

Friedrich-Alexander-Universität Erlangen-Nürnberg
Technische Fakultät, Department Informatik

REBECCA REUTER
MASTER THESIS

The Use of Domain Knowledge for Structuring and Describing Code Systems

Submitted on 11.02.2016

Supervisor: M.Sc. Andreas Kaufmann,
Prof. Dr. Dirk Riehle, M.B.A.

Professur für Open-Source-Software
Department Informatik, Technische Fakultät
Friedrich-Alexander University Erlangen-Nürnberg

Versicherung

Ich versichere, dass ich die Arbeit ohne fremde Hilfe und ohne Benutzung anderer als der angegebenen Quellen angefertigt habe und dass die Arbeit in gleicher oder ähnlicher Form noch keiner anderen Prüfungsbehörde vorgelegen hat und von dieser als Teil einer Prüfungsleistung angenommen wurde. Alle Ausführungen, die wörtlich oder sinngemäß übernommen wurden, sind als solche gekennzeichnet.

Röttenbach, 11.02.2016

License

This work is licensed under the Creative Commons Attribution 4.0 International license (CC BY 4.0), see <https://creativecommons.org/licenses/by/4.0/>

Röttenbach, 11.02.2016

Zusammenfassung

Domänenwissen ist ein häufig gebrauchter, aber selten definierter Ausdruck in der Forschung. Da jedes Forschungsprojekt in mindestens einer dedizierten Domäne durchgeführt wird, scheint der Einfluss von Domänenwissen natürlich, wurde bislang aber nicht untersucht, insbesondere nicht für qualitative Forschung. Durchgeführt wurde eine explorative Studie bei der acht qualitative Forscher aus folgenden Domänen interviewt wurden: Wirtschaftspädagogik, Informatik und Wirtschaftswissenschaften. Die Interviews wurden qualitativ ausgewertet mittels der Grounded Theory Methodology (GTM) als zugrunde liegender Forschungsmethode. Von besonderem Interesse waren möglicherweise vorhandene schematische Prozeduren, welche die Forscher während ihres qualitativen Forschungsprozesses und besonders während ihres Codierungsprozesses anwenden, und, wenn vorhanden, wie sie in der Domäne verwurzelt sind. Mit der Existenz eines domänenspezifischen Code Schemas innerhalb einer Domäne konnte die Existenz von den schematischen Prozeduren bestätigt werden. Weiterhin wurde ein Vorschlag entwickelt, wie auf Basis eines Code Schemas Artefakte für das Requirements Engineering in Form von User Stories generiert werden könnten. Außerdem ist ein Metamodell für ein domänenspezifisches Codesystem erstellt worden, um den Einfluss von Domänenwissen aus einer abstrakten Perspektive zu klären.

Abstract

Domain knowledge is a widely used but seldom defined term in research. Since every research project is located in at least one dedicated domain, its influence to the research work seems to be natural but is not investigated, especially not for qualitative research. An exploratory study was performed interviewing eight qualitative researchers located in different domains: economic education, computer science and business sciences. The interviews were qualitatively analysed using Grounded Theory Methodology (GTM). Of particular interest were possibly existing schematic procedures the researchers apply during their qualitative research process and especially during their coding process and if existing how they are rooted in the domain. With the set-up and reuse of a domain specific code schema in one domain the existence of schemata and its integration to QDA could be confirmed. Furthermore, a suggestion was given on how to generate artefacts for Requirements Engineering (RE) in the form of User Stories based on a code schema. Additionally, a code system meta model has been created to clarify the influence of domain knowledge from an abstract perspective.

Keywords: Qualitative Data Analysis, Grounded Theory Methodology, Requirements Engineering, Domain Knowledge

Content

Versicherung.....	2
License	2
Zusammenfassung.....	3
Abstract	4
Content	5
List of Figures	7
List of Tables	8
1 Introduction	9
1.1 Original Thesis Goals.....	10
1.2 Changes to Thesis Goals	10
2 Research Chapter.....	11
2.1 Introduction	12
2.2 Related Work.....	13
2.3 Research Question.....	14
2.4 Research Approach.....	14
2.4.1 Used Software	14
2.4.2 Data Collection.....	14
2.4.3 Data Analysis (Coding Process).....	15
2.5 Used Data Sources	21
2.6 Research Results	21
2.6.1 Approach towards a “Grounded” Theory as a Result of Selective Coding.....	21
2.6.2 Application in the Context of Requirements Engineering	21
2.6.3 Meta model.....	22
2.7 Results Discussion.....	23
2.7.1 Methodological implications.....	23
2.7.2 Limitations	23
2.7.3 Theoretical Saturation	24
2.7.4 Validation	24
2.8 Conclusion.....	25
3 Elaboration Chapter.....	26
3.1 Grounded Theory Methodology.....	27
3.2 Reflections on Grounded Theory Methodology.....	31
3.3 Code systems.....	32
3.4 Domain Knowledge.....	35

3.5	Application in the Context of Requirements Engineering	37
3.6	Quality criteria for qualitative research.....	37
3.6.1	Triangulation	39
3.6.2	Saturation	40
3.7	Further Work	41
	Glossary.....	42
	Appendices	44
Appendix A	Interview Guide	45
Appendix B	Codebook	47
Appendix C	Rules for Transcription	76
Appendix D	Axial Coding Overview	77
Appendix E	Intercoder Reliability	83
	References	84

List of Figures

Figure 1: The code system after the second round of open coding. The categories are visible. All categories with chevron contain subcategories respectively concepts..... 16

Figure 2: Code Matrix Browser. Viewable are the categories as rows and the grouped interviews as columns. 17

Figure 3: The Coding Paradigm by Strauss and Corbin (Corbin and Strauss, 2014) including the identified categories 18

Figure 4: The Code Matrix Browser in the MAXQDA project showing the most influence of the axial categories to the interview of FZ. 19

Figure 5: Selective Coding as diagram 20

Figure 6: Code system Meta model abstracted and adapted from (Salinger et al., 2008)..... 22

Figure 7: The Coding Paradigm for axial coding by Strauss and Corbin (Corbin and Strauss, 2014)..... 28

Figure 8: Template for the used code Memo..... 30

Figure 9: The process of how a theory is constructed out of codes (Saldaña, 2012). 34

Figure 10: Intercoder Agreement of the Interview with FZ. The appearance of codes was compared. 83

Figure 11: Intercoder Agreement of the Interview with FZ. The frequency of occurrence was compared. 83

List of Tables

Table 1: Coding principles presented by Glaser and Strauss. The four principles are shown in the left column followed by a short description of each principle in the right column. 27

Table 2: The coding process developed by Strauss and Corbin. The three steps are shown in the left column and a short description of each step is given in the right column. 28

Table 3: 10 Coding Families by Glaser presented in (Böhm, 2004). 31

Table 4: First Cycle Methods (grouped) and Second Cycle Methods..... 33

1 Introduction

1.1 Original Thesis Goals

The main goal of this thesis is to build a theory about the influence of domain specific knowledge on the qualitative research process of qualitative researchers in the domain of economic education. Within this field, the thesis focused on the building, use and reuse of schemata for introducing domain knowledge to qualitative researchers' code systems. A further goal is to evaluate whether the techniques of building, using and reusing of schemata can be transferred to the requirements analysis process using qualitative data analysis (QDA) and how this relates to a specialization of a meta model.

In this exploratory research I limit the influence of preconceived notions by starting out only with a clearly defined field of study and an open research question. GTM as underlying research method is chosen, since nobody knows the presence of any domain specific schema and a theory should be generated about this.

1.2 Changes to Thesis Goals

Using Grounded Theory, the research question results from the data.

Hence, the research question changes over the time and the field is expanded to qualitative researchers in the domains of:

- business sciences
- computer sciences

2 Research Chapter

The research chapter presents necessary information about the performed research during the Master's Thesis. It is written in common paper writing style (Heiser, 2014).

2.1 Introduction

Domain knowledge is a widely used but seldom defined term in research. Since every research project is located in at least one dedicated domain, its influence to the research work seems to be natural but is not investigated, especially not for qualitative research. Such a meta- research (research about research) is seldom operated and shows lots of gaps. Qualitative research is widely performed in the research domain of social sciences. More and more this approach is applied in other research domains like economic or computer science.

Qualitative research can be exploratory which means it is used to gain an understanding of underlying reasons, opinions and/or motivation regarding to a specified phenomenon. Collecting qualitative data varies using structured, semi-structured or unstructured techniques. An essential step in analyzing the collected data means to label (code) a statement in the interview text and memos, combining the codes in categories and abstract concepts (understanding patterns and regularity) and leads to a theory in the end. The theory describes the phenomenon of interest grounded in the gathered data. The concrete explanation (manifestation) of the phenomenon is detected while the research.

Any qualitative research is done in a problem domain with the domains own knowledge and the domains own language. Domain knowledge acts as a mental set of properties of a problem domain presented in a domains language. This leads to the area I want to survey: until now there is no investigation on how domain knowledge influences the qualitative research process and especially a qualitative data analysis.

Therefore, the existence, use and integration of domain specific reusable schemata is investigated. Code schemata are any standardized methods, techniques or models and their possible integration in the qualitative research process.

To investigate the survey area and to build a first theory on this, a qualitative research process is chosen.

Mayring described the grounded theory methodology (GTM) as particularly applicable for explorative studies where the object of study is new and unexplored (Mayring, 2002). According to Bortz and Döring GTM is an unprejudiced, inductive and open proceeding methodology which has explicit rules of thumb gaining to a fine linking of categories and subcategories (Bortz and Döring, 2006). I chose GTM proposed by Strauss et al. (1996) due to the following criteria of my research (Mayring, 2002):

- The research is explorative
- There is no research about the influence of domain knowledge to the qualitative research process
- Qualitative research is very individual
- Domains are nearly disjunctive

I contribute to qualitative methodological research on the existence and integration of a domain specific code schema (a standardized structure of codes) and also a theory on the (re)use of structured domain knowledge with the help of a code schema for a specific domain. I will also present a meta- model to a domain specific code system and evaluate the transferability to requirements engineering.

This paper is structured as follows: After an introduction relating research projects is given in chapter 2.2. The next section names the research questions that comes along (see: chapter 2.3).

The research approach describes the used software (chapter 2.4.1) and the data collection method with its implementation in chapter 2.4.2 and the analysis of the collected data (chapter 2.4.3). In chapter 2.5 the used data sources are explained. Having proposed the results (chapter 2.6) and a discussion (chapter 2.7), a short conclusion is given in chapter 2.8.

2.2 Related Work

I know no other research that investigates the influence of domain knowledge in the qualitative research process and also none that examines the influence of domain knowledge to a code system. All presented related work investigated a dedicated domain and not a general methods research.

All handbooks and known books about qualitative research methods and reflections about them are in some way related to the research described in this paper, since they all base their ideas for developing a new research method on special influences that force them to their decisions for new methods (Mayring, 2002), (Flick, 2013), (Strauss et al., 1996), (Saldaña, 2012), (Huberman and Miles, 1994), (Kuckartz, 2007).

The role of general and domain specific knowledge for the reasoning of undergraduate students about problems in genetics was examined by (Duncan, 2007). As survey method interviews were chosen. As a result, it was discovered that both forms influence the reasoning but crucial type of domain specific knowledge seems to affect reasoning by constraining the problem space, defining key components and mechanisms in the domain. Two forms of domain specific knowledge were identified: domain specific heuristics and domain specific explanatory schemas. These two forms were also put into a cognitive model together with mental models and propositional knowledge. This cognitive model is interesting regarding my research question since the existence of schemata was confirmed and I am especially interested in the manifestation of domain specific knowledge in a domain.

Looso et al. (2011) adapt the coding steps of GTM (open, axial and selective coding) to the method engineering process. The aim was to achieve an empirical grounding in the data and not just in literature reviews or construction oriented. The *Grounded Method Engineering* they developed introduced domain knowledge to methods and bridged the gap between methods and their context. They revealed a way to introduce domain knowledge to methods in general but they omitted the special influence to code systems. A concrete meta method is published in a later publication (Börner, 2011). Expert interviews were conducted (Looso et al., 2011).

A *Flexible Expandable Coding Scheme* (AF ECS) was developed in order to make studies comparable to each other. This model was developed using an integrated model of code comprehension and studies about that integrated model (Von Mayrhauser and Lang, 1999). This is an interesting approach to me since the authors must had an understanding about how to describe and structure code systems. A further point is the flexibility of the scheme which would imply that they achieved a code system which is mostly domain independent.

Kunz developed a domain analysis procedure using GTM (Kunz, 2015). Therefore, she used in-depth interviews. The applicability of GTM with some necessary adaptations in the requirements definition phase in software engineering was investigated. As a result, a code system meta model was derived regarding the adapted coding procedure for the domain analysis procedure. She modeled domain specific aspects in the usage of relationships as she specialized in structural and dynamic relationships and furthermore in the specialization of the coded segments in again structural aspects and dynamic aspects. In comparison to my meta model she doesn't consider domain specificity.

Salinger, Plonka and Prechelt developed a coding scheme for analyzing pair programming sessions using GTM (Salinger et al., 2008). This scheme included four practices: The first one was a perspective (here this means a decision what is important) on the data before starting the

coding respectively the analysis. Thereby, they adapted the suggestions from Strauss and Corbin 1990 to get the perspective at the time of selective coding. With the second practice they introduced a structured naming scheme for the codes. The concept name or code scheme was set up as *code* = *<actor>.<description>* whereas the actor could be one of the partners in a pair programming session or both and the description consists of a *<verb>_<object>* construct with defined sets for the verb and the object. The definition of this practice shows the structuring of domain specific knowledge and lastly the influence of domain knowledge to codes. The third practice the set-up of a meta model. They did it primarily for structuring their own insights with the naming in the software and to help themselves guiding through the analysis. This meta model influenced my meta model described in a later section. As last practice they performed pair coding. Therewith they wanted to achieve better understandings in the concepts and lastly better decisions for coding.

2.3 Research Question

This paper aims to build a theory based on the following superordinate research question:

- What is the influence of domain specific knowledge to the qualitative research process and especially QDA?

Moreover, this subordinated research question has been derived:

- Is it possible to derive a reusable code schema using a domain specific vocabulary?

2.4 Research Approach

My investigation focuses on researchers performing qualitative research in different domains and the impact these domains have on the respective research process. Since I didn't want to conduct literature reviews on this but want to ground to the data I chose qualitative interviews and therefore I have to clarify the role of prior existing knowledge first. Some qualitative research method descriptions recommend to have no prior knowledge before the study starts because the open minding to the object of investigation might be influenced. Also among these is the GTM by Glaser and Strauss (Glaser, 1967), (Strübing, 2007).

One of the domains is resided in computer science (my home domain), but the special practice of pair programming investigated only rudimentary known. This also applies for the special area of open source communities, where another computer sciences interviewee has the research focus.

2.4.1 Used Software

For analyzing qualitative data software can be of advantage. Few software products, short CAQDAS (computer assisted qualitative data analysis software), offer computer- assisted QDA, but it is recommended to carefully decide which software is going to be used for the analysis. Thus, the software has to be chosen having regard to the type of analysis and to the working style of the researcher (Flick, 2013).

I chose MAXQDA 12 since the research group has good recommendations on that. Further reasons were the included transcription mode and the export feature of a codebook.

2.4.2 Data Collection

The GTM does not stipulate the method of how the data should be collected. Strauss and Corbin talked exemplary about observations or interviews (Mey and Mruck, 2007).

I chose expert interviews for the data collection. Expert interviews are a sub group of semi-structured guided interviews where the interview partner is an expert and holds special expertise regarding the research question. They can be used under various purposes:

- Explorative: to generate hypotheses and to get orientation in a new field, so that the investigation area can be structured
- Systematic: to provide context information
- Theory-generating: to develop a theory or typology for an object by reconstructing the knowledge of different experts (Flick, 2013)

The interview in itself is managed by the interview guide, whereby the challenge is to limit the interview to the intriguing expertise. Possible failures can occur if the expert blocks the interview, misses the subject, the expert changes the role or also if the expert gives a lecture (Flick, 2013).

The interviews are guided by a catalogue of leading questions. Therefore, the main topics, I want to gather information on, were identified: Domain specific knowledge, qualitative research at chair, schematization and coding.

The leading questions were used as a guideline for the interviewer and were adapted individually to the answer of the expert. Depending on the given answers some questions were excluded or additional targeted questions were asked.

The interview guide (see: Appendix A) was adapted after the first three interviews in order to improve the quality of the interviews. Thereby, the biggest change was to put the section about the domain knowledge to the top in order to get the first ideas of the researchers after an introduction of the thesis.

2.4.3 Data Analysis (Coding Process)

The vital step in QDA is called coding (Saldaña, 2012), (Flick, 2013). The following chapter describes my analysis process referring to the above described GTM approach. I performed the analysis on my own with reflexive talks about the codes to my supervisor.

The GTM terms used are listed and defined in a glossary (see: chapter Glossary).

Starting with conducting the first interviews they were transcribed and open coded. Having theoretical sampling in mind new interviews were conducted and after that the axial coding started. There were some iterations between open coding and axial coding. The theoretical sampling was not performed as prescribed by the GTM (see: chapter 3.1). After the axial coding the selective coding was performed terminating in a grounded theory (GT).

Open Coding

Open Coding is the process of developing categories of information and to be open minded to the data. I started my open coding without any predefined categories.

At the start of the coding the first two interviews have already been conducted. The first round of open coding was performed in the same order as the conducting. As a first step I reread the interview. I coded everything that seems to be interesting- as the open coding of Strauss and Corbin (chapter 3.1) suggests it. I did this procedure for each interview. This resulted in a high number of different codes and a coding tree with only one level which was unstructured. Furthermore, duplicates were included. After this first inductive step I decided to proceed a deductive step as second round of open coding in order to keep an overview of the codes. Therefore, all codes respectively the code tree were refactored constantly. The categories were created and all interviews were recoded when necessary. This refactoring resulted in a leveled code tree, since some concepts could be collected to categories. All in all, 20 categories were identified excluding subcategories. 170 categories were developed and 589 codes were set.

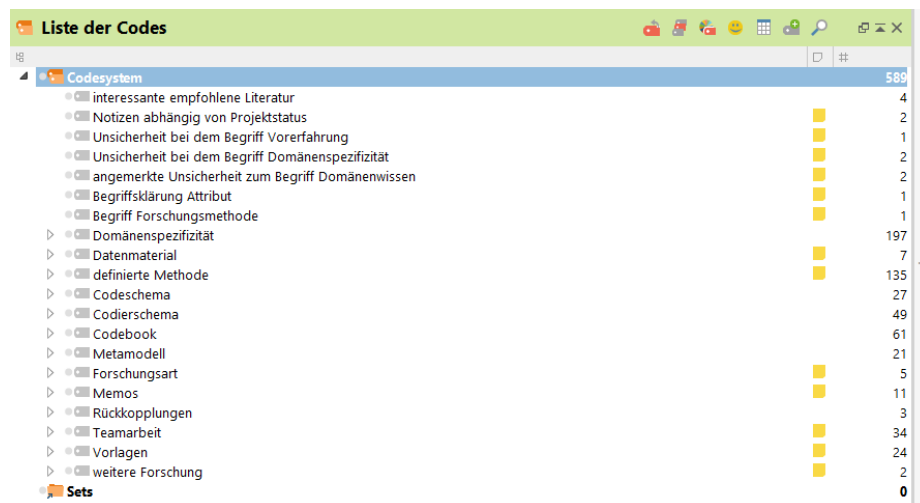


Figure 1: The code system after the second round of open coding. The categories are visible. All categories with chevron contain subcategories respectively concepts

Codebook

During the whole coding I wrote a code memo for each code. Therefore, I used a certain structure as described in chapter 3.1. MAXQDA 12 offers a feature to export a codebook containing all code memos. The codebook improved the quality of the codes since I myself had to define clearly when I wanted the code to be used and when explicitly not and a definition. Sometimes this is quite hard and could be confirmed by more manpower and not in context of a master's thesis. The whole codebook is listed in Appendix B.

Theoretical Sampling

Due to missing access and some other limitations (see: chapter 2.7.2) the theoretical sampling as Strauss and Corbin suggested could not be performed in the same way as I had to do it.

After conducting the first two interviews with interviewees from the economic education domain without any results regarding the existence and integration of schemata in the qualitative research process, I decided to include more domains than one. Following this, the domain of computer sciences was included and three interviews with qualitative researchers in these domains were conducted (AB, 2015), (FZ, 2015), (SS, 2015). All three interviewees confirmed the use of a coding scheme, two of them also the use of a code scheme and the reusability of codes.

The use of a coding scheme means having a schema of how to code (underlying rules or best practices). The use of a code scheme means having a fix structure of codes. A coding scheme might imply the use of a code scheme.

The two interviewees who used a code scheme were located in the domain of software engineering with the research area pair programming. Unfortunately one of the two researchers left the research group already, so that no more interviews could be performed (SS, 2015), (FZ, 2015). In one more interview in the domain of economic education the interviewee had initial approaches that seemed to be a coding scheme and a code scheme (KDS, 2015). These four interviews are mainly used for building the theory.

Two more qualitative researchers in the domain of economic science were interviewed at last but they could not confirm the existence or integration of schemata or reusability of codes (HH, 2015), (JJ, 2015).

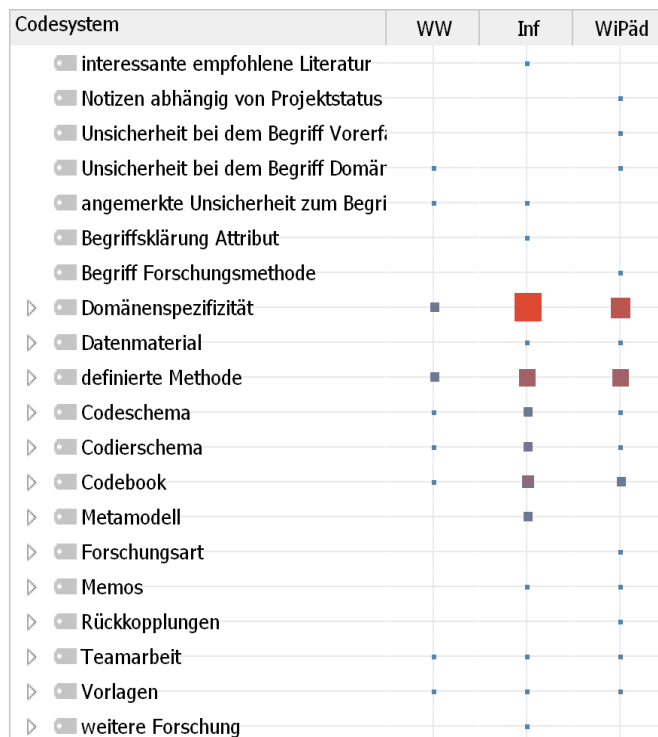


Figure 2: Code Matrix Browser. Viewable are the categories as rows and the grouped interviews as columns.

To verify this statements with the data I used the Code-Matrix Browser feature from MAXQDA (see: Figure 2). The Code-Matrix Browser lists categories (vertical axis) and frequencies of the appearing codes. The bigger each bullet the more codes are set in the interview group (horizontal axis) for each category. With this a comparison of the interview group relating to the categories and category pattern is possible.

Axial Coding

The axial coding was started when all interviews were conducted and the two rounds of open coding were completed. The axial coding is a process for interconnecting the categories by the coding paradigm as described in chapter 3.1.

To get closer to the research question the (re)use of domain knowledge was chosen as central phenomenon. Using the previous findings from the two rounds of open coding the arrangement of the axial categories in the coding paradigm was developed (Figure 3).

Therefore, the code tree has to be revised, some categories must be reorganized. The reason or trigger was the (re)use of a coding schema (nr. 2 in Figure 3), the context was the qualitative research process (3), auxiliary concept was the identification of domain specific attributes (4), the strategy was the use of a code schema and a domain specific meta model (5) and the consequence was determined as the domain specific reuse of a codebook (6).

To confirm the selection of the axial categories, I chose the fourth interview and explained each category according to its statements and interpretations on them. Since semi-structured interviews with different focus depending on the interviewee were conducted, these interviews were very different in their impact. The axial coding for each interview is listed in a table (see:

Appendix D). The axial categories including the paradigm are listed per row, the interview per column. If there was no impact there is always a “-“.

