Domain Ontology for Grounded Theory Qualitative Research: Bridging Interpretation and Structure in Digital Humanities ⋆

Edelweis Rohrer[1*], Regina Motz[1*], Juan Pablo Sandín[1], Emilio Méndez[1], Facundo Fleitas[1] and Alén Pérez[2]

[1]Facultad de Ingeniería, Universidad de la República, Julio Herrera y Reissig 565, 11300 Montevideo, Uruguay
[2]Facultad de Información y Comunicación, Universidad de la República, San Salvador 1944, 11200 Montevideo, Uruguay

Abstract

The study of social and cultural phenomena through qualitative research in Digital Humanities presents a unique epistemological challenge, one that requires reconciling the interpretive, context-rich methodologies of the social sciences with the structured, formalized approaches of computational ontology engineering. This document introduces a computational ontology created specifically to represent the artefacts produced during the research processes of the qualitative grounded theory approach. The ontology aims to improve understanding of the qualitative research domain by systematically organising and examining the artefacts produced, thus enhancing the rigour and clarity of grounded theory studies. The ontology aims to provide a foundational framework that facilitates a more structured approach to analysing and interpreting qualitative data derived from grounded theory research.

Keywords

qualitative research, grounded theory, computational ontology, data interpretation, emergent categories

1. Introduction

Qualitative research constitutes a fundamental methodology for understanding social and cultural phenomena from the subjects' perspectives, emphasizing researchers' contextualized interpretations [1, 2]. This paradigm thrives on methodological diversity, integrating heterogeneous data sources, from textual narratives and oral histories to multimedia artifacts, to capture nuanced humanistic insights. Grounded Theory is a methodology that involves iteratively conducting data collection, categorization, coding, and analysis, with the aim of generating theory. Hence, in Grounded Theory approaches, the research process becomes particularly dynamic: knowledge construction follows an inductive, data-driven trajectory, with new categories and concepts iteratively emerging [3, 4]. However, the inherent fluidity and complexity of this type of research present significant challenges, such as balancing deep interpretive insight with systematic rigor and effectively managing the evolution of conceptual relationships.

To address these challenges, computational ontologies offer a formal framework for structuring disciplinary knowledge without compromising interpretive richness. By codifying shared vocabularies and semantic relationships, ontologies improve interoperability across datasets and methodologies while preserving the contextual specificity essential to humanistic inquiry. Formal ontology languages such as

Proceedings of FOIS 2025 Satellite events co-located with the 15th International Conference on Formal Ontology in Information Systems (FOIS 2025), September 10-12, 2025, Catania, Italy

⋆You can use this document as the template for preparing your publication. We recommend using the latest version of the ceurart style.

*Corresponding author.

†These authors contributed equally.

erohrer@fing.edu.uy (E. Rohrer); rmotz@fing.edu.uy (R. Motz); juanp.sandin@gmail.com (J. P. Sandín); ej.mendez2904@gmail.com (E. Méndez); sfleitasfacundo@gmail.com (F. Fleitas); alen.perez@fic.edu.uy (A. Pérez)

0000-0002-8257-5291 (E. Rohrer); 0000-0002-1426-562X (R. Motz); 0009-0000-5270-4817 (J. P. Sandín); 0009-0006-3346-943X (E. Méndez); 0009-0006-6797-2084 (F. Fleitas); 0000-0002-4050-2481 (A. Pérez)

OWL (Web Ontology Language), grounded in description logic, empower researchers to systematically validate model consistency, infer implicit relationships, and rigorously trace analytical pathways [5, 6].

These capabilities align seamlessly with Grounded Theory's emergent epistemology, where evolving categories and non-linear analytical processes demand both flexibility in representation and precision in logic. By bridging unstructured interpretation with structured semantic modeling, such ontologies enable researchers to navigate the tension between inductive discovery and systematic validation—a critical advancement for qualitative methodologies in the digital age.

Nevertheless, the very strengths of qualitative research, its adaptability and responsiveness to context, pose significant challenges for systematic knowledge organization. The unstructured nature of qualitative data and the constant evolution of analytical categories complicate the retrieval, reuse, and cross-study comparison of data. Traditional computational approaches, designed for fixed schemas and standardized data, often fail to accommodate the fluidity and theoretical depth required in humanistic inquiry. This tension highlights the need for domain ontologies specifically designed to bridge the epistemological foundations of qualitative research with the structured logic of computational representation. To address these challenges, this paper proposes a domain ontology tailored to Grounded Theory research in Digital Humanities. Moving beyond generic classification systems, this ontology incorporates two key dimensions: (1) a formal structure based on OWL's reasoning capabilities to ensure consistency and enable knowledge inference; and (2) flexible, extensible design principles that accommodate emergent categories and theoretical sampling, capturing how codes evolve into conceptual categories and eventually into substantive theories. The ontology aligns with social-theoretical principles, particularly constructivist approaches that view knowledge as situated and relational, while leveraging computational advantages like semantic linking and automated consistency checking. The practical implementation of this ontology serves multiple functions: it structures qualitative artifacts, including codes, memos, and theoretical constructs, while preserving their interpretive context; facilitates collaborative analysis through shared semantic frameworks; and enables sophisticated cross-corpus querying capabilities.

