Hocker, Julian; Schindler, Christoph; Rittberger, Marc

Participatory design for ontologies. A case study of an open science ontology for qualitative coding schemas

Aslib journal of information management 72 (2020) 4, S. 671-685, 10.1108/AJIM-11-2019-0320

Participatory design for ontologies: a case study of an open science ontology for qualitative coding schemas

Received 12 November 2019
Revised 28 March 2020
14 July 2020
Accepted 14 July 2020

Julian Hocker
DIPF | Leibniz Institute for Research and Information in Education, Frankfurt am Main, Germany and Department for Information Science and Natural Language Processing, University of Hildesheim, Hildesheim, Germany
Christoph Schindler
DIPF | Leibniz Institute for Research and Information in Education, Frankfurt am Main, Germany, and
Marc Rittberger
DIPF | Leibniz Institute for Research and Information in Education, Frankfurt am Main, Germany and Darmstadt University of Applied Sciences, Darmstadt, Germany

Abstract

Purpose – The open science movement calls for transparent and retraceable research processes. While infrastructures to support these practices in qualitative research are lacking, the design needs to consider different approaches and workflows. The paper bases on the definition of ontologies as shared conceptualizations of knowledge (Borst, 1999). The authors argue that participatory design is a good way to create these shared conceptualizations by giving domain experts and future users a voice in the design process via interviews, workshops and observations.

Design/methodology/approach – This paper presents a novel approach for creating ontologies in the field of open science using participatory design. As a case study the creation of an ontology for qualitative coding schemas is presented. Coding schemas are an important result of qualitative research, and reuse can yield great potential for open science making qualitative research more transparent, enhance sharing of coding schemas and teaching of qualitative methods. The participatory design process consisted of three parts: a requirement analysis using interviews and an observation, a design phase accompanied by interviews and an evaluation phase based on user tests as well as interviews.

Findings – The research showed several positive outcomes due to participatory design: higher commitment of users, mutual learning, high quality feedback and better quality of the ontology. However, there are two obstacles in this approach: First, contradictive answers by the interviewees, which needs to be balanced; second, this approach takes more time due to interview planning and analysis.

Practical implications – The implication of the paper is in the long run to decentralize the design of open science infrastructures and to involve parties affected on several levels.

The authors would like to thank Christa Womser-Hacker and Thomas Mandl for their input as well as the reviewers. We would also like to thank Gwen Schulte for comments on the text.

Funding: This research was supported by the Leibniz Competition (SAW-2016-DIPF-1) in the project "Abitur examination practices and respective essays from 1882 to 1972".

Aslib Journal of Information Management
Vol. 72 No. 4, 2020
pp. 671-685
Emerald Publishing Limited
2050-3806
DOI 10.1108/AJIM-11-2019-0320

Originality/value – In ontology design, several methods exist by using user-centered design or participatory design doing workshops. In this paper, the authors outline the potentials for participatory design using mainly interviews in creating an ontology for open science. The authors focus on close contact to researchers in order to build the ontology upon the expert's knowledge.

Keywords Ontology engineering, Participatory design, Digital humanities, Semantic web, Open science, Qualitative research, Coding schemas

Paper type Research paper

1. Introduction

While Gruber (1995) defines ontologies as an "*explicit specification of a conceptualization*" and Borst (1999) as a shared conceptualization, Studer *et al.* (1998) combine these definitions: "An ontology is a formal, explicit specification of a shared conceptualization." For the design of an ontology, this definition brings articulation, sharing and agreement of processes to the fore rather than only technological aspects.

This paper presents the application of a participatory design approach to ontology design and argues for a high potential in the creation of ontologies, where a shared conceptualization does not exist yet and has to be built from scratch. Other potentials are mutual learning, increased acceptance and a higher quality input when involving users actively in the process via the use of prototypes at different stages of the development, as well as a user-centered evaluation.

Because of this potential, participatory ontology design can suitably be applied in the field of open science, where transparency and traceability is called for in terms of implicit research practices (see Fecher and Fisieke, 2014). Concretely, this concerns the creation of an ontology for qualitative coding schemas. Qualitative coding schemas are a central output of qualitative research, but in many cases they remain noncomprehensively documented, yet could be reused in the research. Still, there exists no ontology for describing these coding schemas. Earlier research (Faniel *et al.*, 2016) showed that the satisfaction of researchers increases when research data is well-documented. An ontology is therefore the first step to enhance a platform for the sharing of qualitative coding schemas.

The paper is structured as follows: Section two describes the basis of ontology design and evaluation; section three describes a case study for the application of this methodology for creating an ontology, to describe and exchange qualitative coding schemas. Section four describes the outcomes of the case study, whereas section five concludes the paper.

2. Related works for a participatory design of ontologies

The chapter shows the principles of both ontology design as well as participatory design. Both research fields form the basis of the presented method. A special emphasis is put on the evaluation of ontologies in chapter 2.3.