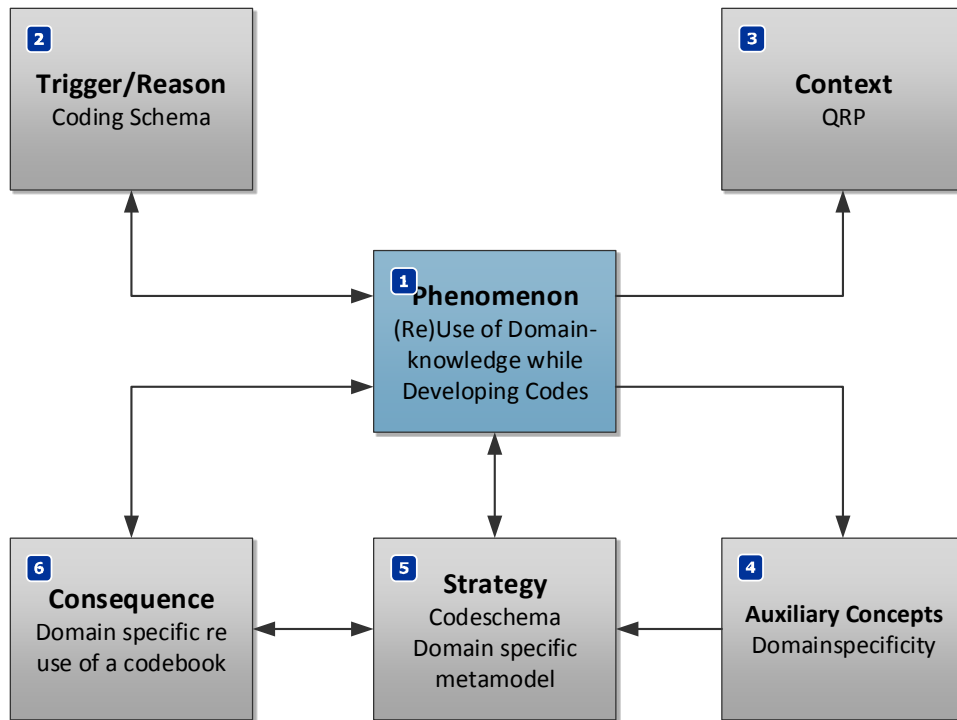


Figure 3: The Coding Paradigm by Strauss and Corbin (Corbin and Strauss, 2014) including the identified categories

The fourth interview was chosen because it had the most influencing codes to the identified categories (Figure 4), (FZ, 2015).

1) The Phenomenon: (Re)use of domain knowledge while developing codes

FZ works in the domain of pair programming and investigates pair programming sessions using a qualitative research process (GTM). Therefore, he uses and reuses a defined vocabulary. This vocabulary is called base concepts and fully standardized. This standardization precipitates in a code schema. This code schema is only applicable for pair programming sessions. For his coding he uses furthermore a coding scheme consisting of four practices including the code schema and a meta model. The base concepts are a kind of codebook and can be reused for analyzing pair programming sessions.

2) The Trigger/Reason: Coding Schema

FZ follows a coding schema consisting of four defined practices. This implies the usage of a code schema and the existence of a meta model which should show him new ways of coding sessions. The existence was a reason for the phenomenon for myself.

3) The context: Qualitative Research Process

FZ does his analysis in the context qualitative research. Therefore, he uses the GTM.

4) Auxiliary Concepts: Domain specificity

FZ's annotation of the base concepts (standardized codes) can only be applied in his special domain where pair programming sessions are analyzed. He knows that and doesn't apply them in other domain analysis.

5) Strategy: Code Schema and domain specific meta model

FZ does the analysis having the meta model in mind. Preparing his analysis he uses the code schema which looks like the following: code = <actor>.<description>. The description itself consists of a <verb>_<object> construct.

6) Consequence: Domain specific reuse of a codebook

The base concepts consist of defining a code and a code memo for each concept. Each code has a domain specific structure, which can only be used for this domain but this enables the reuse of them.

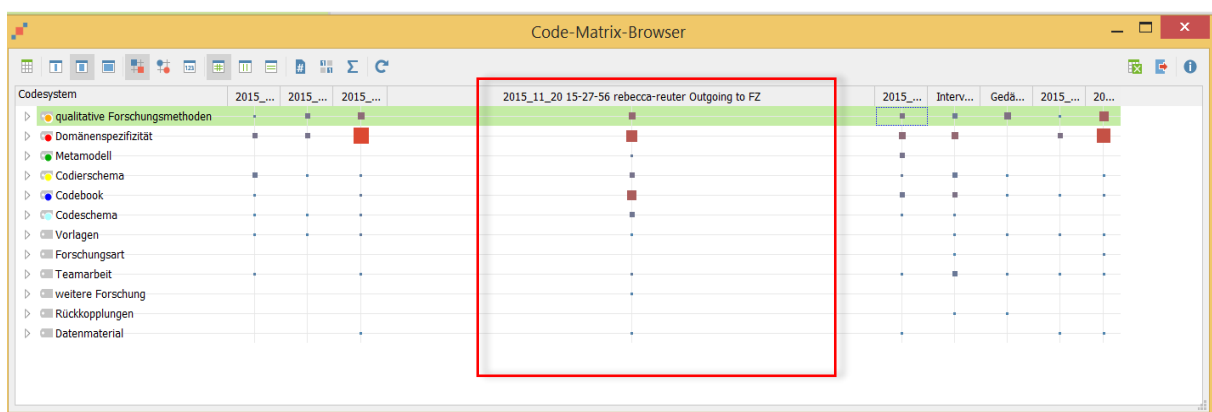


Figure 4: The Code Matrix Browser in the MAXQDA project showing the most influence of the axial categories to the interview of FZ.

Selective Coding

After the axial coding the selective coding is performed.

Selective coding is a process for building a story that connects the categories producing a discursive set of theoretical propositions around a chosen core phenomenon.

For the second part of the selective coding (the influence of domain specific knowledge in QDA) I had to limit the interviews to the two researchers that define and reuse a code schema otherwise this could not be performed.

The resulting core categories of selective coding are illustrated as a diagram in Figure 5 since this a good method to visualize the new relations that arise. Strauss and Corbin suggested that instead or additional to memo writing (Strauss et al., 1990).

During the axial coding “domain specificity” seems to be the auxiliary concept respectively the intervening condition to the core phenomenon “the use of domain knowledge” which is the underlying question for this thesis. But actually this is the most interesting category because it influences the steps before the QDA starts and the QDA process itself (Figure 4 also shows domain specificity as the biggest bullet point). Therefore, the role as core phenomenon is set to domain specificity.

The following section describes the story around the core phenomenon:

Prior knowledge and/or experience of the researcher in the domain is available for any research project. Any researcher has some kind of prior knowledge that influences him in his decisions.

If a researcher wants to experience something new (and this is natural to researchers) he needs to gather more domain knowledge. Therefore, he has several techniques. Some are domain independent like gathering domain specific information about the current state of research- and some are domain specific such as gathering information by mailing lists. After a researcher has completed gathering his domain knowledge, he is able to define a research question depending on his research interests including always domain specific knowledge. To start the investigation regarding its research interests he has to know the field structure and entries to that field. Only if these domain specific properties have been detected he chooses the data collection method. Depending on the data collection method and the research interest he chooses the data analysis method.

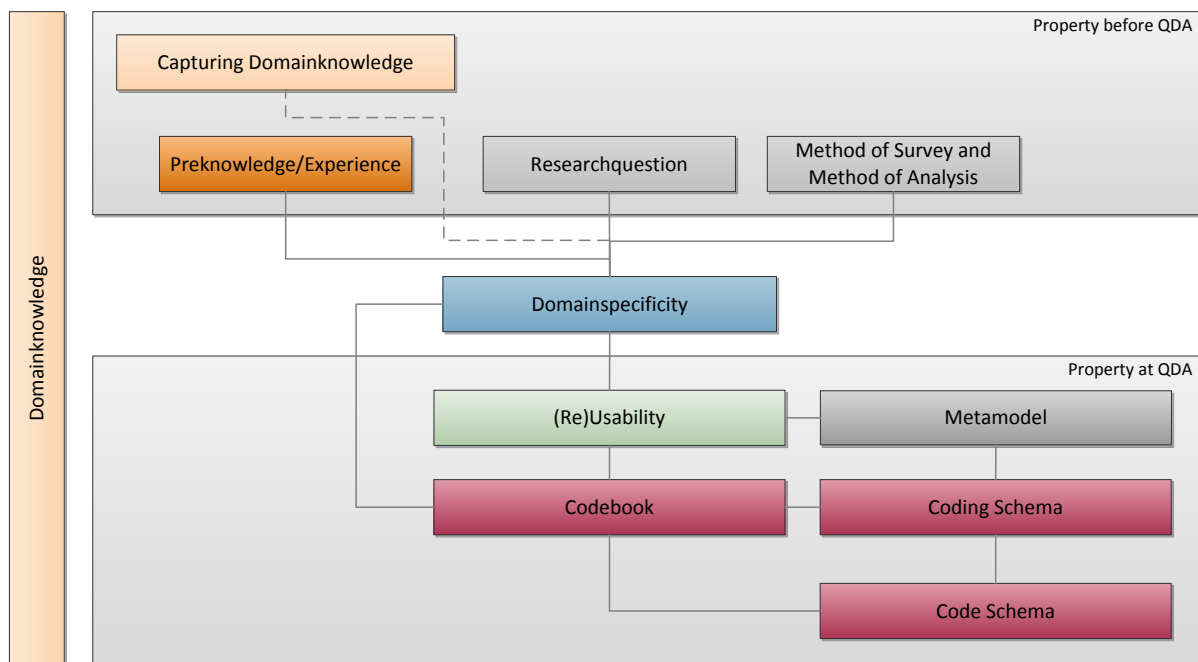


Figure 5: Selective Coding as diagram

After collecting the data, the analysis starts.

For this he annotates codes to data segments. These codes are again domain specific, since these could be in-vivo codes or even if they aren't he uses a special *domain specific vocabulary*. This is the starting point of the theory on the influence of *domain specific knowledge* to QDA: In general there are two ways of constructing codes- the deductive and the inductive approach (Saldaña, 2012). Both ways have in common that the codes are built upon a domain specific knowledge base and based on a *coding schema*. The first researcher I want to consider is the one who developed the standardized vocabulary. He used an inductive approach to create codes by using a defined *coding schema*. His coding schema contained the construction of a meta model and also a *code schema*, both in order to structure the amount of data and in order to gain reusability. All the structured codes resulted in a codebook, including code memos for each code to help further researchers to reuse them correctly.

The second researcher reused the developed codebook to get an entry to the data. Therefore, he annotated the codes as described in the codebook. After annotating the domain specific codes, he was able to start his personal coding. He performed his own coding by using the coding schema. A reuse will be possible since the researcher will work in the same domain.

2.5 Used Data Sources

This section includes information about the physical data collection and transcription.

In order to construct a theory about how researchers construct theories in their research fields and how they deal with domain knowledge I conducted eight semi-structured interviews from different areas in social, economic and computer sciences. All researchers have at least carried out one qualitative research project. The interview length was set to about 45 minutes and varied between 20 and about 60 minutes. All but one interview was conducted via Skype™ and later anonymized. One interview was conducted personally. Due to recording problems one interview is partly transcribed and partly written as verbatim from memory. The interview partners had the choice to do the interview with video or not. All but one interview was conducted in German, one in English.

After the third interview the interviewee got a short guide of the topics of the interview, in order to inform the interviewees beforehand and to achieve quality improvement of the interviews.

The interviews were audio recorded via an open source Skype™ recording tool. The transcription of the interviews was performed by MAXQDA, a common software product for qualitative data analysis.

For the transcription a simplified transcription system was used (see therefore Appendix C). After the first transcription the introducing and the closing section is left out, since this is a common practice and there was no applicable information available for us.

The analysis was started without using literature in order to be open to the collected data. In a later step, for the creation of the meta model further literature was consulted. To keep to the data I used to a recommended paper of one of our interviewees (Salinger et al., 2008).

2.6 Research Results

This section includes the analysis as well as the arising results.

2.6.1 Approach towards a “Grounded” Theory as a Result of Selective Coding

The theory is nothing else than “an expression hidden in the order of the data” (Glaser et al., 2010). But Grounded Theory is not a perfect product, it is understood as to be in “permanent development” (Glaser et al., 2010). Based on this view following an approach to such a GT according to the research question is presented.

Domain specific knowledge influences the steps before QDA and also the QDA itself. The influence of the QDA is especially visible if the used coding is performed, having a noted meta model in mind and having a documented code schema. Having a documented meta model and code schema implies having identified a domain specific vocabulary. These aspects also include that they are reusable for the investigated domain.

2.6.2 Application in the Context of Requirements Engineering

This section presents a first approach to gather and define a standardized vocabulary respectively a standardized code schema for requirements engineering. Therefore, I adapted the code schema that is used and developed by two of our interviews where I derived the theory. Their code schema is presented in (Salinger et al., 2008).

User stories capture the “who”, “what” and “why” questions of a requirement. They will be a first approach that might be used, since they are an interface from verbal to written communication to elicit requirements in agile projects. It is necessary to represent the customer requirements instead of only documenting them. Mike Cohn suggests a template that seems to be applicable respectively adaptable for our standardized vocabulary (Cohn, 2010):

“As a <type of user>, I want <some goal> so that <some reason>”

Based on this two existing approaches I suggest the following code schema:

code = <type of user>_<goal>_<reason>

type of user = {domain specialist, requirements engineer, customer}

goal = I + want + <goal_verb>

goal_verb = {to create, to select, to see + something}

reason = I|we + can + <reason_verb>

reason_verb = {start, participate, estimate, answer}

As Salinger et. al (2008) developed it for their research interests in pair programming sessions and the template of Mike Cohn suggests it, my code is structured as a complete sentence. Such syntactic rules might help at the analysis of workshops etc. to derive user stories in a more formal way to save time since the same user story would be modeled multiple times otherwise.

This is only a suggestion set of the variables (type of user, goal, and reason). The set should be defined in a further work. My approach is valid since Strauss and Corbin suggest the use of literature in the application of GTM (Corbin and Strauss, 2014).

Concluding this section this code scheme can be seen as an artefact which is derived during a RE process.

2.6.3 Meta model

To clarify the influence of domain knowledge during the QDA process I set up a code system meta model. One of my interviewees already developed a meta model for his research and confirmed the adaptability of his meta model to problem solving communication analysis (SS, 2015). To keep grounding in the data I analyzed the interviews again regarding information to the meta model and read the paper recommended by an interviewee in addition (Salinger et al., 2008). I generalized his meta model and implemented the influence of domain knowledge and presented this in Figure 6. It is represented as UML class model in conceptual view (Rumbaugh et al., 2004).

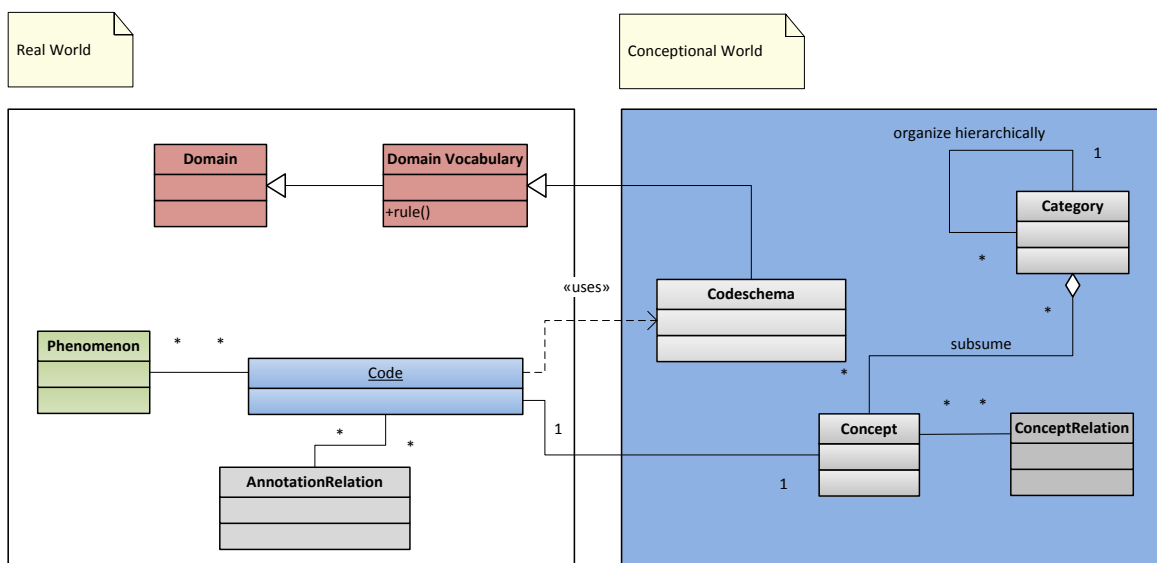


Figure 6: Code system Meta model abstracted and adapted from (Salinger et al., 2008)

I differentiate here between real world and conceptual world whereas the real world contains the text segments respectively the identified phenomena that are labeled by a code. Code is located here since it is directly located to the data.

As I am interested in the influence of a domain this is located in real world. The domain and its domain specific vocabulary take place in real world, whereas the domain vocabulary is derived by the specific domain. Standardizing the domain specific vocabulary using rules makes it possible to derive a code schema. For the setup of a code schema a dependency on the domain vocabulary is necessary.

The code in principle is set up using the code schema. The conceptualization of codes is concept. Concepts can be subsumed to categories. Categories are organized hierarchically.

Considering the coding paradigm used in GTM concepts are in relation. This is represented by ConceptRelation. If such a relationship is not valid for the codes that use these concepts this can be expressed by AnnotationRelation.

My focus on creating the meta model was to point out the influence of a domain or domain knowledge to a code schema and the codes schema rudimentary whereas the code schema is the necessary to build a structured code system. Furthermore, I created the meta model applicable for an inductive, as well as for a deductive or hybrid approach.

2.7 Results Discussion

This section presents general limitations, saturation, validation and methodological implications on the performed research.

2.7.1 Methodological implications

Understanding the influence of domain knowledge and the therewith connected standardized vocabulary has some methodological implications consequently. The standardization of a vocabulary regarding complexity only makes sense if this is reusable. Considering reusability and a possibly derived code schema from the standardized vocabulary this conflicts with the idea of GTM by Glaser and Strauss or the objective hermeneutics. Using the GTM by Glaser and Strauss a researcher has to avoid prior knowledge (and consequently the influence of domain knowledge) and existing theories. This would imply to avoid the existence of a standardized vocabulary. The reuse of a code schema during a GTM by Strauss and Corbin is considered worthy of discussion. Generally, the inductive approach to set up code systems seems to be more problematic than a deductive or hybrid approach where the basic categories are set before the coding starts.

Since the influence of every kind of prior knowledge is worthy of discussion in qualitative research, a possible approach using GTM by Strauss and Corbin and reusing a standardized vocabulary is given by one of my interviewees. He reused a code schema prior to his own coding process in order to get a better entry to the data. After this prior step it is possible that codes keep during the open coding but also that new codes arise.

2.7.2 Limitations

Regarding limitations of this explorative research, one very pragmatic aspect has to be considered. This master thesis is limited to a time frame of six months. This is a very short time to develop a GT using GTM including all coding steps and having saturation and validation in mind, which I describe in chapter 2.7.3 and chapter 2.7.4. Another limiting issue was the number of participants and the actual prescribed theoretical sampling when using GTM by Strauss and Corbin. Theoretical sampling is not applied as described by them due to limited access to interview partners.

After the first two interviews in the domain of economic education I decided being domain independent since the two researchers have no code schemas and no more were available. The domain independence had further consequences. For our theory generation I used two interviews located in one domain basically since our theory shows that the use of codes is very domain specific. Lastly no more researchers using a code schema were accessible to me.

2.7.3 Theoretical Saturation

Strauss and Corbin explain in their FAQs as a request to the answer of how many interviews or observations are enough that the researcher has to collect and analyze data until theoretical saturation is achieved meaning all categories have to be saturated and variations have to be built in. "Sometimes the researcher has no choice and must settle for a theoretical scheme that is less developed than desired" (Corbin and Strauss, 2014). This includes constraints like time and availability of participants. Besides, the skill of the interviewer has an effect on the quality of the data and furthermore in achieving saturation (Morse, 2008), (Guest et al., 2006). The sample size could not be referenced by saturation since the quality of the data is the measurement variable and this depends on the interaction of interviewer and interviewee (Mason, 2010).

A stringent approach of theoretical sampling and coding leads to theoretical saturation at a certain point of the analysis. Due to the named underlying limitations (chapter 2.7.2) a theoretical saturation can only be suspected after the sixth interview where no new insights aroused.

2.7.4 Validation

The qualitative research approach is validated using the quality criteria by Steinke (2004) and a kind of investigator triangulation. Triangulation¹ can be used as a criterion for validation and it was performed here in order to avoid subjective influences of one researcher (Denzin, 1978), (Flick, 2000).

In order to validate my codes, the intercoder reliability method was applied. Since this document is a master thesis, another researcher has coded about 10% of the conducted interviews (i.e. one interview) again using the existing code system in order to show intercoder reliability. Comparing the two coded interviews an intercoder agreement can be derived. The percent agreement of code appearance resulted in 94,15% (see: Figure 10), the percent agreement of the frequency of occurrence resulted in 90,64% (see: Figure 11). Detailed results can be found in Appendix E. Primarily, this method shows a way that leads to codes that are more precise and compact.

Steinke points out that her defined criteria are a catalogue of criteria but the criteria and test procedures should be specific for each application. This may mean to substantiate, modify or supplement the catalogue by other criteria. The catalogue consists of following seven criteria: Intersubjective comprehensibility, indication of the research process, empirical foundation, limitation, coherence, relevance, reflected subjectivity (Steinke, 2004), (Flick, 2013). Following I will take a stand for some of the criteria Steinke has defined here, exemplary (Steinke, 2004), (Steinke, 1999). Each criterion of the catalogue was applied successfully.

Intersubjective comprehensibility

I documented all the necessary information (see: chapter 2.4). Verbal statements of the interviewees are transcribed into a MAXQDA project as well as the records of the meanings of the interviewees' statements.

¹ Triangulation is defined as the view of a research object from at least two points (Flick, 2000).

The second way is the interpretation in groups. This is a discursive form of achieving intersubjectivity and comprehensibility. Therefore, the intercoder agreement of the coding (see above) has been performed (Steinke, 2004).

The third way, called “the use of codified procedures”, is given as the GTM serves as codified procedure (Steinke, 2004).

Indication of the research process

The indication of the qualitative procedure is given as I describe my motivation in the introduction (see: chapter 2.1). As the choice of methods is legitimated and weighed in the research approach also the indication of them is given. The indication of the transcription rules is given since they are simple to use for the transcription, easy to learn and readable. Since the sampling was not performed as prescribed by GTM this is not indicated. Individual methodological decisions are indicated. It is a common practice to analyze semi-structured interviews by GTM. The evaluation criteria are indicated since they are derived from the catalogue.

Empirical foundation

To test the empirical foundation Steinke suggested for example the usage of codified methods. By using them empirical foundation is given (Steinke, 2004). Since I use GTM the empirical foundation is indicated.

Limitation

For the criterion limitation the conditions have to be clarified which have to be fulfilled minimal so that the phenomenon occurs (Steinke, 2004). In my context this is the standardized domain specific vocabulary that has to be deducible to create a code schema. Contrasting the cases, this was true for two interviews. All other interviews don't explicitly show this.

Many of the suggested criteria are applicable for the underlying research. Consequently, I can summarize as having achieved the best possible result.

2.8 Conclusion

This thesis has shown the existence of coding schemata and above all the existence of a domain specific code schema and their integration and influence to the QDA process. To explore this phenomenon, I used the GTM by Strauss and Corbin and due to constraints with deviations in the sampling. I could thus derive an approach to a theory on the influence of domain knowledge to QDA and to the qualitative research process as such. The theory especially says that domain knowledge influences the setting of code systems. If the vocabulary of a specific domain can be standardized, a domain specific code schema can be derived.

I showed an approach to the applicability of a code schema to requirements engineering especially in agile context. Therefore, the user story template by Mike Cohn is used. The code schema is set up in the style of the explored code schema. This has to be validated in further work.

The developed meta model shows a specialization of a code system meta model with the extensions of the domain specific influence viewable in two classes in real world and the derivable code schema in conceptual world. Moreover, the different approaches of setting up code systems (inductive, deductive, hybrid) were considered. Developing the meta model I also want to keep to the data and abstracted the meta model mentioned by two of the interviewees.

