Hocker, Julian; Bipat, Taryn; McDonald, David W.; Zachry, Marc

Evaluating QualiCO. An ontology to facilitate qualitative methods sharing to support open science

Journal of internet services and applications 12 (2021) 5, 29 S., 10.1186/s13174-021-00135-w

Hocker et al. Journal of Internet Services and Applications (2021) 12:5
https://doi.org/10.1186/s13174-021-00135-w

Journal of Internet Services and Applications

RESEARCH Open Access

Evaluating QualiCO: an ontology to facilitate qualitative methods sharing to support open science

Julian Hocker[1*†], Taryn Bipat[2†], David W. McDonald[2†] and Mark Zachry[2†]

*Correspondence:
julian.hocker@dipf.de
†Julian Hocker, Taryn Bipat, David W. McDonald and Mark Zachry contributed equally to this work.
[1]DIPF | Leibniz Institute for Research and Information in Education, Frankfurt am Main, Germany
Full list of author information is available at the end of the article

Abstract

Qualitative science methods have largely been omitted from discussions of open science. Platforms focused on qualitative science that support open science data and method sharing are rare. Sharing and exchanging coding schemas has great potential for supporting traceability in qualitative research as well as for facilitating the reuse of coding schemas. In this study, we present and evaluate QualiCO, an ontology to describe qualitative coding schemas. Twenty qualitative researchers used QualiCO to complete two coding tasks. In our findings, we present task performance and interview data that focus participants' attention on the ontology. Participants used QualiCO to complete the coding tasks, decreasing time on task, while improving accuracy, signifying that QualiCO enabled the reuse of qualitative coding schemas. Our discussion elaborates some issues that participants had and highlights how conceptual and prior practice frames their interpretation of how QualiCO can be used.

Keywords: Ontology evaluation, Open science, Qualitative coding

1 Introduction

The basic ideas behind open science are probably centuries old and are present in the earliest scientific letters and in the establishment of journals for sharing research results. The expansion of science as a productive discipline, and concerns about junk science, have generated renewed interest in the concept of open science. Since roughly 2000, discussions of what constitutes open science have grown [1]. Some clear components of open science include the sharing of scientific data, sharing and explaining scientific methods, as well as sharing research results.

For many open science efforts the accessibility of data is considered the primary means of openness across multiple fields of knowledge, people, and institutions as well as the validation of results. According to the Open Science Collaboration (OSC), prior research "should not gain credence because of the status of authority of their originator but by the replicability of their supportive evidence" [2]. In the midst of the ongoing replication crisis it is important to be able to have access to prior data as well as clearly described methodologies for collecting, cleaning, manipulating and analyzing such data.

Springer Open

Furthermore, sharing data allows researchers to build off the assumptions and efforts of past research. Access to open data enables the broadening of research scope and the ability to diversify perspectives in science [3].

A focus on open data as a road to open science has resulted in concerns about potential unintended outcomes and the differential benefits for scientists of different status [4]. Addressing these challenges is important for facilitating widespread adoption of open science data sharing. However, focusing on data sharing to the exclusion of other, equivalent challenges faced by different scientific traditions may be short sighted. The sharing of methods complements the sharing of data, making all the research process transparent and allowing for traceability and reproducability of research.

To date, the most of the focus on data and methods sharing in open science has been in the quantitative sciences. Scientific disciplines that rely on qualitative methods, in contrast, have largely been overlooked or completely omitted from broad discussions of open science. In scientific disciplines that use qualitative methods, researchers tend to focus on reliability and traceability of data and methods rather than on generalizability and repeatability.

In this study, we evaluate QualiCO, an ontology for qualitative coding schemas, to understand how the ontology might improve the work of open science. We specifically observed how qualitative researchers use QualiCO in an experimental setting. We also want to understand if the definition of the ontology includes the necessary metadata for researchers to apply unfamiliar coding schemas. Overall, we want to better understand how we can improve QualiCO to satisfy the needs of qualitative researchers to reuse qualitative coding schemas and fulfill the needs of open science.

2 The potential for open science in qualitative methods

A potential issue in addressing open science concerns for qualitative methods is that even within that term of art, the range of methodological approaches is quite broad. Given the range of methods, some researchers have rejected the idea of even attempting to define qualitative research [5]. Others who attempt definitions have offered that qualitative research is non-numeric [6] and unstructured; or, simply it is defined as the "opposite" of quantitative research.

The juxtaposition of qualitative methods in contrast to quantitative methods is aligned with the contrasting underlying philosophical stances of interpretivism and positivism. But this simplification obscures the breadth within qualitative traditions; within qualitative methods there is a range of interpretivism to positivism. Some ethnographic methods approach data and their explanation as a purely interpretivist act, seeking to describe human experiences in ways that can be more broadly understood. At the more positivist end of the qualitative method spectrum we find systematic qualitative coding methods. In such systematic methods researchers are data-driven, looking for repeatable, systematic ways to analyze or code the data.

Another way of describing the positivist spectrum of qualitative methods distinguishes between two methods for coding data: Grounded Theory and Qualitative Content Analysis. These methods also differ in the way they treat coding schemas. In Grounded Theory, coding schemas can be seen as the result of the inductive coding process where codes are generated from the data. Meanwhile, in Qualitative Content Analysis, coding schemas can be seen as a tool that tells the researchers how they should code which parts of the texts. However, within these methods there are still wide variations, especially when it comes to developing a coding schema using a Qualitative Content Analysis; this can be done inductively, deductively or as a mixture of both [7].

In the same way that current open science and open data platforms have not addressed all challenges of quantitative science traditions, it seems misguided to believe that all qualitative traditions could be easily supported by a single system. In fact, attempting to address the needs of qualitative researchers by building upon what we already know about open science might best address the needs of qualitative researchers who use methods that are closer to the positivist end of the spectrum. And that is largely the approach taken in our work here.

In this study, we describe and evaluate the ontology for qualitative coding schemas QualiCO [8]. QualiCO is designed to describe a wide range of qualitative coding schemas. In this case a qualitative coding schema is a set of codes, possibly hierarchical, that can be applied to qualitative data by a researcher. A researcher sharing a coding schema would instantiate the ontology with a designated set of codes and code descriptions. The ontology can be elaborated with sample data, publications, and various other metadata information that can make the schema more usable. We describe this more fully below.

Broadly, we believe an ontology for sharing qualitative coding schemas has the potential to make qualitative research more transparent as well as encouraging researchers to share more of their work, making their scientific process more visible. The ontology can make the reuse of qualitative schemas easier by providing a richer context for understanding the schema, how it can be applied, and for which types of data it is effective. However, as our analysis illustrates, for such a system to facilitate open science, the costs of participation need to be integrated into the workflow of researchers at a minimal cost to encourage sharing.

3 Research goal

In this study, we seek to understand the performance of QualiCO, a system that was described in a prior research project focused on defining qualitative coding schemas [8]. A qualitative coding schema is the description of the qualitative codes, which can be seen as a result of the work in grounded theory or as a tool for qualitative coding [7]. QualiCO was implemented into a prototype using Semantic MediaWiki[1] allowing researchers to add coding schemas following the standards of QualiCO. The goal of this research was to see whether a prototype using QualiCO fulfills the requirements for use in open science. We focused on the reuse of qualitative coding schemas. This is an important application since it could support potential generalizability, traceability and better enable replication of qualitative studies. More specifically, we ask the following research questions:

- RQ1: How does researcher performance change by using the ontology across different coding schemas?
- RQ2: Does a nearly complete instantiation of the ontology provide enough metadata for a researcher to reliably apply the coding schema?
- RQ3: How can this ontology be improved to make the reuse of coding schemas easier?

In this paper, we first outline prior work related to how open science researchers share qualitative coding schemas. Our review attempts to highlight some of the similarities and differences between qualitative and quantitative approaches to science and those

[1] https://www.semantic-mediawiki.org

implications for openness. We then review various ontology evaluation methodologies and describe the ontology being evaluated in this study. Following this, we outline our evaluation methods. In our findings we present task performance data and feedback on the ontology based on a semi-structured interview designed to focus participants' attention on the ontology. Finally, we close by discussing the results and recommendations for what a future platform for sharing coding schemas might look like. It is important to note that this study is not a full deployment of QualiCO but an evaluation of one mid-development iteration. The goal of this study is to demonstrate the usefulness of the system and understand how to improve QualiCO to fit the needs of qualitative researchers. Furthermore, the version of QualiCO used in this study is only one iteration in a larger iterative design process [8].

4 Prior work

In our overview of related literature, we begin with the general problem of sharing research data in open science and then move to the more focused issues around sharing qualitative methods and data. This prior work highlights descriptions of research procedures as well as standards for metadata as challenges to sharing. As we review this prior work, it is worth noting that we include as a starting point sources that are primarily about data sharing in open science because they are well studied, and point toward important considerations for our focus on sharing qualitative coding schemas. We consider sharing qualitative schema to be more clearly related to methods sharing than to the broader concern of sharing data. We thus begin with important findings related to research data sharing to consider what they may imply about method sharing. Following this, we describe the state of method sharing in qualitative research as well as archiving qualitative coding schemas.

4.1 Data sharing and reusing in absence of method sharing

In the last few years, several studies examining how open science researchers share their data have been reported [9–12]. Many of these studies notably focus on why researchers do not share their data. When researchers do share data, a major challenge is describing the data. Data description is necessary for information experts to develop adequate insights into the prior research. For this reason, researchers inclined to share their data have need to make it available via easy-to-use platforms [13, 14]. Understanding how data is input into such platforms is also important. To make shared coding schemas possible, those schemas need to be structured following standards that can be encoded into easy-to-use platforms.

When it comes to sharing of research data, there have also been several advances: the FAIR principles have undergone several developments [15] as well as have the platforms that embody CARE principles [16]. The FAIR principles, which address findability, accessibility, interoperability and reuseability, formulate criteria for research data to be shared within a platform. The CARE principles are designed to support collective benefit, authority to control, responsibility and ethics. These principles offer a blueprint for the creation of research data centers, providing guiding principle for the creation of new platforms in open science. With the development of QualiCO, we addressed especially the criteria of findability for a future platform.

Researchers examining the work of open science argue that documentation of resources is key to overcoming common obstacles to data reuse. For example, documentation of data is a key factor in the potential reuse of that data. Kowalczyk & Shankar [17] define this needed documentation as context: "[...] context documents how datasets fit into their physical and technical environments (file formats and field descriptors) as well as into the scientific environment (experiment treatments and applications)." Curty [18] identifies a set of factors that motivates data reuse in the social sciences: 1) the potential expansion of knowledge via re-analysis of data, 2) it is cheaper to reuse data than generate new data, and 3) the pre-endorsement that existing data is perceived to be of high quality when researchers took the time to provide it for other researchers. Curty [18] also names several factors that enable data sharing: 1) the availability of documentation, 2) availability of data repositories, 3) contact with primary investigators, 4) support from the research data centre and 5) skills for analyzing data.