2.1 Designing ontologies

In accordance to Gruber (1995), Borst (1999) and Studer (1998) Guarino *et al.* (2009), there are important factors in ontology: The conceptualization, the formal, explicit specification and the "shared" part of it:

(1) A "conceptualization" can be seen as an abstract, simplified view of the world that we wish to represent for some purpose (Genesereth and Nilsson, 1987).

(2) The "explicit specification" is about choosing the right language and vocabulary for representing the ontology.

(3) "Shared" means that this knowledge is agreed on within a domain.

The design of the ontology can be seen as trying to make these concepts explicit in order to understand the concepts people have in mind about a domain of knowledge. Still, these concepts will not all be the same, but it might be possible to find some common ground everyone can agree on. Since these conceptualizations already exist in people's minds, one could argue that the main purpose of creating ontologies might be the moderation of the process, where users explicate their concepts and creating a space for sharing their conceptualizations.

Several approaches exist for designing ontologies (e.g. the frameworks *Enterprise* (Uschold and King, 1998) or *Methonotology* (Fernández *et al.*, ,1997)). Still, none of these methods involve users in the whole process of creation; rather they bring users in at one or more stages of the development process. Other approaches have introduced a user-centered evaluation (e.g. Zhang and Li, 2008 or Hu *et al.*, 2018) or interviews with users at the beginning of the process in order to grasp the requirements (e.g. Reinhold, 2015; Lam *et al.*, 2015). The goal of this paper is to go beyond these practices and involve researchers earlier in the process of ontology design using participatory design or codesign methods.

2.2 Participatory design for ontologies

The idea of participatory design or codesign is to involve people affected by change in the design process to secure a higher product quality, but also to foster the acceptance of a new technology. In the digital humanities, several authors have advocated for involving actual users more in research (see Warwick, 2012; Kemman and Kleppe, 2015). It has also been noted that users need to actively articulate their needs for new software, especially in the digital humanities (Borgman, 2009).

In digital humanities and open science projects in general, the future users of the software or system generally have a high level of expertise in their research field and a good understanding of what they need for their research. Therefore, it makes sense to involve them at an early state of research to use their knowledge for improving the product, i.e. software or system (see Warwick, 2012; Kemman and Kleppe, 2015; Womser-Hacker and Heuwing, 2015).

Muller (2007) defines participatory design as the "third space in HCI" and argues for using it to create "new settings and experiences that can assist computer professionals to work in partnership with diverse users in improving both computer technology and the understandings that make computer technologies successful in real use." (Muller, 2007, p. 3). This can also be combined with the call for creating software so users also improve their handling of the technologies and also lower the barrier to use these new technologies (Borgman, 2009).

Gregory (2003) describes the following goals to be accomplished with participatory design heuristics:

(1) Improving the knowledge upon which systems are built;

(2) Enabling people to develop realistic expectations and reducing resistance to change;

(3) Increasing workplace democracy by giving the members of an organization the right to participate in decisions that are likely to affect their work (Gregory, 2003, p. 2).

It becomes clear that these goals match the ideas proposed by authors in the digital humanities. For the development of an ontology, it can be added that it is not only about improving the knowledge of the system that is built but also it is a way to find out from the researchers about their research in order to create this shared conceptualization of a domain that is needed to describe and formalize it into an ontology.

Several authors have used participatory design practices for creating ontologies:

Palavatsinidis (2014) focuses on feedback rounds with domain experts. These feedback rounds take place at early stages in the design process and can roughly be described as participatory design. However, the approach only involves domain experts at an early stage of the process, whereas actual users are only considered at the end of the ontology design process.

For designing health services, Ongenae *et al.* (2011) propose a participatory design approach for an ontology based on workshops and observation. Kingsun *et al.* (2018) also suggest using workshops and observation for creating an ontology in the medical domain.

2.3 Ontology evaluation

The last phase of a conventional design process is the evaluation of the developed ontology. Sabou and Fernandez (2012) name several ways to evaluate ontologies. Many of these focus on technical metrics, but they also mention user-centered evaluation as an important approach to evaluate an ontology. Similarly to the user-centered way of designing the ontology, a user-centered way of evaluation is also fruitful.

The user-centered evaluation approaches use qualitative or quantitative methods to assess ontologies: Expert interviews with feedback options are an example of a qualitative approach (e.g. Zhang and Li, 2008 or Hu *et al.*, 2018). Questionnaires are often used to let users provide quantitative ratings of ontologies. Reinhold (2015) used A/B-testing of a new version of an ontology comparing it to a baseline version. Tan *et al.* (2017) used usability/usefulness-tests.

A crucial aspect of user-centered evaluation concerns the criteria that are being evaluated. Palavitsinis (2014, p. 65) uses criteria for information quality developed by Lee *et al.* (2002), based on the information quality criteria by Wang and Strong (1996) to evaluate the developed metadata [1].

However, most of these criteria can only be measured within a system that is already live and contains some real data. This interferes with the approach of testing/evaluation early in order to find problems early thus being able to fix them when this is cheap. Hu *et al.* (2018) describe a different approach, doing feedback interviews as evaluation of a prototype and deriving criteria based on these interviews [2].