Summing up, the validation shows that quality criteria were fulfilled and I achieved an approach to a grounded theory on the existence and integration of code schemata which are applicable in domain specific context.

3 Elaboration Chapter

The elaboration chapter is not a repetition of the research chapter in more detail, but an elaboration of all issues that seem noteworthy but didn't make the cut for the research chapter.

At first the GTM approach as prescribed by Strauss and Corbin is presented (chapter: 3.1) and some reflections on this are given in chapter 3.2 code systems and domain knowledge are described in more detail in chapter 3.3 and 3.4. Chapter 3.5 gives more explanation on the possible application of a code schema in the context of RE. More information on quality criteria used in qualitative research are presented with special focus on triangulation and saturation (as quality criterion) (chapter 3.6). The elaboration chapter is finished with some thoughts on possible further work (chapter 3.7).

3.1 Grounded Theory Methodology

GTM was first developed by Glaser and Strauss during a research project in the domain of medicine sociology but they split up into to schools of GTM. The GTM school by Strauss and Corbin is more straightforward and more expatiated than the GTM by Glaser. For more background information regarding the reasons for the split and the differences Strübing gives elaborations (Strübing, 2011).

Before the GTM by Strauss and Corbin (Strauss et al., 1990) is presented, I first present the original coding principles by Glaser and Strauss since they are often presented in papers using GTM. The coding process includes four principles: iterative proceeding, theoretical sampling, constant comparison and ignoring existent theories. These principles are guidelines for the researcher during the whole research process.

Table 1: Coding principles presented by Glaser and Strauss. The four principles are shown in the left column followed by a short description of each principle in the right column.

Principle	Short Description
Iterative proceeding	There are iterative steps between data collection, data analysis and theory generation to achieve theoretical saturation.
Theoretical sampling	Based on the data analysis the case selection will be approved iteratively and new data will be selected if necessary to push the generation of the theory.
Constant comparison	Constant comparison is a way of maintaining a connection between the code and the data. Every time a new segment is coded, the researcher should think about all the data coded in the same way to push the generation of the theory.
Ignoring existent theories	To be open minded to the data, the researcher should ignore all existing theories regarding the surveyed area.

The following explanations align with the GTM defined by Strauss and Corbin where all but the principle of ignoring existent theories² are included in a structured process of three steps (Strauss et al., 1990), (Corbin and Strauss, 2014):

Table 2: The coding process developed by Strauss and Corbin. The three steps are shown in the left column and a short description of each step is given in the right column.

Principle	Short Description
Open Coding	A process of developing concepts of information, being open minded to the data.
Axial Coding	A process for interconnecting the categories by a coding paradigm.
Selective Coding	A process for building a story that connects the categories around a core phenomenon producing a discursive set of theoretical propositions.

All three steps do not have to be performed in a real sequential way. This is reasonable because of the iterative data collection, since the analysis steers the sampling.

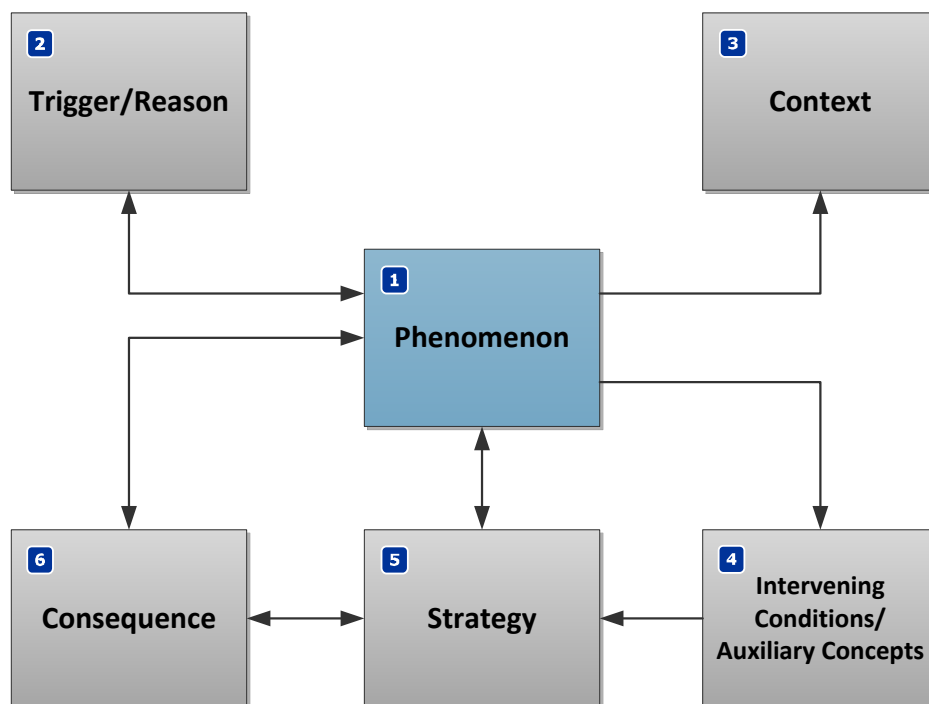


Figure 7: The Coding Paradigm for axial coding by Strauss and Corbin (Corbin and Strauss, 2014)

² Strauss and Corbin have an opposing view regarding this principle. This is explained after in the paragraph about the Recommendations on Theoretical Sensitivity, prior knowledge and the influence of literature.

Open coding

Open coding is the “process of breaking up, investigating, comparing, conceptualizing and categorizing” and a first tree of codes arises (Strauss et al., 1996). The process of breaking up is the segmentation of the data into smaller units, which can be words, sentences, sections or episodes. Investigating means the identification of interesting units (phenomena). The naming of this unit is called conceptualizing. The concepts should be described. The description is put into a code Memo. The process of comparing means the constant comparison of already identified concepts to new identified units, so that similar phenomena get the same concept. Categorizing is the step where the researcher categorizes similar concepts to more abstract categories by comparing existing concepts and categories. This is also called dimensionalization. This step should also help to perform a more effective theoretical sampling. All described steps go hand in hand (Strauss, 1987).

Axial coding

Open coding is followed by the axial coding. During this step the before identified categories are set into relation with the data and are composed in a new way (Strauss et al., 1996). Therefore, Strauss and Corbin suggest to use a coding paradigm (see: Figure 7).

The focus is the phenomenon (here colored in blue and marked with 1) which is described in a category. The leading question for the researchers are: What is at issue?

The phenomenon should be examined regarding all five characteristics. These characteristics are displayed by the axial core categories. For every axial category at least one leading question can help the researcher to identify it in the data. To identify the trigger or reason (numbered with 2) the question “What leads to the investigated phenomenon?” has to be answered. The context (3) can be found by the following questions: What are the manifestations for the current issues? What the conditions for the strategies? The intervening conditions or auxiliary concepts (4) might be identified using the question about the general preconditions about the strategies. The category of the strategy (nr. 5) can be defined by the question, how the phenomenon is handled by the actors. The last category (6), consequence, can be identified by the leading question: What leads to the resulting actions and strategies related to the phenomenon?

The researcher performs the axial coding by questioning regarding the above axial core categories and by making comparisons. This setting of relations offers a systematic inductive and deductive thinking so that a more saturated theory arises in the end.

Selective Coding

To generate a theory, the selective coding has to be performed after the first two steps. A core category is chosen and is related to other categories based on the paradigm. The resulting relations are validated. By the creation of a red thread of the story the researcher tries to translate the results into an analytical story and to generate a theory out of it. This means- alike the open and axial coding- the core phenomenon (core category) has to be named und must be put in to relation to others. The category can already exists and then gets the role of the core category or a core category must be created (Strauss et al., 1996).

Recommendations on Theoretical Sensitivity, prior knowledge and the influence of literature

As already mentioned above, one coding principle in GTM by Glaser and Strauss is the “ignoring of existent literature”. Thereby the role of prior knowledge remains vague. The GTM by Strauss and Corbin on the contrary recommends the integration of prior knowledge. They also name literature and professional and personal experience as source of theoretical sensitivity

(Strauss et al., 1996)³. Thus, they explicitly recommended the use of literature. Besides the improvement of the theoretical sensitivity this literature may be used for a further analysis with GTM methods. The influence of literature is not limited to a special step.

This does not mean to collect interesting theories or models prior to the work. The theory should be discovered and not be verified. Therefore, I initially started the analysis without the extension of prior knowledge of the underlying domains.

The use of Memos and Diagrams

Strauss and Corbin also recommend the use of memos and diagrams. They distinguish between code notes, theoretical notes and operational notes.

Strauss and Corbin don't give a structure of each Memo type but for my code notes I followed a defined structure (Strauss et al., 1996), (Figure 8).

I will first give a definition of the code starting with a “=”. Then I will give a description of the usage of the code, starting with “The code is used if...” and last I will give a description when the code should not be used to contrast to other codes.

This paragraph starts with “The code is not used if...” In general code notes include the results of the three coding steps (Strauss et al., 1996).

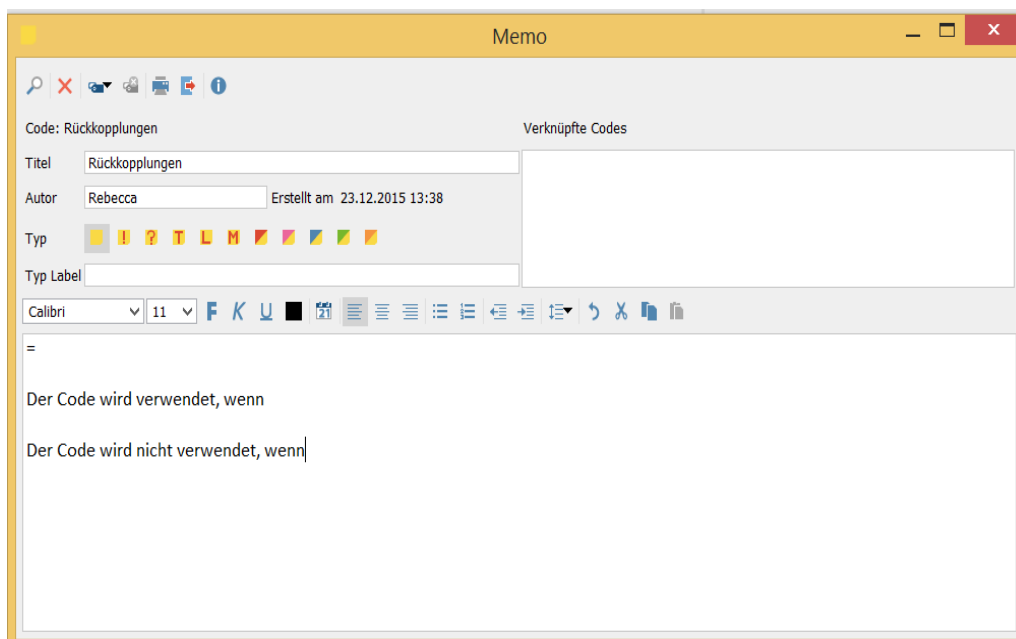


Figure 8: Template for the used code Memo

Theoretical notes are memos that are summarizing for theoretical sensitivity. They can include insights that could arise during deductive and inductive thinking regarding relevant or possible relevant categories and their relations etc. (Strauss et al., 1996).

Planning notes are memos including instructions regarding the sampling, questions, further ideas etc. (Strauss et al., 1996).

³ According to Kelle theoretical sensitivity describes the availability of useful heuristic concepts for the identification of theoretical relevant phenomena in the data (Kelle, 2007).

In contrast to memos which are of textual nature, diagrams are the graphical representation of relations between categories. As they are used to store (intermediate) results. Their usage contributes as well as memos to the improvement of theoretical sensitivity (Strauss et al., 1996).

3.2 Reflections on Grounded Theory Methodology

In this section I expand on Grounded Theory and present some interesting facts on it. As Mey and Mruck detected it is necessary to differentiate between Grounded Theory as a result of an empirical study (the generated theory) on the one hand and the methodology of the grounded theory on the other hand (Mey and Mruck, 2007).

The original version of the GTM was developed by Barney Glaser and Anselm Strauss as described in chapter 3.1 and first presented in 1967 independent from the developed domain (Glaser, 1967). They developed it since they want to provide a qualitative method that has equal significance as quantitative methods, e.g. a method that produces reliable and valid results (Bryant and Charmaz, 2010). Their motivation were the by then developed methods that overvalued the verification of theories in lieu of exploring relevant concepts and hypotheses for the research area first (Glaser, 1967). The presented principles in this book keep widely similar, although some extensions of GTM were developed.

According to Strübing (2011) the publication of the extension of the GTM by Strauss and Corbin (1990) leads to the break between Glaser and Strauss and also to two co existing orientations of GTM that use the same label (Strübing, 2011). The following section describes the differences in the two “schools” in more detail.

The differences Strübing describes between the two schools refer more to the intentions and their epistemological reasons of the authors than to the practical procedures (Strübing, 2011).

The focus of Glaser is on a GTM without any prior knowledge of existent theories and therefore without prior research on literature and specific research question. Influencing the arising concepts of the data by inappropriate concepts arose through existent ideas should be avoided (Glaser, 1967) (Flick, 2013), (Hoda et al., 2012).

Both Glaser and Strauss split their analysis in a coding process, whereas the coding process by Strauss and Corbin has already been presented. Glaser also divided into substantive and theoretical codes. Substantive coding is again divided into open and selective coding (Glaser, 1978).

He suggested 18 coding families to broaden the horizon of the researcher (Strübing, 2011), (Glaser, 1978), (Böhm, 2004). A few of them are presented in Table 3.

Table 3: 10 Coding Families by Glaser presented in (Böhm, 2004).

Coding family	Concepts	Examples
The Six C's	causes, contexts, contingencies, consequences, conditions	... of pain suffering
Process	Stages, phases, phasings, transitions, passages, careers, chains, sequences	Career of a patient with chronic pain
The Degree Family	Extent, level, intensity, range, amount, continuum, statistical average, standard	Extent of pain suffering

Coding family	Concepts	Examples
	deviation	
Type Family	Types, classes, genres, prototypes, styles, kinds	Kinds of pain – sharp, piercing, throbbing, shooting, sting, gnawing, burning
The Strategy Family	Strategies, tactics, techniques, mechanisms, management	Coping with pain
Interactive Family	Interaction, mutual effects, interdependence, reciprocity, symmetries, rituals	Interaction of pain experience and coping
Identity-Self Family	Identity, self-image, self-concept, self-evaluation, social worth, transformations of self	Self-concepts of pain patients
Cutting-Point Family	Boundary, critical juncture, cutting point, turning point, tolerance levels, point of no return	Start of chronification in the medical career of pain patient
Cultural Family	Social norms, social values, social beliefs	Social norms about tolerating pain, ‘feeling rules’
Consensus Family	Contracts, agreements, definitions of the situation, uniformity, conformity, conflict	Compliance

Particularly striking is the C-Family that corresponds to the coding paradigm by Strauss and Corbin consisting of causal conditions, context, intervening condition, action strategy and consequences. Curiously, Glaser accused exactly the coding paradigm to Strauss and Corbin (Strübing, 2011).

A further controversial issue is the role of verification. The position of Strauss is a triad of induction, deduction and verification, whereby verification is a plausibility test and the practical functionality of the theory. This should be part of the theory generation and not a separated step. Glaser took the position that theory generation and verification need not to be proceeded together (Strübing, 2011).

Strübing (2011) came to the conclusion that the position of Glaser is inconsistent and scientific theoretically untenable. Glaser is contradictory regarding the use of prior knowledge and the use of his coding families. The additional waiving of verification is not accepted for a scientific process.

3.3 Code systems

The creation of a code system depends on the choice of the coding method. Choosing the coding method itself may again have several influences. Saldaña (2012) gave advices on how to

decide for the appropriate coding method. He differentiated between the alignment of the research question, paradigmatic, conceptual, and methodological considerations, a priori goals, mixed methods studies, exploratory attributes, generic choices, hybrid schemes and general criteria. All in all, he collected 32 coding methods, 25 first cycle methods, six second cycle methods and one method first to second cycle methods (see: Table 4). First cycle methods are coding processes for the beginning of an analysis that help to split up the data into segments. Second cycle methods are processes for advanced analysis steps that “literally and metaphorically constantly compare, reorganize, or “focus” the codes into categories” (Saldaña, 2012).

Table 4: First Cycle Methods (grouped) and Second Cycle Methods

First Cycle Coding Methods
Grammatical Methods
Elemental Methods
Affective Methods
Literary and Language Methods
Exploratory Methods
Procedural Methods
Themeing the Data
First to second cycle coding method
Eclectic Coding
Second cycle coding method
Pattern Coding
Focused Coding
Axial Coding
Theoretical Coding (Selective Coding)
Elaborative Coding
Longitudinal Coding

Grammatical methods are techniques for improving the organization, granularity and the structure of qualitative data. They are the basic grammatical principles of a technique. Elemental method are approaches for coding qualitative texts. They focus on filters for reviewing and build a basis for coding cycles. Affective methods investigate participants’ emotions, values etc. by directly naming those experiences. Literary and language methods

show up aspects of communications for codes. Thereby they borrow from established approaches to the analysis of literature and a modern approach to the analysis of oral communication. Exploratory methods permit open ended investigation. They are exploratory and preliminary assignments of codes. Often more precise coding methods are proceeded after this kind of method. Procedural methods are standardized ways to code data. They consist of pre-established coding systems or specific ways of analyzing data. Using themeing the data shows that extended passages of code can also capture the essence and essentials. In contrast to a code a theme is a more extended phrase that labels the meaning of a data unit (Saldaña, 2012).

(Saldaña, 2012) stated that eclectic coding can't be categorized but he would categorize it to an exploratory method regarding the profile. An analyst can start here with an array of coding methods first and then transition to strategic recoding decisions based on the previous learnings.

I do not describe the second cycle methods. Suffice is to say that they each require “of linking seemingly unrelated facts logically, of fitting categories with one another” (Morse, 1994), (Saldaña, 2012).

I would like to focus on the general criteria for coding decisions. Flick (2009) provided a checklist for deciding for an analytic method.

He considered first the foundation principles. In there he advised to think about deductive, inductive or hybrid approaches. In general, they differ in their creation time during the QDA process. What they all have in common is the tree- like structure with codes as nodes and the theory as the root (see: Figure 9), (Flick, 2009).

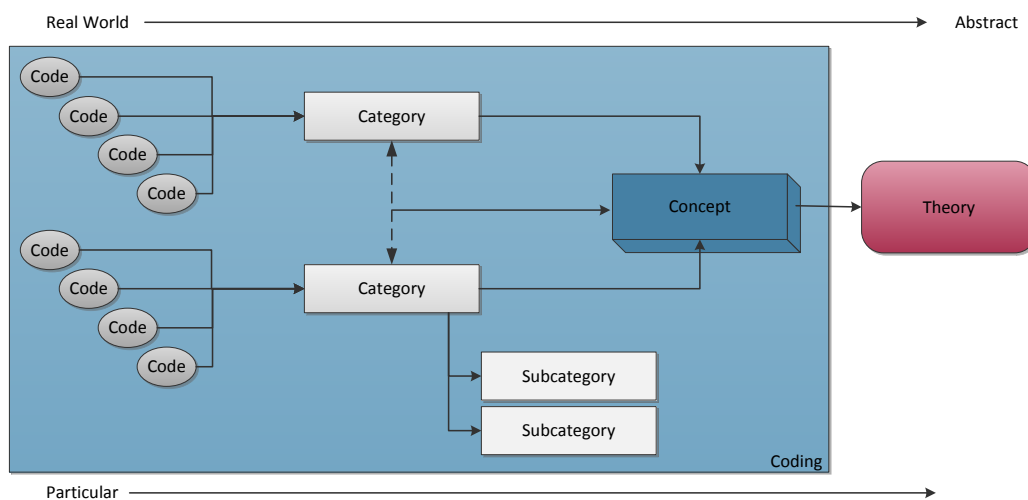


Figure 9: The process of how a theory is constructed out of codes (Saldaña, 2012).

Deductive Approach

Deductive reasoning works from general to a more specific - a "top-down" approach. Beginning with thinking up a theory about the topic of interest, narrowing that down into a more specific hypothesis that can be tested. It is narrowed down even further when observations are collected to address the hypotheses. This ultimately leads to the ability to test the hypotheses with specific data - a confirmation (or not) of the original theories (Saldaña, 2012).

Inductive Approach

Inductive reasoning works from specific observations to broader generalizations and theories - a "bottom up" approach. Beginning with specific observations and measures, one can detect patterns and regularities, formulate some tentative hypotheses that can be explored, and finally

end up developing some general conclusions or theories. Inductive reasoning is more open-ended and exploratory, especially at the beginning of the survey. Deductive reasoning is narrower in nature and is concerned with testing or confirming hypotheses. Even though a particular study may look like, that it is purely deductive, most social research involves both inductive and deductive reasoning processes at some time in the project (Saldaña, 2012).

Hybrid Approach

The majority of qualitative researchers will code their data both during and after collection as an analytic tactic, for coding is analysis (Saldaña, 2012). In Fereday and Muir-Cochrane (2008) the incorporation of an inductive and a deductive approach is shown exemplary. They gave a clear description of their coding process (Fereday and Muir-Cochrane, 2008).

For the initial decision making he suggested to keep open during the initial data collection and to review which method will be most appropriate. Further he recommended to start with generic coding methods (attribute coding⁴, structural or heuristic coding⁵, descriptive coding⁶, in vivo coding⁷ or initial⁸ and/or values coding⁹) and changing them if necessary. If the investigation should generate a new theory, then GTM coding methods are recommended. Lastly, he advised to pilot test the coding choice (Saldaña, 2012).

As a second criterion he considered the coding compatibility. The researcher should think about the types of answers generated by the method. Moreover, it is important that the data form fits to the method and how many coding methods are necessary to capture the whole phenomenon. If more than one method is necessary, then the method should be chosen purposefully and the researcher should be careful about mixing incompatible methods. At last he mentioned the method influence regarding numeric conversions (Saldaña, 2012).

Thirdly, he recommended to mention about the coding flexibility. He explains: “Data are not coded- they’re recoded. Be willing to change your method(s) if your initial choice(s) is not working” (Saldaña, 2012).

Lastly, he recommended thinking about the coding outcomes. In there he stated that the researcher should not feel as if he grasps specificity, insights and new discoveries and not complication (Saldaña, 2012).

3.4 Domain Knowledge

The term “domain” including “domain knowledge” and “domain specificity” seemed to be quite abstract and unknown. At least my first interviewee didn’t know (AR, 2015) the term including an explanation wherefore I decided to change my questions regarding this topic in my interview. Even then my feeling was that the questions about this topic were quite hard to handle for my interviewees. This let me to the hypotheses that domain knowledge is an implicit knowledge base a researcher has acquired which is hard to formulate and hard to set to relation.

In today’s research, domain knowledge plays a role as mental set in any considered domain. But there is no known universal and unique domain independent definition about the term

⁴ Counts to grammatical methods

⁵ Counts to elemental methods

⁶ Counts to elemental methods

⁷ Counts to elemental methods

⁸ Counts to elemental methods

⁹ Counts to affective methods.

domain knowledge. The research that exists about domain knowledge, domain specific knowledge or domain general knowledge is mostly literature review. Also the more recent published papers reference the definition of Alexander and Judy (1988) regarding domain knowledge (Muis et al., 2006), (Khatri et al., 2006), (Muis, 2004), (Buehl et al., 2002).

Alexander and Judy (1988) performed a literature review on the differentiation of strategic knowledge and domain-specific knowledge. They are located in the psychology and investigate domain specific knowledge from a cognitive view and gave a clear definition of how they understand domain-specific knowledge:

They defined “domain-specific knowledge as the declarative, procedural, or conditional knowledge one possesses relative to a particular field of study. Declarative knowledge refers to factual information (knowing what), whereas procedural knowledge is the compilation of declarative knowledge into functional units that incorporate domain-specific strategies (knowing how). Conditional knowledge entails the understanding of when and where to access certain facts or employ particular procedures” (Alexander and Judy, 1988). Since the definition seems to be applicable for different domains, it is used also for the underlying research (Alexander and Judy, 1988).