Another work by Faniel et al. [19] demonstrates that metadata quality is important for the satisfaction of social scientists with data portals. Our exploratory study does not include measures comparing qualitative and quantitative research. We see the factors identified by Curty [18] as potentials for improving the sharing of qualitative coding schemas. We focus on the point of documentation of the coding schemas, which was also named as important for the sharing of data.

Building from these premises, we hypothesize that the quality and completeness of metadata will be crucial to improve the sharing of qualitative methods. Further, this depends on the ontology used to describe those coding schemas. The underlying ontologies describe the rules for documenting coding schemas. Based on this approach, we use the definition of metadata or ontologies as "a formal, explicit specification of a shared conceptualization" [20]. It is an approach we believe has the potential to fill a gap that must be addressed in order to document coding schemas in a meaningful way to better enable open science in qualitative research. We describe this ontology, QualiCO, in Section 5.

4.2 Qualitative coding schemas in the context of open science

Many projects using qualitative methods can also be considered a part of "small science" [21]. Small science is most often conducted by single researchers without published standards or platforms, yielding "dark data" [22], that is data that is functionally impossible to reuse. A part of this reuse challenge is related to limitations in how coding schemas are (or are not) shared. This concern is bound to the pragmatic aspects of how qualitative science is produced. Concerns about descriptions of qualitative data analysis and sampling [23, 24] contribute to problems of method sharing.

Our research is specifically focused on the use of qualitative coding schemas in the work of open science. Broadly speaking, qualitative coding schemas are used with different methods, including Grounded Theory [25], general classification [26], and content analysis [27, 28]. Following these methods, coding schemas include the codes used within the study as well as documentation about when and how to use them. Coding schemas thus include essential documentation integral to their use within a qualitative method. Sharing qualitative coding schemas, consequently, entails method sharing in open science qualitative research.

Notably, the sharing of methods is different from the sharing of data since there is more documentation needed to understand how the researchers worked using their method. In open science data sharing, the important information is the data itself and its origin. For sharing coding schemas, however, it is important to know contextual background information about the research, such as the theoretical assumptions of the researchers. The contribution of this paper builds upon our prior research [29], describing the development of QualiCO as well as first results of an evaluation. Those first results showed that it is important for the acceptance of a future platform that it fits with the workflow of researchers.

4.2.1 Method sharing and re-using in qualitative research

The distinctions between qualitative and quantitative methods are many. One clear aspect is the nature of enumeration and how quantification relates directly to the result. In quantitative methods, one key aspect of finding significance for a result relates to statistical methods. In this context method sharing is simplified through naming the statistical method used (e.g., linear regression plus parameters). This, and other forms of formalization in methods, have helped to simplify and standardize the expression of methods to facilitate sharing. As new methods evolve, they come with more or less formalization and therefore open science here involves communication as well as additional details such as sharing code, additional documentation, and even private contact with researchers [30, 31].

We assert that qualitative research method sharing is more complicated because of less formal definitional foundations of practice. As one example, Saldaña [26] shows 30 methods of 'qualitative coding', each divided into several categories. This illustrates the broad range of qualitative coding methods. Added to this, qualitative coding also involves human interpretation, which means that just naming a code might not make transparent how to perform coding nor explain the criteria for a code application. If codes were derived from an established theory or model, sharing is likely easier, but with inductive coding or mixtures of both, things are less clear. In general, less formality has yielded less clear short-hand for methods descriptions of qualitative research which necessitates more elaborated descriptions of methods.

Sharing qualitative coding schemes is thus inherently challenging from the onset. A known systemic obstacle in the work of open science is that research is hampered by a lack of standards for sharing research materials, including codebooks [32]. Aguinis & Solarino name twelve criteria for making a qualitative study reproducible, which includes methods, data coding practice and first-order codes as well as data analysis and second or higher-order codes [33].

This call is echoed by others who encourage the sharing of intermediate steps of research in open science, such as Grubb and Easterbrook's article [34], which explicitly addresses the "various mechanisms" used for sharing work. One of these mechanisms is pre-registration, which is often missing in qualitative research. Taminen et al. [35] claimed that open science practices are rarely used in the field of sport and exercise psychology (i.e., there were no studies preregistered at the time of their writing). They encourage qualitative researchers to share evidence and artifacts from the research process in order to make their research more transparent. Yet another issue is that open science is rarely taught in qualitative research [36]. In this respect, teaching the practices of open science has huge potential to influence qualitative research.

Individuals and professional organizations have continued to advocate for better open sciences practices from qualitative researchers. Kapiszewski & Karcher [37] advocate for pre-registration of research projects, methodological explanations, annotation, export of codes from QDA-software as well as the sharing of data. Further, the new publication guidelines of APA [38] call for additional information to be provided in scientific articles including:

- Description of selection of participants as well as the relationship to them and how long data collection (e.g., interviews) took
- Description of data analysis, which method was used, whether coding was done inductive or deductive
- Methodological process, including a description of how researchers came to their conclusions
- Results described in as concrete terms as possible, using graphs to describe the most important codes.

Summing, up we can see that description of the data analysis is an important part of an open science practice for qualitative research. For researchers using methods like qualitative coding or Grounded Theory, the coding schema includes most of this information. Therefore it makes sense to make qualitative coding schemas more publicly available and to augment their accessibility with the necessary meta data to make them reusable.

4.2.2 Archiving qualitative coding schemas

Overall, we can see that the issue of data sharing fosters tension within qualitative methods, but the sharing of methods has not yet been a common focus. However, in the discussion about archiving qualitative coding schemas, there has been some development led by the group REFI[2]. REFI took one step towards the sharing of qualitative coding schemes by creating the standard REFI-QDA Codebook [39], an exchange format with which it is possible to transfer a code in one software to the next. With this, the REFI-QDA Codebook also provides a good tool to archive and reuse coding schemas.

In our research, the goal is to reuse these methods. Concretely, we want to see whether the QualiCO ontology implemented into a prototype Wiki platform helps researchers reuse coding schemas for coding sample texts. In other words, we seek to discover whether researchers can reuse a method described by other researchers.

Understanding coding schemas through prior research is challenging. In some cases the creation or development of a schema is the objective of the research. For example, many who follow a Grounded Theory approach, develop coding schemas to reflect a theory of the phenomena they will ultimately attempt to describe [40]. In other qualitative research, the coding schema is a means of getting to the real result. In other words, coding schemas as they are used by qualitative researchers may be either inductive or deductive. Regardless of which approach qualitative researchers use, however, there are few systematic ways of describing and cataloging schemas and how they are applied. Consequently, meta-level discussions of qualitative coding schemas are scarce despite the widespread use of qualitative methods across fields and disciplines.

One way to facilitate the sharing of data is the creation of platforms where researchers can expose how the data was created. This would include the sharing of qualitative coding schemas, but there does not exist such a platform. There needs to be more ways of attribution towards the sharing of coding schemas in order to motivate people sharing their data. Another way to ensure this type of sharing is for journals to require the sharing of coding schemas whenever an article is published [17]. Publishing data and coding schemas together with corresponding articles would also give subsequent researchers

[2] https://www.qdasoftware.org/

the opportunity to easily find data, which would in turn help in the sense of expanding scientific knowledge graphs [41].

Summing up, there is currently no platform that allows the sharing of qualitative coding schemas. If such a system is to be developed, it has to fit into the workflow of the users in order to make it useful and to lower issues for uploading coding schemas to the system.

4.3 Ontology evaluation

Ontology evaluations have been quite diverse and range from technical evaluations with the focus on how the ontology supports reasoning to user-based approaches that focus on the utility or usability of the ontology. A survey by McDaniel & Story [42] categorizes ontology evaluation into five groups: Domain/Task Fit, Error Checking, Libraries, Metrics, and Modularization. Using this categorization a human-centered evaluation might fall into the McDaniel & Story's categories Metrics or Domain/Task Fit.

Work by Palavitsinis et al. [43] offers an overview of the most often used metrics for user studies of ontology systems. The alternative to quantitative metrics based evaluations includes interviews [44], questionnaires [43] or a combination of both [45] that rely on human perceptions of the system.

Critics have argued that ontology evaluations need to be more closely aligned to their intended purposes. A Task-Driven approach as described by McDaniel and Storey [42] focuses on the goal orientation as derived from requirements engineering. Specifically, the goal of an ontology should be to support the information system it is built into. In our case, this means that the ontology should support the upload, search and usage of qualitative coding schemas.

Grey [46] describes this need: "An application ontology should be evaluated against a set of use cases and competency questions which represent the scope and requirements of the particular application. For example, a user query use case may contain the competency question 'what cancer cell line data is there.' This requires sufficient ontological coverage to capture the concept of 'cancer cell line.'" [46]

The user analysis can be done using two ontologies or two versions of ontologies. Reinhold [47] evaluates an old version of the ontology against a new version, Liu et al. [48] evaluate three ontologies in order to pick one to select for a given task. The evaluation by Yu et al. [49] involves users in a browsing task based on the ontology in Wikipedia, which has been enriched in a different way. Another version of task-based evaluation is presented by Pittet et al. [50], where users are able to enrich the ontology.

In the development of the QualiCO ontology, participatory design was used [8], therefore it made sense to focus on user-centered approaches also in the evaluation of the ontology. With this research, we present an approach to ontology evaluation using user tests. Our goal was to create a scenario for a task-driven evaluation of an ontology, wherein the users are able to use the ontology and are given a task they have to solve using the ontology.

5 QualiCO - an ontology for qualitative coding schemas

Our evaluation is focused on the QualiCO ontology, which was developed to describe qualitative coding schemas[3]. These qualitative coding schemas are used during systematic

[3] In this study, we use the term "ontology" to describe QualiCO, the ontology that describes coding schemas and which we want to test in our study. In contrast we use the term "schema" to delineate the coding schemas that are provided by qualitative researchers who are participating in an open science effort.

coding of qualitative data. Qualitative coding is performed in a range of ways, so QualiCO is not aimed at any specific approach. In fact, the QualiCO ontology can likely support a wide range of approaches. It was developed through participatory design methods including interviews and observations. The participatory design methods had a diverse range of qualitative methods practitioners, but most could be described as on the positivist end of qualitative methods (c.f. Section 2). Ontology development included several rounds of feedback on prototypes that represented different aspects of the ontology [8]. In an earlier study, we evaluated whether QualiCO addresses specific open science challenges for qualitative methods [29].