2.4 Summary and outcomes of related works

We can see that the formulation of an ontology and the call from the digital humanities to involve users early at the design process stage of an ontology can be met using methods of participatory design. Reinhold (2015) showed that it makes sense to use expert interviews for requirement analysis. We think this might also make sense for the participatory design since expert interviews allow bringing in a broader range of people than workshops and within the domain of research data it might not be easy for people to convene in one place at the same time for a workshop.

Regarding the evaluation, a qualitative evaluation makes sense, especially when evaluating a prototype that does not have any data in the system yet. The creation of two ontologies for A/B testing is also of limited use if the process before is done via participatory design since this process does not involve creating several versions of the ontology.

3. Participatory ontology design – a case study for qualitative coding schemas

This chapter describes the application of the ontology design in relation to a case study. Section 3.1 gives an overview of the background of the case study; section 3.2 describes the application of the method within this case study.

3.1 Theoretical background of case study

The case study is situated in the field of open science since it deals with the transparency of scientific results. Furthermore, the work is based on discussions within the community of qualitative social scientists.

Coding schemas are an important part of qualitative research in methods like grounded theory (e.g. Glaser and Strauss, 1967) and qualitative content analysis (Mayring, 2015; Saldaña, 2015; Schreier, 2012). Coding schemas refer to the collection of all codes in a qualitative research project. This includes the codes as well as the relationships between these codes and description of what these codes mean and how they are used. Codes can be based on data, theoretical models or reused from other research. Coding schemas can also be seen as a result of the research, especially in grounded theory or as a tool for data analysis (in qualitative content analysis).

Since open science, FAIR data principles (Mons *et al.*, 2017) and the concept of research graphs (Auer, 2018) have become more important topics also in qualitative research; more effort has been invested in opening up qualitative research. FAIR data describe standards for research data. Accordingly, research data should be findable, accessible, interoperable and reuseable. Research graphs describe the connections of research papers, software and data, whereas the goal is to make more transparent how research data and research software was produced and reused.

There exist several research data centers for qualitative data (e.g. UK data service [3], Forschungsdatenzentum Bildung (Germany) [4] or DANS (Netherlands) [5]), as well as exchange formats to exchange coding schemas [6] between commercial software like NVivo [7] or Atlas.TI [8] (Evers, 2018). Furthermore, there is an ongoing discussion about the quality of qualitative research, which also addresses the traceability of results (Strübing *et al.*, 2018), a topic which can also be supported by the structured description and publication of coding schemas.

3.2 Participatory ontology design for case study

Qualitative research is a field where some foundations have been laid, but as of now there is no shared understanding of the description of qualitative coding schemas. Therefore, a participatory approach is useful. Since this shift to open science is also a transformation process, a participatory design first democratizes the process and gives the members of the community the skills and possibilities to add their feedback (Gregory, 2003). The other goal was to create an ontology that works "in action", in the practical research process.

As suggested in chapter 2, a participatory ontology design method was used in this case study. Given the current lack of standards in qualitative coding schemas this also makes sense. So far, approaches have been limited to textbooks or the publication of coding schemas by some researchers in the appendix of a thesis or a paper and there is no standardization yet.

The research from this point focused on two aspects: First, creating a description of coding schemas or qualitative research in general. There was no common concept about what can describe a qualitative project and makes it distinguishable from another qualitative research project.

The second aspect concerns formalizing this concept in order to create an ontology based on this knowledge. For this purpose, semantic web standards were taken on as well as principles for data sharing, mainly the FAIR data principles to use these as tools to have an explicit description (Gruber, 1995) of coding schemas in qualitative research.

The method consists of three parts:

(1) Requirement analysis to grasp the processes in research and to understand the way research works in qualitative methods.

(2) Design phase which allows users to become involved in the process and give feedback to prototypes.

(3) Evaluation phase where the developed ontology is evaluated by users.

The following Figure 1 shows an overview of the process and the methods used in each step:

3.2.1 Requirement analysis. In the requirement analysis, the goal was to analyze what was needed as well as to develop a common understanding from the perspective of the ontology developers for the topic of qualitative coding schemas, as well as from the users' perspective.

This part focused on how qualitative research is done. Due to the affiliation of the researchers, the focus was on educational research in German-speaking countries. The requirement analysis consisted of five parts:

(1) Stakeholder analysis to identify relevant players in the field,

(2) Observation of researchers,

(3) Expert interviews,

(4) Analysis of several published coding schemas and

(5) Analysis of textbooks mentioned in the interviews.

To get an overview of the groups of potential users or of those potentially affected by the ontology in some ways, a stakeholder analysis was conducted. Table 1 shows the identified stakeholders.

The observation was conducted at the beginning of the research. For recruitment purposes, a contact to a research group was used. It was possible to be part of the research process and to see how the researchers were analyzing data. The observation was done in order to delve into the research and into the ideas researchers have when they are working with the material and analyze this material. The results were fed into the interview guidelines.