Patricia Alexander synthesized facts about existing domain knowledge in seemingly separate research programs. She further poses two interesting issues: The conceptualization of domain knowledge and the relation of domain knowledge to other knowledge forms. She particularly stated that the term “domain knowledge” might be used differently by different researchers (Alexander, 1992).

Domain knowledge is seldom explicitly considered in research publications. Searching for papers mostly provides results in technical domains or psychology. I will consider some more research additionally:

- With Hartmann et al. (2006) an approach called “Feature Net” is presented to collect application and domain specific requirements in a structured way. Therewith, a targeted reusability should be enabled. They stated that domain specific knowledge is represented in the relations between features (Hartmann et al., 2006). This paper focusses on technical research that implies domain knowledge.
- All ontology based approaches are also a way to enable domain knowledge for machine processing (see for example (Sure et al., 2002), (Lee and Zhao, 2006), (Kaiya and Saeki, 2006)). This is a research field which explicitly focusses on domain knowledge from technical view.
- Another approach is the investigation of domain knowledge for search tactics (Vibert et al., 2009), (Wildemuth, 2004). Thereby studies are conducted about how researchers or students perform their searches in the web. This is an approach investigated by researchers from psychological view.
- (Buehl et al., 2002) investigated differences between domain specificity and domain generality as one of the few. This research is also located in the psychology domain.

Striking to all considered research work on domain knowledge is the fact that this is mostly conducted by observation or literature review. This supports my hypothesis that domain knowledge is implicit knowledge and people can't hardly formulate it.

3.5 Application in the Context of Requirements Engineering

User stories have been developed for the acquisition of requirements in agile software projects. They describe requirements from the user's perspective in one sentence. All in all, user stories actually consist of three parts: a story card with the written description, discussions about the user story to elaborate it in details and acceptance criteria to define when a user story is finished. The wording of user stories can, for example, take place within in a story-writing workshop. Participants of this workshop are all important stakeholders of the product (i.e. testers, product manager, real users, and design team). By the hand of the product vision, which has been developed before, all types of user roles of the intended product have to be considered. After that, the basic goals are formulated for every user role. Since these would be too big for user stories, they are elaborated as so-called "Epics", that are superordinate user stories and written down on cards. The epics are split up into user stories later. Typically, user stories are written down on so called "story cards". Due to the limited space, the writer has to formulate his user story short. Moreover, this procedure is very uncomplicated and fast, but it is important that customers and developers have discussed the user story and its value before. Additionally, stories should be independent from another so that they could be implemented in any order (Cohn, 2010).

One suggested format writing user stories short is the pattern:

"As a [role], I can [feature] so that [reason]"

The story card contains a distinct identifier, a concise title, acceptance criteria and additional information beside the functional description. A user story needs to be testable. Hence, acceptance criteria are needed to check whether a user story is implemented correctly. Suitable therefore are concrete positive and negative examples, which are written on the backside of the story card. Sometimes additional information is needed to explain the user story (for developers) (Cohn, 2010).

The idea of the exploration of code schemata is that the story-writing workshop can be used in two ways: First group discussions can be recorded, transcribed and coded using a code schema (like the suggested one) and further the story cards might also be used as document for the analysis.

3.6 Quality criteria for qualitative research

As I introduced my work qualitative research has to be high quality and meet as well as quantitative research certain quality criteria. The validation of qualitative research is a discursive issue beyond researchers. One discussed issue is the transferability of quantitative criteria to qualitative criteria in order to define criteria for any research. Central criteria are objectivity, reliability and validity.

Golafshani (2003) discuss the criteria validity and reliability in context of qualitative research. They argue for reliability as well as for validity that there is no clear opinion whether the criteria is appropriate for qualitative research or not (Golafshani, 2003).

Lincoln and Guba explicate that each paradigm (qualitative or quantitative) must have worthwhile criteria but they have to be paradigm-specific for addressing rigor. While rigor for the quantitative paradigm can be reaches by internal validity, external validity, reliability and objectivity, trustworthiness for the qualitative paradigm can be achieved by credibility, transferability, dependability and confirmability (Guba, 1981), (Lincoln and Guba, 1985).

Morse et al. (2008) took the position of reestablishing validity and reliability for qualitative research instead of the substitution to trustworthiness suggested by Guba (1981) or other new

criteria as Leininger (1994) or Rubin and Rubin (1995) suggested it (Leininger, 1994), (Rubin and Rubin, 1995), (Morse et al., 2008).

Such discussions make it hard to take the “right” decision on quality criteria. Lastly I decided for the catalogue defined by Ines Steinke. She stated reasons why objectivity, reliability and validity are not registered in her catalogue. These reasons include different understandings and definitions of the terms, either due to language or due to the origin in quantitative research (Steinke, 1999), (Steinke, 2004). In chapter 2.7.4 the application of some criteria have already been illustrated. Below I describe each criterion more detailed and add the application for my work for them that are not described above.

The first criterion is *intersubjective comprehensibility*, since controllability isn’t possible in qualitative research due to the not existing possibility to reproduce the identical situation for the data collection. It can be checked by three types: The first and the favored type is the documentation of the research process. Therewith the reader can track each step of the study and grade the results. Concretely, the previous understanding, the survey method, the rules of transcription, the data, the evaluation method, the source of information, decisions and problems and the criteria, the study should achieve, should be documented. Based on this the interpretation is comprehensible (Steinke, 2004), (Steinke, 1999).

The second possibility type to check the first criteria is *the interpretation by groups*. This is a discursive form used for objective hermeneutics and for GT (Steinke, 2004), (Steinke, 1999).

The last type is application of codified procedures. This is comparable with the standardization of the methodological procedure. The GT is one example for an already codified procedure. Using codified procedures facilitates the reader to check the study due to additional information. Otherwise the analysis steps have to be elaborated in more detail (Steinke, 2004), (Steinke, 1999).

Indication of the research process is the second criteria. Using this criteria, the whole research process is checked regarding its appropriateness, also named indication. This includes:

- the indication of the qualitative procedure with respect to the research question
- the indication of the selection of methods
- the indication of the rules of transcription regarding manageability readability, learnability and interpretability
- the indication of the sampling strategy
- the indication of the methodological individual decisions in the context of the whole study
- the indication of the grading criteria (Steinke, 2004), (Steinke, 1999)

Empirical anchoring is a further criterion to check the study. The creation and check of the hypotheses and theories should be anchored in the data in qualitative research. By deduction consequences are derived from the theories and verified on empirical data. This is done to check the theory. Falsification tries to disprove the theory. There are five ways to check the empirical anchoring: At first, using codified methods like the GT ensures empirical anchoring. Secondly, there is the question whether there are sufficient textual evidences for the developed theory. A further type is the analytical induction. Thereby, an as far as possible developed theory is checked based on a case. If the theory doesn’t apply, the phenomenon is redefined or the case is excluded from the theory. This is done until a universal relation is established. The fourth type suggests to derive prognosis from the generated theory and to check the occurrence on the text. The last way is called communicative validation. This type validates by reconnecting the theory to the investigated persons (Steinke, 2004), (Steinke, 1999).

The criteria *limitation* finds out and checks the limits of the scope of application, i.e. the generalizability of a theory. Therefore, the minimal conditions, that have to be fulfilled making the findings transferable, have to be clarified. These are the minimal conditions to make the phenomenon happen (Steinke, 2004), (Steinke, 1999).

The criteria *coherence* checks whether the generated theory is internally consistent. Therefore, the generated theory is to be checked regarding coherence and it has to be checked whether all contradictions in data and interpretation are edited. All unresolved questions and contradictions should be referred (Steinke, 2004), (Steinke, 1999). In my opinion there are no unsolved questions or contradictions available at the moment for the underlying paper.

A further suggested criteria is the *relevance*. This criterion checks the pragmatically use of the theory. Hence, Steinke (2004) states the following leading questions and I try to request each shortly (Steinke, 2004), (Steinke, 1999):

- Is the research question relevant?
The question might be relevant for CAQDAS developers to implement a feature that generates codes automatically and gives suggestions to the researchers using the software.
- What contribution is made by the theory developed?
The theory describes the influence of domain knowledge and the connection to the existence of code schemata. Besides, the integration of domain knowledge to the qualitative research process and to QDA is described.
- Does it make new interpretations available?
Two more issues could be considered: The existence and integration of further schemata (not code schemata) for a specific domain and the development, use or reuse of a code schema across domains.
- Does it contain any explanations concerning the phenomenon in question?
Yes. The theory explains the phenomenon in question.
- Does the theory facilitate the solution of problems?
The theory facilitates the solution of problems since it should help to improve to make code systems machine processable.
- Can the results be generalized?
Results can be generalized according to a specific domain. The set-up of a domain specific code schema is conceivable for a specific domain.
- Is the presentation of the theory comprehensible?
The theory is derived by GTM and parts of them also presented in a meta model in order to increase comprehensibility.

The last criteria of the catalogue is the *reflected subjectivity*. This criterion checks how the role of the researcher as part of the social reality is taken into account regarding the theory building. Therewith, the self-observation, the personal preconditions, relations between researcher and informant, reflections during the entry in the field are checked (Steinke, 2004), (Steinke, 1999).

3.6.1 Triangulation

To achieve trustworthiness, it is necessary to consider more than one perspective. Lincoln and Guba (1985) and also Flick (2000) explicated the importance of triangulation, meaning the improvement of rigor of the analysis by weighting the results from more than one perspective.

This can either performed by the use of multiple data sources, multiple researchers, multiple theoretical perspectives, and/or multiple methods (Schwandt, 2001), (Denzin, 1978), (Leech and Onwuegbuzie, 2007).

Following explanations refer to Jenner et al. (2004) who translated the handbook by Flick, Steinke and von Kardorff (Jenner et al., 2004), (Flick, 2013). Flick referred to Denzin (1978), who stated triangulation as first as a validation strategy and later as a strategy for justifying and underpinning knowledge by gaining additional knowledge (Denzin, 1978), (Denzin, 1994). He differentiates five types:

Data Triangulation

Beside the usage of interviews also visual data or electronical process data could be considered as own source. The can be triangulated with common verbal data. Leech and Onwuegbuzie (2007) stated that triangulation should be extended to data analysis tools.

Investigator Triangulation

Investigator triangulation is characterized by the use of different observers or interviewers to compensate subjective influences from one interpreter meaning to extend, to correct or to check the view.

Theory Triangulation

Theory triangulation means the approximation to a research object using different perspectives and hypotheses.

At this type of triangulation, the researcher has to keep in mind that each method has been developed with different theoretical backgrounds. "In concrete situations of triangulation the partially incompatible epistemological assumptions about the research issue, or about (qualitative) research in these different theoretical backgrounds, are carried over by the methods."(Jenner et al., 2004).

Within- Method Triangulation

Using the within- method triangulation the researcher tries to combine for example a guided interview with narrative elements on a systematic way. Thereby, an additional perspective on the object on the research issue of the interviewee should be opened up and different aspects of the subjective discussion to the research issue should be illustrated.

Between-Method Triangulation

This type of triangulation combines different methods. This can be either the combination of qualitative and quantitative methods or the combination of reactive procedures with non-reactive procedures. Reactive procedures are methods where the researcher is part of the research situation, non-reactive procedures are methods where data that is not set up for the investigation is analyzed. Further, using between-method triangulation helps understanding different areas of the research issue. Moreover, the concept of triangulation means to understand the combined methods as equal and not one as pre procedure of other(s).

3.6.2 Saturation

Saturation is mostly associated with an adequate sample size and originally developed within GT (Francis et al., 2010), (O'Reilly and Parker, 2012), (Mason, 2010), (Guest et al., 2006).

Nevertheless, the application of saturation beyond the origin in GT receives limited attention. With the renaming to data saturation instead of theoretical saturation the meaning has changed as well. According to O'Reilly and Parker (2012) these new meanings in other qualitative research are less developed. They argue that data saturation is an inappropriate generic quality criterion for qualitative research.

There are at least two forms of saturation beside the theoretical saturation, I found during my research. One is the already named data saturation suggested exemplarily by Guest et al. (2006) and Francis et al. (2010), furthermore there is thematic saturation suggested by Guest et al. (2006) and lastly simply saturation (Starks and Trinidad, 2007), (O'Reilly and Parker, 2012).

Thematic and data saturation in general mean to continue data collection until no new idea arise (Green and Thorogood, 2013). In contrast theoretical saturation means, as I describe in chapter 2.7.3 that categories have to be fully developed and variations must be built in and relationships between them are tested (Corbin and Strauss, 2014).

To come back to the introducing sentence of this section sample size is often justified with saturation. Where quantitative research needs as many information as possible to achieve high quality, in qualitative research there is the need of a sample size that is large enough to explore a phenomenon but not to have repetitions. The procedure of quantitative research is not applicable to qualitative work as more data does not necessarily lead to more information (Mason, 2010). Mason (2010) named saturation as the guiding principle to aim this. Thereto Guest et al. (2006) states that there is the trend to use saturated samples as quality indicator, but currently there is no clear guidance for researchers to determine when saturation has been achieved. Additionally, there are some constraints (as in my work too) researchers have to have in mind in reality. Starting with the requirement of defining the size of the sample in advance up to time and money restrictions. Also unforeseen participant availability must not be underestimated. I agree with O'Reilly and Parker (2012) who stated saturation as a convincing concept but as concept with a number of practical weaknesses. They clarify by referring to Morse (1995) that "transparency about these limitations on reaching saturation does not necessarily invalidate the findings. If saturation is not reached this simply means that the phenomenon has not yet been fully explored rather than that the findings are invalid (Morse, 1995). It is acceptable, therefore, that any limitations of sampling adequacy are transparently reported."

3.7 Further Work

Thus I proposed an approach to a GT in this thesis, an interesting aspect would be the exploration of the detected phenomenon in more domains. A research about the different approaches to construct a code schema could help to improve the knowledge of domain knowledge influence.

Further, the data showed an interesting phenomenon according to teamwork. Although, it is clear that student thesis has to be performed alone, coding in teams should be an alternative to achieve a deeper analysis of the data with more saturated results. Actually, Strauss and Corbin suggest approaches to that. Two of my interviewees proceed the practice of pair coding which could also be an approach instead of validating the codes by peers at the end of the analysis process. An additional thought could be, that teams need rules to code, which means in turn that they might create a standardized code schema.

I developed a basic possible code schema for RE in agile context. With this I showed the possible transferability to RE. The suggested code schema is oriented towards the user story template and towards the detected code schema by two of my interviewees. Here again is further research conceivable and necessary. The explored code schema has been built upon defined sets of verbs and objects. This is also conceivable for our suggested one.

Glossary

Term	Definition
Concept ¹⁰	A concept is a term, to which single statements, occurrences or other examples for phenomena are assigned to.
Category ¹¹	Concepts are classified in categories by comparing concepts with each other where they relate to the same phenomenon. Concepts are therefore classified to concepts of higher order. The hierarchy is organized in core category, category and sub category.
Property	The property is an attribute or characteristic that belongs to a category.
Code ¹²	The code is a labeled statement in the interview text which is defined as concept and assigned to a category.
Codetree	The code tree represents the hierarchical order of categories and properties.
Coding	Basically the coding is the process of the data analysis and consists of open coding “process of breaking up, investigating, comparing, conceptualizing and categorizing of the data to create a code tree. During the axial coding relations between categories are created and composed in a new way. The selective coding is the process of building the grounded theory.
Coding Paradigm	The coding paradigm helps thinking systematically about the data and to work out complex relations by analyzing the data regarding causes, context, conditions, strategies and consequences.
Phenomenon	The phenomenon is the central idea, the event, incident, occurrence that is investigated at the axial coding regarding its causal characteristics.

¹⁰ The terms code, concept and category are used widely synonym by Glaser and Strauss. This is due to the fact that the transitions of the termini are flowing.

¹¹ The terms code, concept and category are used widely synonym by Glaser and Strauss. This is due to the fact that the transitions of the termini are flowing.

¹² The terms code, concept and category are used widely synonym by Glaser and Strauss. This is due to the fact that the transitions of the termini are flowing.

Appendices

Appendix A Interview Guide

Einleitung

[Aufnahme starten]

Hallo vielen Dank, dass Sie an dem Interview teilnehmen. Wir benötigen Ihre Daten für unsere Forschung am Lehrstuhl. Ihre Angaben werden vertraulich behandelt und anonymisiert.

Zunächst habe ich eine organisatorische Frage:

Ich würde das Interview gerne aufnehmen, damit ich mich während der Auswertung darauf beziehen kann. Ist das in Ordnung für Sie?

Dann starten wir jetzt mit dem eigentlichen Interview! Sie können zu jeder Frage alles was Ihnen einfällt erzählen. Es geht grundsätzlich darum, wie Sie an ihrem Lehrstuhl Theoriebildung betreiben.

Qualitative Forschungsmethoden

- Welche Art qualitativer Forschung haben Sie bereits durchgeführt?
- Haben Sie an Ihrem Lehrstuhl einen festen Prozess (spezifizierte FM) zur Theoriebildung etabliert?
 - Nein: Suchen Sie nach einem solchen Prozess?
 - Gibt es Guidelines oder Best Practices die angewandt werden?
 - Können Sie mir ihr Vorgehen anhand eines Beispiels erläutern?
 - *ToDo: eigenes Beispiel konstruieren*
 - Ja: Ist die Prozessdefinition frei verfügbar?
- Welche Auswertungsmethoden verwenden Sie?
 - Warum?
 - Von welchen Punkten machen Sie abhängig, welche Methode Sie wählen?
- Arbeiten Sie zur Auswertung im Team?
 - Haben Sie Regeln?
 - Ja: Wie erfolgen die Absprachen?
 - Notieren Sie die Absprachen?
 - Ja: Wie? Verwenden Sie die Absprachen beim nächsten Projekt wieder?

Bereichsspezifisches Wissen (Domänenspezifizität)

- Welchen Einfluss hat Domänenspezifizität auf Ihre Forschung?
- Wie grenzen Sie Domänen voneinander ab?
- Wie erkennen Sie wenn zwei Domänen sich ähneln oder gar gleich sind?
- Haben Sie eine Vorgehensweise wie sich Domänenwissen erfassen lässt?
 - Ja: Können Sie mir diese beschreiben?

Schematisierung

- Haben Sie Vorlagen (Templates), die Sie in Ihrem Auswertungs- bzw. Forschungsprozess immer wieder verwenden?
 - Verwenden Sie Vorlagen (für die Codierung)?
 - Wie entstehen Vorlagen?
 - Sind diese Vorlagen zielgruppenspezifisch?
 - Wie beeinflussen Vorlagen die Bildung Ihrer Code systeme?
 - Gibt es Rückkopplungen oder Feedbackloops, sodass die Vorlagen aktualisiert werden, wenn neue Erkenntnisse entstehen?
 - Wie fließt Ihre Vorerfahrung und vorherige Erkenntnisse in eine neue Analyse ein?
 - Sind diese Erfahrungen bzw. Erkenntnisse schriftlich festgehalten?
 - Welche Rolle spielen Memos für Ihre Analyse?

Codierung

- Können Sie mir den Ablauf beschreiben wie Sie Codieren?
 - Worauf achten Sie bei der Erstellung von Codes?
 - Folgen Sie Codierregeln? (wie bspw. In der qualitativen Inhaltsanalyse?)
 - Wie bilden Sie Kategorien?
 - Nach einer bestimmten Methodik?

Nun wären wir am Ende des Interviews. Vielen Dank, dass Sie sich zur Verfügung gestellt haben. Haben Sie noch Fragen an mich?

[Aufnahme stoppen]

Appendix B Codebook

B.1 qualitative Forschungsmethoden

= *die gewählte Forschungsmethode ist eine definierte qualitative Methode; das bedeutet, es gibt eine vorgeschriebene Vorgehensweise, die in einschlägiger Literatur verfügbar ist.*

Der Code wird verwendet, wenn die Forschungsmethode in wissenschaftlicher Literatur verfolgbar/nachzulesen ist. Sie somit wissenschaftlich bekannt.

Der Code wird nicht verwendet, wenn der Forscher keine festgeschriebene Forschungsmethode benennt.

B.1.1 etablierte Methode

= *die gewählte Methode ist schon etabliert, da sie schon sehr lange existiert.*

Der Code wird verwendet, wenn die Methode in der Literatur (Handbuch) beschrieben ist.

B.1.2 Vorgehen zur Methodenwahl

= *der Weg, den der Forscher beschreibt um eine Methode auszuwählen.*

Der Code wird verwendet, wenn ein schrittweises Vorgehen beschrieben wird, um eine Methode auszuwählen. Dabei geht es auch darum, welche Einflüsse eine Rolle spielen, eine Methode nicht auszuwählen.

Der Code wird nicht verwendet, wenn Gründe angegeben werden, warum Methoden ausgewählt wurden aber keine schrittweise Vorgehensweise angegeben wird.

B.1.2.1 allgemeine Gründe zur Methodenwahl

= *die Gründe, weshalb eine Methode ausgewählt wird.*

Der Code wird verwendet, wenn in irgendeiner Art und Weise (auch abstrakt) die Wahl einer Methode begründet wird.

Der Code wird nicht verwendet, falls ein spezieller Grund genannt wird

Beispiel: Vorherige Analyse der Situation und anschließende Wahl.

B.1.3 Methodenkombination

= *die Verwendung von zwei unterschiedlichen Erhebungsmethoden oder zwei unterschiedlichen Auswertungsmethoden.*

Der Code wird verwendet, wenn sich herausstellt, dass zwei bereits definierte Methoden beispielsweise im Auswertungsprozess kombiniert wurden. Das kann bedeuten, dass beide Methoden parallel ausgeführt wurden, oder dass Teile der einen Methode durch die andere Methode aus bestimmten Gründen ausgetauscht wurden.

Der Code wird auch verwendet, wenn eine unklare Aussage getroffen wird, ob nun eine oder mehrere Methoden herangezogen werden.

Der Code wird nicht verwendet, wenn eine Methode nur angepasst wurde bspw. auf den eigenen Stil oder Mixed Methods eingesetzt werden.

B.1.3.1 Begründung für Methodenkombination

= *der Grund/ die Gründe weshalb mehrere Methoden eingesetzt werden.*

Der Code wird verwendet, wenn Gründe für die Zusammenführung oder Mischung von Methoden angegeben werden.

Der Code wird nicht verwendet, wenn nur das Faktum angegeben wird, dass Methoden gemischt wurden, aber nicht warum.

B.1.3.2 Beschreibung der Methodenkombination

= *der Ablauf WIE die beiden Methoden kombiniert wurden wird beschrieben. Werden Teile der einen Methode durch die andere ersetzt? Werden zwei Methoden parallel abgearbeitet?*

Der Code wird verwendet, wenn das Vorgehen, wie Methoden kombiniert eingesetzt werden, beschrieben wird.

Der Code wird nicht verwendet, wenn das Vorgehen unklar bleibt und/oder Gründe für die Kombination angegeben werden.

B.1.4 gewählte Forschungsmethode

= *der Rahmen oder die Methodik für die der Forscher/die Forscherin sich entschieden hat.*

Der Code wird verwendet, wenn es einen standardisierten Forschungsprozess gibt, den diese Methode beschreibt. Mit standardisiert ist gemeint, dass die Methode bereits definiert war.

Der Code wird nicht verwendet, wenn die gewählte Methodik nicht vordefiniert war oder keine spezielle Methodik angegeben wird.

B.1.4.1 Wiederverwendbarkeit und gewählte Forschungsmethode

= *Die Wiederverwendbarkeit eines Codierschemas im Zusammenhang mit der Forschungsmethode wird begründet. Es gibt Forschungsmethoden, die zu Beginn der Auswertung bzw. vor dem Beginn des Codierens keine Konzepte vorab bilden (induktives Vorgehen)*

Der Code wird verwendet, wenn Codierschema und Forschungsmethode in Zusammenhang gebracht werden.

Der Code wird nicht verwendet, wenn hinsichtlich der Wiederverwendbarkeit und der Forschungsmethode kein Zusammenhang hergestellt wird bzw. Wiederverwendbarkeit nicht in Betracht gezogen wird.

B.1.4.2 fixiertes Vorgehen unabhängig von Forschungsmethode

= *Ein Schritt oder Vorgehensweise, die der Forscher immer durchführt, egal welche Forschungsmethode er gewählt hat.*

Der Code wird verwendet, wenn sich eine bestimmte Vorgehensweise oder eine immer gleiche Handlung feststellen lässt, die der Forscher unabhängig vom Projekt oder der Forschungsmethode praktiziert. Dieses wird auch erklärt.