QualiCO consists of five main categories: publications, research data, study descriptions, coding schemas and codes. Each of these categories contains further information about these types of information. The development process of the ontology is described in [51] and can be found on GitHub[4]. The ontology is divided into four main classes: *coding schema* (23 attributes), *study* (ten attributes), *research data* (15 attributes) and *publications* (seven attributes). There is also the class *code*, which is a subclass of coding schema and contains eight attributes. The following graphics shows the structure:

The goal of the ontology was to provide enough metadata to understand the background of the research activity as well as providing a detailed description of the coding schema. At the main level of the description we use the term "project."[5] Naturally, not all projects will contain all potential artifacts, but the ontology was designed to represent key artifacts related to a project in order to give a good understanding of the coding schema.

The class "coding schemas" describes the metadata for the coding schemas contained by any project in the ontology. This includes metadata to describe the methods used for analyzing data, the theoretical background of the work, research questions as well as the process of schema creation. It also supports the upload of coding schemas using REFI standards [39] allowing researchers to easily export and import coding schemas into their commercial QDA software.

The class "coding schema" provides a link to each "code" in the schema. Each code is described using the following metadata:

- Name of the code
- Description of the code
- Example
- Counter example
- Connection between codes
- Provenance
- Count

The fields "Name of the code," "Description of the code," and "Example" are required, the other fields are optional. The ontology is designed to support systematic qualitative coding, but may support grounded theory methods as well, making connection between codes more important for Grounded theory and the number of codes more relevant for more reproducibility-oriented qualitative coding.

[4]https://github.com/julianhocker/Quali-Codes-Ontology.

[5]The common peer-contribution model has a "project" added to QualiCO by the researchers who conducted the study. Those researchers contribute their coding schema and other artifacts of QualiCO. At this point the coding schema is a required artifact. One goal of the research is to understand what other data and relationships (e.g. research data, publications) are required to help other researchers reuse coding schemas available through a system implementing QualiCO.

The evaluation described in this paper is focused on the potential effectiveness of QualiCO to support reuse of specific coding schemas. While any system that would use the full ontology must represent and facilitate the effective navigation of multiple projects, we will not address the search and navigation question in this evaluation. Our work here focuses on the potential to reuse coding schemas and has restricted the specific number of projects represented in the prototype for this evaluation.

For the evaluation, a prototype implementation of QualiCO was built using Semantic MediaWiki and populated with information about coding schemas, research data, publications and project background for two research projects that our research participants were able to browse. The research participants were able to navigate freely in the system and choose the information they needed for the specific study tasks. The MediaWiki navigation was as follows: on the main page there were links to overview pages for publication, study, research data and coding schema. If the participant clicked on one of these main categories, they found a page with all content within these main categories, e.g., all coding schemas in the system. From this, participants could navigate to a specific item, e.g. coding schema. On these item pages, participants found all metadata as well as links to codes, publication, study and research data, which belong to this coding schema. This navigation largely reflects the links present in the ontology diagram of Fig. 1.

We set up the prototype by describing two projects using QualiCO including information about publications, research data, the coding schemas and each code: The first project was about the coding of Barnstars, which are awards in Wikipedia given to recognize specific work performed by contributors. The coding schema used in this project is described in prior work [52, 53]. The other coding schema was based on a project analyzing reviews of Amazon's voice assistant system Alexa. A version of the coding schema had been previously used to code online discussions about Sony's Aibo [54].

Both of these coding schemas have been used in prior research projects. However, while the Barnstars data had only been used for related Barnstar data, the Alexa coding schema had been previously used for a Sony Aibo study and then modified to fit comments about Amazon Alexa. This may in turn impact how the coding schema is used within the ontology because the data being coded for the Alexa task was not naturally created for that data.

Both coding schemas were also used in different ways. The Alexa schema included two coding parts, first the researcher had to find the presence of a particular anthropomorphization characteristic and then either choose whether the comment yielded an affirm or negate statement. The Barnstar task presented a hierarchical coding schema. In

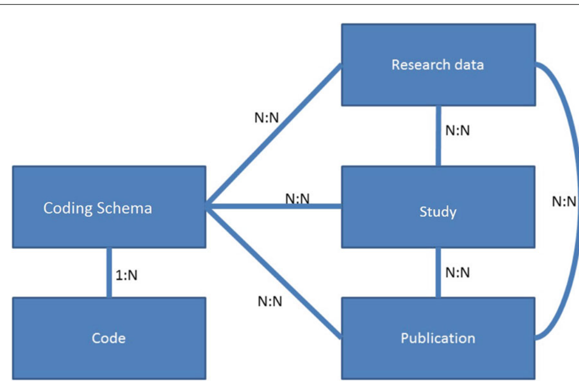


Fig. 1 Structure of the ontology

Barnstars - coding schema

Author	David W. McDonald
Research questions	How are Barnstars used in Wikipedia? How are people awarded with Barnstars?
Method	Qualitative codin
Theoretical background	The codebook contains three groups of codes: (1) the scope of the barnstar, such as whether it points to a specific edit or comments on general contributions to Wikipedia, (2) the genre of the barnstar, such as whether it is antagonistic, and (3) the work that the barnstar acknowledges.
Process of creation	codebook by open coding a random sample of 200 barnstars. We then applied the codebook to a second random sample of 200 barnstars. Through our attempt to systematically code the second barnstar sample, we iteratively refined our codes and agreed upon their application to the barnstars.
Dimensionalization and coding cycles	We used this codebook to iteratively code a random sample of 2400 barnstars. The barnstars were divided randomly into six bins. Two of the authors independently coded each bin. One coder reviewed the codes and noted discrepencies. Discrepencies were iterated upon until there was consensus. After coding was completed, we performed extensive consistency checking of commonly correlated codes. A barnstar often suggests multiple legitimate codings, either because a particular phrasing calls out multiple dimensions or because the barnstar contains multiple independent statements. We therefore chose to apply multiple codes for every code group instead of force-fitting a dominant code. For example, the following barnstar provides evidence of a prior encounter and serves as a peace offering (genre), while acknowledging general administratative actions, committment to the project, and conflict mediation (work)
Date	2008
Research area	Information Science
Inter-coder-reliability	85%
Keywords	Wikipedia, CSCW
Visualizations	
Software	NVivo
Coding schema	Datei:Codes.xml
Project data	Datei:Codes.xml
Research data	Barnstars - research data
Publication	Articulations of WikiWork, Finding Patterns in Behavioral Observations by Automatically Labeling Forms of Wikiwork in Barnstars
Study	Barnstar Study

Fig. 2 Overview of metadata for coding schema for the task "Barnstars"

the coding task, the researcher first chose an overarching code that represents the data and then chose subcodes that fit within the larger theme. The two coding schemas were distinct making them difficult to compare directly.

Both coding schemas were described in the prototype using QualiCO. The goal in the description was to provide as much information as possible. Therefore, in the level of publication, study, research data, project background and coding schema all metadata fields were filled out. There was also a graphical overview of the structure of the codes provided. On the coding schema for the Barnstars, for each top-level code a description and an example were provided. For the lower-level codes only the description was provided. The pictures of Figs. 2, 3, 4 and 5 show the instantiation:

	Definition	Example	Has Successor
Collaborative Action	Collaborative Actions and Disposition is differentiated from Social and Community Support Actions by the direct implication of collaborative activity, such as conflict mediation on talk pages. Although this category points to work activities that involve others, it is important to recognize that the dependencies are not clearly present in the texts of the barnstars	It is my honor to award Anonymous Editor this...green cucumber with sunglasses, for being cool when the editing isn't. For always keeping a level head. If you can keep your head when all around you are losing theirs and blaming it on you	Root
Meta Content	The last category is work related to meta-content. Meta-content work includes the acknowledgement of tool creation (programming), creation of templates, creation and management of categories or category tags, and work on formal Wikipedia policies.	I award you this fine barnstar for the work you did in improving the Taxobox template by adding conservation status shortcuts. I was thinking of doing that myself, but you beat me to the punch!	Root
Social/Community	This category includes welcoming newcomers, initiating or leading new projects, rewarding individuals who give out barnstars, and general social support. In the examples below we see individuals receiving barnstars for their personal characteristics and willingness to help other Wikipedians	I hereby ordain Andhrimnir to "The Order of the Smiley", for his dedication in helping newcomers like me. Without people like you, we would be lost. Awarded by Hermod	Root
Administrative	The barnstars in this category pertain to the actions taken by administrators and acknowledge participation in formal processes. Common processes include Editor Reviews and Featured Article Reviews	I hereby award you The Working Man's Barnstar for repeatedly notifying relevant parties during featured article reviews, as it's an oft-forgotten task.	Root
Editing	Acknowledgments for copy editing, general editing, and for contributions to specific articles	Mani, you have contributed a great deal of Estonian articles and done major and useful copyedits in a short time. You are a very productive user and deserve recognition. I award you, Magni, this Barnstar for keeping on top of this Wikipedia article and for being the "Master Editor". Keep up the fine work!	Root
Misc/Unknown	There are some barnstars that are not very specific about the actual work completed. Many such barnstars include references to "Janitorial Services" and "mop and bucket", reflecting a general work ethic of cleaning and maintaining various aspects of Wikipedia.	I, Gerd, award you this barnstar for your hard work here at Wikipedia, and to let you know I feel your wiki-stress. Gefion awards this Barnstar to Frigg for dedicated hard work with the mop and bucket, making Wikipedia better for everyone.	Root
Border Patrol	Detecting and disciplining such behavior constitutes a wide range of punitive work in Wikipedia. The most acknowledged border patrol activity is fighting vandalism by reverting or repairing the damage.	For standing with me this morning to revert all of those AOL vandals hell-bent on causing as much trouble as possible, I award you this barnstar.	Root

Fig. 3 Table representation of top-level codes in the task "Barnstars"

