly defined groups of people that are of interest for many researchers. The creation of samples is guided by statistical measures and usually based on other secondary non-survey statistics that allow for efficient sampling strategies. In these large scale surveys, respondents are expected to answer questions independently from those unobserved context conditions, which are not controlled for by the survey's questions or sampling strategy. [1]

However, for the treatment of many innovative, more detailed, rather explorative or new questions, researchers have to rely on small scale surveys that deliver specific data for the question at hand. These data are usually meant to inform

one single research project and are therefore interpreted within a given context, which is well known to those who conduct both: the survey and the analysis. If researchers only plan to deal with the one specific case at hand, requirements concerning the control of framing conditions or unobserved and potentially influential characteristics are lower. This applies even more for qualitative data, which are usually collected for in-depth analyses of selected cases rather than for the purpose of (quantitative) comparison of many cases or for the purpose of generalization on a population or sub-population. [2]

Therefore, if the same data were to be re-used by other users or at a later time, a considerable amount of additional information would have to be provided that is not routinely collected in small scale surveys. This situation impedes the secondary use of data from small scale surveys and specifically from qualitative data. While considerable progress has been made recently in the development of nationwide, trans-organizational facilities that serve the storage and re-use of qualitative data (see for example IASSIST, 2010 or SLAVNIK, 2013), considerable skepticism prevails among researchers concerning the re-use of qualitative data (KUULA, 2010). A larger share of the relevant research community does not participate actively either as supplier or as user of data from these repositories (SLAVNIK, 2013). [3]

Researchers' resistance to sharing their data in larger pools is at least partly caused by considerable general deficits in the documentation of surveys and survey data (SMIOSKI, 2013). Despite the fact that surveys and subsequent data transformation steps are costly, many social research projects and institutions do not invest much effort in archiving qualitative data. Nevertheless, a lack of appropriate meta-data not only hinders the reuse of survey data but deficits in the documentation of the different steps in surveying, processing and analyzing data create problems in compliance with even some of the most fundamental standards of scientific practice. These standards enable transparency, traceability and the reproducibility of analyses and results. The relevance of meta-data for the usage and interpretation of qualitative data is stressed, for example, by the perspective of "interviews as social interaction" (DEPPERMAN, 2013). The lack of informative meta-data therefore often hinders a secondary use of qualitative data, even within the same institution (MEDJEDOVIC & WITZEL, 2010). [4]

The frequent deficits in documentation are explainable and might even be rational from the perspective of the individual researcher if (transaction) costs are considered. Documentation of analytical steps and intermediary results, as well as the substantiation of results and an assessment of their robustness, would be necessary to give recipients an impression of a study's rigor. However, even today, many journals do not demand even a minimum level of documentation from qualitative analyses, for example, in terms of coding-schemes, coding examples or analytical principles. The development of techniques for documentation could be fostered, if reviewers from qualitative analyses would routinely demand summaries of qualitative data and detailed information concerning the context of the data's gathering in order to enable an assessment of their reliability. Under such circumstances the value of good documentation

would become apparent immediately as it would alleviate the frequent need to satisfy a reviewer's specific demands via the provision of belatedly developed measures. [5]

However, the meta-data needed to document the survey, its data, and subsequent analyses are complex. The heavy demands on comprehensive meta-data reflect the complexity of the research process itself. Each phase of a project with qualitative data comes along with its own demands and obstacles, which often bear consequences for subsequent steps as well. Project phases have been differentiated as "thematising, designing, interviewing, transcribing, analysing, verifying and reporting" (FINK, 2000, §12). This breakdown clarifies the complexity and interdependence of decisions to be taken in a survey-based research process. Researchers' decisions are often made unconsciously or the relevance of specific events and conditions becomes apparent only from an ex post perspective. A lack of awareness of the complexity of qualitative research projects, even on the researchers' side, becomes apparent, for example, in the common confounding of "data collection methods" and "data analysis methods" (BERGMANN & COXON, 2005, §10). Disciplined and detailed documentation poses high demands on the individual under these circumstances. Therefore, documentation costs might be prohibitive if they rest on the shoulders of individual researchers alone. [6]