The interviews were conducted as semi-structured expert interviews based on Gläser and Laudel (2010). The interviewees were selected using theoretical sampling. The criterion was that people needed experience in either qualitative content analysis or grounded theory in a larger project, e.g. a doctoral thesis, and should work in the field of educational research in a broad sense. It turned out that interviews with researchers at the end of their PhD or at post-doc level were more fruitful, so many interviews with people from this scientific level were conducted. Recruitment was done via a snowball system, using personal contacts, contacts from our affiliation as well as instruction workshops for qualitative research methods. The number of participants at this stage was $n = 10$. The interviews were transcribed and analyzed using qualitative content analysis (inhaltlich-strukturierende Codierung) based on Kuckartz (2018, p. 77).

Based on the interviews, recommended textbooks from workshops as well as published coding schemas were analyzed. Different textbooks from grounded theory as well as qualitative content analysis were analyzed in order to grasp the actual practice of publishing and documenting qualitative coding schemas. Based on these findings, a first prototype was developed.

3.2.2 Design phase. The design phase was meant to get feedback for the developments. Interviews bear the advantage of receiving a broader range of feedback from different people as opposed to workshops. From the requirement analysis, it became clear that qualitative methods are used differently, depending on the method researchers are using. The goal was therefore to include researchers using grounded theory as well as qualitative content analysis into the research. The goal was also to include other groups identified in the stakeholder analysis in this phase. Thus, an approach with three phases was used:

(1) Feedback interviews based on paper prototypes;

(2) Presentations and feedback at workshops run by research data centers as well as developers for QDA software;

(3) Feedback interviews based on an implementation of the ontology in Semantic MediaWiki [9].

For the feedback interviews, the same researchers were interviewed in both rounds. These researchers had been recruited from the previous interviews so they already knew the project and their feedback was assumed to be of higher quality than with new people, and it was clear that these interviewees already endorsed the project. To add to the range, another interviewee from a research data center was introduced. The ultimate number of participants was in the interviews was $n = 5$.

In the first phase, a paper prototype was presented to the interviewees. During the interview, the experts received an overview of the complete ontology and were encouraged to give feedback to every metadata item. The questions focused on the following quality criteria: usefulness, uniqueness, completeness and naming. Furthermore, open-ended questions were used to give the interviewees the chance to also give feedback on other points. The paper prototype had the advantage of showing a first version of the ontology with descriptions of each item to the researchers as well as showing the structure of the metadata.

To get feedback from a wider range of stakeholders, the prototype was introduced at two workshops with a research data center as well as REFI (Rotterdam Exchange Format Initiative) [10], a group involved in standardization and exchange of QDA projects between several proprietary software tools. Owing to these settings, it was possible to talk to a wide range of experts from the field of research data management as well as the development of qualitative research software. The workshops allowed for presentation of the ideas and feedback from these experts, which had not been included in the research before. This was also done to see if these people were interested in implementing the ideas of the ontology in their products. It was also possible to gain hints on how to develop the ontology in a way consistent with the data structures used in research data centers.

In the second design phase, the ontology was implemented in an environment based on Semantic MediaWiki, and example coding schemas were implemented allowing browsing of the information. The participants were asked to complete a task that involved searching for a coding schema for a given future research project. The goal here was to get feedback on whether the items in the ontology were relevant for their research as well as if there was information missing and the naming appropriate. At the end, participants got a questionnaire in order to rate the relevance of the items.

3.2.3 Evaluation phase. In the evaluation phase, the ontology was evaluated by possible future users. A prototype of the ontology was evaluated since this still gives the opportunity to get feedback from a broader audience than during the design phase and to do changes before actually going live with the system.

Based on the literature, the following criteria were selected: naming, relevance and completeness. This research thus follows a pragmatic approach, where the goal is not to actually test the abstract ontology "by itself" but rather "in action". Based on Kuhlen's definition of information as knowledge in action (Kuhlen, 1995, p. 34), it is tested whether the ontology is relevant for the user's needs and whether it supports the research process. Therefore, relevance for the user's research is one quality criterion. Two other quality criteria for the ontology are a) the completeness of the ontology, so whether information is missing and b) the naming of the metadata items. In the case study, completeness in this sense means that all relevant information is given in the metadata schema so users find all the information they are searching for [11].

The completeness of the metadata schema is important since the goal is to provide a system for two scientific methods and all information should be included that is important for both methods. The naming of metadata is in this case especially important since the goal is to bring together scientists who use two different research methods and therefore also are likely to name things differently.

The evaluation consists of two parts, and in each case the goal was to check if the ontology fits the researchers' needs in one step in their research cycle as well as the metrics described above. In each evaluation phase, the quality criteria relevance for research completeness as well as naming is investigated:

(1) First, a user-centered test is conducted in order to analyze how the ontology supports the reusage of coding schemas.

(2) In the second phase, a qualitative evaluation is performed in order to test whether the use cases of searching for a coding schema are met. This was also done to select the most relevant metadata items, which will then become mandatory.