Der Code wird nicht verwendet, wenn die Vorgehensweisen bezogen auf die Forschungsmethode durchgeführt werden. Hier kommt es auf eine Vorgehensweise an, die unabhängig von Projekt oder FM durchgeführt wird.

Beispiel:

One-Page-Description

B.1.4.2.1 empfohlene Auswertungsmethode

= *die Auswertungsmethode, die vom Lehrstuhl zwar nicht vorgeschrieben aber empfohlen wird.*

Der Code wird verwendet, wenn es zwar keine Vorschrift aber eine Empfehlung für eine spezielle Auswertungsmethode gibt.

Der Code wird nicht verwendet, wenn keine bestimmte Auswertungsmethode empfohlen wird oder wenn es eine Lehrstuhlvorschrift gibt.

B.1.4.3 Forschungsmethode keine Lehrstuhlvorgabe

= *die gewählte Forschungsmethode war weder Empfehlung noch Vorgabe vom Lehrstuhl*

Der Code wird verwendet wenn die Forschungsmethode (Vorsicht hier scharf unterscheiden zwischen Forschungsmethode, Erhebungsmethode und Auswertungsmethode) selbst gewählt wurde

Der Code wird nicht verwendet, wenn die Forschungsmethode aufgrund von Lehrstuhlempfehlungen gewählt wurde.

B.1.4.4 Gründe für Auswahl der Forschungsmethode

= *Die Bedingungen, die erfüllt sein müssen oder Gründe, die abgewogen wurden, damit sich ein Forscher für eine Forschungsmethode entschieden hat.*

Der Code wird verwendet, wenn Einflüsse und Gründe für die Auswahl einer Forschungsmethode gegeben werden.

Der Code wird nicht verwendet, wenn es um die Auswahl einer Erhebungs- oder Auswertungsmethode geht

B.1.4.4.1 Begründung der Wahl der Forschungsmethode

= *Der Forscher erklärt, warum genau diese Methode für genau dieses eine Projekt ausgewählt wurde*

Der Code wird verwendet wenn eine Art von Begründung für die Wahl der Forschungsmethode geliefert wird.

B.1.4.4.2 Wahl der Forschungsmethode domänenunabhängig

= *Die Forschungsmethode wird unabhängig vom Feld ausgewählt. Ein Feld kann durch verschiedene Methoden erforscht werden.*

Der Code wird verwendet, wenn explizit erläutert wird, dass eine Methode unabhängig von der Domäne angewendet werden kann.

Der Code wird nicht verwendet, wenn es dabei um Gründe für die Wahl der Forschungsmethode geht.

B.1.4.4.3 Wahl der Forschungsmethode durch positive Erfahrung

= *wenn eine Forschungsmethode ausgewählt wurde, weil der Forscher bereits gute Erfahrungen mit dieser gemacht und deshalb keine weiteren Forschungsmethoden in Betracht zieht.*

Der Code wird verwendet, wenn der Forscher keine weiteren Methoden gegenüberstellt für sein Projekt.

Der Code wird nicht verwendet, wenn eine Gegenüberstellung von Methoden und darauffolgend eine Auswahl der Forschungsmethode aus bestimmten Gründen getroffen wird.

B.1.4.5 grobe Erläuterung der Forschungsmethode

= die verwendete Forschungsmethode an sich wird erläutert

Der Code wird verwendet, wenn eine Beschreibung der verwendeten Forschungsmethode gegeben wird.

Der Code wird nicht verwendet, wenn die Methode im Kontext des Forschungsfelds der Forscherin erläutert.

B.1.4.6 frei verfügbare Forschungsmethode

= die Forschungsmethode ist frei verfügbar, d.h. auch, dass sie außerhalb des Lehrstuhls verwendet wird und wissenschaftlich anerkannt und etabliert ist (in der Literatur)

Der Code wird verwendet, wenn die verwendete Forschungsmethode in der Wissenschaft bekannt ist und wenn auch Forscher außerhalb des Lehrstuhls bzw. derjenigen Person, diese Methode anwenden.

Der Code wird nicht verwendet, wenn der Lehrstuhl oder der Forscher/die Forscherin eine eigene, respektive eine nicht definierte Methode verwendet, oder eine nicht explizit benannte Methode.

B.1.4.7 Adaption der Forschungsmethode

= die Methode wurde aus dem Zusammenhang, in dem sie bekannt ist herausextrahiert und in einem anderen Zusammenhang eingesetzt

Der Code wird verwendet, wenn die Durchführung der Methode abgewandelt wird.

Der Code wird nicht verwendet, wenn die Methode definitionsgemäß verwendet wird.

B.1.4.8 empfohlene Forschungsmethode

= die Forschungsmethode, die der Lehrstuhl oder die Forschungsgruppe derzeit bevorzugt bzw. viele Mitarbeiter dort einsetzen.

Der Code wird verwendet, wenn es eine Forschungsmethode gibt, die empfohlen wird.

Der Code wird nicht eingesetzt, wenn eine solche Empfehlung nicht existiert oder wenn diese Empfehlung keine Empfehlung mehr ist, sondern einer Vorschrift.

B.1.5 durchgeführte Erhebungsmethoden

= die verschiedenen Erhebungsmethoden, die der/die Forscherin bereits durchgeführt hat.

Der Code wird verwendet, wenn es um **mehrere** tatsächlich durchgeführte Erhebungsmethoden geht.

Der Code wird nicht verwendet, wenn es sich bei den verwendeten Methoden um Forschungsmethoden oder Auswertungsmethoden handelt und wenn der Forscher über eine konkrete Erhebungsmethode spricht und keine weiteren nennt.

B.1.5.1 Gründe für die Wahl der Erhebungsmethode

= *warum wird welche Erhebungsmethode ausgewählt. (Datenzugang, Domäne, usw.)*

Der Code wird verwendet, wenn es um die Auswahl der Erhebungsmethode geht und die Gründe dafür.

Der Code wird nicht verwendet, wenn nur allgemein die Wahl von Methoden erklärt wird oder die Wahl von Forschungs- oder Auswertungsmethoden begründet wird.

B.1.5.2 durchgeführte Erhebungsmethode

= *beschreibt die Methode, mit der die Daten erhoben wurden.*

Der Code wurde verwendet wenn der Forscher von **einer konkreten** Erhebungsmethode berichtet, die er in seinem Forschungsprojekt durchgeführt hat.

Der Code soll nicht verwendet werden, wenn der Forscher mehrere Erhebungsmethoden erwähnt und miteinander in Beziehung setzt.

B.1.5.3 Adaption der Erhebungsmethode

= *die Art und Weise, wie eine Erhebungsmethode durch den Forscher abgewandelt wird.*

Der Code wird verwendet, wenn eine Erhebungsmethode verändert angewandt wird/ wurde.

Der Code wird nicht verwendet, wenn Methoden kombiniert werden

B.1.6 durchgeführte Auswertungsmethoden

= *die verschiedenen Auswertungsmethoden die der Forscher bereits durchgeführt hat.*

Der Code wird verwendet, wenn es um tatsächlich durchgeführte Auswertungsmethoden geht.

Der Code wird nicht verwendet, wenn es sich bei den verwendeten Methoden um Forschungsmethoden oder Erhebungsmethoden handelt oder wenn es um Gegenüberstellungen von Auswertungsmethoden zur Auswahl einer Geeigneten geht.

B.1.6.1 Adaption der Auswertungsmethode

= *die Art und Weise wie eine Auswertungsmethode durch den Forscher abgewandelt wird.*

Der Code wird verwendet, wenn eine Auswertungsmethode verändert angewandt wird/wurde.

Der Code wird nicht verwendet, wenn Methoden kombiniert werden

B.1.6.2 Gründe für die Wahl der Auswertungsmethode

= *Die Gründe, warum sich der Forscher/die Forscherin für genau diese Auswertungsmethode entschieden hat.*

Der Code wird verwendet für die Gründe, die angegeben werden, warum die Auswertungsmethode gewählt wurde.

Der Code wird nicht verwendet, wenn die Wahl der Auswertungsmethode nicht begründet wird oder die Auswertungsmethode erklärt wird.

B.1.6.2.1 Wahl der Auswertungsmethode durch positive Erfahrung

= *wenn die Auswertungsmethode eingesetzt wird, weil bereits positive Erfahrungen mit ihr*

gemacht wurden und deshalb keine weitere in Betracht gezogen werden.

Der Code wird verwendet, wenn die Wahl der Auswertungsmethode auf positive Erfahrung zurückzuführen ist und deshalb keine weiteren in Betracht gezogen werden.

Der Code wird nicht verwendet, wenn es nicht um die Wahl der Auswertungsmethode handelt oder wenn andere Methoden zuvor gegenübergestellt wurden.

B.1.6.3 Vorgehen vor Auswertung

= die Vorarbeit, die geleistet wurde bevor der Auswertevorgang beginnt.

Der Code wird verwendet, wenn vor dem Beginn der Auswertung die Daten "prepariert" werden.

Der Code wird nicht verwendet, wenn nach der Erhebung sofort die Auswertung beginnt ohne dass die Daten vorher strukturiert o.ä. werden

B.1.6.3.1 Einfluss des Pre-Coding auf eigenes Coding

= nach den annotierten Basisaktivitäten müssen nicht unbedingt neue Codes/Konzepte beim offenen Codieren entstehen und es genügt einem bestehenden Konzept neue Eigenschaften hinzuzufügen oder das Konzept komplett so beizubehalten. Insofern ist der Einfluss der annotierten Basiskonzepte eventuell sogar so groß, dass sie kopiert bzw. in das eigene Coding übernommen werden.

Der Code wird verwendet, wenn der Übergang zum offenen Codieren und das offene Codieren beschrieben werden.

B.1.6.4 Gegenüberstellung von Auswertungsmethoden

= Gegenüberstellung von mindestens zwei Auswertungsmethoden.

Der Code wird verwendet, wenn verschiedene Methoden zur Auswertung in Betracht gezogen werden.

Der Code wird nicht verwendet, wenn nur verschiedene Methoden benannt wurden.

B.1.6.5 Quellenwahl zur Auswertungsmethode

= warum eine bestimmte Quelle ausgewählt wurde.

Der Code wird verwendet, wenn begründet wird warum sich der Forscher/die Forscherin für diese Quelle entschieden hat.

Der Code wird nicht verwendet, wenn das Vorgehen erläutert wird.

B.1.6.6 Quelle der Auswertungsmethode

= die Literatur in der die Auswertungsmethode erläutert wird.

Der Code wird verwendet, wenn die Quelle die verwendet wurde, explizit genannt wurde.

Der Code wird nicht verwendet, wenn die Quelle nicht explizit genannt wird sondern nur erläutert wird.

B.1.6.7 definierte Auswertungsmethode

= die verwendete Auswertungsmethode ist definiert und somit wissenschaftlich bekannt.

Der Code wird verwendet, wenn die Auswertungsmethode wissenschaftlich anerkannt ist, d.h.

wenn bereits Literatur existiert, die die Methodik für andere Wissenschaftler überprüfbar gestaltet.

Der Code wird nicht verwendet, falls der Forscher/ die Forscherin eine eigene Methode verfolgt und somit keine feste Methode benennt.

B.1.6.8 grobe Erläuterung der Auswertungsmethode

= es wird erläutert, wie der methodisch gesteuerte Auswertungsprozess abläuft.

Der Code wird verwendet, wenn die Auswertungsmethode erklärt wird.

Der Code wird nicht verwendet, wenn die Wahl der Auswertungsmethode begründet wird.

B.2 Domänenspezifität

B.2.1 nötiges domänenspezifisches Wissen abh. von Forschungsmethode

= das domänenspezifische Wissen, das ein Forscher benötigt, um die Forschung durchführen zu können, ist abhängig von der Forschungsmethode, die er wählt.

Der Code wird verwendet, wenn domänenspezifisches Wissen bzw. Vorwissen, das für ein Projekt, als abhängig von der Forschungsmethode angegeben wird.

Der Code wird nicht verwendet, wenn es um die Erfassung von Domänenwissen geht.

B.2.2 Vorwissen in Domäne

= Fach und Methodenwissen, das vor der Durchführung des Projektes erworben wurde.

Der Code wird verwendet, wenn von Wissen die Rede ist, dass vor dem durchgeführten Projekt schon vorhanden war.

Der Code wird nicht verwendet, wenn es um die Erfassung von Wissen geht.

B.2.2.1 Einfluss von Vorerfahrung

= wenn erfahrenere Forscher über den negativen oder positiven Einfluss von vorher gemachten Erfahrungen und daraus resultierenden Veränderungen in Vorgehensweisen in Forschungsprojekten berichten.

Definition erfahrene Forscher: Forscher, die bereits mehr als ein Projekt durchgeführt haben.

Der Code wird verwendet, wenn der Forscher von gemachten Vorerfahrungen berichtet und aufgrund dieser verändertes Verhalten in seiner Forschung ausmacht. Dies kann auch beispielhaft sein oder schriftlich aus vorhandenen Forschungsberichten resultieren.

Der Code wird nicht verwendet, wenn keine Vorerfahrung aus vorherigen Projekten existiert oder wenn der Forscher keinen Einfluss für seine Forschung darin sieht.

B.2.2.1.1 spezielle Veränderung durch Vorerfahrung

= die spezielle Veränderung, die der Forscher wegen seiner gemachten Erfahrung nun vollzieht bzw. diese neue Handlung, die der Forscher einführt aufgrund der gemachten Erfahrung.

Der Code wird verwendet, wenn Veränderungen speziell benannt werden, die aufgrund von Vorerfahrung vollzogen werden.

Der Code wird nicht verwendet wenn keine konkreten Veränderungen benannt werden, die durch die Vorerfahrung hervorgerufen wurden.

B.2.2.2 Vorwissen in spezifischer Domäne

= *Vorwissen in der spezifischen Domäne, die sich total unterscheidet von der übergeordneten Domäne ist nötig um das Forschungsprojekt durchzuführen.*

Der Code wird verwendet, wenn die Gegebenheit gewährleistet ist, dass eine spezifische Domäne im Forschungsprojekt eingebettet ist und dazu Vorwissen nötig ist.

Der Code wird nicht verwendet, wenn die spezifische Domäne nicht definiert wurde oder nur eine "allgemeine" Domäne erkennbar ist.

B.2.3 Forschungsfeld des Forschers/der Forscherin

= *beschreibt das Forschungsfeld, die Thematik mit der sich der Forscher/die Forscherin auseinandersetzt.*

Der Code wird verwendet, wenn das Forschungsfeld des Forschers klar definierbar ist oder der Forscher es selbst definiert.

Der Code wird nicht verwendet, wenn das Forschungsfeld nicht explizit erwähnt wird.

B.2.3.1 Forschungsarbeit

= *der Inhalt des Forschungsprojekts wird erläutert.*

Der Code wird verwendet, wenn spezielle Inhalte der geleisteten Forschungsarbeit erläutert werden.

Der Code wird nicht verwendet, wenn die Erläuterungen verallgemeinert werden.

B.2.3.2 forschungsspezifisches Feld

= *gemeint ist nicht die übergeordnete Domäne des Forschers wie Informatik oder Wirtschaftspädagogik als solche, sondern die spezifische Domäne beispielsweise die Sprachkompetenz bei Schülern oder Paarprogrammierung. Das konkrete Thema, das der Forscher untersucht.*

Der Code wird verwendet, wenn der Forscher/ die Forscherin das spezielle Forschungsgebiet benennt.

Der Code wird nicht verwendet, wenn der Forscher/die Forscherin Beispiele erwähnt oder nur die übergeordnete Domäne, beispielsweise Wirtschaftspädagogik.

B.2.4 Erfassung von Domänenwissen

= *die Art und Weise wie Vorwissen über die Domäne erfasst wird; dabei geht es sowohl um die Erfassung des Vorwissens zur übergeordneten Domäne, als auch darum das Spezialwissen zur spezifischen Domäne zu erfassen.*

Der Code wird verwendet, wenn es die verschiedenen Möglichkeiten zur Erfassung von Wissen zur übergeordneten Domäne geht (d.h. die Domäne in der das Projekt verankert ist) und wenn es um Vorwissen in der speziellen Domäne geht.

Der Code wird nicht verwendet, wenn Methodiken zur Erfassung von Domänenwissen gegenübergestellt werden.

B.2.4.1 Gegenüberstellung von Methoden zur Erfassung von Domänenwissen

= verschiedene Methoden zur Erfassung von Domänenwissen werden angegeben und deren Ziel und Zweck gegenübergestellt.

Der Code wird verwendet, wenn verschiedene Methoden gegenübergestellt werden um Domänenwissen zu erfassen.

Der Code wird nicht verwendet, wenn es nur generell um die Erfassung von Domänenwissen geht.

B.2.4.2 Erfassung des Domänenwissens projektspezifisch

= die Erhebung des Vorwissens in einer Domäne die man neu erforschen möchte hängt vom Projekt ab.

Der Code wird verwendet, wenn die Art und Komplexität zur Erfassung von Domänenwissen abhängig vom jeweiligen Forschungsprojekt ist.

Der Code wird nicht verwendet, wenn der Forscher sich Wissen projektunabhängig aneignet.

B.2.4.3 Erfassung von Domänen durch Umfrage

= der Forscher eignet sich das Basiswissen für eine ihm unbekannte Domäne durch Umfragen an.

Der Code wird verwendet, wenn Umfragen angegeben werden, um sich Wissen in oder über eine Domäne anzueignen. Dabei ist egal zu welchem Zeitpunkt dieses erworben wird.

Der Code wird nicht verwendet, wenn Umfragen nicht im Rahmen der Aneignung von Domänenwissen benannt werden, sondern als Erhebungsmethode.

B.2.4.4 Erfassung von Domänenwissen durch spezielle Modelle

= es existieren spezielle Modelle um Domänenwissen zu erfassen.

Der Code wird verwendet, wenn spezielle Modelle genannt werden, um Domänenwissen zu erfassen. Dazu muss der Begriff "Modell" explizit erwähnt werden.

Der Code wird nicht verwendet, wenn Domänenwissen nicht über spezielle Modelle erfasst wird.

B.2.4.5 Erfassung von Domänenwissen durch Literaturrecherche

= der Forscher eignet sich das Basiswissen durch Literaturrecherche an. Dies gilt für die initiale neue Suche, wie auch für eine Recherche wenn der Forscher sich bereits neues domänenspezifisches Vokabular angeeignet hat und somit Basiswissen besitzt.

Der Code wird verwendet, wenn Literaturrecherche angegeben wird, um sich Wissen in oder über eine Domäne anzueignen. Dabei ist egal zu welchem Zeitpunkt dieses erworben wird.

Der Code wird nicht verwendet, wenn Recherchetätigkeit nicht im Rahmen der Aneignung von Domänenwissen benannt wird.

B.2.4.6 Erfassung von Domänenwissen durch Konferenzen

= der Forscher informiert sich über eine Domäne durch Konferenzen.

Der Code wird verwendet, wenn Konferenzen angegeben werden, um sich Wissen in oder über eine Domäne anzueignen. Dabei ist egal zu welchem Zeitpunkt dieses erworben wird.

Der Code wird nicht verwendet, wenn Konferenzen nicht im Rahmen der Aneignung von Domänenwissen benannt werden.

B.2.4.7 Erfassung von Domänenwissen durch Interviews

= *der Forscher eignet sich das Wissen für eine ihm unbekannt Domäne durch Interviews (Vorgespräche, Reflexionsgespräche...) an.*

Der Code wird verwendet, wenn Interviews angegeben werden um sich Wissen in oder über eine Domäne anzueignen. Dabei ist egal zu welchem Zeitpunkt dieses erworben wird.

Der Code wird nicht verwendet, wenn Interviews nicht im Rahmen der Aneignung von Domänenwissen benannt werden, sondern als Erhebung.

B.2.4.8 Erfassung von Domänenwissen durch Beobachtung

= *der Forscher eignet sich das Basiswissen für eine ihm unbekannt Domäne durch Beobachtungen an.*

Der Code wird verwendet, wenn Beobachtungen angegeben wird um sich Wissen in oder über eine Domäne anzueignen. Dabei ist egal zu welchem Zeitpunkt dieses erworben wird.

Der Code wird nicht verwendet, wenn Beobachtungen nicht im Rahmen der Aneignung von Domänenwissen benannt wird, sondern als Erhebungsmethode.

B.2.4.9 Erfassung von Domänenwissen durch allgemeinen Überblick

= *wenn der Forscher sich einen generellen ersten Überblick über die Domäne durch nicht-wissenschaftliche Literatur verschafft.*

Der Code wird verwendet, wenn der Forscher eine Suche durchführt zur Erfassung des Domänenwissens um sich einen allgemeinen Überblick zu verschaffen.

Der Code wird nicht verwendet wenn die Verschaffung des allgemeinen Überblicks nicht explizit als Schritt zur Erfassung des Domänenwissens genannt wird.

B.2.4.10 domänenspezifische Erfassung von Domänenwissen

= *wenn eine für die Domäne spezifische Methode zur Erfassung von Domänenwissen gewählt wird.*

Beispiel:

Mailing Liste im Feld Open Source

Der Code wird verwendet, wenn eine spezielle Methode zur Erfassung von Domänenwissen geschildert wird.

Der Code wird nicht verwendet, wenn die Erfassung des Domänenwissens generell beschrieben wird und keine besonderen Eigenschaften aufweist.

B.2.5 Domänendifferenzierung

= *die Attribute durch die Domänen differenziert oder voneinander abgegrenzt werden.*

B.2.5.1 Abgrenzung eigener Domäne durch bekannte Literatur

= *dadurch, dass der Forscher das Fach studiert hat, ist er mit fachspezifischer Literatur konfrontiert worden und kann anhand derer entscheiden, wenn er neue Literatur kennenlernt, ob ihm diese thematisch bekannt ist, oder nicht. Wenn ihm diese bekannt ist, ist sie auch seiner*

Domäne zuzuordnen, anderenfalls nicht.

Der Code wird verwendet, wenn es darum geht die eigene Domäne zu anderen Domänen anhand Literatur abzugrenzen.

Der Code wird nicht verwendet, um zwei unbekannte Domänen voneinander abzugrenzen.

B.2.5.2 Abgrenzung eigener Domäne durch Vorwissen

= Der Forscher kann die Domäne, in der er tätig ist durch sein eigenes Vorwissen (beispielsweise durch Studium erworben) von einer anderen abgrenzen (dadurch dass er gewisses Material kennt).

Der Code wird verwendet, wenn die eigene Domäne zu einer anderen Domäne mit Hilfe von Vorwissen (beispielsweise in spezieller Literatur) abgegrenzt wird.

Der Code wird nicht verwendet, wenn Vorwissen zur Abgrenzung keine Rolle spielt.

2.5.3 Abgrenzung zu weiterer Subdomäne

= wenn eine Subdomäne durch eine weitere Subdomäne abgegrenzt wird.

Der Code wird verwendet, wenn eine Subdomäne mit einer weiteren Subdomäne unterschieden oder durch sie abgegrenzt wird.

Der Code wird nicht verwendet, wenn nicht explizit weitere Domänen zur Abgrenzung herangezogen werden.

B.2.5.4 Domänendifferenzierung anhand Literatur

= wenn das Forschungsprojekt interdisziplinär durchgeführt wird, kann pro Domäne spezielle Literatur ausgemacht werden.

Der Code wird verwendet, wenn der Interviewpartner eine Differenzierung von Domänen vornehmen kann und wenn diese Differenzierung anhand der Literatur vorgenommen wird.

Der Code wird nicht verwendet, wenn gar keine Domänendifferenzierung vorgenommen wird oder nicht anhand Literatur Domänen unterschieden werden.

B.2.5.5 Domänendifferenzierung anhand Theorien

= wenn verschiedene Domänen anhand ihrer Theorien, die sie aufstellen, differenziert werden können.

Der Code wird verwendet, wenn verschiedene Domänen anhand ihrer Theorien differenziert werden. Dies ist vor allem an interdisziplinärer Forschung auszumachen.

Der Code wird nicht verwendet, wenn gar keine Domänendifferenzierung vorgenommen wird oder nicht anhand Literatur Domänen unterschieden werden.

B.2.5.6 Domänendifferenzierung anhand Vokabular

= Anhand der Begriffe oder Konzepte, die in einer spezifischen Domäne verwendet werden, kann eine Domäne zu einer anderen abgegrenzt werden, da hier die Begriffe oder Konzepte anders verwendet werden.