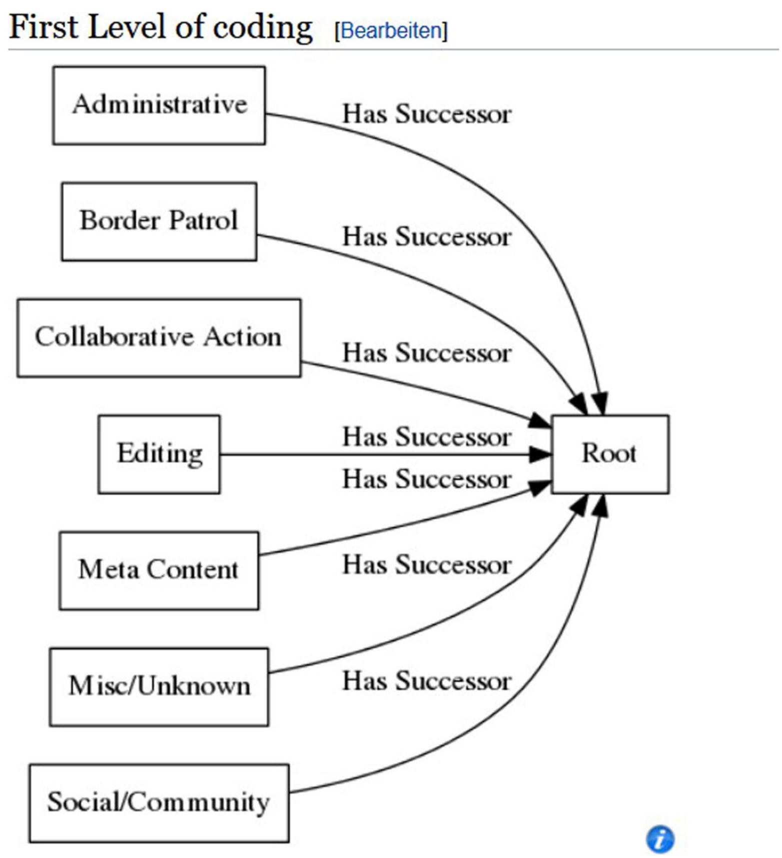


Fig. 4 Graphical representation of top-level codes in the task "Barnstars"

article editing

EG	general editing
EMAJ	contributions-major (MAJOR CHANGES TO TEXT; some kind of large refinement or improvement to the article)
EMED	contributions-media (images, audio, graphic design work, photos, latex)
EINI	initiative (starting an article, work on article stubs)
ERED	redesign / restructuring (refactoring, merging pages)
EMIN	minor (cleanup, copy-editing, maintenance, work on captions, work on a single article)
EACH	achievement (FA, FP, GA, DYK, readable)
ECLA	classification (categorization, disambiguating links, redirects)
EATT	attribution (adding references; removing unattributable information)
ETRA	translation (to another language)
EEXT	external research (doing outside research in order to make quality contributions)

collaborative actions/disposition (individual attribute)

CG	general collaborative
CDA	diplomatic-action (working toward concensus, mediation)
CDD	diplomatic-disposition (staying cool, civility, maintaining sense of humor, accepting of criticism, cooperative/collaborative)
CEXP	explanation (instance of explanation; to newcomers, of particular area of expertise, intricacy of policy)
CADH	adherence (to policy, maintaining NPOV, balancing policies, academic integrity)

meta content

MG	general meta
MPRO	programming artifacts (external tools, bots)
MCLA	classification (category creation, refactoring)
MPOL	policy / formal process (creation, refactoring, elaboration)
MTEM	template (creation, refactoring, barnstars, page templates, userboxes, DYK)
MFOR	forums / portals (creation, maintenance)
MARC	archiving (user talk pages, discussion pages)

Fig. 5 Detailed description of codes in the task "Barnstars"

Author	Taryn Bipat
Research questions	How do consumer's use the echo? Were expectations of using the Echo met? What terminology do consumers use to address the Alexa and the Echo?
Method	Coding of the data with pre-defined and refined coding schema
Theoretical background	[[Theoretischer Hintergrund::Coding Schema based on Coding schema for Hardware Companions; Value Sensitive Design (VSD) methods.]]
Process of creation	A code book was produced using the pilot data. This code book included explanations on how to code this corpus of data and the characteristics of each code that could be found in the qualitative data After the generation of the code book, the systematic codes were used to analyze the remainder of the verified product reviews. The coding was fixed in order to get correct percentages if a user used the same code multiple times. Many of the posts included more than one code and were coded accordingly so that each code was only counted once per participant.
Dimensionalization and coding cycles	[[Dimensionalisierung::The codes from Coding schema for Hardware Companions was adopted to the needs of the study.]]
Date	2018
Research area	Information Science
Inter-coder-reliability	85%
Keywords	
Visualizations	
Software	Atlas TI
Coding schema	Datei:Codes.xml
Project data	Datei:Codes.xml
Research data	Research data for Voice Conversational Agents
Publication	Voice Conversational Agents: A Value Sensitive Approach
Study	Voice conversational agents - study

Fig. 6 Overview of metadata for coding schema for the task "Alexa"

For the coding schema supporting the Alexa task, we provided a description of each code as well as examples. However, we did not provide examples for affirmation and negation for every code, though we did so for some of them. The graphics of Figs. 6, 7 and 8 show the instantiation of this coding schema. Compared with the Barnstar coding schema, the Alexa coding schema was more flat, which means that the participants had

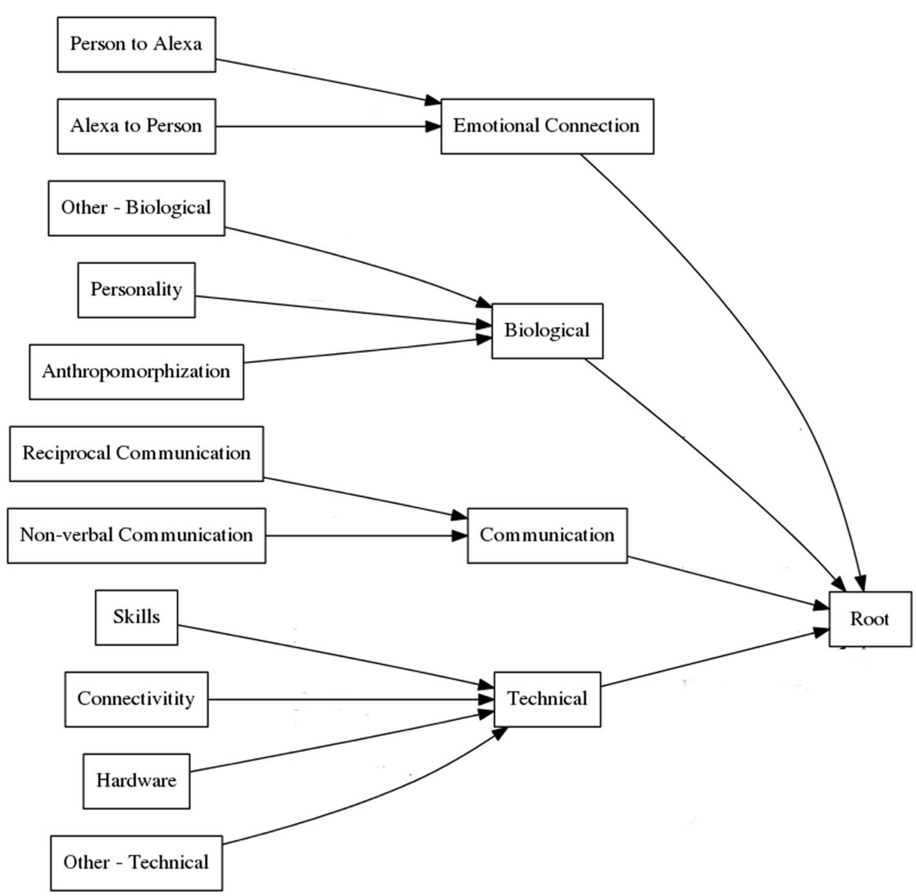


Fig. 7 Graphical representation of top-level codes in the task "Alexa"

	Definition	Example	Subcode of
Anthropomorphization	Are users attributing human characteristics to Alexa, especially in the way they address Alexa. Do they call the device her/she versus it/device/thing. (only what they are calling Alexa) Always have a code. Every line of code should have an antrophomorphization code.	Y: Alexa is overrated and she doesn't know anything you have to constantly repeat yourself. N: Gave this as a gift hopefully they are enjoying it. Y/N: I love playing with this device, she is so funny.	Biological
Personality	Are they describing the personality of Alexa? Is the device funny, sassy, etc..	I love speaking to Alexa, she is so funny. Sometimes she makes some snark remarks that I do not expect.	Biological
Other - Biological	Gives Alexa any other human traits that does not to seem to fit into any other human traits category		Biological
Reciprocal Communication	Any other issues or situations where the owners have conversations with Alexa.	We do have words; from time to time when I ask a simple question and get "I don't have an answer to that" or simple verbiage. After perhaps the third time of rephrasing the question, i may get the right answer.	Communication
Non-verbal Communication	The light at the top of the device, shutting down instead of giving a response.	Some times I know Alexa is responding to me or thinking about how to respond because the light at the top will be flashing.	Communication
Missing Alexa's company	The person using Alexa feels the absent of the device when they are not using it	When I am away from my home, I forget that Alexa is not there and I still try to talk to her	Companionship
Alexa as Family Member	Considers Alexa as part of the family	Had the Echo and Alexa in our lives almost a year now and she's like a member if the family.	Companionship
Alexa as Friend	Considers Alexa a friend	Alexa is my new bestie.	Companionship

Fig. 8 Detailed description of codes in the task "Alexa"

to keep all codes mentally active, while at the Barnstar coding schema, they only needed to have the top-level codes mentally active.

6 Methodology

Our goal was to understand potential user's experiences with this type of open data system, focusing on how potential users might see this type of system as part of their qualitative research process. We structured our evaluation process around a pair of qualitative coding tasks that would require that participants in our study explore and use the various portions of the system. This was used to test not only the ontology by itself, but rather test an application of the ontology. Since QualiCO will be used as an ontology for open science, we can follow the schema by [13] dividing the tasks in open science into three points: willingness to share, locating shared data, and the reuse of shared data. In this evaluation, we want to focus on the reusing of qualitative coding schemas, therefore we gave the participants tasks where they had to apply a coding schema to new data.

We recruited participants for our study through flyers, targeted mailing lists, and personal contacts at the University of Washington in Seattle. All of our participants had prior experience as a researcher or research assistant on a project that used systematic qualitative coding as part of the data analysis. While we recruited broadly at our university, all participants had a background in human-centered design or information science. Our participants spanned undergraduates, graduate students and professional research staff. Undergraduate student participants are characterized as beginners with only one hands-on project experience using systematic coding. The graduate students and professional researchers all had additional training and one or more project experiences. In total we had 20 participants complete our study; 4 undergraduates, 14 graduate students, and 2 professional research staff members.

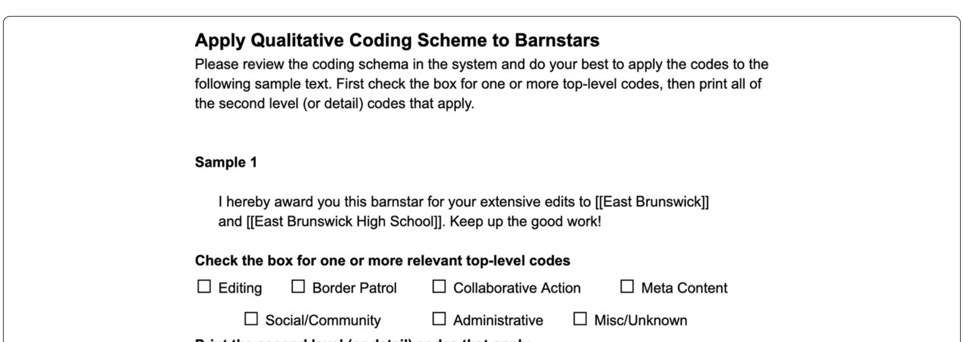
The study was structured in several distinct phases. The first phase included an introduction to the system. The introduction consisted of a comprehensive video tutorial that had been prepared to illustrate key aspects of the prototype and ontology. An 8 minute long video highlighted each area of the ontology, describing each area in detail and demonstrated how to navigate the prototype system. The participant was then asked to take a "system comprehension quiz," which gave the participants a chance to familiarize themselves with the system. Participants were told they could use the prototype to answer the questions in the quiz. After the participant completed the quiz, the researcher running the study went over the answers and clarified any answers that were wrong. The video tutorial and comprehension quiz helped ensure that all participants received a consistent presentation and that their basic understanding of the ontology and prototype were all at a similar level.