This article proposes that the introduction of small scale solutions, which allow for an efficient in-house documentation and the re-use of data within research organizations, would represent a valuable first step on the way to a more general secondary use of data. The proposed small scale solutions for in-house use do not solve the difficult questions of anonymization and data protection that are related to data re-use by third parties; instead they support the development of standards, routines and other capacities and capabilities, which prepare even very small projects and institutions for compliance with increasing demands concerning research documentation. Respective small-scale solutions could be implemented by single researchers, or within single projects and institutions, until common standards for meta-data storage have been delivered and commonly accepted. In order for the implementation of the support system to become feasible and valuable, only a minimum level of support from the institution and a minimum interest in the potential re-use of data and meta-data on the researchers' side is necessary. Provided the interest of research institutions in the creation of a common pool of information and meta-data that relate to the institutions' specific research questions, concepts, surveys and analyses, the proposed small-scale meta-database approach is inexpensive and technically easy to implement once it has been developed successfully.¹ [7]

1 The data base SuPER described by this article is meant to inspire the development of valuable survey and data management support tools. SuPER itself represents a prototype rather than a perfect product. Nevertheless, I provide it without charge upon request together with a short documentation, but no further service or support can be provided. Support by Lorenz HÖLSCHER (CLS Software-Service GbR, Aachen) would be conditional on resources and come with costs. If the data base is used for or described in a publication, acknowledgment of creatorship is expected. Please note that both, data base and documentation, are in German and that the Access data base is sensitive with respect to computers' operating's system and its local specificities. I use the data base in its current form but I believe that it will be necessary to

The data-base solution also needs to be flexible and adaptable to researchers', research groups' or institutions' specific needs. This adaptability is important if the database is to be perceived as a support, rather than as a nuisance, by its active users because most researchers follow specific scientific paradigms with assumptions, concepts and methods that often differ fundamentally from each other. As BERGMANN and COXON (2005, §11) write, "these meta theoretical approaches make different assumptions about human thought and action and, thus, are likely to accumulate, code, analyze, and interpret data differently." Accordingly, a database-based support system needs to be able to capture and reflect these differences. Therefore, a tractable solution for the problem of documentation could, and in fact, should, serve two purposes at one time: It should support 1. the structured collection of comprehensive and clearly represented meta-data and 2. the conscious development of a comprehensible research concept from survey planning to secondary analyses of gathered data. [8]

One aim of the article is to draft an outline of the complex structure of a database that serves most of a researcher's demands for documentation. In doing so, I simultaneously propose that local small scale solutions of survey data archives for individual institutes, projects or researchers might considerably reduce individual researchers' costs of documentation and at the same time raise the perceived benefits of documentation. The proposed small scale solutions adhere to common standards and at the same time deliver additional value to the user in order to motivate application. Attractiveness is gained by elements like master-detail interfaces and tree-views that provide the user with immediate overviews and by a structure that allows for simultaneous work-flow organization and meta-data entry (IVERSON, 2009). [9]

I demonstrate the information needs for the comprehensive description of a survey-based scientific project with the help of an exemplary tool, which allows for simultaneous work-flow organization and documentation. This exemplary meta-database's structure mirrors the information needs surrounding activities related to scientific surveys. In the following sections, I first discuss the general properties of this meta-database. Afterwards, I describe its structure along the typical steps in survey-planning, survey-conduct and analysis in subsequent subsections. I conclude with a final assessment of the value and restrictions of the proposed small-scale solution. [10]

2. Saving Data Along the Work-Flow

The meta-database SuPER² has been developed in order to enable a single researcher or a scientific team to capture all the relevant information that is

translate the data base into another data base environment for larger applications. Nevertheless, SuPER's general structure provides a valuable template for further developments in this respect.

2 "SuPER" (Survey Program for Enquiry Reporting) is a database based on Access (compatible with versions 2007 and later) for the documentation and planning of surveys. It was designed for use at the Thünen-Institute for Rural Studies and consists of a front end- and a backend database and a help file in Word. The database was developed by Lorenz HÖLSCHER (CLS Software-Service GbR, Aachen) in collaboration with me.

necessary to understand the collection and analysis of survey-data in retrospect. It is supposed to support

- the provision of an immediate and later overview of survey based projects,
- the (in-house) re-use of contacts,
- the (in-house) re-use of data,
- development of capabilities for survey documentation,
- compliance with documentation requirements, and
- planning and structuring of the research process. [11]

Using the database supports the planning process because it fosters self-reflection and raises awareness of the multiplicity of steps of a complete research process. It does so by proposing standards and steps or delivering other information that is already included in the database's tables. [12]