For instance, OWL-based reasoning can help identify connections between seemingly descriptive and analytical categories or trace the development of theoretical concepts across different stages of research. This approach does not constrain the inductive nature of Grounded Theory but rather provides a systematic way to document and build upon its analytical processes [7]. By integrating the representational power of computational ontologies with the epistemological foundations of qualitative inquiry, this work advances methodological management in Digital Humanities. It demonstrates how formal knowledge representation can enhance rather than reduce qualitative analysis, making interpretive processes more transparent while enabling new forms of knowledge synthesis and reuse.

This paper proposes a domain ontology specifically tailored to Grounded Theory research. Unlike generic classification systems, this ontology is designed to capture the emergent, relational, and theory-building aspects of qualitative analysis. It offers a structured yet adaptable representation of research artifacts, such as codes, memos and theoretical categories, enabling researchers to trace the evolution of concepts while maintaining interoperability across studies. Furthermore, by aligning ontological design with key principles from grounded theory (e.g., emergent concepts), the framework ensures that computational formalization does not reduce qualitative data to rigid taxonomies but instead amplifies its analytical potential.

The proposed ontology serves a dual purpose: it improves the organization and retrieval of qualitative data while fostering collaboration among researchers through a shared semantic framework. By bridging the epistemic divide between interpretive social science and computational knowledge representation, it enables more transparent communication of findings and more efficient synthesis of insights across projects. Ultimately, this structured approach does not constrain the inductive spirit of Grounded Theory but rather provides a scaffold to make its implicit analytical processes explicit, scalable, and reusable, advancing both transparency rigor and methodology management in Digital Humanities.

Beyond their reasoning capabilities, which include detecting logical contradictions and deriving new insights, ontologies serve as valuable artifacts for domain representation. As noted in the *Ontologies in the Behavioral Sciences*, they help expert users identify weaknesses, omissions, and conceptual gaps in knowledge representation, fostering consensus on standardized definitions and terminologies within a field [8].

This work advances these foundations by introducing a formal ontological framework that bridges Grounded Theory's emergent logic with computational knowledge representation. Key contributions include:

- A methodological bridge between social-theoretical constructs and machine-interpretable semantics, maintaining fidelity to qualitative research principles while enabling computational validation.
- A practical OWL implementation that enhances analytical transparency (e.g., through traceable coding hierarchies), enables automated knowledge inference, and facilitates collaborative domain model refinement. The ontology is openly accessible at: [https://technoportal.hevs.ch/ontologies/ONTOINVQUALI].

By harmonizing epistemological flexibility with formal rigor, the framework addresses a critical gap in integrating inductive methodologies with ontology-driven research.

The following sections detail the ontology's design principles, its theoretical alignment with Grounded Theory methodology, and its applications in facilitating rigorous yet flexible qualitative research.

2. Related work

Recent advances in computational qualitative methodologies have established critical foundations for ontological approaches to mixed-methods research. A seminal contribution by Hanchard and Merrington (2019) systematically addresses the epistemological and practical challenges of deriving formal ontologies from qualitative research paradigms [9]. Their framework demonstrates how interdisciplinary teams can successfully bridge methodological divides to develop computational representations while preserving what they term the *"essential epistemological rigor of qualitative approaches."* This work provides particular value in demonstrating: (1) the feasibility of maintaining qualitative integrity during formalization processes, and (2) effective collaboration models between social scientists and computer scientists. While their focus remained primarily on database applications, their methodological insights directly inform our work's core challenge of balancing Grounded Theory's emergent nature with computational formalism.

Building on these foundations, Hocker and Zachry's (2021), QualiCO ontology, represents a significant advance in computational support for qualitative analysis [10]. Their work addresses a critical gap in open science by developing the first open ontology dedicated for qualitative coding schemes. QualiCO's key innovations include: (1) a formal structure for preserving coding hierarchies and relationships, (2) mechanisms to facilitate the reuse and comparison of coding frameworks across studies, and (3) an approach to standardisation that respects qualitative research's inherent methodological diversity. Notably, their findings reveal how ontological structures can enhance transparency in qualitative analysis without imposing artificial rigidity—a balance central to our framework's design. However, QualiCO's exclusive focus on coding schemes leaves unaddressed the broader need for domain ontologies that encompass the full research lifecycle, from data collection to theoretical development. Our work extends this by embedding coding within a comprehensive ontological framework while retaining QualiCO's strengths in interoperability and reuse.

However, it is important to recognize that many existing approaches rely primarily on static representations or simple organizational schemas without employing formal reasoning mechanisms. While Bryda et al.'s (2024) [11] domain ontology offers a valuable and data-driven systematization of the qualitative research field, mapping its vast methodological diversity into 369 ontological classes and revealing distinct subfields, it notably does not incorporate formal reasoning mechanisms. This limitation means that, although the ontology effectively organizes and visualizes relationships between qualitative research practices, it function primarily as a descriptive classification or semantic network rather than a reasoning-enabled framework. Moreover, the absence of reasoning mechanisms can hinder the ontology's ability to support tasks such as consistency checking or the integration of datasets.