In the second phase, the participants were challenged to use the prototype to explore qualitative coding schemas and apply them to sample data. The prototype contained two research coding schemas for this phase of the experiment. The two schemas chosen were selected from several potential schemas that the research team had used on prior research projects. Both schemas had been applied to different data sets on at least two occasions. The versions of the coding schemas loaded into the prototype were simplified to facilitate a timely completion of the coding task, but still reflected the key analytical questions raised by the original schemas.

The schemas were applied to data as separate timed tasks and participants were allowed up to 25 minutes per task. The task order was counter-balanced to control for possible learning effects. In each task, the participant needed to identify and review one schema and apply the schema codes to four sample texts. The participant was given a paper based coding sheet with the four text samples and had to circle or write down the code(s) that best fit the sample text. Figure 9 shows one sample from one coding task. If a participant exceeded the allotted time they were to be stopped. In fact, only one participant was stopped and a majority of the participants completed each coding task well within the allotted time.

In the third phase of our evaluation, we conducted semi-structured interviews to obtain qualitative data about our participant's experiences with the prototype and their strategies for applying the schemas. These closing interviews lasted 10-20 min and were structured around the following questions:

1. Can you briefly describe your strategy for applying the codes?
2. What was different about your strategy between the first coding scheme and the second?
3. What questions did you have about the coding schemes that you could not answer with the system?
4. What aspect of the system was most helpful to your work to apply the codes?



Apply Qualitative Coding Scheme to Barnstars

Please review the coding schema in the system and do your best to apply the codes to the following sample text. First check the box for one or more top-level codes, then print all of the second level (or detail) codes that apply.

Sample 1

I hereby award you this barnstar for your extensive edits to [[East Brunswick]] and [[East Brunswick High School]]. Keep up the good work!

Check the box for one or more relevant top-level codes

☐ Editing ☐ Border Patrol ☐ Collaborative Action ☐ Meta Content

☐ Social/Community ☐ Administrative ☐ Misc/Unknown

Print the second level (or detail) codes that apply

Fig. 9 Example of a coding task. Participants were given instructions and two samples before beginning the task. They had the ability to freely use the system to complete all tasks

5 What meta-data do you feel is missing from the system that would be helpful for these tasks?
6 Is there anything else you would like to share about your experience with the system?
7 Do you have any questions for us about this system or this study?

The interviews were recorded and transcribed. We then conducted a thematic analysis on the transcripts. The research team met weekly to extract and refine the themes in the transcripts. The main findings from each participant were listed and then reviewed by all authors of the paper. The findings reflect common themes. In a prior publication [29] we focused only on the qualitative analysis on the interviews using the themes that related to the open science framework by Birnholtz & Bietz [13]. This analysis focuses on the task performance characteristics and their specific strategies for using schemas represented within this ontology.

7 Findings

Our findings are based on the quantitative and qualitative data that we collected during our structured trials. All of our participants should be considered experienced qualitative researchers. In particular, all met the condition that they had participated in at least one research project where they had performed a systematic qualitative coding. Many of our participants also participated in the development of one or more coding schemas during research projects. Four participants in the study met only the minimum threshold; the other 16 participants had more experience. Given our small number of participants who might be considered "novice" we do not breakout the results of the study by experience levels.

We structure the results of our study based on the performance of the participants in the two coding tasks. We provide both quantitative and qualitative data to demonstrate participants performance using the prototype and their perceptions of the ontology. The quantitative results demonstrate the time taken on each coding task and their accuracy at the coding task. Our qualitative results illustrate some of the users' perceptions of the prototype and how they approached the different coding tasks. Before we dive into the findings, we briefly explain some distinctions between our specific experiment and the way most qualitative coding tasks are performed.

7.1 Qualitative coding in an experimental setting - some caveats

In this study, we asked participants to participate in two coding tasks using schemas that were instantiated into our prototype. Participants were given a 25 minute time limit to use the prototype to code four data samples. Conducting the experiment in this way puts the prototype, the ontology, and our participants at a disadvantage. Like many experimental conditions, these conditions do not accurately mirror the way that qualitative coding is performed.

First, it is quite rare that one would simply pick up a qualitative coding schema and immediately start coding data. Often there is some time spent with the coding schema to try and understand something of what is to be coded, code inclusion or exclusion criteria, and boundary conditions between codes. As well, in many cases there is often someone to talk with about the codes; an advisor, collaborators, a trusted colleague. Even tangential discussions about a qualitative phenomena can provide insight or clarity to the concept. Our experiment did not provide for any form of in-situ clarification.

Second, it is fairly rare for researchers to apply two distinct coding schemas back-to-back. Certainly, it can be the case that research supervisors may work with different individuals or teams who are applying different coding schemas. But in those cases a research supervisor is not likely to be applying the codes individually and challenged for their accuracy in code application. Applying two distinct qualitative coding schemas is a form of "switching gears" that is somewhat rare.

Perhaps a third distinction is that applying a coding schema as a timed task is quite uncommon. Most often individuals are allowed to compare, consider and reflect on the way that the codes are applied. This can include multiple passes through the data. We set a time limit of 25 min for exploring and applying each of the two schemas to sample data. While there was only one participant who "timed out", making the 25 min limit appear reasonable, coding as a timed task, with no opportunity to return to the data, is not a realistic qualitative coding practice.

We state these caveats here, upfront, because the disadvantages of this experimental condition may not be well understood by readers who are not familiar with qualitative coding. An alternative evaluation method, such as a full field deployment, could mitigate these disadvantages, but would introduce another set of complex evaluation challenges. In this work we focused on a more experimental approach, while recognizing the implications for qualitative practice as mentioned above.

7.2 Time on task

During the study, each participant was asked to complete two coding tasks using two different schemas. These tasks were counterbalanced across the participants to control for potential learning effects, Table 1 shows these results. Participants were given 25 min to complete each task. The table below presents the time on task data, measured to the minute. No participant used more than 25 min. The findings show that the Alexa coding task took approximately 2.5 min more to complete than the Barnstar task. Uniformly across both tasks, time on task was reduced from the first to second coding task.

We performed a t-test to determine whether the decrease in time from first to second coding tasks were significantly different. While the tests were not significant at $p < 0.05$ levels, they were close. In particular, we point out that there were only 10 participants in each condition, making the likelihood of significance low.

7.3 Task accuracy

We use a formula to calculate an accuracy score for each coding task. The accuracy scores were calculated using the formula below.

$$AccuracyScore = (CodesApplied - CodesCorrect) + CodesOmitted$$

Table 1 Time on task. The table indicates differences whether a task was taken first or second

	All	First	Second
Barnstar	10.1	10.9	9.3
Alexa	12.45	13.4	11.5

We calculated each participants accuracy score as the sum of applying too many codes and omitted codes. We subtracted the number of correct codes from the number of codes applied and added the codes that were completely missed. A participant with an accuracy score of 0 (zero) would have a perfect score, indicating complete agreement with the prior coding of the sample data.

Our findings, shown in Table 2, reflect an increased number of incorrect responses in the Alexa coding task compared to the Barnstars coding task. This implies that the Alexa task where participants had to first select if a code applies and afterwards do a rating about affirmation or negation, was harder than the Barnstars task, which was a straightforward two-level coding schema task.

Uniformly, participants' accuracy improved from the first task to the second task (lower numbers for accuracy scores are better). This result, in conjunction with the improved times, suggest that the ontology and the specific prototype are not impeding the performance of the coding tasks. While the counterbalanced experimental design cannot completely eliminate all possible learning effects, that both time and accuracy improved suggest that QualiCO provided benefits for the coding tasks. Next, we consider insights from the semi-structured interviews.

7.4 The need for additional metadata

In the post-study interview, participants were asked if they felt that any metadata was missing from this instantiation of the ontology. Most participants felt that the metadata presented covered most of the information needed to complete the coding tasks. However, they did comment on a few additional needs that could help address the understandability and discoverability of the coding schemas.

Participants asked for additional examples of the codes with explanation of how the codes were applied to a specific example. P16 had difficulty understanding the examples for the Alexa coding task:

> *For the Alexa one, it would have been nice to have a positive and negative example, I don't know if that would have been a lot of work. Sometimes whenever I was reading the examples I had to decide what was the negative of that or the positive of that, so I feel like if I had an example of that, it would be helpful. Other metadata (scrolls through system) I cannot think of any, I think it was fairly straight forward.* (P16)

Furthermore, participants wanted additional metadata to support their understanding of the coding schema. Specifically, they wanted additional contextual knowledge to better apply the codes. P3 asked for more domain specific knowledge related to Wikipedia because they did not understand the domain-specific language for the Barnstars coding task:

Table 2 Performance on each task based on whether it was taken first or second

	Alexa			**Barnstar**		
	First	**Second**	***p*-value**	**First**	**Second**	***p*-value**
Sample 1	9	3	0.03	1.1	0.8	0.26
Sample 2	8.6	3.9	0.07	1.2	1	0.39
Sample 3	7.65	5.2	0.19	3.4	2.5	0.09
Sample 4	9.1	4.9	0.11	5.1	4.8	0.35

> *"In some cases, especially for the Barnstars example, there is a lot of specific domain related knowledge that I did not have related to what people do on Wikipedia. It would be helpful to have more explicit things to onboard a research. There was some stuff of the structure of the data but more onboarding to explain how to use the coding schema.(P3)*

The Barnstars coding example had our study participants code the comments that Wikipedia editors gave to justify giving an award to another editor. These editors sometimes use specific language related to Wikipedia behavior that may not be completely understood by readers who do not have experience in editing Wikipedia. Related to this issue of contextual familiarity, P5 noted that the language from the Alexa conversation task had language that was more clear and familiar to them:

> *"For the second task, I was more thorough going through because these are more straightforward language than acronyms. It helped me to go through and say yes/yes/no. If I did not know then I would go back and re-read. It was easier for me because in the Alexa one the language was more clear and it was helpful cause all of it was right there."* (P5)

P19 gives an example from their own profession that demonstrates the need for contextualization of the coding schema. P19 discussed their own field site of oceanographers. In particular, they noted that oceanographers have different subdomains so it is important to understand how a person interprets and is connected to the data:

> *"I think having some other things like having a connection to the person who did the analysis, is something I want to see. But if the same person who collected the data is also doing the coding and the analysis and stuff, then obviously you are aware of all that stuff"* (P19)

Another issue is information about how to apply a code. In most academic research papers, codes are explained but not necessarily with details about the process of applying the codes. Participants requested this additional metadata:

> *"The Alexa one I did not understand as well when you described it to me at first. I heard that you said that it was positive and negative examples but with each one having a yes/no or I felt like I had to make a call on all of them but then I realized that was not how it was. I think that is the most frequent error people make."* (P2)

Similarly to P2, participants asked for more explicit instructions. One participant noted that the Alexa coding task needed additional information:

> *"There were a couple of questions in the Alexa task because the definition felt like they had examples but I was not sure if this falls within the definition or not within the specific code. So some codes were a little bit more confusing. I wish there was more explicit instructions."* (P3)

In our coding tasks, we did provide participants with instructions on how to complete the coding tasks external to the ontology. However, this participant requested additional instructions on how to apply the coding schema within the prototype.