A pre-condition for these achievements is the immediate filling of the database in the course of the research project, i.e., during the process of planning, conducting and processing of surveys and survey results. In order to motivate the immediate and disciplined completion, the database should be clearly arranged, provide structured forms for data entry that are easily accessible and reward the user's effort immediately via provision of information already included in the database and/or via provision of a good summary of the information that has been entered. In SuPER, the immediate overview of entered information concerning a specific survey is provided via so called tree-views. A tree-view is a graphical user interface that presents information in hierarchical order. In our meta-database these tree-views are integrated in master-detail interfaces, which allow for convenient meta-data entry in comprehensive data entry forms without losing track of the current position in the database's structure as represented by the tree-view. The divided screen with the tree-view to the left and the detailed data entry form to the right is shown in Figure 1. In the open form on the right, the user will find detailed information, for example concerning the type, the form and the medium of the survey, the time period in which it was conducted and the project it belongs to, while the tree-view on the left provides an overview of specific important characteristics like the survey's themes, regions/cases, dates, respondents and analytical steps. The tree-view grows in the course of data-entry and not only provides a first overview but also supports the user in her navigation between data-entry forms in the right window.

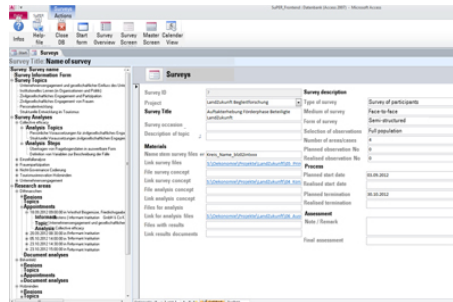


Figure 1: Screen-shot of the master-detail interface with tree-view (left) and data entry form (right). Please click [here](#) for an increased version of Figure 1. [13]

One important additional value of the database for its users stems from the fact that it provides access to information that has been entered before. Therefore, entered information needs to be saved systematically, i.e., in self-explanatory orderings and hierarchies, and data consistency needs to be guaranteed. Consistency means that data are always entered within pre-defined structures, can be referenced unambiguously and show a unique relationship to other data and tables in the database. Technically, consistency is guaranteed in SuPER by the separation of entry of two classes of data in two different screens: The "survey screen" is used in order to enter ongoing data from the current research process; the "master screen" serves the description and entry of so-called master data and reference data. Reference data include permissible values to be used in other data fields, specifically for the description of the ongoing process on the survey screen. Reference data are rarely changed or supplemented. Master data contain key data for the research process, such as the names of interviewees that might be accessed for other surveys as well but are also frequently amended and changed. In order to guarantee consistency, reference data and master data may not be entered directly on the survey screen but need to be entered on the master screen. [14]

As the data themselves are not stored with the meta-data, it is recommended not only to use the database for documentation, but to develop a complementary system of systematic folder-structures and file-naming conventions. These naming-conventions allow files to be detected, even if they have been moved in the folder-structure and the data-path documented in the meta-database is no longer correct. I propose the following convention for naming files that comprise primary data from a survey in different regions or cases: Combine, separated by an underscore character, 1. the name of the case/region, 2. the last name or an informative alias of the interviewee or of the group of people, 3. the project's code, 4. the serial number of the survey within the project (usually double-digit), 5. a code for the survey's type (personal interview, postal questionnaire etc.) from a common list and 6. the serial number of the observation within the survey (two-, three- or more digit, depending on the survey's size). 1. and 2. can be omitted for the sake of anonymization, they only serve as immediate orientation and are not necessary in combination with the meta-database. Of course there are numerous other possibilities for effective naming conventions but I strongly recommend that

a common denomination stem be used. In this case, each single file does not need to be named separately in the database and files' names can be reconstructed and files detected with a minimum of prior knowledge, even if meta-information has been lost and folder-structures been changed. For the folder structure I propose to create separate folders for each project. A project's folder structure should include, among other things, separate folders storing 1. concepts of a project, a project's surveys and surveys' analyses, 2. information concerning contact persons and participants as well as schedules and dates, 3. the primary data raised in the project's surveys, 4. routines and raw-results of data analyses and 5. outputs like papers and presentations that systematically document, edit and interpret the project's results. The meta-database references all of this information as the proposed folder structure roughly describes the most fundamental steps of the survey-based research process. The following sub-sections describe those steps that are paralleled by data-entry in the meta-database. [15]

3. Defining a Survey

Surveys are the central units around which meta-data are collected in the meta-database. If others later screen the meta-data for surveys that could be helpful for their question at hand, they might be interested in the survey's themes, in the cases or regions that have been in the survey's focus and in the survey's type or methodology. This information enables users to quickly understand the surveys' content. In order to enable efficient data entry and effective search routines, all descriptions need to be standardized as far as possible. Standardized values are provided via lists. Lists refer to master data that can only be defined and changed on the master screen (see Section 2). The database provides an overview of surveys that already collects some of the most fundamental information (Figure 2). These and other characteristics are entered more completely on the survey's main form, which is approached via the survey-screen. If a survey is created, either in the overview or on the survey-screen, it is automatically assigned a unique ID.