3.2.3.1 User-centered tests. In the first test phase, the goal was to see whether the further descriptions of the ontology help to understand how coding schemas are applied to new data. To test this, 20 usability tests with qualitative interviews were conducted. The tests were conducted with students of information science or user-centered design and engineering at one university in the US. The participants were at Master's or PhD level, and they had used qualitative coding before. At the beginning of the tests, participants were introduced to the ontology implemented in a system including example coding schemas. Then, they were given a practical coding task and a time limit to solve it. The coding tasks were derived from other projects and the coding schemas were represented in the system with all the information that was given in coding guidelines or papers.

Afterward, semi-structured interviews were conducted with the participants to find out about their impression of the system, how they felt the information was presented and whether some information was missing. This was done to find out if the ontology met the researchers' need to get all necessary information to apply coding schemas. We also asked questions about the acceptance of the system, e.g. if participants would upload their coding schemas and would use the platform as a source to find coding schemas for their research.

3.2.3.2 Qualitative evaluation. The second part of the evaluation focuses on qualitative evaluations of the complete ontology with an emphasis on the search and selection of coding schemas. This evaluation is projected for researchers from educational research in German-speaking countries. The participants will be at PhD or post-doc level and have experience in at least one qualitative research project using grounded theory or qualitative content analysis. The goal is to get a saturation of the answers users of qualitative content analysis as well as grounded theory. Therefore an equal number of users for each method were targeted. The overall number of interviewees was $n = 10$. The search and selection of coding schemas is an important part of the research cycle if coding schemas are meant to be reused. Furthermore, the ontology can help to standardize the description of qualitative research in order to provide guidelines for the quality of this research. This approach is similar to Hu *et al.* (2018), the main difference is that the quality criteria have been defined before rather than being derived from the interviews.

In this evaluation, the ontology was implemented in a system, making it fully clickable and navigable. The researchers were given the scenario of having to select a coding schema for a new research project. They navigated through the systems, and they were asked to give their opinion on all metadata in the ontology. The focus was to find out if the metrics relevance for the research, naming and completeness are met. At the end, questions were asked concerning the acceptance of the platform, if people would share their coding schemas and if they thought

the ontology is useful for selecting coding schemas, which can be reused for their own research.

4. Results

This chapter describes the results of the research. Since it is still work in progress, there are preliminary results for the project. The chapter is split into two parts, i.e. the experiences when using the participatory design (5.1) and the results for the actual research (5.2).

4.1 Preliminary results of applying a participatory design process

This chapter discusses the applicability of the participatory ontology design and the experiences made during the research process. We can see four main results in the participatory design process, i.e. commitment of users, mutual learning, quality of feedback during the process and quality of overall ontology.

During the design phase, it was clear that the participants felt a strong commitment to the project and were happy to be involved in creating standards for "their" research. Participatory ontology design allowed for the users' strong commitment during the interviews, observation and design phases. The further development of the ontology benefitted from the many perspectives and opinions. For instance, users mentioned that better descriptions of coding schemas could also serve teaching of qualitative methods. Given the participants' teaching experience, they can also function as multipliers of the ontology. This can be also seen when the interviewees referred to prior interviews saying things like "we already talked about this before". During the research project, a shared frame of reference was thus created between the users and the developers of the ontology.

Throughout the whole design process, mutual learning could be seen. For the project manager, the learning of the methods of qualitative research was evident during the whole research process. This is also evident for the interviewees. At the beginning, they had little experience with standardization and metadata. During the design phase, they learned the perspective of the ontology engineer and were also able to add their thoughts to the proposed concepts. At the beginning, the ontology engineer also found it hard to access the field because of a lack of knowledge. Participatory research also helps build capacities regarding the usage of qualitative research methods, so this point of mutual learning was also achieved. The interviews during the design process also showed that the capacities increased on both parts, so the interviews were conducted on a higher level.

The quality of feedback also increased during the process: Interviewees in the design phase were more focused on several points as well as being more specific on what information is needed for the ontology. This makes this approach superior to only involving interviewees in the requirement analysis. It also became evident that when doing interviews in the design phase, the interviews were more and more focused, people had also made their minds up about the ideas as well as remembering things from the previous interviews. The other advantage was that during the design phase the researchers' understanding of the concepts and the ideas improved a lot compared to the initial interviews during the requirements analysis.

The design phase also showed that as soon as the interviewees are given a prototype, they understand the idea better and also give better feedback. This raises an issue. On the one hand, it is sensible to show early prototypes in order to grasp user requirements (see Warwick, 2012). Therefore it is essential to show early prototypes in the process instead of waiting until the end of the process to involve users. On the other hand, the more concrete examples are given, the more concrete the feedback is. We addressed this issue by giving the

user as many opportunities as possible in the process phase to give feedback, which got more concrete in the course of development.

Two examples highlight the advantages of user involvement in the participatory design with respect to the quality of the ontology, introducing new ideas: The first example is the possibility to upload a visualization of a coding schema: In grounded theory, it is important to connect codes in a meaningful way. But this is not very well described in textbooks and only a few examples are given, which are – according to the interviewees – not sufficient. Therefore, a metadata visualization was added to get this information and to allow the researchers to upload and publish these figures.