Addressing this gap, Operationalizing Scholarly Observations in OWL [12] demonstrates how leveraging formal ontology languages such as OWL and rule languages like SWRL enables not only structured representation but also powerful reasoning over scholarly data. By encoding ontologies in OWL, researchers can perform tasks such as classification, data sharing, and logical inference, which facilitate discovery of implicit relationships and ensure consistency within complex knowledge bases. This reasoning capability is crucial for advancing beyond static taxonomies toward dynamic, interoperable frameworks that can handle the complexity of qualitative research data.

In contrast, although the Semantic Web Rule Language (SWRL) extends ontological expressiveness through the use of rules, its application introduces significant computational complexity, which can negatively impact the scalability and decidability of reasoning processes. In dynamic and extensible domains such as qualitative research, where flexibility and maintainability are paramount, SWRL may complicate ontology management and evolution. Nonetheless, SWRL proves valuable when complex and domain-specific rules are required that cannot be easily expressed using description logic alone. For example, it supports particular deductions or detailed validations that complement deductive reasoning. Therefore, its use should be carefully evaluated based on domain requirements and the trade-off between expressiveness and computational efficiency.

Furthermore, advances in OWL reasoning frameworks, including distributed and rule-based reasoning approaches, have enhanced scalability and efficiency, enabling reasoning over large and heterogeneous datasets (e.g., biological knowledge networks). Techniques such as MapReduce-based property chain reasoning and precomputed specialized rules improve performance, making reasoning feasible in real-time applications. These developments underscore the potential of OWL, based ontologies to support sophisticated qualitative research workflows by enabling automated hypothesis generation, error detection, and semantic integration.

Complementing these ontological approaches, Computer-Assisted Qualitative Data Analysis Software (CAQDAS) tools like ATLAS.ti [13] demonstrate the practical implementation of computational support in qualitative research [14]. These platforms have evolved beyond basic coding functionality to offer: (1) sophisticated multimedia data management, (2) visualisation tools for complex qualitative relationships, and (3) features supporting team-based analysis. While not ontologically grounded, CAQDAS systems provide crucial empirical evidence that qualitative researchers will adopt computational tools when they: (a) align with existing workflows, and (b) enhance rather than constrain interpretive flexibility. Our framework incorporates these lessons by ensuring seamless integration with standard analysis practices while adding unique value through ontological reasoning capabilities that current CAQDAS tools lack.

Together, these strands of research position our work at an important inflection point—where emerging computational approaches must simultaneously:

- Respect qualitative research's epistemological foundations
- Address concrete analysis challenges through formal representation
- Achieve practical adoption through thoughtful design

Our ontological framework advances beyond prior work by unifying these requirements through a *grounded yet computable* knowledge representation, offering both methodological fidelity and novel computational affordances.

3. An ontology to describe qualitative research

This section presents the design of the qualitative research ontology. The development of the ontology has been guided by the activities proposed by the NeOn ontology design methodology, following the agile development approach eXtreme Design Metodology [15, 16]. That is, once the high-level requirements were gathered and a preliminary high-level abstraction model was created, the activities of requirements specification, conceptualization, formalization, implementation, verification, and validation were carried out in short cycles. Each cycle extended the implementation from the previous one with small increments, which were validated by the domain expert Alén Pérez. We followed a bottom-up approach in the sense that we do not reuse foundational ontologies terms; since the qualitative research domain is narrowly scoped and characterized by a highly specific vocabulary, we deem it unsuitable to reuse overly general terms such as those defined in foundational ontologies.

Taking into account the particularities of the domain, a model is constructed that describes the main artifacts and actors involved in a qualitative research project, as well as the relationships between them in a way that reflects the flexibility of the research process. Since the primary role of the qualitative researcher is to understand the structure of the subject's or object's knowledge, the design of this ontology seeks to represent a flexible classification of the collected data and the artifacts produced during the project, within different categories—some of which emerge from the researcher's interpretation of the data.

The ontology model has three main classes: *ProjectFormulation*, *DataCollection* and *Finding*, that play the role of grouping data, artifacts and actors, according to the project phase they are defined or produced: *ProjectFormulation* models artifacts and actors involved in the project formulation phase, *DataCollection* for data collected from the subject or object of study, and *Finding* representing the results reached by the project.

Classes *ProjectFormulation*, *DataCollection* and *Finding* are indeed superclasses connected by complex relations and also intermediate classes, which serve as the link between instances of subclasses belonging to different superclasses.

Figure 1 shows the main classes and properties in the internal structure of the class *ProjectFormulation*. In this diagram, ovals symbolize classes and arrows depict properties. In the model, besides classes *Project* and *Researcher*, connected by the property *hasResearcher*, project objectives are represented through the class *Objective*, whereas the theory that supports the research and corresponding literature are modelled by classes *TheoreticalFramework* and *Bibliography*. The class *MethodologicalStrategy* represents the strategy researchers decide applying based on the theoretical framework, which is related by the property *appliesTechnique* to the class *Technique*, that represents how the strategy is implemented. *Technique* has some subclasses *Interview*, *DocumentAnalysis* and *Observation*, which do not intend to be an exhaustive or complete classification, to keep the model flexible. Therefore, there is no axiom that restricts *Technique* to be the union of its subclasses. *Interview* is connected to the class *Question*, representing questions' interviews, by the property *hasQuestion*.