Survey ID	Project Name	Survey Title	Survey occasion	Type of survey	Medium	Survey form	Selection of obs.
	Project XY	Activities in region	Initial competition	Part. observation	Face-to-face	Externally d.	
	Project XY	Super-region activity	Policy implementation	Part. observation	Face-to-face	Externally d.	
	Project XY	Phone conference	Policy implementation	Part. observation	Phone	Externally d.	
	Project XY	Kickoff survey	Policy establishment	Survey in participant	Face-to-face	Structured I	Full population
	Project XY	Web pages	Project development	Internet research			
	Project Z	Official documents	Policy overview	Document analysis			
	Project Z	Village study	Important project	Case study	Face-to-face	Structured I	Show ball
	Project W	Project applications	Policy assessment	Document analysis			Full population
	Project W	Selection of regions	Planning of case site	Document analysis			

Figure 2: Screen-shot of the overview of surveys the database offers to users for orientation. Please click [here](#) for an increased version of Figure 2. [16]

On the survey's main form, the user selects the project that the survey belongs to from a list field. If the project has not yet been defined it needs to be added in the master data on the master screen. Moreover, file links and file names, respectively denomination stems, are being entered for different kinds of files. File links and denomination stems are entered for survey concepts, for survey

data, for concepts for data analysis and for results. Other information to be entered contains the dates of the planned as well as actual start and beginning of the survey, the planned and actual number of observations, interviewees, respondents and the number of regions/cases of the survey; there is also room for a free note concerning the survey. [17]

While this information so far mainly serves documentation purposes, other aspects are more interesting considering the aim of simultaneous support of a conscious development of a comprehensible research concept. Important in this respect are for example the options for the survey's qualitative description. Four characteristics can be defined on the survey's main form: 1. the survey type, 2. the survey medium, 3. the form of the survey and 4. the way that interviews, observations or respondents are selected. These qualities are defined via list-boxes, i.e., only choices that are predefined in an existing list can be entered. As outlined before, these so-called "reference data" can only be changed or entered in the "master screen"; normally reference data should not be changed at all. With respect to the database at hand, I propose that each user initially adjust these lists to his/her specific needs. This process is well-suited for the initialization of a process of self-reflection concerning methods and approaches used by single researchers, groups of researchers or within a scientific institution. This initial adjustment of reference data is proposed for those lists that refer to kinds of subjects, which are usually applied and maybe even conceptually developed within a specific institution. [18]

The "type of survey" field (1) serves the rough classification of the survey. In the database's current version, there are ten permissible types to select: Survey among concerned persons, survey among participants, (representative) random survey, analysis of documents, group discussion, expert consultation, case study, internet research, situation analysis, participative observation. Obviously, this list is neither exhaustive nor systematic, changes and amendments in these specific "reference data" might become necessary if someone begins working with the database. Additionally, the user of the database, who defines a survey, can select among five media in the current version of the database (2): E-mail, Internet, postal/in writing, phone and face-to-face. Possible forms of the survey (3) are standardized, semi-standardized, open questions, structured interview, and participatory observation. Obviously, these choices, too, are neither exhaustive nor consistent. A new user of the database will initially have to adapt them to his/her needs, depending on the type of surveys and the type of data that are used regularly. Finally, respondents or observations might represent (4) a full population or a sample. A sample may either be externally determined or selected by the researcher. Selection might happen arbitrarily, randomly or as a stratified random sample, by criteria like best or worst performers, in a snow-ball process or in a matching approach. [19]

A user of the database who wants to find out, for example, whether expert consultations, internet based surveys, semi-standardized interviews or snow-ball selection have been carried out before by a colleague, enters the "master screen" and chooses the form "type of survey," "survey medium," "form of survey" or

"survey observation selection." There, for each of the options described above, the user finds an integrated table with all surveys that belong to the selected category. [20]