Another example is information about the sampling of a study: in qualitative research sampling is often done via theoretical sampling. But this is not easy to formalize like in quantitative research, where the most important information is whether the sampling was representative. Therefore, a nonformalized field was added, where people can describe how they did their sampling and why.

Two problems persisted when using a participatory design, especially with interviews: contradicting comments of the interviewees and the higher effort involved in the participatory design.

For one, the interviewees submitted contradictory comments. In these cases a compromise was sought during the design process, by showing ideas and then talking to the interviewees again to find out how they felt with respect to their opinion from the last feedback round. Therefore it is important to balance in the design process between interested and positive participants. It is important to find participants who are open to the proposed ideas. Otherwise, the interviews will not get a great outcome. At the same time, this selection might also cross out critical voices that should still be considered. Using theoretical sampling is helpful in hearing different voices. It can also be useful to involve nonmainstream opinions in the research.

One limitation is the higher effort that a participatory and iterative approach needs, but this time is well invested because feedback quality is good in the development process. Qualitative interviews and observations are also time-consuming, yet especially in the case study with researchers not familiar with ontology engineering, this was useful in order to also build capacities. Qualitative methods also may yield contradictive results; therefore, moderating skills are needed to find a compromise.

4.2 Preliminary results on the case study

Another main result of the research is the ontology itself, which can be found on Github [12]. The requirements for the ontology were that it is crucial to add further metadata, e.g. the theoretical background, the type of data and the scientific discipline. The researchers mentioned it should be possible to get an overview of the coding schemas without needing to read the complete methods section of the publication.

Figure 1.
Participatory ontology design process

The ontology is split into five categories: publications, studies, research data, coding schema and codes. This allows including all information that is important for describing coding schemas. The following Figure 2 shows the structure of the ontology. The relation between the coding schema, research data, project and publications is N:N, which denotes that as many instances of one can relate to the other, e.g. when for one project there were several publications. Only the codes depend on one coding schema, so one coding schema can have an unlimited number of codes, but not vice versa:

Users identified several areas where the ontology can help to make qualitative research more open: while searching for coding schemas, a standard would help find coding schemas and researchers would not have to read complete studies, reusage would also be easier. Publication could also be easier because many coding schemas have not been published yet and owing to a publication, researchers can get credit for their coding schemas, similar to the sharing of data or research software. Furthermore, this process might expand in two other ways, i.e. the validation of other researcher's work and usage for teaching because textbook examples are often rather short.

Two critical points were raised: first, it might be an issue that the amount of work might increase. All except for one interviewee were positive about sharing, but mentioned as a potential problem for many of the researchers. Therefore it is crucial to also involve developers of research software to provide basic tools for this in their software, so researchers can do more of the work during their research. One interviewee also mentioned that she would train her students to use the qualitative methods in a more traceable manner, e.g. she would teach the students to create code-memos when they created a code so they would not have to

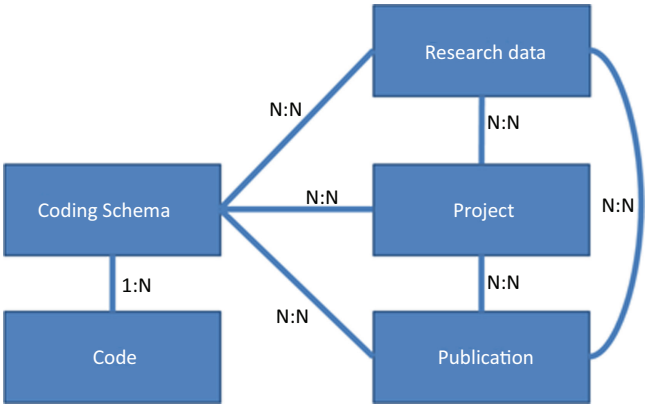

Figure 2.
Structure of the ontology

Group	Importance
Researchers	Most important group works with the data. Often also teaches methods
Professors/supervisors	Important, might or might not be part of the research process, but can be ambassadors for opening up research and are often also involved in teaching
Developers of QDA (qualitative data analysis) software	Less important, but via the functions of software, data-sharing can be made easier
Research data centers	Less important. Responsible for infrastructure and therefore important when it comes to the research of data
Funding bodies	Least important. Can enforce data sharing via funding

Table 1.
Result of stakeholder analysis

do this as a time-consuming task at the end. It was also mentioned that a structure of rewards within the scientific community (e.g. to treat the publication of a coding schema as meaningful as the publication of data) is needed.

Another concern raised by one person was the fear of mainstreaming research when publishing schemas. This was described as the problem of researchers just copying codes from other researchers rather than thinking for themselves and questioning if these codes fit the actual research.

5. Conclusion

Based on this experience, it can be claimed that participatory design in ontology engineering has a huge potential, although more research is needed in order to validate the methods and the ideas better. The positive outcomes can be narrowed down to four points:

(1) The users' commitment toward to project got stronger when they were asked multiple times referring to prior interviews.