Below we illustrate the model of Figure 1 with some instances for a project example inspired by a real project [17].

> The project "Study and reading practices of communication students", has the objective of "identifying the most commonly used study practices among students", is supported by the theoretical framework cited by "Ciesielski et al., 2017" and defines conducting the study using an "exploratory" strategy. This strategy is implemented through the application of an interview (*I*) to students, that has the question:
> "What resources, networks, and platforms do you use to get informed, share information, and organize team tasks?".

During the data collection phase, raw data from the subject or the object of study are collected and recorded on different supports. This is represented by the superclass *DataCollection*. Figure 2 shows the structure that links the set of subclasses of *DataCollection*, and also shows how *ProjectFormulation* and *DataCollection* are related. The class *SubjectOrObject* has subclasses *Subject* and *Object*, and possible participants, represented by the class *Participant*, related to *Subject* by the property *isParticipantOf*. *Record* represents the set of records of the raw data collected, classified into (not exhaustive) subclasses *FieldNote*, *EnrichedDocument* and *InterviewAnswer*. As data can be recorded onto different supports,

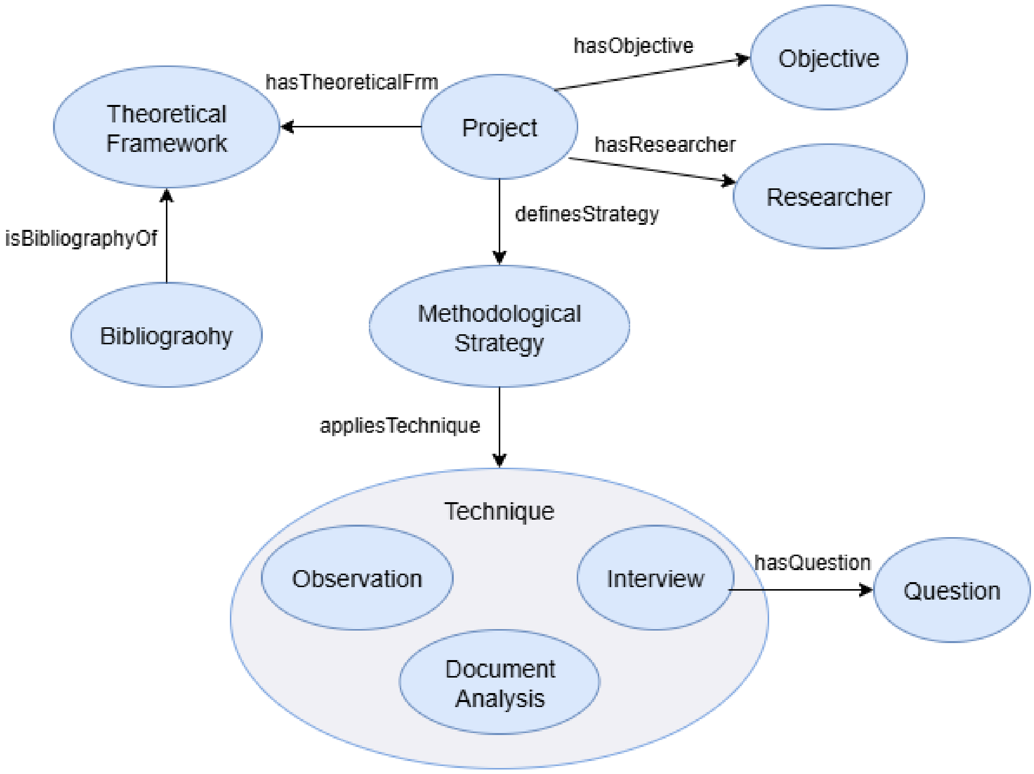

Figure 1: Internal structure of the class *ProjectFormulation*

Record is connected by the property *supportedBy* to the class *Support*, with different kinds of supports *Video*, *Audio* and *Text* as subclasses of *Support*.

The connection between *ProjectFormulation* and *DataCollection* lies in the application of techniques to subjects or objects, and the recording of collected data. This is ternary relation among *Technique*, *SubjectOrObject* and *Record*, which is solved by the property *isAppliedIn* from *Technique* to the union of *SubjectOrObject* and *Record*. As Figure 2 shows, *isAppliedIn* composes two pairs of properties, with the intermediate class *TechniqueOverSubjectObject*, enabling that, from a given technique, the reasoner infers subjects and objects to which it is applied and the records that are generated. Property chain axioms in description logics are as follows:

$$hasSubjectObjectApplication \ o \ toSubjectObject \sqsubseteq isAppliedIn$$
$$hasSubjectObjectApplication \ o \ generates \sqsubseteq isAppliedIn$$

The class *TechniqueOverSubjectObject*, that represents the application of a technique stated in the project formulation to subject or object, generating data records, is neither modelled as a subclass of *ProjectFormulation* nor as subclass of *DataCollection*, because it is needed to connect them. Instances that populate *TechniqueOverSubjectObject* make it possible to identify which subject or object the records generated as a result of applying a technique correspond to.