A survey is not only defined by its form and its structure but also by its topic. Accordingly, themes are selected on a separate form belonging to the survey. These themes will later be visible in the tree-view in order to give a first impression of the survey. Themes are selected in list-fields, while lists can only be changed and amended in the master screen. As there are many possible wordings for identical topics, themes are defined in a hierarchical manner: groups of topics, topics and sub-topics. Only sub-topics can be selected on the "survey screen." Like types of surveys and forms of surveys, topics need to initially be defined carefully by researchers when they start working with the database. A process of intense reflection is necessary in order to develop an adequate hierarchy that fits the researchers' field of interest. Nevertheless, this process has proven to be a valuable investment in terms of reflection of the relations between topics, projects and surveys that are being conducted in an institution, a research group or by a single researcher. [21]

An important group of topics in the database's current application is "innovation and entrepreneurship." Topics that have been defined on the master-screen within this group so far are "economic innovation," "political and administrative innovation," "social innovation," "business foundation" and "social entrepreneurs." "Political and administrative innovation" is repeated as a sub-topic to its own category, such that it may be selected on the survey-screen if further differentiation is not adequate. Otherwise more detailed sub-topics to the topic "political and administrative innovation" are, so far, "personal pre-conditions for political and administrative innovation," "structural pre-conditions for political and administrative innovation," "characteristics of political and administrative innovation," "development of political and administrative innovation" and "effects of political and administrative innovation." The other topics and sub-topics are differentiated similarly. [22]

Finally, if applicable, the survey is defined by the different regions, or, more generally, by the different cases it covers. The user of the database assigns a telling name to these "research areas" in a separate form. If a differentiation is not indicated, a single "research area" needs to be defined by an arbitrary identification in order to proceed. If the database is used for a spatially defined survey in Germany, in the database's current version it is possible to assign the administrative regions to the survey's regions that fit best to the spatial design. The reference to administrative units enables later searches for surveys that have taken place in specific regions or on specific spatial scales. It is possible to assign specific topics to specific research areas additionally to the definition of topics on the survey level. As before, research area topics are selected from the pre-defined list of sub-topics described above. The possibility to select more specific or differentiated topics on the level of the single research area has the potential to provoke additional reflection concerning the areas' specificities and comparative values and respective needs and demands in survey design. [23]

After the survey has been defined by the dimensions described, the most important information is captured in the tree-view. Up to this point the tree-view shows the survey's name, its topics, its research areas, however defined, and specific topics to be addressed in specific research areas:

- Survey title
- Survey-level topics
- Research areas (regions or cases)
 - Area-level topic [24]

Important additional information concerning the survey can be seen on the same screen if the survey's main form is opened on the screen's right hand side. [25]

4. Planning and Organizing a Survey

The definition of a survey's scope and aim is usually followed by the survey's realization. If the survey relies on informants, these informants need to be identified and contacted. In this respect, the database not only provides the possibility for documentation, but it increasingly facilitates the identification of informants and their contact data, if surveys are conducted repeatedly within the same region. The user simply writes the initials of a name and the database provides him/her with information on informants already entered in the database. Nevertheless, initially, the user will have to feed the database with his/her contacts. He/she will do so on the master screen. Here he/she selects the type of institution the contact represents, and then he/she selects the specific institution, if it already exists, or creates a new institution and enters the names of all contacts for the survey that belong to this institution. The hierarchical structure of institutional and personal level information in the database prompts reflection concerning the function of interviewees in the survey on the researcher's part. Before a person is added, the question needs to be answered, whether its role in the survey is that of private person or that of a representative of a specific institution. Once in the database, the respondents can be selected from a list or they are entered directly, possibly supported by the auto-fill function. [26]

After entering contact data, the user generates a separate form for each appointment or observation on the survey screen in the course of making appointments and organizing the survey. In the database's current version, each observation is identified in the tree-view by time and place of its occurrence. Usually, each of the corresponding forms relates to one appointment with one or more respondents, but the observations defined can refer to any unit that seems appropriate for the survey at hand. If applicable, for example in face-to-face surveys, the informants or respondents belonging to the observation are defined separately "below" the observation or appointment form in the tree structure. In the database's current version, informants that belong to an observation are identified by their name and their institution in the tree-view. Once they are selected, information concerning them becomes visible in a separate form on the survey screen if it has been entered on the master-screen before. [27]