Overall, participants felt that there was enough metadata to support the coding tasks but some of the current metadata was not needed for the coding tasks. P5 noted that consolidating some of the current information in the ontology might be more beneficial to their coding process:

> *"I felt like I was clicking through a bunch of things. There is an opportunity to consolidate the data. The bit I care about is to code so really interested in only the coding information."(P5)*

7.5 Switching strategies between the first and second coding task

Switching between two unrelated coding tasks also presented some challenges for participants. In our prototype, the coding schemas were presented in two ways: 1) using a table with a list of the codes, and 2) a visualization to show the relationship between the codes. During the tasks, participants used different pieces of the system to understand what was happening. Additionally, the coding schemas included detailed descriptions and a lot of metadata. Moving back and forth between the coding schemas presented a challenge.

Most participants mentioned that their coding strategy was similar from the first task to the second. In the two coding tasks there were two different ways to apply the coding schema. The Barnstars coding schema had a hierarchical structure, where the participant first had to choose an overarching theme and then choose sub-codes that could fit under that theme. In the Alexa coding task, the participant had to first find if the anthropomorphic character existed and then had to choose for every code whether it was presented in a positive or negative way in the comment. These differences in reusing the coding schema led to minor changes in the strategy used by the participants to complete the coding task.

Some participants noted that it was easier to remember the Alexa coding scheme since it did not have a two layer structure. P7 completed the Barnstars coding task first. They noted their strategic changes helped handle the differences between coding schemas, especially since the Alexa coding task required not just the presence of a code but also if it was a positive or negative representation:

> *"A little bit, in the sense that the tasks were different. Here [Barnstars] I was just seeing if the code was present or not, in the second one [Alexa] I was seeing a positive or negative reaction. In the first one I was just trying to see if a code was present. In the second one, I was going by the category after having read the quote but in the other one I was going through them"* (P7)

In constrast, P15 noted that the Alexa coding required much more processing since there were more codes and layer, that require more maneuvering back and forth from the prototype to the coding questions. P15 completed the Barnstars task first:

> *The first one [Barnstars] I think is really easy to remember all the codes and all the codes in the coding schema like two tasks. However, the second one [Alexa] I needed to go back to see the codes every time. Because it was it was more complex and it was super long.* (P15)

P12 completed the Alexa task first and noted that the Barnstars coding task required more work since they had to work through 10 different codes for a layer of the coding task but for the first layer of Alexa, there was only 2 choices, positive or negative:

> *In the first coding schema [Alexa] there was nesting in the sense of you have to identify whether it affirms or negates, which is a very constrained type of coding scheme in a way. But it was one who did not require reference back to Alexa coding schema. If I identify that something is anthropormophized or making a statement about anthropomorphization, whether that is affirmative or negative is not a thing that I have to got often to check in the coding scheme. There is only two options in my head. If I determine something that is an edit, it could be any number of up to ten subcodes, so that is a thing where I was slower and I had to get back and take a look at them.* (P12)

Their strategy was also highly dependent on how much knowledge they already had on the particular topic or research. P16, completed the Barnstars task first, mentioned that their strategy changed because they had better understanding of one coding task over the other:

> *Yeah, I think so. Maybe because. One of this had more codes and the other one was more straightforward and I think I could have coded it without necessarily having to read all the explanations in depth, but this one was more nuanced, I had to understand the difference between connectivity and skill.* (P16)

P16 spent more time analyzing the codes for the Alexa coding task because they needed to understand context around the device. For the Barnstars task, it was not necessary to read all of the detailed explanations of codes but for the Alexa coding task it was needed to complete the task.

P17 completed the Barnstars task first and also noted that he needed to understand more about the research study in the Barnstars coding task which required him to look deeper into the research question for that particular coding task:

> *" I wanted to get a better understanding what the study was about and why - what - So I guess when we are talking about social vs. collaborative, is this a paper that is looking at collaboration or is it more looking at complementing somebody or describing value to certain behavior or something. So specifically that required me to go beyond just the coding and actually go to the research question."* (P17)

7.6 Understanding how the ontology fits in a research workflow

After testing the prototype through two coding tasks, we asked the participants questions about problems they had with the ontology and QualiCO. The experimental environment of the ontology led to difficulties in the coding task. Participants reported obstacles completing both qualitative coding tasks using the ontology. There was no clear distinction between either task related to their preference for completing the tasks. Six participants preferred the Alexa task, five preferred the Barnstars task and the remaining participants did not declare a preference.

One of the biggest challenges participants faced was understanding the purpose of QualiCO. Most of the concerns about the system centered around the process of coding rather than actually reusing the coding schema. Participants faced tensions about whether the prototype/ontology are to help build/develop the coding process or whether it is a place (repository) for getting/reusing some of the code scheme. Participants were interested more in "how to do it now" and less in using it as a platform to exchange coding schemas.

The prototype system we developed only allowed for a concrete representation of the final codes. Participants noted that creating a coding schema is a dynamic process and codes can shift. Currently, our ontology does not accommodate that need:

> *"Once you have a project that is mature, you have a stable coding schema but it takes a great deal of work to get up to that point and what we really have not seen or tested is the extent which it can support a living document that has this kind of history of codes that get subdivided or merged or codes that definitions had.* (P10)

P11 also similarly asked for more detailed information that demonstrates the multiple stages of the coding process and the researchers involved. They noted that the interpretation of a coding schema might be different and it is important for someone using a particular coding schema to understand its historical lineage.

> *Some of these have very lose or limited information about the development process. You know, sometimes it is just a couple of words. A lot of times it is helpful to get more detailed information about that. So it is not the kind of thing where you can force people to give extremely detailed information, but if I am being involved I do wanna know - you know - did they go through multiple stages of affinity diagramming? Did they - when they say they started with the data, where they borrowing anybody else's schema to begin with? You want a lot of rich content when you approach these things and there is only so much information you can present in a table like this. And you want your interpretations of the codes to be appropriately in line with what the original creators.. And if you find problems, another thing, I don't know if this is built-in, but a kind of feedback or comment process, if you find issues with the schema, like "I think it would be helpful to merge these or to add this other one. That could be helpful, if it is gonna be a collaborative process.* (P11)

This quote from P11 shows that the participant is thinking of the prototype and ontology more as a tool for the coding process. Being able to integrate more of the coding process into the prototype would also be helpful for qualitative researchers to share their codes across a team. However, the QualiCO system was not developed to create coding schemas but rather to store coding schemas that had already been developed and used in prior research.

In the following section, we will connect the summary of our findings to our initial research questions and demonstrate how these findings have implications for the future development of QualiCO.

8 Answering the research questions

For the research questions, we can state the following:

RQ1: *How does researcher performance change by using the ontology across different coding schemas?* Our evaluation showed that participants who had not worked with the coding schemas before were able to use the coding schemas in an assigned coding task. Improved time on task as well as improved accuracy suggest that QualiCO helped the participants to reuse the coding schemas in our coding task.

RQ2: *Does a nearly complete instantiation of the ontology provide enough metadata for a researcher to reliably apply the coding schema?* When we asked the participants whether information was missing, they only mentioned the usage of a code, more examples and versioning of the coding schemas. Apart from this, there were no issues mentioned. We conclude that a relatively full instantiation is sufficient to represent schemas for potential reuse.

RQ3: *How can this ontology be improved to make the reuse of coding schemas easier?* We saw that many participants had issues with understanding special domain knowledge related to Wikipedia or Amazon's Alexa system. Participants were also interested in richer data about positive (application) and negative (omit) coding examples, and information about code frequency. We note that some of this information is commonly in publications that detail a qualitative coding research result.

Participants also mentioned that they were interested in understanding schema changes, potentially as part of the schema development, and tracking how schemas are reused. These are important aspects of qualitative schema sharing, but were outside the scope of our initial work to develop QualiCO. A future prototype that leveraged the QualiCO ontology could also implement versioning features, perhaps like GIT, allowing researchers to have a versioning structure and facilitating the possibility of forks when coding schemas are reused by other researchers. But we note these features would not require changes to the QualiCO ontology itself.

9 Discussion

In this section, we discuss the findings of our evaluation in detail and next what we learned from applying our method to an evaluation of the ontology. The section closes with the limitations of the research.

9.1 Discussion of findings

Next we reflect on our findings. We present some design recommendations and future directions for integrating these types of infrastructures into current approaches to fostering open science. Our results illustrate the potential for such platforms to give researchers new perspectives on coding schemas, better access to qualitative methods, better documentation for schema reuse, as well as potential support for team coding and improved traceability for qualitative research.

Directly comparing the Alexa and Barnstar coding tasks is not really possible. They were adopted as our experimental tasks because we had access to previously coded data that had been consensus coded providing a highly reliable standard for the correct codes. The tasks had distinctly different structures that aligned with the underlying conceptual models of our different participants. This is clear from the variety of preference statements we got from participants with no clear preference for one task over the other.

However, it is clear from our quantitative performance metrics that the Alexa task was "harder" in some sense. That is, regardless of whether the Alexa task was first or second, it took the participants more time and they were less accurate. We believe this is mostly a function of the constrained conditions of the experiment rather than a reflection of relative qualitative validity differences between the tasks.

Rather than comparing the specific tasks, we need to compare the experimental conditions. Looking at the time on task, we can see that the participants consistently performed faster in the second task, it did not matter whether this was the Barnstar or the Alexa task. While the faster performance was not statistically significant, we also note that the task accuracy metric shows a similar improvement trajectory. That is, our participants improved their coding accuracy regardless of which task was second. Again, overall, these accuracy improvements were not statistically significant. The combination of these two metrics, in the same positive direction, with the associated counter balanced control, is highly suggestive that the prototype and the ontology has some positive impact in support of reuse. If the prototype and QualiCO ontology were too complex, one or both of these metrics might be otherwise oriented.