There is also other property that relates *ProjectFormulation* to *DataCollection*, which is the simple property *hasAnswer* from *Question* in *ProjectFormulation* to *InterviewAnswer* in *DataCollection*. *hasAnswer* is populated only in case of applying the interview technique in the project. A property chain *hasInterviewAnswer* is also defined to directly obtaing answers from interviews. Continuing with the example about the project "Study and reading practices of communication students", below some instances are presented to illustrate the model of Figure 2.

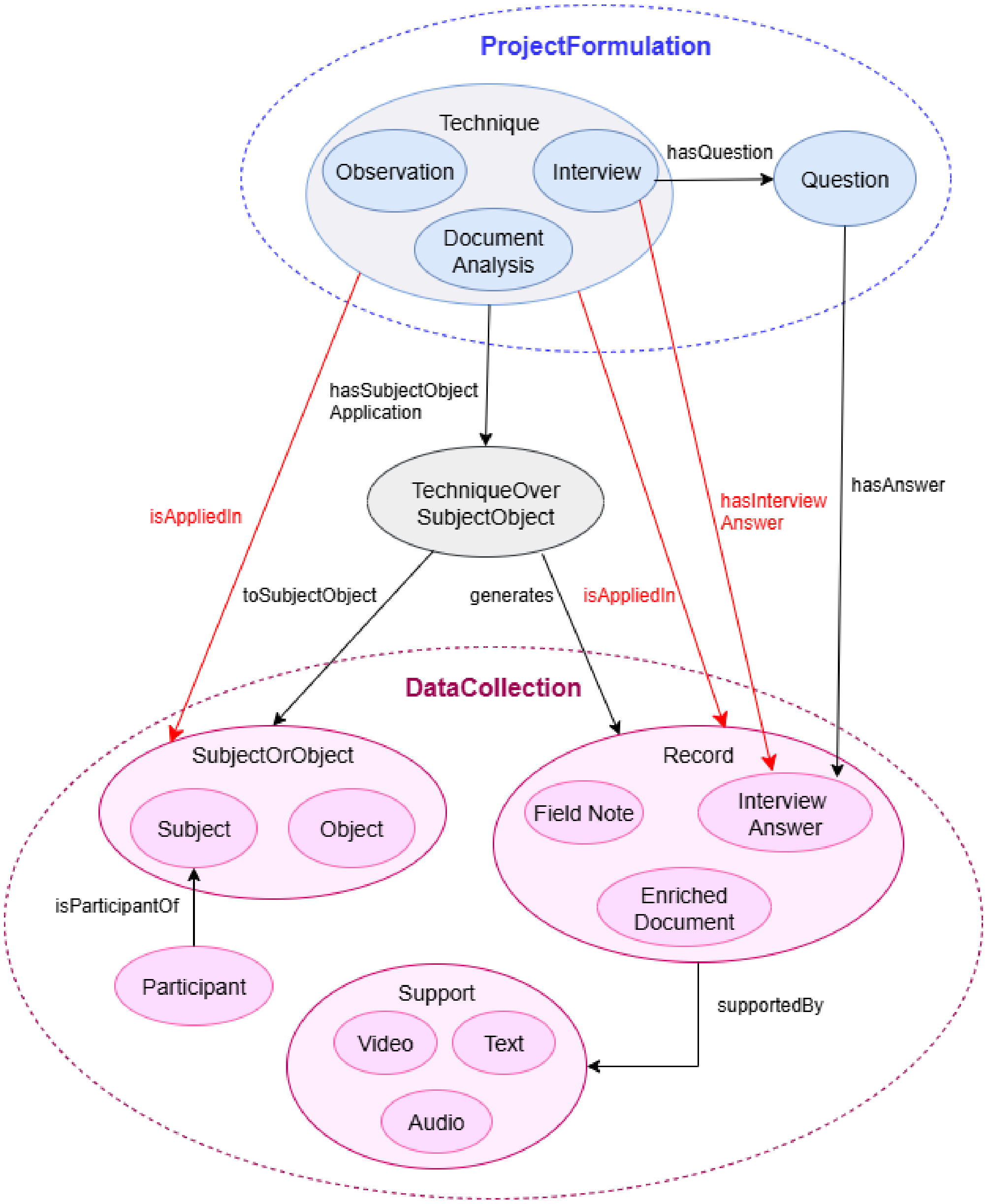

Figure 2: Internal structure of the class *DataCollection*

> The interview *I* is applied to the subject "Students of the course Society, Culture and TICs, from the Communication program at FIC" (*S1*), that answers the question: "What resources, networks, and platforms do you use to get informed, share information, and organize team tasks?" (*Q1*).
> Among answers, one registered answer is "To share information, I use Gmail or WhatsApp (mostly to share summaries or bibliography texts). To organize tasks, Google Drive is the main tool; we create a Google document which we share via WhatsApp or Gmail" (*A1*).
> Then, there is an instance *AIS1* of *TechniqueOverSubjectObject* that connects the interview *I* to the subject *S1* and the answer *A1*. If the same interview *I* is applied to other subject "Students of the course Science History, from the Communication program at FIC" (*S2*), with answers *A2*, *A3* to the question *Q1*, another instance *AIS2* will relate *I* to *S2*, *A2* and *A3*.