Der Code wird verwendet, wenn Vokabular/Begriffe als Möglichkeit zur Differenzierung oder Abgrenzung von Domänen genannt werden.

Der Code wird nicht verwendet, wenn die Abgrenzung anhand anderer Attribute einer Domäne

vorgenommen wird oder nicht abgegrenzt werden kann.

B.2.5.7 Domänendifferenzierung durch Praktiker und Forscher

= *Die Personen, die in diesem Feld tätig sind oder Forscher bestimmen die Grenzen einer Domäne und können sich damit auch von anderen Domänen unterscheiden.*

Der Code wird verwendet, wenn es um die Abgrenzung einer Domäne zu einer anderen Domäne geht und diese durch Praktiker und Forscher der Domäne bestimmt wird.

Der Code wird nicht verwendet, wenn eine Differenzierung oder Abgrenzung zwischen zwei Domänen durch andere Faktoren als Praktiker oder Forscher im Feld getroffen wird oder gar nicht.

B.2.5.8 Domänendifferenzierung durch weitere Domänen

= *Anhand anderer Domänen kann eine Domäne durch Parallelen oder Unterschiede zu einer anderen abgegrenzt werden.*

Der Code wird verwendet, wenn weitere Domänen zur Differenzierung oder Abgrenzung von Domänen genannt werden.

Der Code wird nicht verwendet, wenn die Abgrenzung anhand anderer Attribute einer Domäne vorgenommen wird oder nicht abgegrenzt werden kann.

B.2.6 Domäneneinfluss

B.2.6.1 Einfluss von spezifischen Domänen auf eigene Forschung

= *spezielle Domänen, nicht unbedingt die, in die die Forschungsarbeit eingebettet ist, haben Einfluss auf das Forschungsprojekt. Dies können auch mehrere Domänen sein.*

Der Code wird verwendet, wenn verschiedene Subdomänen Einfluss auf die Forschung haben.

Der Code wird nicht verwendet, wenn nur von der eigenen Forschungsdomäne gesprochen wird.

B.2.6.1.1 Einfluss von mehreren Domänen auf Forschung

= *auf das spezielle Forschungsgebiet haben mehrere Domänen Einfluss*

Der Code wird verwendet, wenn eindeutig mehrere Domänen benannt werden, die die Forschung beeinflussen.

Der Code wird nicht verwendet, wenn keine weiteren Domänen benannt werden.

B.2.6.1.2 Beispiel für Einfluss der spezifischen Domäne auf Projekt

= *Erklärung wie Wissen in der speziellen Domäne Einfluss auf das Forschungsvorhaben haben.*

Der Code wird verwendet wenn ein Beispiel genannt wird, wie Wissen in der speziellen Domäne sich auswirkt auf den Forschungsprozess.

B.2.6.2 Einfluss der übergeordneten Domäne

= *die Domäne, in die die Forschungsdisziplin eingebettet ist. (i.e. WiPäd, Informatik) und ohne deren Einfluss und Vorwissen in dieser die Untersuchung nicht stattfinden würde.*

Der Code wird verwendet, wenn der Einfluss der Domäne verdeutlicht wird.

Der Code wird nicht verwendet, wenn es um spezielle Domänen geht (i.e. Pair Programming, Schulentwicklung usw.).

B.2.6.3 Einfluss der spezifischen Domäne(n) auf übergeordnete Domäne

= *es wird erklärt, wie sich die spezifischen Domänen zur übergeordneten Domäne verorten lassen. Welche Unterschiede und Gemeinsamkeiten haben die Domänen?*

Der Code wird verwendet, wenn die spezifischen Domänen getrennt betrachtet werden von der übergeordneten Domäne und hinsichtlich Schnittmengen und Unterschieden differenziert werden.

Der Code wird nicht verwendet, wenn nicht zwischen Domänen unterschieden wird.

B.2.6.4 Einfluss von domänenspezifischem Vokabular auf Forschung

= *wenn speziell Vokabular als domänenspezifischer Einfluss angegeben wird.*

Der Code wird verwendet, wenn verwendetes Vokabular als spezifischer Einfluss angegeben wird.

Der Code wird nicht verwendet, wenn mit Vokabular Domänen differenziert werden sollen.

B.2.6.5 Einfluss der Domäne auf Vorgehensweise

= *die Domäne hat Einfluss auf den Vorgehensprozess.*

Der Code wird verwendet, wenn angegeben wird, dass die Domäne Einfluss auf die Vorgehensweise hat, also auf den Forschungsprozess.

Der Code wird nicht verwendet, wenn der Einfluss der Domäne nicht konkret auf die Vorgehensweise bezogen wird.

B.2.6.6 Einfluss der Domäne auf die Wahl der Methode

= *Die Rolle, die eine Domäne spielt bei der Wahl einer Methode.*

Beispiel: Domänen favorisieren spezielle Methoden

Der Code wird verwendet, wenn der Forscher den Einfluss der Domäne auf die Wahl der Methode bestätigt.

Der Code wird nicht verwendet, wenn die Wahl der Methoden als domänenunabhängig bezeichnet wird.

B.2.6.7 Einfluss der Domäne auf Auswertungsmethode

= *die Domäne hat Einfluss auf das Auswertungsformat.*

Der Code wird verwendet, wenn angegeben wird, dass die Domäne Einfluss auf die Auswertungsmethode hat, also die Methodenwahl durch die Domäne gesteuert wird.

Der Code wird nicht verwendet, wenn der Einfluss der Domäne auf Wahl der Auswertungsmethode nicht genannt wird.

B.2.6.8 allgemeiner Einfluss von spezifischem Domänenwissen auf Projekt

= *spezielles Fachwissen hat Einfluss auf die Art der Durchführung des Projektes und/oder die Entscheidungen die gefällt werden.*

Der Code wird verwendet, wenn der Einfluss von spezifischem Domänenwissen bestätigt und

erläutert wird.

Der Code wird nicht verwendet, wenn die Definitionen der Begriffe vorher unklar waren.

B.2.6.9 Einfluss der Domäne auf Erhebungsmethode

= *die Domäne hat Einfluss auf die Wahl der Erhebungsmethode.*

Der Code wird verwendet, wenn angegeben wird, dass die Domäne Einfluss auf die Wahl der Erhebungsmethode hat.

Der Code wird nicht verwendet, wenn der Domäneneinfluss auf die Erhebungsmethode nicht explizit angesprochen wird.

B.2.6.10 Einflussdomäne

= *eine spezielle Domäne, die in das Forschungsprojekt eingebettet ist.*

Der Code wird verwendet, wenn die Domäne explizit benannt wird.

Der Code wird nicht verwendet, wenn keine weitere Domäne benannt wird.

B.2.6.10.1 Subdomäne

= *die Eingrenzung der übergeordneten Domäne auf eine "Unterkategorie", sozusagen eine Spezifizierung der übergeordneten Domäne.*

Der Code wird verwendet, wenn eine Subdomäne von einer übergeordneten Domäne abgegrenzt wird.

Der Code wird nicht verwendet, wenn nicht klar wird, wie die Übergeordnete Domäne und die Subdomäne zusammenhängen.

B.2.7 Codes

= *die entwickelten Codes/Konzepte sind spezifisch für jede Domäne.*

Der Code wird verwendet, wenn innerhalb eines Projekts entwickelte Codes oder Konzepte explizit als domänenspezifisch bezeichnet werden.

Der Code wird nicht verwendet wenn Codes oder Konzepte als nicht domänenspezifisch bezeichnet werden.

B.2.8 Material

= *Material wird als spezifisch für eine Domäne erachtet.*

Material ist beispielsweise Literatur.

Der Code wird verwendet, wenn beispielsweise Literatur als domänenspezifisch bezeichnet wird.

Der Code wird nicht verwendet, wenn es um die Erfassung von Domänenwissen oder Differenzierung von Domänenwissen geht.

B.2.9 spezifisches Domänenwissen

= *es werden Inhalte genannt, die spezifisches Wissen darlegen.*

Der Code wird verwendet, wenn domänenspezifische Inhalte erläutert werden.

Der Code wird nicht verwendet, wenn allgemeine Inhalte zu einem speziellen Vorgehen

erläutert werden.

B.2.9.1 standardisiertes Vokabular

= wenn es schriftlich festgehaltenes und standardisiertes Vokabular gibt, das speziell in der untersuchten Forschungsdomäne vorhanden ist.

Der Code wird verwendet, wenn der Forscher solches Vokabular verwendet und dieses schriftlich festgehalten ist.

Der Code wird nicht verwendet, wenn kein standardisiertes Vokabular existiert.

B.2.9.2 standardisiertes domänenspezifisches Vokabular

= wenn es schriftlich festgehaltenes und standardisiertes Vokabular gibt, das speziell in der untersuchten Forschungsdomäne relevant ist.

Der Code wird verwendet, wenn der Forscher solches Vokabular angibt und dieses schriftlich festgehalten ist.

Der Code wird nicht verwendet, wenn über den allgemeinen Einfluss von spezifischem Vokabular die Rede ist.

B.2.10 Wiederverwendbarkeit

= die Wiederverwendbarkeit von Artefakten (Codes, Codebook, Codeschema) ist domänenspezifisch.

Der Code wird verwendet, wenn die Wiederverwendbarkeit von Artefakten eingeschränkt ist auf die Domäne.

B.2.11 spezieller Fachspezifischer Einfluss auf Projekt

= eine spezielle Bedingung, die das Projekt beeinflusst.

Der Code wird verwendet, wenn ein fach-spezifischer Grund oder eine Bedingung genannt wird, der/die das Projekt beeinflusst.

Der Code wird nicht verwendet, wenn eine allgemeine Erklärung über den vorhandenen Einfluss von fach- resp. Domänenwissen geliefert wird.

B.2.12 Forschungsfrage

= die Forschungsfrage ist spezifisch für jede Domäne

Der Code wird verwendet, wenn die Forschungsfrage als domänenspezifisch bezeichnet wird.

Der Code wird nicht verwendet, wenn es nicht um Domänenspezifika geht.

B.2.13 Informationspool

= es existiert ein Informationspool, der aus vorangegangenen Projekten in das Aktuelle überführt wird, wenn die spezifische Domäne identisch ist. Wenn die spezifische Domäne nicht identisch ist, aber Schnittmengen existieren wird diese übernommen.

Der Code wird verwendet, wenn ein domänenspezifischer Informationspool existiert, der über verschiedene Projekte mitgenommen wird.

Der Code wird nicht verwendet, wenn kein Informationspool existiert.

B.2.14 Hierarchisierung des Datenmaterials

= es werden verschiedene Arten von Datenmaterial erhoben (Audio, Video...) und diese werden unterschiedlich gewichtet für die Auswertung.

Der Code wird verwendet, wenn das erhobene Datenmaterial gewichtet wird.

Der Code wird nicht verwendet, wenn eine solche Hierarchisierung nicht stattfindet.

B.2.15 Domänenidentifizierung anhand Vokabular

= eine spezifische Domäne verwendet für ihre Domäne spezifisches Vokabular oder Begrifflichkeiten. Sie definiert spezielle neue Begriffe oder verwendet Begriffe mit anderen Definitionen als andere Domänen.

Der Code wird verwendet, wenn der Interviewpartner angibt, dass er eine bzw. seine Domäne anhand eines speziellen Vokabulars identifiziert werden kann.

Der Code wird nicht verwendet, wenn Vokabular nicht als Identifikationsmechanismus genannt wird.

B.2.16 Codes domänenspezifisch

= die entwickelten Codes/Konzepte sind spezifisch für jede Domäne

Der Code wird verwendet wenn innerhalb eines Projekts entwickelten Codes oder Konzepte explizit als domänenspezifisch bezeichnet werden.

Der Code wird nicht verwendet wenn Codes oder Konzepte nicht explizit als domänenspezifisch bezeichnet werden.

B.2.17 Material domänenspezifisch

= Material (i.e. Literatur) wird als spezifisch für eine Domänen erachtet.

Der Code wird verwendet, wenn vor allem Literatur als domänenspezifisch angesehen wird.

Der Code wird nicht verwendet, wenn es um die Erfassung von Domänenwissen geht.

B.2.18 Wiederverwendbarkeit domänenspezifisch

= die Wiederverwendbarkeit von Artefakten (Codes, Codebook, Codeschema) ist domänenspezifisch.

Der Code wird verwendet, wenn die Wiederverwendbarkeit von Artefakten eingeschränkt ist auf die Domäne.

B.2.19 Forschungsfrage domänenspezifisch

= die Forschungsfrage ist spezifisch für jede Domäne

Der Code wird verwendet wenn die Forschungsfrage explizit als domänenspezifisch bezeichnet wird.

Der Code wird nicht verwendet, wenn es bei der Nennung des Begriffs "Forschungsfrage" nicht darum geht, was an dem Forschungsprozess domänenspezifisch ist.

B.2.20 domänenspezifischer Informationspool

= es existiert ein Informationspool, der aus vorangegangenen Projekten in das aktuelle

überführt wird wenn die spezifische Domäne identisch ist. Wenn die spezifische Domäne nicht identisch ist, aber Schnittmengen existieren wird diese übernommen.

Der Code wird verwendet, wenn ein domänenspezifischer Informationspool existiert, der über verschiedene Projekte mitgenommen wird.

Der Code wird nicht verwendet, wenn kein Informationspool existiert.

B.2.21 domänenspezifische Hierarchisierung des Datenmaterials

= es werden mehrere Komponenten erhoben (Audio, Video...) und diese werden unterschiedlich gewichtet für die Auswertung.

Der Code wird verwendet, wenn das erhobene Datenmaterial gewichtet wird.

Der Code wird nicht verwendet, wenn eine solche Hierarchisierung nicht stattfindet.

B.3 Metamodell

- Eigenschaften des Metamodells
- Aufbau und Vorgehensweise bei der Erstellung des Metamodells

B.3.1 Übertragbarkeit des Metamodells

= grundsätzlich ist das Metamodell übertragbar. Es werden aber Einschränkungen genannt.

Der Code wird verwendet, wenn der Forscher angibt, dass das aufgestellte Metamodell grundsätzlich übertragbar ist.

Der Code wird nicht verwendet, wenn die Übertragbarkeit ausgeschlossen wird.

B.3.2 Aufbau des Metamodells

= der grundlegende Aufbau des Metamodells wird beschrieben.

Der Code wird verwendet, wenn der Aufbau des Metamodells beschrieben wird.

Der Code wird nicht verwendet, wenn es nicht um den Aufbau des Metamodells geht, sondern um Eigenschaften dessen.

B.3.3 Metamodell domänenspezifisch

= das Metamodell, das erstellt wurde, ist nur für die spezifische Domäne aufgestellt.

Der Code wird verwendet, wenn deutlich wird, dass das aufgestellte Metamodell nur für die Domäne, in der es erstellt wurde, verwendbar ist.

Der Code wird nicht verwendet, wenn es um Code den Aufbau des Metamodells geht.

B.3.4 Erstellung eines Metamodells

= Bildung eines Metamodells

Der Code wird verwendet, wenn ein Metamodell aufgestellt wird.

Der Code wird nicht verwendet, wenn es keines gibt oder es nicht als solches bezeichnet wird.

B.3.4.1 Zweck der Erstellung des Metamodells

= *Ziel und Zweck, warum das Metamodell aufgestellt wurde, werden erläutert.*

Der Code wird verwendet, wenn der Zweck zur Erstellung des Metamodells erläutert wird.

Der Code wird nicht verwendet, wenn es um den Aufbau des Metamodells geht.

B.3.4.2 Gründe für das Aufstellen des Metamodells

= *es werden Gründe genannt, warum das Metamodell überhaupt aufgestellt bzw. entworfen wurde.*

Der Code wird verwendet, wenn Gründe benannt werden warum das Metamodell entworfen wurde.

Der Code wird nicht verwendet, wenn der Aufbau des Metamodells begründet wird.

B.4 Codierschema

= *eine Verschriftlichung der Art und Weise wie der Codiervorgang vollzogen wird.*

- die Regeln, die beim Codieren beachtet werden
- Coding im Team
- verwendete SW und deren Begründung

B.4.1 Codierregeln

= *wenn der Forscher unabhängig von dem Codierverfahren Regeln angibt, die er oder wenn es mehrere sind, sie sich auferlegt haben.*

Der Code wird verwendet, wenn Codierregeln existieren.

Der Code wird nicht verwendet, wenn keine speziellen Codierregeln abgesehen von dem gewählten Vorgehen existieren.

B.4.1.1 festgelegte Codierregel

= *eine Praktik, die für den Codierprozess zugrunde gelegt wurde, unabhängig vom vorher definierten Codierverfahren (Auswertungsmethode)*

Der Code wird verwendet, wenn es unabhängig vom gewählten Codierverfahren (bspw. Codieren nach GT) Regeln gibt, die sich der oder die Forscher selbst geben.

Der Code wird nicht verwendet, wenn es abgesehen von dem ausgewählten Auswertungs- oder Codierverfahren keine Regeln gibt, die sich die Forscher auferlegt haben auch informell.

B.4.1.2 keine Codierregeln

= *wenn der Forscher unabhängig von dem Codierverfahren keine Regeln angibt, die er oder wenn es mehrere sind, sie sich auferlegt haben.*

Der Code wird verwendet, wenn keine Codierregeln existieren.

Der Code wird nicht verwendet, wenn spezielle Codierregeln abgesehen von dem gewählten Vorgehen existieren.

B.4.1.3 Coding ohne Software

= zum Auswerten der Daten wird keine Software verwendet. Diese Art Auswertung wird begründet.

Der Code wird verwendet, wenn keine Auswertungssoftware benutzt wird.

Der Code wird nicht verwendet, wenn Software benutzt wird.

B.4.1.4 Coding mit MAXQDA

= die verwendete Software zum Auswerten der Daten ist MAXQDA.

Der Code wird verwendet, wenn die Auswertungs-bzw. SW MAXQDA verwendet wird.

Der Code wird nicht verwendet, wenn andere SW oder gar keine SW benutzt wird.

B.4.1.5 Coding mit ATLAS.ti

= die verwendete Software zum Auswerten der Daten ist ATLAS.ti. Die Wahl wird begründet.

Der Code wird verwendet, wenn die Auswertungs-bzw. SW ATLAS.ti verwendet wird.

Der Code wird nicht verwendet, wenn andere SW oder gar keine SW benutzt wird.

B.4.1.5.1 Gründe für die Wahl von ATLAS.ti

= warum ATLAS.ti als SW eingesetzt wird.

Der Code wird verwendet, wenn der Einsatz von ATLAS.ti begründet wird.

Der Code wird nicht verwendet, wenn keine Gründe angegeben werden.

B.4.1.6 Coding im Team

= wenn mehrere Forscher im Team Material codieren und sich untereinander absprechen bzgl. der Auswahl der Codes.

Der Code wird verwendet, wenn das Coding im Team stattfindet und Absprachen bezüglich der Codes erfolgen.

Der Code wird nicht verwendet, wenn die Forscher alleine codieren.

B.4.2 Erstellung eines Codierschemas

= der Forscher erstellt ein eigenes Codierschema, das für weitere Forscher wiederverwendbar gestaltet ist.

Der Code wird verwendet, wenn ein eigenes Codierschema erstellt wird, das auch für die zukünftige Forschung wiederverwendbar ist.

Der Code wird nicht verwendet, wenn ein Forscher ein bereits existierendes Codierschema wiederverwendet.

B.4.3 weitere Codierschemata

= es existieren Arbeiten, die auch Codierschemata aufstellen.

Der Code wird verwendet, wenn der Forscher die Existenz weiterer Codierschemata (auch wenn nicht in ihrer Domäne) bestätigt. Dies muss nicht unbedingt belegt werden.

Der Code wird nicht verwendet, wenn sich der Forscher unsicher ist, ob weitere

Codierschemata existieren oder dies nicht bestätigt.

B.4.4 (Wieder-)Verwendung eines Coding Schemas

= *der Forscher verwendet oder verwendet ein Codierschema wieder. Die Wiederverwendbarkeit eines Codierschemas ist also gewährleistet bzw. funktioniert.*

Der Code wird verwendet, wenn explizit von einem Codierschema gesprochen wird, das der Forscher prinzipiell verwendet oder wiederverwendet. Das kann aus fremden Projekten sein. Es gibt verschiedene Szenarien der Wiederverwendung des Schemas (d.h. auch nur Teile können wiederverwendet werden)

Der Code wird nicht verwendet, wenn keine explizites Codier Schema definiert werden kann oder der Begriff unbekannt ist.

B.5 Codebook

= *synonym zu Basisvokabular, Rahmenwerk. Die Zusammenstellung aller Codes und deren Verwendung.*

B.5.1 Memos

= *zeigt die Verwendung von Memos an. Diese können auch als "Ergebnismemo" verwendet werden, d.h. wenn nicht die Verwendung eines Codes begründet wird, sondern das Ergebnis, warum man einen bestimmten Code verwendet hat.*

Der Code wird verwendet wenn, eine Art von Memos verwendet wurde. (Pro Datei/ für Codes etc.)

Der Code wird nicht verwendet, wenn auf die Verwendung von Memos verzichtet wird.

B.5.1.1 Notizen abhängig von Projektstatus

= *die Art und Weise, wie Notizen während einer Informationssammlung/Erhebung gemacht werden ist abhängig vom Fortschritt oder Status des Projektes.*

Der Code wird verwendet, wenn Notizen gemacht werden und diese vom Status des Projekts abhängig sind.

Der Code wird nicht verwendet, wenn die Notizen unabhängig vom Projekt gemacht werden oder gar nicht.

B.5.1.2 keine Memos

= *wenn weder Code noch Datei noch Projektmemos verwendet werden.*

Der Code wird verwendet, wenn der Forscher keine Memos benutzt.

Der Code wird nicht verwendet, wenn der Forscher Codememos, Dateimemos oder eine andere Art von Memos für seine Analysearbeit verwendet.

B.5.1.3 Einfluss von Memos

= *wenn der Forscher Memos schreibt und für sich feststellt, dass diese in irgendeiner Art und Weise Einfluss auf seine Forschung haben. Das kann geringerer oder größerer Einfluss sein.*

Der Code wird verwendet, wenn der Forscher die Memos als nützlich empfindet und diese im Forschungsprozess immer wieder anschaut, aktualisiert und neue generiert.

Der Code wird nicht verwendet, wenn die Memos keinen Einfluss auf den Forschungsprozess

haben oder als unwichtig erachtet werden.

B.5.1.4 Memos für das Ergebnis der angewendeten Regeln

= *die Begründung warum ein spezieller Code verwendet wurde und der Weg dorthin und warum nicht ein anderer Code verwendet wurde, wird in Memos festgehalten.*

Der Code wird verwendet, wenn nicht die eigentlichen Codememos verwendet werden, sondern Begründungen für die Wahl eines bestimmten Codes darin beschrieben werden.

Der Code wird nicht verwendet, wenn die Memos nicht das regelbasierte Ergebnis für einen bestimmten Code enthalten.

B.5.1.5 Begründung des Nutzens von Memos

= *wenn die Verwendung von Memos begründet wird.*

Beispiel: die Memos, die angelegt werden um eine Entscheidung für einen Code zu erläutern legt der Forscher für sich selbst, va. zur Zeitersparnis an um bestimmte Überlegungen nicht doppelt zu machen.

Der Code wird verwendet, wenn Memos angelegt werden und deren Nutzen oder Nützlichkeit begründet wird.

Der Code wird nicht verwendet, wenn der Nutzen und die spätere Verwendung von Memos unklar bleiben.

B.5.2 Kategorienbildung

= *das Vorgehen, nach dem die Kategorien gebildet wurden, falls es ein definiertes gab. Dabei sind Kriterien wie der Zeitpunkt der Erstellung und Vorwissen dazu wichtig.*

Der Code wird verwendet, wenn die Art und Weise wie Kategorienbildung geschieht, geschildert wird.

Der Code wird nicht verwendet, wenn es sich um die Erhebungsmethode oder die Forschungsmethode handelt und diese nicht ein Vorgehen zur Kategorienbildung beinhalten.

B.5.2.1 Kein Codierprozess

= *Der Forscher hat keine spezielle Vorgehensweise oder Art und Weise wie er Codes vergibt, oder wie diese aufgebaut sind abgesehen von der Auswertungsmethode.*

Der Code wird verwendet, wenn der Forscher angibt, keinen speziellen Codierprozess zu haben bzw. seine Codes immer auf eine andere Art und Weise (projektabhängig) zu vergeben.