Our qualitative interview data revealed a number of interesting aspects with regard to the specifics of the tasks, aspects of the QualiCO ontology, and, perhaps most interestingly, how the participants conceptualized the prototype.

One issue with the experimental tasks was that both of them required some domain specific knowledge to be performed well. Differences in this domain knowledge might explain why some participants preferred one task over the other. This shows that our instantiation of the coding schemas had some metadata gaps. For example, including metadata on specific publications that used the schema and metadata on the research data, sources and context, would be helpful. This also illustrates a specific challenge when using a task-based evaluation of an ontology system. The participants familiarity with the task domain influences the outcomes for the ontology system.

Embedded in some of the quotes from our participants are underlying assumptions about what our prototype tool was designed to achieve. That is, despite our efforts to focus the participants on the underlying ontology, what it represented and what it contained, participants still answered questions by "framing" their answers in their experiences of prior qualitative coding work similar to the phenomena of technological perceptions described in [55]. Many participants framed the platform as a tool that they could use for conducting qualitative research rather than for sharing qualitative schemas or finding an open schema on which to build.

One frame was to think of our prototype as a tool for conducting their qualitative coding activity. Some of our participants had used commercial QDA-software before, while others only had experience coding through spreadsheets. We made an explicit attempt to discourage this framing by having the participants specifically code on paper based coding sheets. With this frame, the prototype and the QualiCO ontology provided access to the schema and some examples, but did not provide a way to apply codes to data. And, during interviews participants noted that they probably would not use the tool for conducting their coding. This does not actually illustrate a flaw with the QualiCO ontology. It does illustrate some conceptual barriers for incorporating open science practice among research disciplines that leverage qualitative methods.

A similar framing of the tool is as a collaboration platform. This framing is closer to how the prototype was explained. Finding a shared coding schema or sharing a coding schema is a type of collaboration. That the participants would see collaboration as present in the prototype seems understandable. However, this was still framed through the participants' experiences of practice. Several thought of the prototype as a potential tool for collaboratively developing a schema. This is evident in interview answers where the participants mention refinement or versioning of the individual codes. Some participants were framing the tool as a mechanism for tracking the work of developing coding schemas. This framing has a potential positive side effect. If a tool facilitated the development of the schema, then there would be little effort necessary to share that final version.

The prototype we tested explained a model where researchers upload and publish their coding schema after the development of the coding schema and associated research was completed.

We saw another important conceptual framing of the prototype that may not be obvious. The experimental tasks were specifically a methods task; applying codes to data. However, we noticed that some participants thought the prototype was more about sharing data than the sharing of method. The rhetoric and practices around open science has been focused on data sharing as that perspective is dominant in much of the prior literature. The participants in our study are at least tacitly aware of this. Several participants stuck to this "open science is data sharing" frame of reference even when we asked them about sharing methods. This is not explicitly a flaw in QualiCO nor our research design. This does illustrate that among qualitative research practitioners that constituents of open science are shaped and framed by the popular or dominant conversations.

For research in open science, our study shows that reuse of coding schemas is possible given reasonably rich metadata for the schemas. QualiCO bridges an important gap between pure archiving of coding schemas and allowing for the sharing of methods within qualitative research. Professional science societies have called to fill this gap (e.g., APA guidelines [38]) through data and method sharing that aligns with the information researchers would need to provide QualiCO. QualiCO could help address the calls of these professional societies. Still, documenting qualitative practices and creating incentive structures to publicly share the details of methods and schemas will require significant community effort [29].

9.2 Usability testing ontologies

In Section 4.3 we described several approaches to evaluating ontologies. We conducted a task-based evaluation that focused the participants on a methods reuse problem. Given the range of methods that fall under the umbrella of qualitative methods, our approach might not be the only strategy to test an ontology. However, prior work shows that it is hard to test an ontology without its instantiation in a prototype. Our specific approach attempted to focus participants on the QualiCO ontology and background the prototype as much as possible. Based on the interviews, we believe that our approach worked. However, as we point out above, there can still be conceptual issues that influence the participants' responses and that shift the focus of the participant back on the prototype software. Part of the success may have been the choice of using a wiki (i.e., Semantic MediaWiki) for the prototype implementation. In general, the wiki provides a minimal interface and the QualiCO ontology was therefore rather explicit in the linking structure that connected the metadata for the tasks.

We claim that ontologies like QualiCO can and should be tested against real-world tasks, especially when they describe data or methods that are meant to be reused. However, we recognize that there are important methodological issues in these evaluations that are not simple to resolve. For example, our specific experimental design put QualiCO at a disadvantage in comparison to real-world coding practices. Given the time limits, participants were not able to read prior publications, talk to the researchers who developed the coding schema or follow up on task domain knowledge that might have been helpful. Some participants explicitly mentioned these challenges as distinct from their standard practice. Still, our participants were able to complete the assigned tasks, suggesting that the metadata supplied was at least minimally sufficient, and demonstrating that a real-world task-based evaluation is reasonable.

9.3 Limitations

We note that there are several limitations to this research. First, we have focused on one particular qualitative research method. We focused on systematic qualitative coding because it occupies a philosophical space of research that is relatively close to the positivist stance of most research methods that have been the focus of the majority of prior research in open science. Evaluating whether QualiCO can address other methods in the qualitative tradition is open for future research.

As in many experiments, the experimental conditions do not exactly mirror the real-world conditions in which this type of qualitative coding activity would happen. However, we believe that our experimental conditions create a type of deficit for the ontology that is not explained away through experimental demand characteristics nor simple learning. While we do not have statistical significance in the experiment, we believe the direction of the implications point to some benefits derived from QualiCO.

Our evaluation relied on one specific prototype that represented the ontology in one specific way. One can imagine that other prototyping techniques or a full-blown application might reflect the QualiCO ontology in a slightly different way. Focusing on just one prototype was important to begin a systematic evaluation. Future evaluations might approach the implementation of QualiCO in a prototype differently.

Another specific limitation is how we instantiated the ontology to reflect the two experimental conditions. In this work we specifically biased toward providing rich metadata for the schemas. We filled as many of the QualiCO fields and relations as possible. This reflects a particular baseline. An alternative experiment could vary the amount of metadata for each experimental schema. For example, it might have been interesting to see at what minimum of metadata the prototype becomes unusable. However, we believe that would not be helpful as a first evaluation to illustrate basic utility of the ontology and prototype. Our focus was to evaluate the ontology itself and not some range of instantiations of the ontology.

Lastly, our participants largely came from one disciplinary perspective - Human Centered Design. While we solicited in several different academic and research units, we were not able to garner broad participation across those units. We believe that systematic qualitative coding is largely taught and practiced in similar ways across a number of disciplines, but it is possible that there are factors among the set of participants that generated our results.

10 Conclusion

In this evaluation we tested the QualiCO ontology. Our goal was to understand whether QualiCO could be used to understand a previously unknown systematic qualitative coding schema. We tested this by creating two coding tasks and having participants attempt to code sample data using schemas entered into a prototype system.

QualiCO was designed to bridge the gap between simply archiving coding schemas to sharing of supplemental information to research, like research data. With our prototype implementation, we show the potential for method sharing in qualitative research using

QualiCO. A description of a coding schema using QualiCO enables researchers to share their coding schemas and therefore also their methods.

The next steps will be the implementation of QualiCO within research data centers. The complete ontology can be found on GitHub[6], anyone interested in implementing the ontology is welcome to get in touch with us.

Acknowledgements
The authors would like to thank Mark Rittberger, Christoph Schindler, Christa Womser-Hacker and Thomas Mandl for providing feedback throughout the whole research project. We also want to thank DIPF | Leibniz Institute for Research and Information in Education for supporting Julian's stay at University of Washington. We also want to thank our study participants.

Authors' contributions
The authors read and approved the final manuscript.

Authors' information
Not applicable.

Funding
Open Access funding enabled and organized by Projekt DEAL.

Availability of data and materials
The datasets supporting the conclusions of this article are available in the Zenodo repository, https://doi.org/10.5281/zenodo.4442523. In the repository you can find all raw data about the performance on the tasks, including the calculation on how we rated the tasks as well as t-tests and time on task. In the repository, you can also find the system comprehension quiz, task description and interview guidelines we used to conduct the tests. In our Github repository, https://github.com/julianhocker/Quali-Codes-Ontology, you can find the complete description of QualiCO. A representation of QualiCO as OWL can be found at https://github.com/julianhocker/Quali-Codes-Ontology/blob/master/QualiCO.owl

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Author details
[1]DIPF | Leibniz Institute for Research and Information in Education, Frankfurt am Main, Germany. [2]Human-Centered Design & Engineering, University of Washington, Washington, US.

Received: 15 January 2021 Accepted: 8 July 2021
Published online: 09 August 2021

References
1. Google. Google Books Ngram Viewer. Search Term: "Open Science". 2020. https://books.google.com/ngrams/graph?content=open+science&case_insensitive=on&year_start=1800&year_end=2008&corpus=15&smoothing=3&share=&direct_url=t4%3B%2Copen%20science%3B%2Cc0%3B%2Cs0%3B%3Bopen%20science%3B%2Cc0%3B%3BOpen%20Science%3B%2Cc0%3B%3Bopen%20Science%3B%2Cc0%3B%3BOpen%20science%3B%2Cc0#t4%3B%2Copen%20science%3B%2Cc0%3B%2Cs0%3B%3Bopen%20science%3B%2Cc0%3B%3BOpen%20Science%3B%2Cc0%3B%3Bopen%20Science%3B%2Cc0%3B%3BOpen%20science%3B%2Cc0. Accessed: 26 July 2021.
2. Collaboration OS, et al. Estimating the reproducibility of psychological science. Science. 2015;349(6251):4716.
3. Fischer BA, Zigmond MJ. The essential nature of sharing in science. Sci Eng Ethics. 2010;16(4):783–99.
4. Louis KS, Jones LM, Campbell EG. Macroscope: sharing in science. Am Sci. 2002;90(4):304–7.
5. Madill A. Qualitative research is not a paradigm: Commentary on Jackson (2015) and Landrum and Garza (2015). Qual Psychol. 2015;2(2):214–20. https://doi.org/10.1037/qup0000032.
6. DuBois JM, Strait M, Walsh H. Is it time to share qualitative research data? Qual Psychol. 2018;5(3):380.
7. Bücker N. How to code your qualitative data—a comparison between grounded theory methodology and qualitative content analysis. In: Forum Qualitative Sozialforschung/Forum: Qualitative Social Research, vol. 21; 2020.
8. Hocker J, Schindler C, Rittberger M. Participatory design for ontologies. a case study of an open science ontology for qualitative coding schemas. Aslib J Inf Manag. 2020;72(4):671–85. in print.