Additionally, each observation or appointment can also be characterized by a specific topic in the tree-view. Specific topics can be added on the observation level as well as on the survey level and on the level of the research area. Observation level topics will only become relevant for qualitative surveys where researchers put strong emphasis on individual observations. If this is the case, the possibility to define specific topics on the observation level provokes researchers' reflection concerning expected individual observations and factual contributions to the question at hand. Expected contributions can be described in the planning phase of the survey; factual contributions may be added in the survey's wrap-up phase, i.e., in the time of reflection that follows the individual observation before the formal analysis starts. This information might be important in order to structure the material in the upcoming analysis or to inform potential secondary users of the qualitative material. It might also support reflections on whether or not additional information is to be expected from additional interviews, and if the criterion of data saturation (GUEST, BUNCE & JOHNSON, 2006) is used for sample size determination. [28]

In the course of planning and in the wrap-up phase, the user may enter further information in each observation's main form. Entered meta-information may concern the appointments' date and place, possible activities or events framing the appointment, information concerning the initiation and planning of an appointment and assessments and remarks concerning the appointment's tenor. With respect to the latter, there are a number of list-fields with pre-defined options like "high," "medium," "low," or "none" for assessments of items like "willingness to participate in another interview," "knowledge of topic," "courtesy," "frankness" and "overall assessment." There are also fields for unrestricted notes concerning "explorative observations and considerations" and "further comments." Very important for possible secondary use of survey data is the field that asks the user to enter information concerning the data protection agreement reached with the respondent. These fields guide the researcher's reflection concerning specific appointments and their surrounding conditions if the database is filled in the course of the survey. [29]

Survey data needs to be filed in the wrap-up phase. Usually, the file link in the folder structure and the files' nomination stem are previously defined on the survey level. Nevertheless, information concerning the filing of data relating to specific observations can be entered on the observation's main form as well. If a consistent convention for naming observations and their data has been generated (see above), it might suffice to enter the observation's running number on the appointment level. [30]

Alternatively or supplementary to primary data, the researcher might rely on secondary data from documents and other sources. Such "document analyses" can be defined within each research area (region or case). They are characterized in a corresponding form by the document-analysis' working title, which later on appears in the tree-view, by the name of the responsible researcher, the beginning and end of the analysis, by file names and file links and by an assessment of the documents' information content. The specific topic of the

document analysis may be entered in a separate form and in this case appears later on in the tree view. The documents for each document analysis are defined in new branches in the tree view as well. A separate form with bibliographic information is created for each document. This branch for document analysis is in so far of importance, as the considerable contribution of documents to qualitative research is often under-valued by researchers. Documents, whether forms, web-pages, brochures or business reports, potentially deliver important supplementary information; this information may concern the question at hand, the characteristics of informants and their institutions, or the external conditions. Nevertheless documents need to be treated as scientific data in gathering, analysis and documentation in order to use their content appropriately. The database might help to create awareness of this fact. [31]

After entering the information concerning the organization of the survey and the survey's wrap-up phase, the tree-view not only shows the survey's main characteristics and its research areas as described above but also the observations within each research area sorted by date and identified by place, the informants related to each observation identified by name and institution and the specific topic treated at this specific observation. Similar information is provided for potential document analyses related to the survey:

- Survey title
- Survey-level topics
- Research areas (regions or cases)
 - Area-level topic
 - Observation / appointment
 - Observation level topic
 - Respondent / informant
 - Document analysis
 - Document analysis level topic
 - Documents [32]

5. Data Processing and Analysis

Finally, the database offers a possibility to keep track of data-processing and analytical steps applied to the data at hand. Initially, each analysis needs to be defined on the survey level. This step guarantees ex-post consistency of data and provokes the researcher's reflection on whether the analytical steps applied belong to one single analysis. If the material at hand is being approached in different ways or with different questions, separate analyses should be defined. Each separate analysis is later characterized by its working title in the tree-view and further described in a separate form. The name of the responsible researcher is entered as well as starting and end date of the analysis. File names and file links are provided in the form. [33]