Der Code wird nicht verwendet, wenn ein Muster in der Art und Weise wie der Forscher codiert erkennbar ist. Es kann sein, dass der Forscher einen impliziten Codierprozess verfolgt.

B.5.2.2 Teilcoding

= *es wurde nur der erste Teil des Codings (das open Coding) durchgeführt. Dieses wird auch begründet.*

Der Code wird verwendet, wenn beispielsweise beim Auswerten nach Strauss und Corbin, wo offenes, selektives und axiales Codieren nacheinander angewandt wird, nur offenes Codieren angewandt wird. Das Teilcoding wird begründet.

Der Code wird nicht verwendet, wenn nicht explizit deutlich gemacht wird, dass nur ein Teil des Codierprozesses ausgeführt wird.

B.5.2.3 grobe Erläuterung zum Vorgehen der Kategorienbildung

= *die Methode, wie man auf die Kategorien kommt, wird erläutert.*

Der Code wird verwendet, wenn das Vorgehen zum Bilden von Kategorien erläutert wird.

Der Code wird nicht verwendet, wenn die Methode nur benannt aber nicht näher darauf eingegangen wird.

B.5.3 Zweck des Codebooks

= *der Zweck des Codebooks wird erläutert.*

Beispiel: Nach dem Annotieren der Codes aus dem Codebook, hat der Forscher einen "besseren" Zugang zu den Daten hinsichtlich seiner Forschungsfrage.

Der Code wird verwendet, wenn der Zweck des Codebooks erläutert wird: Es soll dem Forscher dabei helfen die Daten im Blick seiner Forschungsfrage besser betrachten zu können.

Der Code wird nicht verwendet, wenn nur der Aufbau des Codebooks erläutert wird.

B.5.4 Validierung eines Codebooks

= *wenn die erstellten Codes nach einer Methodik validiert werden.*

Beispiel: Nach der Erstellung eines Codebooks zu bestimmtem Material wird ein Kollege gebeten einen kleinen Teil eigenständig zu vercoden. Unterschiede im Coding werden diskutiert und anschließend das Codebook damit aktualisiert und eben validiert.

Der Code wird verwendet, wenn ein Codebook existiert und dieses validiert wird.

Der Code wird nicht verwendet, wenn keine Validierung stattfindet.

B.5.5 Aufbau des Codebooks

= *wie das Codebook strukturiert ist.*

Der Code wird verwendet, wenn die Struktur und der Aufbau des Codebooks erläutert werden.

Der Code wird nicht verwendet, wenn es um den Aufbau der einzelnen Codes geht.

B.5.5.1 Aufbau des standardisierten Vokabulars als Codebook

= *die Struktur der konkreten Begriffe (auch Konzepte), Phänomene. Diese sind als Codes zu verstehen, die insgesamt ein Codebook ergeben.*

Der Code wird verwendet, wenn der Aufbau resp. der Inhalt des Basisvokabulars beschrieben wird.

Der Code wird nicht verwendet, wenn es um die Entstehung des Vokabulars geht.

B.5.6 (Wieder-)verwendung des Codebook

= *das Codebook, das aufgestellt wurde, ist wiederverwendbar gestaltet und wird auch praktisch wiederverwendet. Dabei kann auch beschrieben werden, wie man das Codebook verwendet bzw. es bereits verwendet wird.*

Der Code wird verwendet, wenn das Codebook, das der Forscher generiert während der Auswertung, für weitere Forscher wiederverwendbar gestaltet ist. Es wird in der Praxis eingesetzt. Auch der praktische Einsatz kann beschrieben sein.

Der Code wird nicht verwendet, wenn die eindeutige Wiederverwendbarkeit nicht gewährleistet ist.

B.5.6.1 Wichtigkeit der Eigenständigen Codierung

= *es wird hervorgehoben, dass es wichtig ist, keine Codes zu übernehmen, da erst hierbei die neuen Perspektiven entstehen.*

Der Code wird verwendet, wenn dieses Phänomen der eigenständigen Codierung angesprochen wird und deren Wichtigkeit begründet wird.

Der Code wird nicht verwendet, wenn nicht explizit auf die eigenständige Codierung eingegangen wird.

B.5.7 Erstellung des Codebooks

= *Die Art und Weise wie das Codebook erstellt wurde.*

Der Code wird verwendet, wenn der Forscher beschreibt, wie das Codebook erstellt wurde resp. wie dabei vorgegangen wurde.

Der Code wird nicht verwendet, wenn der Forscher gar kein Codebook als solches schreibt oder nicht klar ist wie es entstanden ist.

B.5.7.1 erweiterbares Codebook

= *das Codebook ist erweiterbar gestaltet und neue Konzepte können hinzugefügt werden.*

Der Code wird verwendet, wenn es um die Erweiterbarkeit des Codebooks geht und den Zusammenhang zu bestimmten Auswertungsmethoden.

Der Code wird nicht verwendet, wenn es um den konkreten Aufbau des Codebooks geht.

B.5.7.2 Entstehung des domänenspezifischen Vokabulars

= *die Art und Weise, wie das grundlegende Vokabular identifiziert wurde und entstanden ist.*

Der Code wird verwendet, wenn es explizite Entstehungsweisen gibt hinsichtlich der Identifizierung des domänenspezifischen Vokabulars.

Der Code wird nicht verwendet, wenn die Entstehung des Vokabulars unklar bleibt

B.5.7.3 Erläuterung einer Codegruppierung

= *die vergebenen Codes sind gruppiert in Klassen und diese wiederum gruppiert. Es wird erläutert, welche Klassen zusammengehören und eine Konzeptgruppe ergeben.*

Der Code wird verwendet, wenn eine Gruppe oder Zusammenfassung von Codes erläutert wird.

Der Code wird nicht verwendet, wenn es um Codierschemata geht.

B.5.7.4 Aufnahme-strategie von Codes in das Codebook

= *Wann und warum Codes in das Codebook aufgenommen werden.*

Der Code wird verwendet, wenn der Forscher beschreibt, welche Codes in das Codebook aufgenommen werden.

Der Code wird nicht verwendet, wenn es um den Aufbau des Codebooks geht.

B.5.7.5 Erstellung des Codebooks durch Verwendung von Codierschema

= die Erstellung des Codebooks wurde zusätzlich zu eigenen Codes mit einem Codierschema durchgeführt, das aus einer ähnlichen Arbeit resultierte.

Der Code wird verwendet, wenn ein Codebook zusätzlich mit der Hilfe eines Codierschemas aus einem anderen Projekt erstellt wird.

Der Code wird nicht verwendet, falls das Codebook nur durch eigene Codes erstellt wird.

B.6 Codeschema

= der standardisierte Aufbau eines Codes.

B.6.1 Wiederverwendung von Codes

= einige Codes sind wiederverwendbar.

Der Code wird verwendet, wenn der Forscher Codes aus einem bestehenden Codebook wiederverwendet für sein Projekt.

Der Code wird nicht verwendet, wenn es um die Verwendung des Codierschemas, d.h. um die Art der Codierung geht.

B.6.2 Wiederverwendung des Codeschemas

= das Codeschema ist wiederverwendbar.

Der Code wird verwendet, wenn der Forscher ein Codeschema aus einem fremden oder vorherigen Projekt wiederverwendet. Dabei ist Art und Aufbau der Codes gemeint.

Der Code wird nicht verwendet, wenn es nicht um die Verwendung von konkreten Codes geht.

B.6.3 Aufbau der Codes

= Es existiert ein Muster bzw. Schema nach dem Codes aufgebaut werden. (Codeschema)

(Zu einem Objekt existiert eine definierte Verbenmenge, die einem speziellen Objekt aus der Objektmenge zugeordnet werden können. Hierzu wird ein Beispiel gegeben.)

Der Code wird verwendet, wenn der Forscher erklärt, wie ein Code aufgebaut ist. (Ein Code besteht aus einem Objekt und einem Verb. Dabei geht er beispielhaft darauf ein, welche Verben für ein spezielles Objekt zur Verfügung stehen.)

Der Code wird nicht verwendet, wenn es um das Codierschema geht.

B.6.3.1 kein spezieller Aufbau der Codes

= die vergebenen Codes haben keinen speziellen Aufbau (wie bspw. 2 Worte)

Der Code wird verwendet, wenn die Codes die der Forscher vergibt keinen speziellen Aufbau aufweisen, es kein Codeschema gibt.

Der Code wird nicht verwendet, wenn die Codes einen speziellen Aufbau haben, wie beispielsweise zwei Worte oder Verb und Objekt.

B.6.3.2 Codes projektspezifisch

= die entwickelten Codes/Konzepte sind spezifisch für jedes Projekt.

Der Code wird verwendet, wenn die innerhalb eines Projekts entwickelten Codes oder Konzepte explizit als projektspezifisch bezeichnet werden.

Der Code wird nicht verwendet, wenn Codes oder Konzepte nicht explizit als projektspezifisch bezeichnet werden.

B.6.3.3 Beispielcode mit Erläuterung

= ein Code, der aus dem Code Schema entstanden ist und aus einer definierten Menge an Begriffen gewählt wurde.

Der Code wird verwendet, wenn der Forscher einen Beispielcode und dessen Bedeutung erläutert.

Der Code wird nicht verwendet, wenn der Code nicht aus einer definierten Begriffsmenge gewählt wurde.

B.6.3.4 Begründung für den Aufbau der Codes

= der Aufbau bzw. das Muster nach dem Codes gebildet werden, wird begründet.

Der Code wird verwendet, wenn die Struktur der Codes begründet wird.

Der Code wird nicht verwendet, wenn der Aufbau erläutert wird.

B.7 Vorlagen

= der Forscher verwendet Vorlagen.

Eine Vorlage ist dann eine Vorlage, wenn sie mehr als nur einmal verwendet wird. Diese Vorlage benötigt kein spezielles Format oder Struktur.

Der Code wird verwendet, wenn Vorlagen generell verwendet werden.

Der Code wird nicht verwendet, wenn die Verwendung von Vorlagen verneint wird.

B.7.1 formales Template

= ein Template das verschriftlicht ist.

Der Code wird verwendet, wenn der Forscher/die Forscherin ein Template oder einer Vorlage für irgendetwas im Forschungsprozess verwendet.

Der Code wird nicht verwendet, wenn keine verschriftlichte Form für eine Vorlage vorliegt.

B.7.1.1 empfohlenes Template

= Ein Template, das innerhalb des Lehrstuhls verwendet wird und gängige Praxis ist.

Der Code wird verwendet, wenn es am Lehrstuhl ein Template gibt, das immer wieder eingesetzt wird und empfohlen wird.

Der Code wird nicht verwendet, wenn das Template nur individuell eingesetzt wird.

B.7.1.2 informelle Vorlage

= eine Vorlage, die nicht so genannt wird (also nicht dokumentiert), sondern eher ältere bereits vorhandene Dokumente, die für das aktuelle Dokument, angepasst wird.-> Best Practice

Der Code wird verwendet, wenn keine formale Vorlage, die auch als Vorlage benannt ist, existiert, sondern über bereits vorhandene ähnliche Dokumente, die sozusagen als informelle

Vorlage gelten, das neue Dokument erstellt wird. -> Best Practices

Der Code wird nicht verwendet, wenn eine formale Vorlage existiert.

B.7.2 keine Guideline

= *der Lehrstuhl oder die Forschungsgruppe hat keine Guideline, an die sich der Forscher hält, wenn er ein neues Projekt beginnt. Es gibt keinen Ablauf mal abgesehen vom Forschungsprozess, den er für sein Projekt wählt.*

Der Code wird verwendet, wenn es keinen Vorgehensprozess, abgesehen vom gewählten Forschungsprozess respektive der Forschungsmethode gibt.

Der Code wird nicht verwendet, wenn es in der Forschungsgruppe oder dem Lehrstuhl ein gewisses Schema zum Ablauf oder Prozedere eines Projektes gibt.

B.7.3 keine Vorlagen

= *wenn am Lehrstuhl oder der Forschungsgruppe oder auch persönlich keine Vorlagen verwendet werden. Das können auch informelle Vorlagen sein.*

Der Code wird verwendet, wenn am Lehrstuhl/in der Forschungsgruppe oder persönlich keine Muster oder Vorlagen (wieder)verwendet werden.

Der Code wird nicht verwendet, wenn persönliche Vorlagen existieren (durch vorherige Projekte) oder wenn der LS oder Kollegen Vorlagen zur Verfügung stellen (Wenn beispielsweise immer wieder der gleiche Aufbau bei Interviewbögen verwendet wird). Die Vorlagen können auch informell existieren.

B.7.4 Guideline

= *es existiert ein (informeller) Prozess für ein Vorgehen innerhalb des Forschungsprozesses.*

Der Code wird verwendet, wenn eine Art von Guideline innerhalb des Forschungsprozesses existiert (beispielsweise für das Coding oder die Erfassung von Domänenwissen). Diese muss nicht formal definiert sein.

Der Code wird nicht verwendet, wenn es sich nicht um die gewählte Forschungsmethode handelt.

B.7.4.1 Beschreibung der Guideline

= *es existiert ein (informeller) Prozess für ein Vorgehen innerhalb des Forschungsprozesses. Dieser wird beschrieben.*

Der Code wird verwendet, wenn eine Art von Guideline innerhalb des Forschungsprozesses existiert (beispielsweise für das Coding oder die Erfassung von Domänenwissen). Diese muss nicht formal definiert sein. Der Forscher beschreibt sein Vorgehen.

Der Code wird nicht verwendet, wenn die Forschungsmethode beschrieben wird.

B.8 Forschungsart

= *die Art der Forschung.*

Beispiel: qualitative Sozialforschung. Generell bedeutet qualitative Sozialforschung die Erhebung und Auswertung nicht standardisierter Daten.

Der Code wird verwendet, wenn der Forscher die Forschungsart anspricht.

Der Code wird nicht verwendet, wenn es um die Forschungsmethode geht.

B.8.1 Forschungsart am LS

= *die Art von Forschung, die am Lehrstuhl oder der Forschungsgruppe hauptsächlich betrieben wird.*

Der Code wird verwendet, wenn die Art von Forschung definiert wird.

Der Code wird nicht verwendet, wenn die Forschungsart nicht konkret benannt wird.

B.9 Teamarbeit

= *Teamarbeit wird so verstanden, dass man beispielsweise tatsächlich gemeinsam auswertet und sich mit Kollegen über die verwendeten Codes austauscht oder gemeinsam ein Projekt durchgeführt wird und dementsprechend Absprachen jeglicher Art und Weise erfolgen (Arbeitsweise, Umgang mit Ressourcen).*

Der Code wird verwendet, wenn beispielsweise die Daten-Auswertung im Team erfolgt oder generell ein Projekt aus mehreren Personen besteht.

Der Code wird nicht verwendet, wenn keine Teamarbeit stattfindet.

B.9.1 informeller Erfahrungs-/Vorgehens Austausch

= *Der Forscher/ die Forscherin holt sich Informationen bzw. berät sich informell mit Kollegen über mögliches Vorgehen/gängige Praxis o.ä.-> Peer Review*

Der Code wird verwendet, wenn zwar keine Teamarbeit als solche stattfindet, trotzdem aber Lehrstuhlmeetings oder lockerer Austausch/ Gespräche/Diskussionen unter Kollegen. Dabei können auch bereits vorhandene Vorlagen ausgetauscht werden; sie werden aber nicht gemeinsam bearbeitet. Es existieren auch keine verschriftlichen Protokolle, Guidelines o.ä. der Gespräche.

Der Code wird nicht verwendet, sobald während des Gesprächs/Meetings schriftliche Artefakte entstehen oder Entscheidungen nicht mehr individuell, sondern im Team getroffen werden.

B.9.2 Keine Teamarbeit

= *Teamarbeit wird so verstanden, dass man beispielsweise tatsächlich gemeinsam auswerten und sich mit Kollegen über die verwendeten Codes austauscht oder gemeinsam ein Projekt durchgeführt wird.*

Der Code wird verwendet, wenn keine Projektarbeit mit mehreren Teilnehmern pro Projekt gegeben ist und Forscher alleine auswerten.

Der Code wird nicht verwendet, wenn beispielsweise die Daten-Auswertung im Team erfolgt oder generell ein Projekt aus mehreren Personen besteht.

B.9.3 Regel beim Teamcoding

= *eine Regel oder Absprache (auch informell), die sich das Team für das Codieren im Team gibt.*

Der Code wird verwendet, wenn eine formelle oder informelle Regel zum Codieren/Auswerten im Team existiert.

Der Code wird nicht verwendet, wenn keine Regeln bestehen.

B.9.4 Vor/Nachteile im Teamcoding

= es werden Vor und Nachteile im Teamcoding vorgestellt.

Der Code wird verwendet, wenn Vor- und Nachteile angesprochen werden, die das Codieren im Team betrifft.

B.9.5 Vorgehensweise zum Coding im Team

= die Art und Weise wie im Team bzw. im Paar codiert wird.

Der Code wird verwendet, wenn im Team codiert wird und die Vorgehensweise dazu beschrieben wird.

Der Code wird nicht verwendet, wenn das Vorgehen im Team nicht näher beschrieben wird.

B.10 weitere Forschung

= Forschung, die neben dem Hauptforschungsgebiet noch durchgeführt wurde und eventuelle Unterschiede, die sich ergeben haben.

Der Code wird verwendet, wenn explizit weitere Forschung erwähnt wird, beispielsweise um Unterschiede hervorzuheben.

Der Code wird nicht verwendet, wenn keine weiteren Forschungsprojekte erläutert werden.

10.1 Unterschied zwischen Projekten

= ein Unterschied, wie sich die durchgeführten Projekte konkret unterscheiden.

Beispiel: hinsichtlich der Erhebung und der daraus folgenden Codiergranularität

Der Code wird verwendet, wenn konkret mindestens 2 Projekte unterschieden werden und diese anhand einem Merkmal herausgestellt werden.

Der Code wird nicht verwendet, wenn der Unterschied nicht konkret hervorgehoben wird.

B.11 Rückkopplungen

B.11.1 schriftliche Rückkopplungen

= gemachte Erfahrungen fließen in das nächste ähnliche Projekt mit ein und zwar durch Projektberichte.

Der Code wird verwendet, wenn der Forscher sein Vorgehen in einem Bericht reflektiert und/oder evaluiert und dieses in schriftlicher Form festgehalten wird.

Der Code wird nicht verwendet, wenn informelle Feedbackloops oder Rückkopplungen stattfinden.

11.2 nicht schriftliche Rückkopplungen

= gemachte Erfahrungen fließen über Erfahrung in das nächste ähnliche Projekt mit ein, werden aber nicht in schriftlicher Form festgehalten.

Der Code wird verwendet, wenn der Forscher zwar sein Projekt reflektiert aber keine Protokolle/Berichte o.ä. existiert.

Der Code wird nicht verwendet, wenn eine formale Art von Feedbackloops stattfindet oder wenn Feedbackloops resp. Rückkopplungen verneint werden.

B.11.3 nicht schriftlichen Rückkopplungen

= *gemachte Erfahrungen fließen zwar in das nächste ähnliche Projekt mit ein, werden aber nicht in schriftlicher Form festgehalten.*

Der Code wird verwendet, wenn der Forscher zwar sein Projekt reflektiert aber keine Protokolle o.ä. existiert.

Der Code wird nicht verwendet, wenn eine formale Art von Feedbackloops stattfindet oder wenn Feedbackloops resp. Rückkopplungen verneint werden.

B.12 Datenmaterial

= *die Art von Daten, die bei der Forschung erworben und ausgewertet wird.*

Der Code wird verwendet, wenn das Datenmaterial, das für die Auswertung erhoben wird, benannt wird.

Der Code wird nicht verwendet, wenn es nicht um das Datenmaterial für die durchgeführte Forschung/das Projekt geht.

B.12.1 Beobachtungsbogen

= *das verschiedenartige Aussehen eines Beobachtungsbogen abhängig von der jeweiligen Absicht wird grob beschrieben.*

Der Code wird verwendet, wenn der Aufbau eines Beobachtungsbogens erläutert wird.

Der Code wird nicht verwendet, wenn es sich nicht um einen Beobachtungsbogen handelt oder nicht der Aufbau bzw. Beschaffenheit eines Beobachtungsbogens beschrieben wird.

Appendix C Rules for Transcription

- Both, the spoken words of the interviewer as well as the spoken words of the interviewee are transcribed.
- Every tenth row a time code is set.
- Rows are numbered automatically.
- Paragraphs are set at least when the speaker changes.
- The text of the interviewer is set italic and bold.
- Lower and uppercase as well as punctuation characters are observed.
- Question marks were only used if they are obvious.
- Exclamation marks are not used.
- Verbatim quotes are set to double quotes.
- Dialectal characteristics and colloquial language are transferred to standard language.
- For the first interview fillers were transcribed in order to get the self-consciousness of the interview partner.