Figure 3 shows the model within the superclass *Finding*, and the relations with *ProjectFormulation* and *DataCollection*. The structure of *Finding* models the information that results from the interpretation that researchers make on data collected by the application of techniques, e.g. the interpretation of data records corresponding to interviews'answers. This is represented by the concept *Information*, that has associated the conclusion reached by the researcher about a given data. This information along with conclusions is part of the final report containing the project results, modeled by the class *Report*. To reach these conlusions, data records are categorized by researchers into categories that emerge from a reflexive process, when interpreting the data. The class *DescriptiveCategory* models these categories, which are indeed a hierarchy of categories, each related to its immediate higher category by the property *hasHigherCategory*. This resulting information is also annotated with analytic categories, modeled by the class *AnalyticCategory*, that groups descriptive categories. The relation between *AnalyticCategory* and *DescriptiveCategory* is represented by the property *correspondsToDescriptiveCategory*. The information together with the conclusions are cited in the project's final report, modeled by the class *Report*. In this ontology model, the interpretation process is not modelled within the class *Finding* because is considered a central concept which connects the three main classes *ProjectFormulation*, *DataCollection* and *Finding*. It represents the interpretation made by a researcher of a data record, generating a piece of information, which is modelled by the class *Interpretation*, and is related to classes *Researcher* in *Formulation*, *Record* in *DataCollection*, and *DescriptiveCategory* and *Information* in *Finding*, as Figure 3 shows.