(2) Mutual learning did happen during the process, as can be seen in the comments of the interviewees as well as in our knowledge about the working process of the interviewees.

(3) We saw that the quality of feedback increased during the design process, the more involved people got and the more concrete the prototypes were.

(4) The quality of the ontology increased due to the design. Without expert knowledge, it is difficult to develop ontologies. The participatory approach also helped because the goal was to create an ontology that supports the research process in qualitative research. The users articulated ideas from their practical experience that are not mentioned in other sources.

Two obstacles have to be considered: when using interviews, it can often happen that the interviewees give contradicting comments which have to be moderated and the whole process or participatory and iterative process takes more time.

There is a great potential for participatory ontology design in cases where no ontologies to build upon are available as well as fields where users have high expertise, e.g. in science. There is another high potential in the creation of ontologies "in action" in fields, where the conceptualization is newly created like in the proposed case study.

Gregory (2003) [13], mentioned criteria that have been met with these outcomes. The involved persons showed a great commitment to the process and fed unexpected ideas into the design process. We can also see that participatory design supports the creation of this explicit shared conceptualization (Studer *et al.*, 1998) based on implicit conceptualizations in the minds of the researchers before.

At the point of writing, the design phase as well as the first evaluation phase is finished, the next and final step will be the second evaluation phase. The ontology will then be ready to be implemented in research data centers. The prototype of the ontology can be found on Github [14].

Notes

1. The criteria are: Accuracy, Correctness, Completeness, Appropriateness, Consistency, Objectiveness.
2. The criteria are: Interoperability, Completeness, Order arrangement, Usefulness in information retrieval, Difficulty in cataloging, Terminology Interviewee, Redundancy Interviewee.

3. https://discover.ukdataservice.ac.uk//QualiBank.
4. https://www.fdz-bildung.de/.
5. https://dans.knaw.nl/nl.
6. https://www.qdasoftware.org/codebook-exchange/.
7. https://www.qsrinternational.com/nvivo/nvivo-products.
8. https://atlasti.com/.
9. https://www.semantic-mediawiki.org.
10. https://www.qdasoftware.org/.
11. Palavitsinidis (2014) defines completeness as how complete a metadata item has been filled out. Hu *et al.* (2018) use the definition that is also used here.
12. https://github.com/julianhocker/Quali-Codes-Ontology.
13. Improving the knowledge of people affected by systems as well as involving them in the design process and giving them a say.
14. https://github.com/julianhocker/Quali-Codes-Ontology.

References

Auer, S. (2018), January 22, "*Towards an open research knowledge graph (Version 1)*", Zenodo, doi: 10.5281/zenodo.1157185.

Borst, W.N. (1999), *Construction of Engineering Ontologies for Knowledge Sharing and Reuse*.

Borgman, C.L. (2009), "The digital future is now: a call to action for the humanities", *Digital Humanities Quarterly*, Vol. 3 No. 4.

Evers, J.C. (2018), "Current issues in qualitative data analysis software (QDAS): a user and developer perspective", *Qualitative Report*, Vol. 23 No. 13, pp. 61-73.

Faniel, I.M., Kriesberg, A. and Yakel, E. (2016), "Social scientists' satisfaction with data reuse", *Journal of the Association for Information Science and Technology*, Vol. 67 No. 6, pp. 1404-1416.

Fecher, B. and Friesike, S. (2014), "Open science: one term, five schools of thought", in *Opening Science*, Springer, Cham, pp. 17-47.

Fernández-López, M., Gómez-Pérez, A. and Juristo, N. (1997), *Methontology: From Ontological Art Towards Ontological Engineering*.

Genesereth MR, N.N. (1987), *Logical Foundations of Artificial Intelligence*.

Gläser, J. and Laudel, G. (2010), *Experteninterviews und qualitative Inhaltsanalyse*, Springer-Verlag, Wiesbaden.

Glaser, B.G. and Strauss, A.L. (1967), *The Discovery of Grounded Theory, Strategies for Qualitative Research*, Routledge, London.

Gregory, J. (2003), "Scandinavian approaches to participatory design", *International Journal of Engineering Education*, Vol. 19 No. 1, pp. 62-74.

Gruber, T.R. (1995), "Toward principles for the design of ontologies used for knowledge sharing?", *International Journal of Human-Computer Studies*, Vol. 43 No. 5-6, pp. 907-928.

Guarino, N., Oberle, D. and Staab, S. (2009), "What is an ontology?", in *Handbook on Ontologies*, Springer, Berlin, Heidelberg, pp. 1-17.

Heuwing, B. and Womser-Hacker, C. (2015), "Zwischen Beobachtung und Partizipation–nutzerzentrierte Methoden für eine Bedarfsanalyse in der digitalen Geschichtswissenschaft", *Information-Wissenschaft and Praxis*, Vol. 66 Nos 5-6, pp. 335-344.

Hu, X., Ng, J. and Xia, S. (2018), "User-Centered evaluation of metadata schema for nonmovable cultural heritage: murals and stone cave temples", *Journal of the Association for Information Science and Technology*, Vol. 69 No. 12, pp. 1476-1487, doi: 10.1002/asi.24065.