Appendix D Axial Coding Overview

Coding paradigm	Category	Interview							
		AR	FB	AB	FZ	KDS	SS	HH	JJ
Phenomenon	(Re)use of domain knowledge while developing codes	-	-	-	<p>He explicitly reuses a defined code schema that was set up before. This code schema implies domain knowledge. <i>Example:</i> "FZ: Ja genau. Also ich setze das konkret ein um besser zu verstehen, was eigentlich in so einer Sitzung passiert.[...]" FZ: ...und deswegen, also ich benutze die Basisaktivitäten um interessante Stellen quasi durchzucodieren, aber das durchcodierte Ergebnis ist für mich gar nicht interessant... "FZ: ...sondern die Entscheidung, die ich zwischendurch bei den ganzen Grenzfällen getroffen habe, die sind für mich interessant, weil die mir dabei helfen zu verstehen, was da wirklich in der Sitzung gerade passiert ist."</p>	<p>In some extent domain knowledge is reused so far as codes are set up influenced by experience <i>Example:</i> "DS: Wir haben das bisher so in einem ersten Schritt gemacht. Aber das muss ja nicht immer so gemacht werden."</p>	<p>He explicitly uses domain knowledge to create his codes. <i>Example:</i> "SS: Vielleicht einen Satz vorweg. Ich denke, dass ich meine Analysetätigkeit ohne mein Fachwissen gar nicht hätte machen können."</p>	-	-

Coding paradigm	Category	Interview							
		AR	FB	AB	FZ	KDS	SS	HH	JJ
Trigger/Reason	Coding schema			<p>For one project a coding scheme was used. It has been reused from another extern project. <u>Example: "Appendix B specifically list up coding scheme. And since the task that they were trying to do is similar to the tasks that we are trying to do, in this case a literature review, part of what they used was they coded some of the metadata of the papers and then did a statistical analysis on it. So, for example, they looked at the research method that was used in the paper. [...]</u> <u>AB: They looked at the domain that the paper was published in and some other things like this and data collection method. [...]</u> <u>AB: And these categories and the codes within them we felt would be or I felt would be broadly applicable to most literature reviews that you could do.</u></p>	<p>He follows a specified coding schema, including a meta model, a code schema etc. <u>Example: "I: Ok. Und das setzen Sie praktisch ein? FZ: Ja genau. Also ich setze das konkret ein um besser zu verstehen, was eigentlich in so einer Sitzung passiert. [...]</u> <u>FZ: Also immer wenn ich ein Bauchgefühl habe, da ist irgendeine eine Episode, die finde ich irgendwie interessant, da wird, jetzt aus meiner Forschungsfrage her, da wird viel Wissen transferiert... [...]</u> <u>FZ: ...was für mich interessant ist. Dann nehme ich die Basiskonzepte her und versuche die zu annotieren, und das ist nämlich nicht immer leicht, weil es sehr viele Grenzfälle gibt..."</u></p>	<p>The codes are always created hybrid: First starting with the topic over the interview guide to the transcript. "theory driven" <u>Example: "DS: Theoriegeleitet die ersten Kategorien anhand des Leitfadens entwickelt, die ersten Interviews kodiert. Und dann festgestellt, es fehlen verschiedene Kategorien, oder einzelne Kategorien müssen erweitert werden, oder 2 Stück auch mal zusammengefasst weil es ganz einfach Sinn macht."</u></p>	<p>He creates his own coding schema implying a meta model, a code schema, a defined perspective on the data and a coding practice (Pair Coding) <u>Example: "Und ja, dann bin ich halt mit einer Kollegin damals, weil wir gesagt haben, ok wir versuchen das mal gemeinsam in sogenannten Paarcodierungssitzungen. Das ist dann so ein Methode, die ist nicht Standard, auch nicht nach Strauss, aber da habe ich mir gedacht das könnte sinnvoll sein; die direkte Auseinandersetzung über ob das sozusagen ein elementarer Bestandteil ist oder nicht. Und das habe ich dann gemacht, mit der Kollegin zusammen."</u></p>		

Coding paradigm	Category	Interview							
		AR	FB	AB	FZ	KDS	SS	HH	JJ
Context	Qualitative research process	<p>The Critical Incident Technique has been applied <i>Example : "AR: Die Methode dazu, die dahinterstand, war die Critical Incident Technique."</i></p>	<p>The Design Best Research approach has been applied. <i>Example: "Als etabliert kann man keinen Prozess bezeichnen, aber der Design Best Research Ansatz wird derzeit empfohlen. Dieser Prozess ist noch in der Etablierungsphase. Davor gab es keinen solchen Ansatz. "</i></p>		<p>GTM has been applied. <i>Example: "FZ: Ja. Also, die Forschungsmethode, die damals vor vielen Jahren ausgewählt wurde, ist die Grounded Theory."</i></p>		<p>GTM has been applied. <i>Example: "SS: Ja genau. Wir haben bisher fast ausschließlich mit der Grounded Theory gearbeitet."</i></p>	<p>For some projects the data is collected in a qualitative research process. <i>Example: "Alle Prozesse die wir durchführen basieren in der Regel auf einer pilotierenden Vorgehensweise."</i></p>	<p>For some projects the data is collected in a qualitative research process- the process is project dependent. <i>Example: "[...]leine grundsätzliche Frage und zwar welche Art von Forschung führen Sie denn durch? JJ: Im Bezug auf qualitative Forschungsmethoden?" "[...] gibt es eine feste Forschungsmethode, ok ein Beispiel wäre jetzt Grounded Theory oder nehmen wir die Critical Incident Technique, ja eben eine feste Forschungsmethode die Sie verfolgen? JJ: Nein das wird immer an das Forschungsprojekt angepasst"</i></p>

Coding paradigm	Category	Interview							
		AR	FB	AB	FZ	KDS	SS	HH	JJ
Auxiliary Concepts	Domain specificity	<p>The term wasn't understood. Example: AR: <u>Können Sie Domänenspezifizität definieren?</u></p>	<p>Domain Knowledge is collected depending on the status and structure of the project and the accessibility of the state of the art. The methods of gathering domain knowledge vary from observation over interviews to literature review Example: "FB: Also das <u>liegt natürlich daran, was an Informationsquellen zur Verfügung steht, also wenn ich natürlich, wenn ich entsprechend, nehmen wir nochmal das Beispiel Lehrkräfte: Wenn die verfügbar sind, wenn ich die identifizieren kann, die ich befragen möchte, wo ich auch weiß okay die können mir ein bisschen was dazu erzählen. [...] dann gehe ich natürlich auf die zu; wenn ich die aber nicht identifizieren kann [...]oder wenns auch um, oder wenns um, bleiben wir noch im Bereich der Schule, wenns beispielsweise eben auch um Unterrichtsentwicklung oder so geht [...] Dann kann es durchaus auch Sinn machen, dass ich mich einfach mal in den Unterricht reinsetze und zu gucke"</u></p>	<p>Domain Knowledge is collected depending on the status and structure of the project and the accessibility of the state of the art. Also a kind of step-by step process is observable. The gathering is domain specific since a possible method is the observation of a topic using mailing lists. Example: "<u>[...] I e-mailed her and we chatted for a bit and she just gave me this list of mailing list, website, other resources that she recommended and some online groups--"</u></p>	<p>Domain knowledge is naturally given based on his studies. Example: "<u>FZ: [...] Und deswegen brauchen wir auf jeden Fall so einen Informatikhintergrund, weil wir ansonsten halt höchstens deren Sprechakte verstehen könnten, aber nicht was da inhaltlich passiert.[...] FZ: So in diesem Sinne ist also das Domänenwissen, was wir als Informatiker haben, unersetzlich für die Analyse, weil wir sonst gar nicht wüssten, was wir da vor uns haben an bloßen Phänomenen.[...] Na ja, das ist im Prinzip bei jeder Äußerung, die der Entwickler macht. Um die zu verstehen, was das eigentlich ist, ist das jetzt eher ein, ist es vielleicht sogar ein Witz, den er da macht, oder ist es ein konstruktiver Design-Vorschlag, den er da jetzt gerade unterbreitet; diese Entscheidung zu treffen, das geht im Prinzip nur mit so einem Informatikerhintergrund."</u></p>	<p>A literature pool is collected and tagged "[...] <u>Natürlich sind das die Erfahrungen und Erinnerungen die man persönlich mitbringt und zum anderen stehen uns auch die ganzen Projekte noch zur Verfügung in Form von Dateien und auch die ganze Literatur. Die ganze Literaturverwaltung wurde über Citavi erfasst und ist nach wie vor abrufbar. Das ist natürlich etwas schwierig, denn die Art wie das in Citavi eingepfleat wurde war nach Autoren und da muss man schon genau drauf achten und nochmal sortieren und untersortieren, und da bin ich auch zum Teil dabei welcher Autor hat genau mit welchem speziellen Bereich nochmal innerhalb Fachbereichs zu tun."</u></p>	<p>His developed coding schema is only applicable for his specific domain.</p>	<p>Domain Knowledge is collected depending on the structure of the project. The methods of gathering domain knowledge depend also on the project and the accessibility of data. Example: "<u>Das ergibt sich aus der Struktur der Daten die ich haben möchte und aus der Struktur des Feldes mit dem ich konfrontiert bin. Das sind für mich die 2 großen Anker"</u></p>	

Coding paradigm	Category	Interview							
		AR	FB	AB	FZ	KDS	SS	HH	JJ
Strategy	Domain specific Code schema/Domain specific meta model		-	-	<p>Within the coding schema a domain specific meta model is included and implicitly applied as well as a code schema which is explicitly applied.</p> <p><u>Example: "[...]und die vierte Praktik, die ist ziemlich technisch, nämlich sich ein Metamodell vorher zu überlegen, welche Klassen, nicht welche Klassen[...]wie im Prinzip die Ergebnisse strukturiert sind.[...] Also es gibt zum Beispiel so etwas wie Konzepte. Konzepte haben Beziehungen zueinander, Konzepte können Eigenschaften haben..."</u></p>	<p>In some extent: A code consists of a maximum of two words</p> <p><u>"DS: Jetzt nicht den Code an sich. Also der Code ist ein maximal zwei Worte."</u></p>	<p>He defines a domain specific code schema and developed a meta model for his research and also for further researchers, so that a explicit reuse can happen.</p> <p><u>Example: "I: Okay. Jetzt hat Herr XY letzte Woche angesprochen, dass Sie daraus auch ein Metamodell entwickelt haben.</u></p> <p><u>SS: Ach ja, ja habe ich."</u></p>	<p>A domain specific code schema was explicitly declined since everything depends on the underlying object of investigation.</p>	-

Coding paradigm	Category	Interview							
		AR	FB	AB	FZ	KDS	SS	HH	JJ
Consequence	Domain specific reuse of a codebook			<p>A codebook was created and two students reused it. <u>Example: "So, that was for a literature review of the episodic volunteering literature. [...] AB: And there are 2 students that are working on this. [...] AB: And the idea was that we have a kickoff meeting where we would develop the code book either way through doing the coding of one document and then syncing with one another to be sure that we had a shared understanding--"</u></p>	<p>He reuses codes setup by another team member of his project domain. The codes are collected and the usage of the codes are described very detailed in a codebook (Here named Basiskonzepte) <u>Example: "FZ: Ja genau. Also ich setze das konkret ein um besser zu verstehen, was eigentlich in so einer Sitzung passiert.[...]"</u></p>		<p>A very detailed codebook was created for reuse. <u>Example: " I: Ok. Das ist im Prinzip... oder kann man sich das so vorstellen, oder ich habe das so kennengelernt, muss man noch dazu sagen natürlich, ich habe den Namen des Codes, dann habe ich eine Beschreibung des Codes also Definition wie ich ihn mir vorstelle, und dann habe ich Beispiele dazu, die den Code beschreiben und ich habe vielleicht noch zusätzlich eine Beschreibung wann der Code verwendet wird und wann er eben nicht verwendet wird. SS: Ja aber so genau können Sie sich das vorstellen. ja."</u></p>		

Appendix E Intercoder Reliability

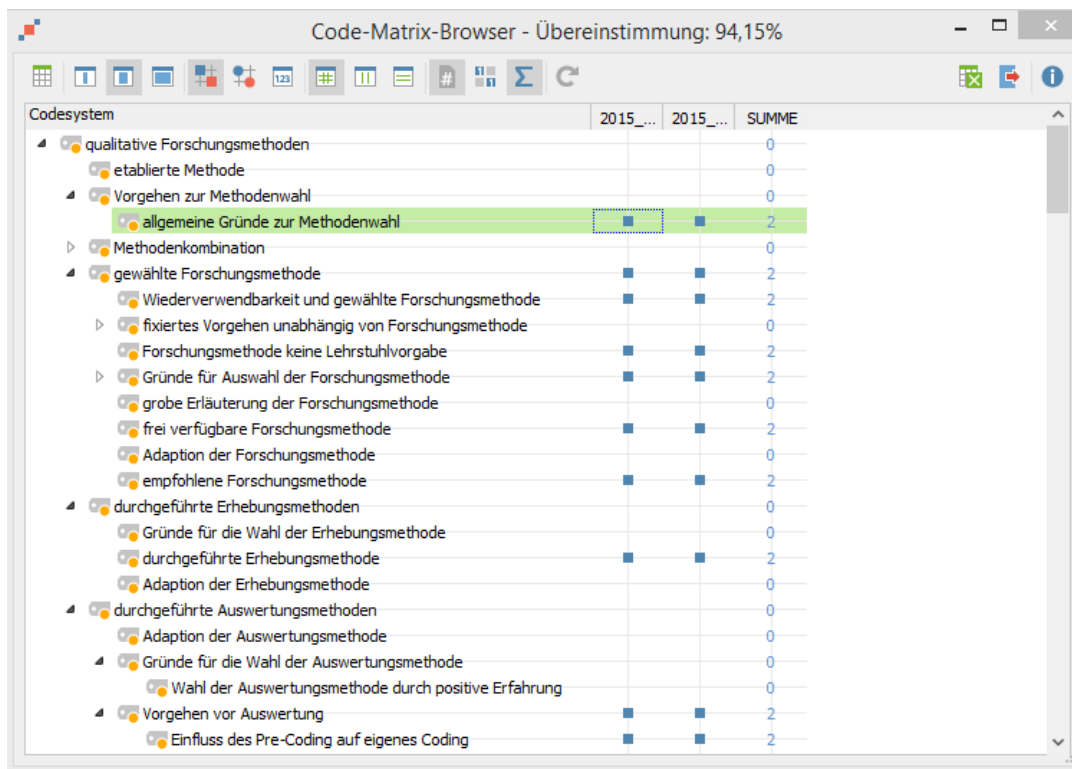


Figure 10: Intercoder Agreement of the Interview with FZ. The appearance of codes was compared.

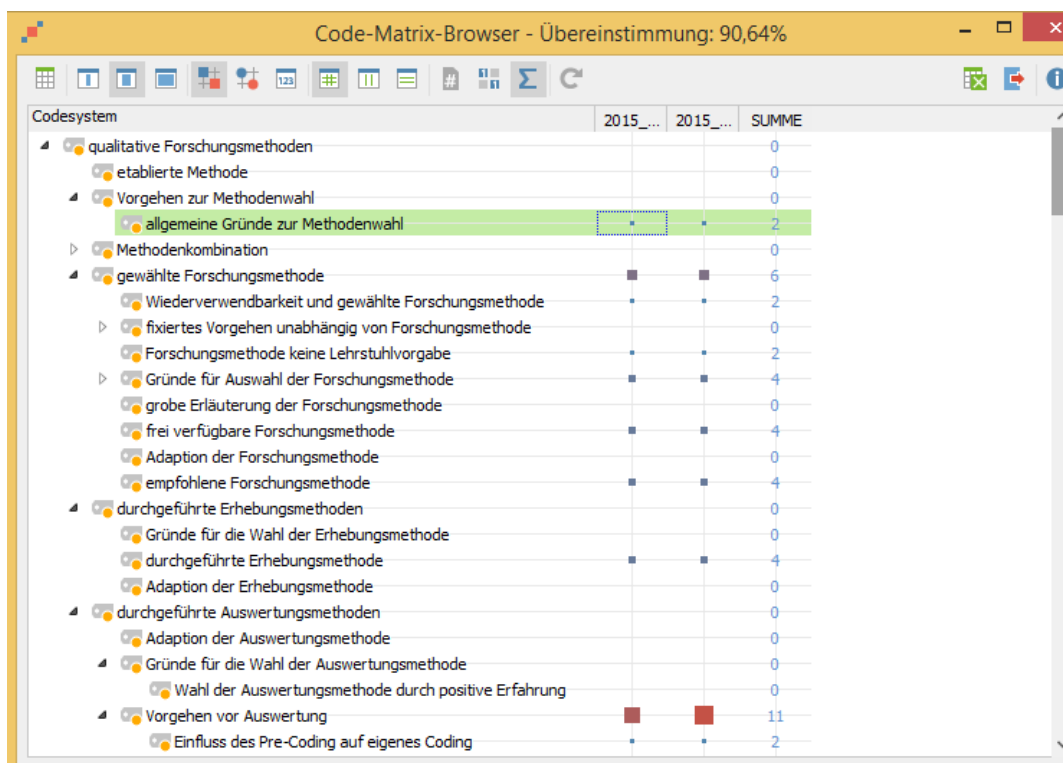


Figure 11: Intercoder Agreement of the Interview with FZ. The frequency of occurrence was compared.

References

- AB (2015). conducted by the author via Skype on 2015-11-04.
- Alexander, P. A. (1992). Domain knowledge: Evolving themes and emerging concerns. *Educational psychologist*, 27(1):33–51.
- Alexander, P. A. and Judy, J. E. (1988). The interaction of domain-specific and strategic knowledge in academic performance. *Review of Educational research*, 58(4):375–404.
- AR (2015). conducted by the author via Skype on 2015-10-27.
- Böhm, A. (2004). Theoretical coding: Text analysis in grounded theory. *A companion to qualitative research*, pages 270–276.
- Börner, R. (2011). Towards construction of situational methods for service identification. In *Engineering Methods in the Service-Oriented Context*, pages 204–218. Springer.
- Bortz, J. and Döring, N. (2006). *Forschungsmethoden und Evaluation für Human-und Sozialwissenschaftler. 4. überarbeitete Auflage*. Springer Medizin Verlag, Heidelberg.
- Bryant, A. and Charmaz, K. (2010). Grounded theory in historical perspective: An epistemological account. *Handbook of grounded theory*, pages 31–57.
- Buehl, M. M., Alexander, P. A., and Murphy, P. K. (2002). Beliefs about schooled knowledge: Domain specific or domain general? *Contemporary educational psychology*, 27(3):415–449.
- Cohn, M. (2010). User stories: für die agile software-entwicklung mit scrum. *XP ua mitp*.
- Corbin, J. and Strauss, A. (2014). *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory*. SAGE Publications.
- Denzin, N. (1994). Introduction: Entering the field of qualitative research. *Handbook of qualitative research*, pages 1–17.
- Denzin, N. K. (1978). The research act: A theoretical orientation to sociological methods.
- Duncan, R. G. (2007). The role of domain-specific knowledge in generative reasoning about complicated multileveled phenomena. *Cognition and Instruction*, 25(4):271–336.
- Fereday, J. and Muir-Cochrane, E. (2008). Demonstrating rigor using thematic analysis: A hybrid approach of inductive and deductive coding and theme development. *International journal of qualitative methods*, 5(1):80–92.
- Flick, U. (2000). Triangulation in der qualitativen forschung. s. 309-318. *Ders.; Ernst von Kardorff; Ines Steinke (Hg.): Qualitative Forschung. Ein Handbuch. Reinbek. Rowohlt*.
- Flick, U. (2009). *An introduction to qualitative research*. Sage.
- Flick, U. (2013). *Qualitative Forschung*. Rowohlt.
- Francis, J. J., Johnston, M., Robertson, C., Glidewell, L., Entwistle, V., Eccles, M. P., and Grimshaw, J. M. (2010). What is an adequate sample size? operationalising data saturation for theory-based interview studies. *Psychology and Health*, 25(10):1229–1245.
- FZ (2015). conducted by the author via Skype on 2015-11-20.
- Glaser, B. G. (1967). The discovery of grounded theory. strategies for qualitative research.

- Glaser, B. G. (1978). *Theoretical sensitivity: Advances in the methodology of grounded theory*. Sociology Pr.
- Glaser, B. G., Strauss, A. L., and Paul, A. T. (2010). *Grounded theory: strategien qualitativer forschung*. Huber.
- Golafshani, N. (2003). Understanding reliability and validity in qualitative research. *The qualitative report*, 8(4):597–606.
- Green, J. and Thorogood, N. (2013). *Qualitative methods for health research*. Sage.
- Guba, E. G. (1981). Criteria for assessing the trustworthiness of naturalistic inquiries. *ECTJ*, 29(2):75–91.
- Guest, G., Bunce, A., and Johnson, L. (2006). How many interviews are enough? an experiment with data saturation and variability. *Field methods*, 18(1):59–82.
- Hartmann, J., Fleischmann, A., Pfaller, C., Rappl, M., Rittmann, S., and Wild, D. (2006). Feature net-ein ansatz zur modellierung von automobilspezifischem domänenwissen und anforderungen. In *GI Jahrestagung (1)*, pages 761–765.
- Heiser, G. (2014). Tips and guidance for students writing papers and reports.
- HH (2015). conducted by the author via Skype on 2015-12-07.
- Hoda, R., Noble, J., and Marshall, S. (2012). Developing a grounded theory to explain the practices of self-organizing agile teams. *Empirical Software Engineering*, 17(6):609–639.
- Huberman, A. M. and Miles, M. B. (1994). Data management and analysis methods.
- Jenner, B., Flick, U., von Kardoff, E., and Steinke, I. (2004). *A companion to qualitative research*. Sage.
- JJ (2015). conducted by the author via Skype on 2015-12-08.
- Kaiya, H. and Saeki, M. (2006). Using domain ontology as domain knowledge for requirements elicitation. In *Requirements Engineering, 14th IEEE International Conference*, pages 189–198. IEEE.
- KDS (2015). conducted by the author personally on 2015-11-24.
- Kelle, U. (2007). Theoretisches vorwissen und kategorienbildung in der „grounded theory“. In *Qualitative Datenanalyse: computergestützt*, pages 32–49. Springer.
- Khatri, V., Vessey, I., Ramesh, V., Clay, P., and Park, S.-J. (2006). Understanding conceptual schemas: Exploring the role of application and is domain knowledge. *Information Systems Research*, 17(1):81–99.
- Kuckartz, U. (2007). Einführung in die computergestützte analyse qualitativer daten.
- Kunz, K. (2015). Developing a domain analysis procedure based on grounded theory method. Master’s thesis, FAU Erlangen-Nuernberg.
- Lee, Y. and Zhao, W. (2006). Domain requirements elicitation and analysis-an ontology-based approach. In *Computational Science–ICCS 2006*, pages 805–813. Springer.

- Leech, N. L. and Onwuegbuzie, A. J. (2007). An array of qualitative data analysis tools: A call for data analysis triangulation. *School psychology quarterly*, 22(4):557.
- Leininger, M. (1994). Evaluation criteria and critique of qualitative research studies. *Critical issues in qualitative research methods*, pages 95–115.
- Lincoln, Y. S. and Guba, E. G. (1985). *Naturalistic inquiry*, volume 75. Sage.
- Looso, S., Borner, R., and Goeken, M. (2011). Using grounded theory for method engineering. In *Research Challenges in Information Science (RCIS), 2011 Fifth International Conference on*, pages 1–9. IEEE.
- Mason, M. (2010). Sample size and saturation in phd studies using qualitative interviews. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, 11(3).
- Mayring, P. (2002). Qualitative sozialforschung. *Eine Anleitung zu qualitativen Denken*, 5.
- Mey, G. and Mruck, K. (2007). Grounded theory methodologie—bemerkungen zu einem prominenten forschungsstil. *Historical Social Research/Historische Sozialforschung. Supplement*, pages 11–39.
- Morse, J. M. (1994). Emerging from the data: The cognitive processes of analysis in qualitative inquiry. *Critical issues in qualitative research methods*, 346:350–351.
- Morse, J. M. (1995). The significance of saturation. *Qualitative health research*, 5(2):147–149.
- Morse, J. M. (2008). Styles of collaboration in qualitative inquiry. *Qualitative Health Research*, 18(1):3–4.
- Morse, J. M., Barrett, M., Mayan, M., Olson, K., and Spiers, J. (2008). Verification strategies for establishing reliability and validity in qualitative research. *International journal of qualitative methods*, 1(2):13–22.
- Muis, K. R. (2004). Personal epistemology and mathematics: A critical review and synthesis of research. *Review of educational research*, 74(3):317–377.
- Muis, K. R., Bendixen, L. D., and Haerle, F. C. (2006). Domain-general and domain-specificity in personal epistemology research: Philosophical and empirical reflections in the development of a theoretical framework. *Educational Psychology Review*, 18(1):3–54.
- O'Reilly, M. and Parker, N. (2012). 'unsatisfactory saturation': a critical exploration of the notion of saturated sample sizes in qualitative research. *Qualitative Research*, page 1468794112446106.
- Rubin, H. J. and Rubin, I. S. (1995). *Qualitative interviewing: The art of hearing data*. Sage.
- Rumbaugh, J., Jacobson, I., and Booch, G. (2004). *Unified Modeling Language Reference Manual*, The. Pearson Higher Education.
- Saldaña, J. (2012). *The coding manual for qualitative researchers*. Number 14. Sage.
- Salinger, S., Plonka, L., and Prechelt, L. (2008). A coding scheme development methodology using grounded theory for qualitative analysis of pair programming.
- Schwandt, T. A. (2001). *The Sage dictionary of qualitative inquiry*. Thousand Oaks, CA: Sage.
- SS (2015). conducted by the author via Skype on 2015-11-25.

- Starks, H. and Trinidad, S. B. (2007). Choose your method: A comparison of phenomenology, discourse analysis, and grounded theory. *Qualitative health research*, 17(10):1372–1380.
- Steinke, I. (1999). Kriterien qualitativer forschung. *Ansätze zur Bewertung qualitativ-empirischer Sozialforschung, München*.
- Steinke, I. (2004). Quality criteria in qualitative research. *A companion to qualitative research*, pages 184–190.
- Strauss, A. L. (1987). *Qualitative analysis for social scientists*. Cambridge University Press.
- Strauss, A. L., Corbin, J. M., et al. (1990). *Basics of qualitative research*, volume 15. Sage Newbury Park, CA.
- Strauss, A. L., Corbin, J. M., Niewiarra, S., et al. (1996). *Grounded theory: Grundlagen qualitativer sozialforschung*. Beltz, Psychologie-Verlag-Union.
- Strübing, J. (2007). Glaser vs. strauss? zur methodologischen und methodischen substanz einer unterscheidung zweier varianten von grounded theory. *Historical Social Research/Historische Sozialforschung. Supplement*, pages 157–173.
- Strübing, J. (2011). Zwei varianten von grounded theory? zu den methodologischen und methodischen differenzen zwischen barney glaser und anselm strauss. In *Grounded Theory Reader*, pages 261–277. Springer.
- Sure, Y., Erdmann, M., Angele, J., Staab, S., Studer, R., and Wenke, D. (2002). Ontoedit: Collaborative ontology development for the semantic web.
- Vibert, N., Ros, C., Bigot, L. L., Ramond, M., Gatefin, J., and Rouet, J.-F. (2009). Effects of domain knowledge on reference search with the pubmed database: An experimental study. *Journal of the American Society for Information Science and Technology*, 60(7):1423–1447.
- Von Mayrhauser, A. and Lang, S. (1999). A coding scheme to support systematic analysis of software comprehension. *Software Engineering, IEEE Transactions on*, 25(4):526–540.
- Wildemuth, B. M. (2004). The effects of domain knowledge on search tactic formulation. *Journal of the american society for information science and technology*, 55(3):246–258.