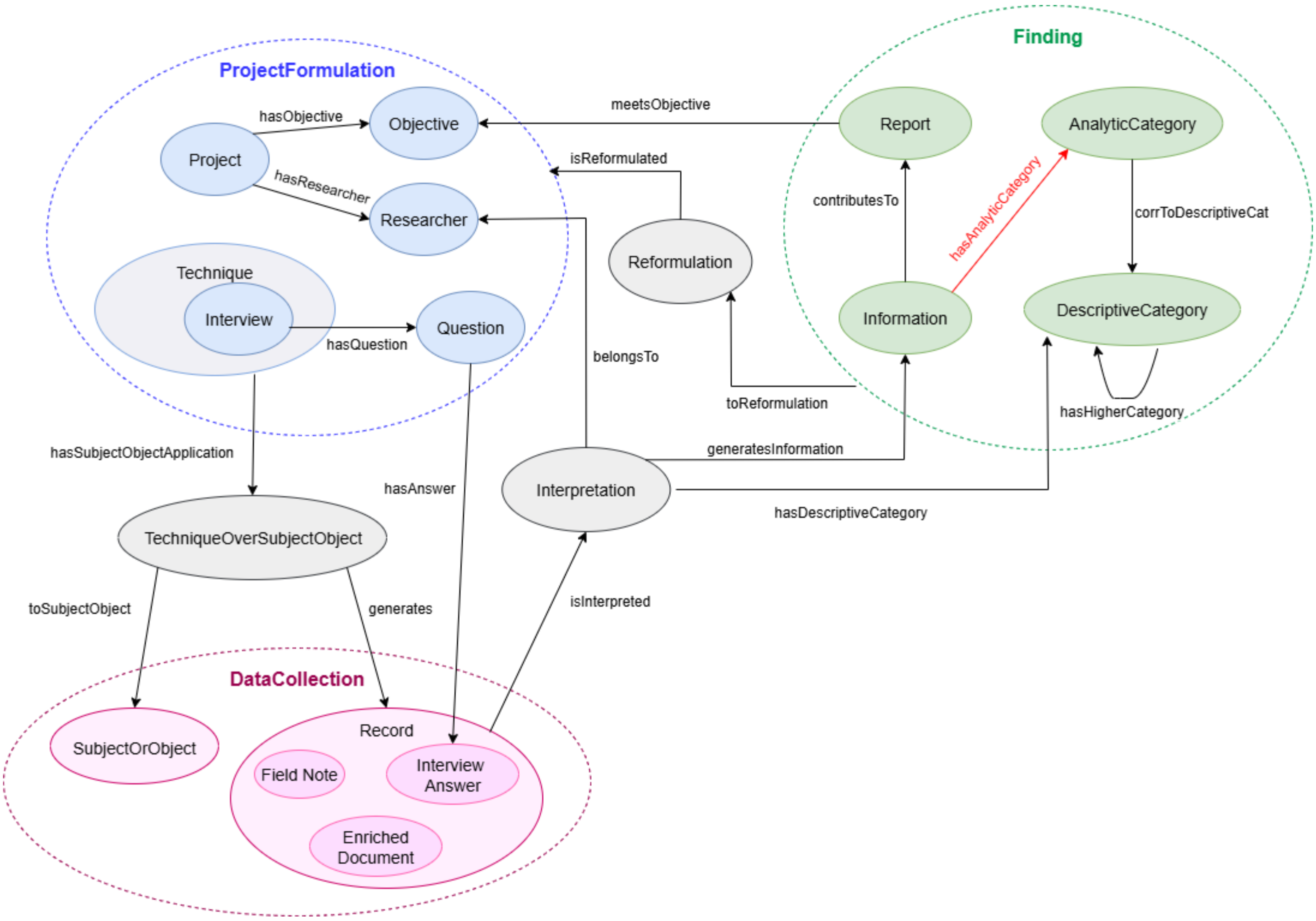

Figure 3: Internal structure of the class *Fnding*

To complete the illustration of the model, we provide the example below.

Which is the interpretation of the data record corresponding to the answer *A1*:
"To share information, I use Gmail or WhatsApp (mostly to share summaries or bibliography texts). To organize tasks, Google Drive is the main tool; we create a Google document which we share via WhatsApp or Gmail"?
It is an instance of *Interpretation* that is related to:
• The researcher "Soledad Morales" by the property *belongsTo*
• The information with the conclusion "There are, therefore, certain criteria that help define the need for a meeting, and these are clearly identified by the students. They are able to determine what technology to use for which purpose. This understanding has been built over time through experience and the development of agreements—sometimes explicit and sometimes implicit, resulting from the acceptance of certain ground rules that gradually emerge, for example, within working groups."
• The descriptive category "Whatsapp, Zoom y Discord. Google drive, Google meet, Gmail, Zoom, EVA, Navegadores (Chrome)." (*C111*).

Regarding the previous example, it is worth noting that *C111* has the higher category "Resources, networks, and platforms used to organize team tasks" (*C11*), that has the higher category "Study practices and resources" (*C1*). Then, the interpretation of the answer *A1* is also related to descriptive categories *C11* and *C1*, due to the property *hasDescriptiveCategory* is a property chain formalized by the axiom below.

$$hasDescriptiveCategory\ o\ hasHigherCategory \sqsubseteq hasDescriptiveCategory$$

In addition to leveraging the deductive reasoning capabilities of descriptive logic to classify data interpretation into higher-level categories, it is important to highlight that this model allows the researcher to introduce new emerging categories as instances at any time, without the need to change the structure of the ontology, that is, the TBox axioms.

The property *hasAnalyticCategory* from *Information* to *AnalyticCategory* is also a property chain that enables the inference of analytic categories corresponding to instances of *Information*, based on descriptive categories associated to the records' interpretation that generates the information. The definition of *hasAnalyticCategory* is as follows:

$$generatesInformation^{-}\ o\ hasDescriptiveCategory\ o\ corrToDescriptiveCat^{-} \\ \sqsubseteq hasAnalyticCategory$$

Moreover, two properties connect *Finding* to *ProjectFormulation.* The property *meetsObjective* from *Report* to *Objective* relates the report containing project results to those objectives that were met, from the set of objectives outlined in the formulation. The other relation models the situation that arises when obtained findings lead to a reformulation of the project, e.g. with different objectives or strategy. For this, properties *toReformulation* and *isReformulated*, and the intermediate class *Reformulation* are defined. *Reformulation* has associated (as a data property) the rationale behind the reformulation.

To obtain a flexible model, we intentionally do not characterize any of the defined properties with cardinality restrictions; i.e., no property is defined as functional, inverse functional, or with any other cardinality constraint.

The implemented ontology is available in https://technoportal.hevs.ch/ontologies/ONTOINVQUALI.

4. Conclusions and future work

This paper presents an ontology designed to describe qualitative research methodologies. Specifically, our model encompasses all key artifacts, including raw data, processed data, research objectives, strategies, and the corresponding techniques or methods used to implement these strategies, as well as the actors involved in the Grounded Theory qualitative methodology.

Qualitative research activities inherently lack a fixed structure because, fundamentally, they involve understanding and analyzing a particular issue or situation from the perspectives of the individuals or sources involved. Consequently, qualitative researchers require the ability to organize and classify their data within a flexible structure, more precisely, a structure that evolves dynamically as the project progresses. Our work places particular emphasis on addressing this critical need for flexibility.

To accommodate this requirement, we deliberately avoided introducing axioms that restrict the union of subclasses to be equivalent to their superclass, as well as property (inverse) functionality or cardinality constraints. Given that structural flexibility demands a model supporting a wide range of alternatives and complex relationships, many involving arity greater than two, so we defined multiple intermediate classes and composite relations to enable more direct and nuanced connections between model instances. Additionally, our ontology supports the dynamic creation of categories that typically emerge during the analysis phase of a project.