More interesting in its potential to provide conceptual guidance is the expected characterization of the analysis via list fields. In the database's current form, the aim and the type of the analytical step are defined as well as the degree of its structuration. Initially, the new user of the database might need to adapt the lists that help to characterize analyses like the lists that previously served the characterization of surveys. The necessary adaption depends on the active user's dominant research paradigm. This should not be judged as a nuisance but rather as a possibility for some fundamental reflection concerning the specific research concepts applied. In the current version of the database, the content of the lists is mainly based on the qualitative content analysis and its description by MAYRING (2000, 2010). In light of this underlying paradigm, the analysis' aim may be described as the identification of "information and topics," of "mental states," or of "effects" in the database's current version. "Content analysis," "narrative analysis," "discourse analysis" or "thematic analysis" may be chosen as types of analysis. Content analysis may further be differentiated in its "explicating," "structuring" and "summarizing" forms. The content of the material at hand needs to be structured for an analysis. Form and degree of structuring may be described in the corresponding list-fields as "formally," "with regard to content," "scaling," or "typifying." Formal structuring may further be differentiated into "semiotic" and "rhetoric," and structuring according to content into "ideological," "logical" and "interpretative." Again, the need to select among the offered choices potentially provokes a desired reflection on the researcher's side. This reflection should be part of every research process, as the validity of questions and of analytical approaches, as well as the interpretation of answers and results, clearly depend on whether research centers on "objective" information or on the collection of world views, respectively "mental states," for example. Nevertheless, at least in applied research, awareness of these sometimes subtle differences is lacking in many cases, casting doubt on the validity of a researcher's conclusions. [34]

Each analysis may further be characterized in the tree-view by the topics of analysis and by the different steps that are taken in its course. Steps of data processing and analysis are defined in a separate form, stating the name of the responsible processor and starting and end time of the step. The step's content itself is selected in a list field. The list contains steps relating to the fields of data processing, of case summary and description, of text summary, of coding and code revision as well as of qualitative and quantitative or contextual analysis. The user is free with respect to the details he/she provides in the steps of analysis, but the possibility to document single steps can provoke reflection concerning the research process and sensitizes to the necessities of documentation. An assessment of the analysis, its potential and success, is provided in a list-field with pre-defined values, additionally there is room for notes concerning the analysis. [35]

The information that an observation of the survey has been analyzed within a specific analysis can be entered below the observation level in the tree-view. If all observations of the survey are analyzed in a specific analysis, this step can be omitted, if only a subset of observations is analyzed, the observation level definition adds important information to the database for the recapitulation of an

analysis. Document analyses that have been added to the tree-view can be assigned to specific analyses as well. If document analyses are related to specific analyses, the document analysis has to be defined anew if it provides information for another analysis of the survey as well. I expect this case to be the exception. Usually, each analysis will be supported by specific accompanying documents. On the document level, only the document analysis steps can be defined in the tree view. [36]

After the definition of the analyses and the data processing steps, the tree view summarizes the database' information as follows:

- Survey title
- Survey-level topics
- Survey-level analyses
 - Analysis-level topic
 - Steps of the analysis
- Research areas (regions or cases)
 - Area-level topic
 - Observation / appointment
 - Observation-level topic
 - Respondent / informant
 - Observation-level analysis (replicated for each step if desired)
 - Document analysis
 - Document analysis level topic
 - Documents
 - Steps of document's analysis [37]

6. Concluding Remarks

Based on the example of a specific small-scale database solution, the article showed how much information needs to be collected and systematically filed in order to provide a comprehensive documentation of the full research process. Comprehensive documentation is a necessary pre-condition for objective assessment, traceability and to support the reproducibility of the research process by a third person. While comprehensive documentation obviously comes at a cost, the article tried to show that it can also be of value for researchers, projects and institutions when an adequate support system is offered. Different personnel, administrative or technical support systems could in principle support researchers in their daily survey-related work, specifically in structuring and documenting their activities. The proposed small-scale meta-database in-house solution is inexpensive, flexible and easily to implement and could therefore be seen as a first step in raising awareness on a research process' factual complexity, the many decisions that need to be taken on its way and the corresponding need for detailed documentation. The discussion of the tool

showed that guidance in the documentation of research processes potentially delivers the following benefits:

- Initializing systematic reflection concerning the researcher's, the project's or the institution's scientific paradigms and general approaches;
- support in structuring an ongoing research process;
- enabling efficient documentation;
- support in the provision of effective, standardized information for secondary users. [38]

As the support system is also meant to raise researchers' awareness of the demands of a survey-based research project, the necessary initial adaption of the database's reference data should not be seen as a nuisance but as an opportunity to reflect on the own methodological and paradigmatic orientation, on central fields of research and on types of analyses typically applied. In order to adapt the database, researchers need to decide upon the types and forms of surveys they might happen to use in advance. Researchers using the database for documentation also need to classify and systemize their topics in such a way that the database afterwards allows them to concisely describe their field of research and the contribution of surveys, observations and analyses. Finally, researchers need to find classifications that allow them to characterize their analyses and analytical steps, for example by describing the aim and the type of an analytical step and the degree of structuration applied to the data at hand. [39]