[6]https://github.com/julianhocker/Quali-Codes-Ontology

9. Huang X, Hawkins BA, Lei F, Miller GL, Favret C, Zhang R, Qiao G. Willing or unwilling to share primary biodiversity data: results and implications of an international survey. Conserv Lett. 2012;5(5):399–406. https://doi.org/10.1111/j.1755-263X.2012.00259.x. https://conbio.onlinelibrary.wiley.com/doi/pdf/10.1111/j.1755-263X.2012.00259.x.
10. Melero R, Navarro-Molina C. Researchers' attitudes and perceptions towards data sharing and data reuse in the field of food science and technology. Learned Publ. 2020;33(2):163–79. https://doi.org/10.1002/leap.1287. https://onlinelibrary.wiley.com/doi/pdf/10.1002/leap.1287.
11. Kim Y, Nah S. Internet researchers' data sharing behaviors. Online Inf Rev. 2018;42(1):124–42.
12. Tenopir C, Christian L, Allard S, Borycz J. Research data sharing: Practices and attitudes of geophysicists. Earth Space Sci. 2018;5(12):891–902. https://doi.org/10.1029/2018EA000461 https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2018EA000461.
13. Birnholtz JP, Bietz MJ. Data at work: supporting sharing in science and engineering. In: Proceedings of the 2003 International ACM SIGGROUP Conference on Supporting Group Work. New York: Association for Computing Machinery; 2003. p. 339–48.
14. Rodrigues J, Castro JA, da Silva JR, Ribeiro C. Hands-on data publishing with researchers: Five experiments with metadata in multiple domains. In: Italian Research Conference on Digital Libraries. Cham: Springer; 2019. p. 274–88.
15. Wilkinson MD, Dumontier M, Aalbersberg IJ, Appleton G, Axton M, Baak A, Blomberg N, Boiten J-W, da Silva Santos LB, Bourne PE, et al. The fair guiding principles for scientific data management and stewardship. Sci Data. 2016;3(1):1–9.
16. Alliance GID. Care principles for indigenous data governance. 2019. GIDA https://www.gida-global.org/care. Accessed: 26 July 2021.
17. Kowalczyk S, Shankar K. Data sharing in the sciences. Ann Rev Inf Sci Technol. 2011;45(1):247–94.
18. Curty RG. Factors influencing research data reuse in the social sciences: An exploratory study. Int J Digit Curation. 2016. https://doi.org/10.2218/ijdc.v11i1.401.
19. Faniel IM, Kriesberg A, Yakel E. Social scientists' satisfaction with data reuse. J Assoc Inf Sci Technol. 2016;67(6): 1404–16.
20. Studer R, Benjamins VR, Fensel D. Knowledge engineering: principles and methods. Data Knowl Eng. 1998;25(1-2): 161–97.
21. Borgman CL. Big Data, Little Data, No Data: Scholarship in the Networked World. Amsterdam: Elsevier; 2015.
22. Heidorn PB. Shedding light on the dark data in the long tail of science. Libr Trends. 2008;57(2):280–99.
23. Mays N, Pope C. Quality in Qualitative Research: Wiley; 2020, pp. 211–33. Chap. 15. https://doi.org/10.1002/9781119410867.ch15. https://onlinelibrary.wiley.com/doi/abs/10.1002/9781119410867.ch15.
24. LONG AF, GODFREY M. An evaluation tool to assess the quality of qualitative research studies. Int J Soc Res Methodol. 2004;7(2):181–96. https://doi.org/10.1080/1364557032000045302.
25. Glaser BG, Strauss AL. Discovery of Grounded Theory: Strategies for Qualitative Research. Miltion Park, United Kingdom: Routledge; 2017.
26. Saldaña J. The Coding Manual for Qualitative Researchers. Los Angeles: Sage; 2015.
27. Mayring P, Brunner E. Qualitative inhaltsanalyse. In: Friebertshäuser B, Langer A, Prengel A, editors. Handbuch Qualitative Forschungsmethoden in der Erziehungswissenschaft. Juventa-Verl., Weinheim; 2013. p. 323–34.
28. Schreier M. Qualitative Content Analysis in Practice. Los Angeles: Sage Publications; 2012.
29. Hocker J, Bipat T, Zachry M, McDonald DW. Sharing your coding schemas: Developing a platform to fit within the qualitative research workflow. In: Proceedings of the 16th International Symposium on Open Collaboration. New York: Association for Computing Machinery; 2020. p. 1–10.
30. Vegas S, Juristo N, Moreno A, Solari M, Letelier P. Analysis of the influence of communication between researchers on experiment replication. In: Proceedings of the 2006 ACM/IEEE International Symposium on Empirical Software Engineering, ISESE '06. New York: Association for Computing Machinery; 2006. p. 28–37. https://doi.org/10.1145/1159733.1159741.
31. González-Barahona JM, Robles G. On the reproducibility of empirical software engineering studies based on data retrieved from development repositories. Empir Softw Eng. 2012;17(1):75–89.
32. Scheliga K, Friesike S. Putting open science into practice: A social dilemma? First Monday. 2014;19(9):. https://doi.org/10.5210/fm.v19i9.5381.
33. Aguinis H, Solarino AM. Transparency and replicability in qualitative research: The case of interviews with elite informants. Strat Manag J. 2019;40(8):1291–315. https://doi.org/10.1002/smj.3015.https://onlinelibrary.wiley.com/doi/pdf/10.1002/smj.3015.
34. Grubb AM, Easterbrook SM. On the lack of consensus over the meaning of openness: An empirical study. PLOS ONE. 2011;6(8):1–12. https://doi.org/10.1371/journal.pone.0023420.
35. Tamminen KA, Poucher ZA. Open science in sport and exercise psychology: Review of current approaches and considerations for qualitative inquiry. Psychol Sport Exerc. 2018;36:17–28. https://doi.org/10.1016/j.psychsport.2017.12.010.
36. Steinhardt I. Learning open science by doing open science. a reflection of a qualitative research project-based seminar. Educ Inf. 2020;36(3):263–79.
37. Kapiszewski D, Karcher S. Transparency in practice in qualitative research. APSA Prepr. 2019. https://doi.org/10.33774/apsa-2019-if2he-v2.
38. Levitt HM, Bamberg M, Creswell JW, Frost DM, Josselson R, Suárez-Orozco C. Journal article reporting standards for qualitative primary, qualitative meta-analytic, and mixed methods research in psychology: The apa publications and communications board task force report. Am Psychol. 2018;73(1):26.
39. Evers J, Caprioli MU, Nöst S, Wiedemann G. What is the refi-qda standard: Experimenting with the transfer of analyzed research projects between qda software. In: Forum Qualitative Sozialforschung/Forum: Qualitative Social Research, vol. 21; 2020.
40. Strauss A, Corbin J. Grounded theory methodology. Handb Qual Res. 1994;17:273–85.

41. Auer S, Kovtun V, Prinz M, Kasprzik A, Stocker M, Vidal ME. Towards a knowledge graph for science. In: Proceedings of the 8th International Conference on Web Intelligence, Mining and Semantics. New York: Association for Computing Machinery; 2018. p. 1–6.
42. McDaniel M, Storey VC. Evaluating domain ontologies: Clarification, classification, and challenges. ACM Comput Surv. 2019;52(4):. https://doi.org/10.1145/3329124.
43. Palavitsinis N. Metadata quality issues in learning repositories: Dissertation at Universidad de Alcalá; 2014. http://hdl.handle.net/10017/20664.
44. Hu X, Ng J, Xia S. User-centered evaluation of metadata schema for nonmovable cultural heritage: Murals and stone cave temples. J Assoc Inf Sci Technol. 2018;69(12):1476–87. Online verfügbar unter: https://doi.org/10.1002/asi.24065.
45. Lee JH, Clarke RI, Perti A. Empirical evaluation of metadata for video games and interactive media. J Assoc Inf Sci Technol. 2015;66(12):2609–25. https://doi.org/10.1002/asi.23357. https://asistdl.onlinelibrary.wiley.com/doi/pdf/10.1002/asi.23357.
46. Malone J, Parkinson H. Reference and Application Ontologies. 2010. http://ontogenesis.knowledgeblog.org/295. Accessed: 26 July 2021.
47. Reinhold A. Forschungsdaten in der Videobasierten Unterrichtsforschung: Benutzerzentrierte Modellierung und Evaluierung Einer Domänen-Ontologie. Glückstadt: Hülsbusch; 2015.
48. Liu D, Bikakis A, Vlachidis A. Evaluation of semantic web ontologies for modelling art collections. In: Kirikova M, Nørvåg K, Papadopoulos GA, Gamper J, Wrembel R, Darmont J, Rizzi S, editors. New Trends in Databases and Information Systems. Cham: Springer; 2017. p. 343–52.
49. Yu J, Thom JA, Tam A. Requirements-oriented methodology for evaluating ontologies. Inf Syst. 2009;34(8):766–91. https://doi.org/10.1016/j.is.2009.04.002. Sixteenth ACM Conference on Information Knowledge and Management (CIKM 2007).
50. Pittet P, Barthélémy J. Exploiting Users' Feedbacks: Towards a Task-based Evaluation of Application Ontologies throughout Their Lifecycle. In: International Conference on Knowledge Engineering and Ontology Development. Proceedings of the International Conference on Knowledge Engineering and Ontology Development, part of the 7th International, Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management (IC3K 2015), vol. 2. Lisbonne; 2015. https://hal.archives-ouvertes.fr/hal-01459827. Accessed: 26 July 2021.
51. Hocker J. Creating a domain ontology for qualitative coding schemas - qualico. Dissertation, University of Hildesheim, unpublished. 2021.
52. Kriplean T, Beschastnikh I, McDonald DW. Articulations of wikiwork: Uncovering valued work in wikipedia through barnstars. In: Proceedings of the 2008 ACM Conference on Computer Supported Cooperative Work, CSCW '08. New York: Association for Computing Machinery; 2008. p. 47–56. https://doi.org/10.1145/1460563.1460573.
53. McDonald DW, Javanmardi S, Zachry M. Finding patterns in behavioral observations by automatically labeling forms of wikiwork in barnstars. In: Proceedings of the 7th International Symposium on Wikis and Open Collaboration, WikiSym '11. New York: Association for Computing Machinery; 2011. p. 15–24. https://doi.org/10.1145/2038558.2038562.
54. Friedman B, Kahn Jr PH, Hagman J. Hardware companions? what online aibo discussion forums reveal about the human-robotic relationship. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. New York: ACM; 2003. p. 273–280.
55. Orlikowski WJ, Gash DC. Technological frames: Making sense of information technology in organizations. ACM Trans Inf Syst. 1994;12(2):174–207. https://doi.org/10.1145/196734.196745.

Publisher's Note