To the best of our knowledge, no existing ontology describes and categorizes, with comparable flexibility, the primary artifacts and actors involved in qualitative research projects employing the Grounded Theory methodology. Moreover, our domain expert regards this ontology as a valuable resource for teaching qualitative research. While there are vocabularies and ontologies (e.g., CIDOC) that share a limited subset of terms with our model, they do not comprehensively cover all the concepts or the domain structure that our ontology addresses. We plan to align our ontology with these established vocabularies where appropriate to enhance interoperability.

Beyond Grounded Theory, we plan to extend our ontology to encompass additional qualitative research methodologies and to integrate it with ontologies that represent qualitative research processes as well as other domains within Open Science. Currently, we are evaluating neuro-symbolic approaches to support and enhance ontology design. Building on these findings, we aim to leverage existing domain corpora to predict and establish meaningful relationships between our qualitative ontology and other ontologies in the Open Science ecosystem.

Declaration on Generative AI

During the preparation of this work, the authors used ChatGPT in order to: Grammar and spelling check, Paraphrase and reword. After using these tool/service, the authors) reviewed and edited the content as needed and take full responsibility for the publication's content.

References

[1] D. Silverman, Qualitative research, SAGE Publications Ltd, 2020.

[2] M. Q. Patton, Qualitative research, Encyclopedia of statistics in behavioral science (2005).

[3] B. G. Glaser, A. L. Strauss, E. Strutzel, The discovery of grounded theory; strategies for qualitative research, Nursing research 17 (1968) 364.

[4] A. Strauss, J. Corbin, Grounded theory methodology: An overview., Sage Publications, Inc, 1994.

[5] F. Baader, Description Logics, volume 5689 of *Lecture Notes in Computer Science*, Springer, Heidelberg, 2009, pp. 1–39. doi:10.1007/978-3-642-03754-2_1.

[6] J. Tao, Integrity constraints for the semantic web: an OWL 2 DL extension, Rensselaer Polytechnic Institute, 2012.

[7] M. Menárguez-Tortosa, J. T. Fernández-Breis, Owl-based reasoning methods for validating archetypes, Journal of biomedical informatics 46 (2013) 304–317.

[8] M. A. Musen, J. Bian, B. F. Chorpita, V. L. Patel, C. Tao, Ontologies in the behavioral sciences: Accelerating research and the accessibility and use of knowledge, in: American Medical Informatics Association Annual Symposium, 2022. URL: https://api.semanticscholar.org/CorpusID:256666034.

[9] M. Hanchard, P. Merrington, Developing a computational ontology from mixed-methods research: A workflow and its challenges, in: L. Pitcher, M. Pidd (Eds.), Proceedings of the Digital Humanities Congress 2018, The Digital Humanities Institute, Sheffield, UK, 2019. URL: https://www.dhi.ac.uk/openbook/chapter/dhc2018-hanchard, available online at: https://www.dhi.ac.uk/openbook/chapter/dhc2018-hanchard.

[10] J. D. W. M. Hocker, T. Bipat, M. Zachry, Evaluating qualico: An ontology to facilitate qualitative methods sharing to support open science, Journal of Internet Services and Applications 12 (2021) 5. URL: https://doi.org/10.1186/s13174-021-00135-w. doi:10.1186/s13174-021-00135-w.

[11] G. Bryda, J. Gajda, N. Martini, D. Płatek, Domain ontology: The new method of mapping the field of qualitative research practices, Przegląd Socjologii Jakościowej (2024).

[12] E. M. Sanfilippo, C. Masolo, A. Mosca, G. Tomazzoli, Operationalizing Scholarly Observations in OWL, in: CEUR workshop proceedings - https://ceur-ws.org/Vol-3809/, volume 3809 of *Proceedings of the fourth edition of the International Workshop on Semantic Web and Ontology Design for Cultural Heritage*, Tours (FR), France, 2024. URL: https://hal.science/hal-04771107.

[13] S. Friese, Qualitative data analysis with ATLAS. ti, SAGE Publications Ltd, 2019.

[14] M. Wickham, M. Woods, Reflecting on the strategic use of CAQDAS to manage and report on the qualitative research process, University of Tasmania, 2005.

[15] M. C. Suárez-Figueroa, A. Gómez-Pérez, E. Motta, A. Gangemi (Eds.), Ontology Engineering in a Networked World, Springer, 2012. URL: https://doi.org/10.1007/978-3-642-24794-1. doi:10.1007/978-3-642-24794-1.

[16] E. Blomqvist, K. Hammar, V. Presutti, Engineering ontologies with patterns - the extreme design methodology, in: P. Hitzler, A. Gangemi, K. Janowicz, A. Krisnadhi, V. Presutti (Eds.), Ontology Engineering with Ontology Design Patterns - Foundations and Applications, volume 25 of *Studies on the Semantic Web*, IOS Press, 2016, pp. 23–50. URL: https://doi.org/10.3233/978-1-61499-676-7-23. doi:10.3233/978-1-61499-676-7-23.

[17] R. Winocur, M. Cabrera, S. Morales, C. Rojas, Estrategias de estudio en jóvenes universitarios de la udelar. propuestas y desafíos teórico-metodológicos para abordar su estudio, in: V Jornadas de Investigación FIC, Montevideo, Uruguay, 2024.