Kemman, M. and Kleppe, M. (2015), "User required? On the value of user research in the digital humanities", in *Selected Papers from the CLARIN 2014 Conference, October 24-25, 2014*, Linköping University Electronic Press, Soesterberg, No. 116, pp. 63-74.

Kingsun, M., Myers, T. and Hardy, D. (2018), January, "C-DOM: a structured co-design framework methodology for ontology design and development", in *Proceedings of the Australasian Computer Science Week Multiconference*, pp. 1-10.

Kuckartz, U. (2018), *Qualitative Inhaltsanalyse: Methoden, Praxis, Computerunterstützung*, BeltzJuventa, Weinheim.

Kuhlen, R. (1995), *Informationsmarkt: Chancen und Risiken der Kommerzialisierung von Wissen.*

Lam, J., Abdullah, M.S. and Supriyanto, E. (2015), "Architecture for clinical decision support system (CDSS) using high risk pregnancy ontology", *ARPN Journal of Engineering and Applied Sciences*, Vol. 10 No. 3, pp. 1229-1237.

Lee, Y.W., Strong, D.M., Kahn, B.K. and Wang, R.Y. (2002), "AIMQ: a methodology for information quality assessment", *Information and Management*, Vol. 40 No. 2, pp. 133-146.

Mayring, P. (2015), *Qualitative Inhaltsanalyse. Grundlagen und Techniken*, 12th ed., Beltz, Weinheim/ Basel, 2015, ISBN 978-3-407-25730-7.

Mons, B., Neylon, C., Velterop, J., Dumontier, M., da Silva Santos, L.O.B. and Wilkinson, M.D. (2017), "Cloudy, increasingly FAIR; revisiting the FAIR data guiding principles for the European open science cloud", *Information Services and Use*, Vol. 37 No. 1, pp. 49-56.

Muller, M.J. (2007), "Participatory design: the third space in HCI", in *The Human-Computer Interaction Handbook*, CRC Press, Boca Raton, pp. 1087-1108.

Ongenae, F., Bleumers, L., Sulmon, N., Verstraete, M., Van Gils, M., Jacobs, A., De Zutter, S., Verhoeve, P. and Ackaert, A. (2011), "Participatory design of a continuous care ontology-towards a user-driven ontology engineering methodology", *Knowledge Engineering and Ontology*, Proceedings of the International Conference on Knowledge Engineering and Ontology Development, KEOD, (IC3K 2011), ISBN 978-989-8425-80-5, Vol. 1, pp. 81-90, doi: 10.5220/0003654700810090.

Palavitsinis, N. (2014), "Metadata quality issues in learning repositories", available at: http://hdl.handle.net/10017/20664.

Reinhold, A. (2015), "Das experteninterview als zentrale methode der Wissensmodellierung in den digital humanities", *Information-Wissenschaft and Praxis*, Vol. 66 Nos 5-6, pp. 327-333.

Sabou, M. and Fernandez, M. (2012), "Ontology (network) evaluation", in *Ontology Engineering in a Networked World*, Springer, Berlin, Heidelberg, pp. 193-212.

Saldaña, J. (2015), *The Coding Manual for Qualitative Researchers*, Sage, Newcastle upon Tyne.

Schreier, M. (2012), *Qualitative Content Analysis in Practice*, Sage Publications, Newcastle upon Tyne.

Strübing, J., Hirschauer, S., Ayaß, R., Krähnke, U. and Scheffer, T. (2018), "Gütekriterien qualitativer sozialforschung. Ein diskussionsanstoß", *Zeitschrift für Soziologie*, Vol. 47 No. 2, pp. 83-100.

Studer, R., Benjamins, V.R. and Fensel, D. (1998), "Knowledge engineering: principles and methods", *Data and Knowledge Engineering*, Vol. 25, Nos 1-2, pp. 161-197.

Tan, H., Adlemo, A., Tarasov, V. and Johansson, M.E. (2017), "Evaluation of an application ontology", in *Proceedings of the Joint Ontology Workshops 2017 Episode 3: The Tyrolean Autumn of Ontology Bozen-Bolzano*, September 21–23, 2017, Vol. 2050, CEUR-WS.

Uschold, M., King, M., Moralee, S. and Zorgios, Y. (1998), "The enterprise ontology", *The Knowledge Engineering Review*, Vol. 13 No. 1, pp. 31-89.

Warwick, C. (2012), "Studying users in digital humanities", *Digital Humanities in Practice*, Vol. 17 No. 5, pp. 1-21.

Wang, R.Y. and Strong, D.M. (1996), "Beyond accuracy: what data quality means to data consumers", *Journal of Management Information Systems*, Vol. 12 No. 4, pp. 5-33.

Zhang, Y. and Li, Y. (2008), "A user-centered functional metadata evaluation of moving image collections", *Journal of the American Society for Information Science and Technology*, Vol. 59 No. 8, pp. 1331-1346.

Corresponding author
Julian Hocker can be contacted at: julian.hocker@dipf.de