If reference data have been adapted accordingly, they have the potential to effectively guide the structuring of an ongoing research process. Here, reflection is provoked, for example on whether documents may be used as equivalent data sources or as valid supplements for primary survey data with respective analytical steps and documentation. Reflection may also be motivated on whether specific research areas or observations deliver specific information or whether all observations form a single common pool for the analysis of the same questions. The possibility to add information for each appointment in a survey potentially fosters awareness of changing circumstances and their potential influence on data and data validity. [40]

The standardization of information in the database via list fields and reference data supports efficient information entry but it also supports effective searches for information on surveys, research areas and informants by secondary users. This might be a first step in enabling the secondary use of qualitative survey data, albeit initially within common institutions. This could pave the way for more general secondary use of qualitative data, if awareness is only once raised for the necessities and benefits of intensive meta-data documentation. [41]

References

- Bergman, Manfred Max & Coxon, Anthony P.M. (2005). The quality in qualitative methods. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, 6(2), Art. 34, <http://nbn-resolving.de/urn:nbn:de:0114-fqs0502344> [Accessed: March 4, 2014].
- Deppermann, Arnulf (2013). Interviews as text vs. interviews as social interaction. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, 14(3), Art. 13, <http://nbn-resolving.de/urn:nbn:de:0114-fqs1303131> [Accessed: February 12, 2014].
- Fink, Anne Sofia (2000). The role of the researcher in the qualitative research process. A potential barrier to archiving qualitative data. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, 1(3), Art. 4, <http://nbn-resolving.de/urn:nbn:de:0114-fqs000344> [Accessed: March 4, 2014].
- Guest, Greg; Bunce, Arwen & Johnson, Laura (2006). How many interviews are enough? An experiment with data saturation and variability. *Field Methods*, 18(1), 59-82.
- IASSIST (2010). Special issue: Qualitative and qualitative longitudinal resources in Europe. *IASSIST Quarterly*, 34(3/4)/35(1/2), <http://www.iassistdata.org/iq/issue/34/3> [Accessed: August 5, 2014].
- Iverson, Jeremy (2009). Metadata-driven survey design. *IASSIST Quarterly*, 33(1/2), 7-9, <http://www.iassistdata.org/iq/metadata-driven-survey-design> [Accessed: August 25, 2014].
- Kuula, Arja (2010). Methodological and ethical dilemmas of archiving qualitative data. *IASSIST Quarterly*, 34(3/4)/35(1/2), 13-17, <http://www.iassistdata.org/iq/methodological-and-ethical-dilemmas-archiving-qualitative-data> [Accessed: August 25, 2014].
- Mayring, Philipp (2000). qualitative content analysis. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, 1(2), Art. 20, <http://nbn-resolving.de/urn:nbn:de:0114-fqs0002204> [Accessed: February 17, 2014].
- Mayring, Philipp (2010). *Qualitative Inhaltsanalyse. Grundlagen und Techniken*. Weinheim: Beltz.
- Medjedović, Irena & Witzel, Andreas (2010). Sharing and archiving qualitative and qualitative longitudinal research data in Germany. *IASSIST Quarterly*, 34(3/4)/35(1/2), 42-46, <http://www.iassistdata.org/iq/sharing-and-archiving-qualitative-and-qualitative-longitudinal-research-data-germany> [Accessed: August 25, 2014].
- Slavnik, Zoran (2013). Towards qualitative data preservation and re-use—Policy trends and academic controversies in UK and Sweden. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, 14(2), Art. 10, <http://nbn-resolving.de/urn:nbn:de:0114-fqs1302108> [Accessed: August 5, 2014].
- Smioski, Andrea (2013). Archivierungsstrategien für qualitative Daten. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, 14(3), Art. 5, <http://nbn-resolving.de/urn:nbn:de:0114-fqs130350> [Accessed: August 5, 2014].

Author

Anne MARGARIAN is senior researcher at the Thünen-Institute for Rural Studies. Her interest focuses on firm-centered economics, innovation and entrepreneurship, structural change in rural areas and the critical assessment of rural policy approaches.

Contact:

Dr. Anne Margarian

Thünen-Institute for Rural Studies
Bundesallee 50
38116 Braunschweig
Germany

Tel.: +49 (0)531 596-5511

E-Mail: anne.margarian@ti.bund.de

Citation

Margarian, Anne (2014). One Bird in the Hand ...: The Local Organization of Surveys and Qualitative Data [41 paragraphs]. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, 15(3), Art. 13, <http://nbn-resolving.de/urn:nbn:de:0114-fqs1403130>.