

# Establishing semantic research graphs in humanities' research practice

## Open, participatory, and agile software engineering for educational research

Julian Hocker,<sup>1</sup> Cornelia Veja,<sup>2</sup> Christoph Schindler,<sup>3</sup> Marc Rittberger<sup>4</sup>

**Abstract:** This paper presents approaches and challenges to developing research software for qualitative humanities and social sciences research. While these methods so far lack a digital enhancement, we realized a range of research environments based on semantic graph technologies in research practice. In order to understand the researchers' needs and take the existing research practice seriously, a participatory and agile design approach was used by which several researchers were involved in realizing concrete research projects. A user-centered agile approach with fast prototyping was included. To materialize the needed interaction at an early stage of the research life cycle, this was accompanied by trainings, workshops, and joint paper publications in the different research communities.

**Keywords:** Software Engineering; Agile methods; Digital Humanities; Semantic Web; Participatory Design

## 1 Introduction

Educational research is a highly diverse field and therefore well-suited to create new methods for developing research software. Our focus lies on qualitative research, which is still under-represented in Digital Humanities [Dr12]. For this paper, we focus on the aspect of prototyping and agile development for creating research software and virtual research environments (VRE) for qualitative methods.

Qualitative methods are very divers, spanning methods where data are coded based on rules or coding frames, like Grounded Theory, but also interpretative methods like objective hermeneutics, where the understanding and interpretation of one text or certain parts of the text is focused.

This paper outlines the main software engineering approach, the setup of the projects, and the challenges to establishing the VREs in research practice.

---

<sup>1</sup> DIPF | Leibniz Institute for Research and Information in Education, julian.hocker@dipf.de

<sup>2</sup> DIPF | Leibniz Institute for Research and Information in Education, veja@dipf.de

<sup>3</sup> DIPF | Leibniz Institute for Research and Information in Education, schindler@dipf.de

<sup>4</sup> DIPF | Leibniz Institute for Research and Information in Education, rittberger@dipf.de

## **2 Co-evolutionary and participatory design approach**

All qualitative research methods are grounded in the interpretation of certain sources, which are set into relation with each other, forming a network of different entities. Still, these networks stay implicit and are only explained as text in a publication. This leads to issues concerning openness and confirmability. From a technical point of view, these networks can be made explicit as semantic networks and modeled using semantic web standards. This approach has the advantage of transferring implicit knowledge from researchers explicitly to readers. Additionally, the data representation as a semantic graph overcomes the ongoing struggle between qualitative and quantitative research. Qualitative researchers can describe the graph in detail, while quantitative researchers can serialize these data working on the same data platform.

Given that research software and especially semantic web technologies have not been established in a broad sense in the qualitative research communities, the technological engineering needed to be accompanied by workshops and meetings in order to improve the attitude of humanities researchers towards digital methods. As stated in [RS18] waterfall-models are not very helpful with regard to creating this kind of software.

For the software development an agile and participatory approach was used to include researchers and their disciplinary needs at an early stage (as proposed by [Wa12]). Accordingly, observations, joined meetings, rapid prototyping and workshops were used to stay in contact with the researchers and to exchange knowledge in order to find out about the needs of the research field and built tools that are tailored to the researcher's needs. The research community is thus encouraged to articulate their needs for digital research [Bo10]. In all pilot projects presented in this paper, the technological tools were developed in parallel to concrete research projects.

Apart from the user-centered approach in the development, a secondary aim was to feed the results back to the community. In order to achieve this aim, we organized workshops, published papers in several communities and published the code under a free software licence.

Another aspect covered by the research projects concerns infrastructure and support. We provided several instances of the VRE, according to the researchers' requirements and allowing for use in projects or seminars. Since qualitative researchers are often not experts in hosting infrastructure, this point is crucial as well. Only making software free does not bring non-technical users closer to using it.

## **3 Settings of semantically enhanced research environments for qualitative research**

In contrast to quantitative research with large corpora, in small qualitative research projects the methods are often tailored to researchers' needs. This also means that a supporting tool

has to be flexible enough to be adjusted by the researcher [SE16]. This was a main reason to use Semantic MediaWiki<sup>5</sup> as a research platform due to its modularity and openness. Additionally, it offers the re-use of a well-established software and for the development team the possibility to manage time and financial constraints of pilot projects. While the main functionalities are developed and maintained by the software community, concrete facilities for the researchers can be focused by developing software extensions. The following VREs have been engineered with this setting:

- Virtual research environment for serial-iconographic analysis of classroom photographs in educational journals connecting semantically bibliographic data, pictures, and research classifications. (Joint work with H. Kemnitz at TU Braunschweig)
- Citizen Science platform Interlinking Pictura<sup>6</sup>. Creating a research graph while semantically linking pictures, texts and bibliographical data of Friedrich Justin Bertuchs *Bilderbuch für Kinder*, an encyclopaedia for children from the 18th century. This project includes a citizen science component that allows non-researchers to add information to the graph and answer their own research questions. (Joint work with BBF | Research Library for the History of Education at DIPF, S. Kollmann)
- Virtual research environment for objective hermeneutics: Development of a virtual research environment for the method of objective hermeneutics with the focus on time and place-independent collaborative work using semantic web technology to make links between annotations explicit. (Joint work with H. Kminek, Goethe University, Frankfurt and M. Meier, University of Flensburg)
- Virtual research environment for high school essays: VRE for high school essays with the enrichment of bibliographic data and in combination with laws and historical research discussions. (Joint work with BBF | Research Library for the History of Education at DIPF, S. Reh)

In the following, the approach used in the project *VRE for objective hermeneutics* is presented.

## 4 Case study VRE for objective hermeneutics

In this pilot project a virtual research environment for the qualitative method objective hermeneutics [We13] was created. The team of developers collaborated with a research group distributed across different cities in Germany. This group was already using collaborative analysis to carry out their research using video conferencing software and interpreting sources on paper. The goal was to create a research environment that allows the researchers to work time-independently and make the research process more transparent.

---

<sup>5</sup> <https://www.semantic-mediawiki.org>

<sup>6</sup> <https://interlinking.bbf.dipf.de>



that the software would be also a good tool for teaching, since it makes the tacit knowledge by the experts explicit.

The software is now used in several seminars for teaching the method of objective hermeneutics at the universities of Frankfurt and Flensburg. Due to the explicitness of the interpretation the lecturer can now see how the interpretation by the students developed and which problems the students encountered when interpreting a text. The software was also used for another research project that used the method of objective hermeneutics [St18].

The design process helped the researchers to articulate their needs for research. This does not simply mean to translate from paper-based practices to digital practices, but also to think about the research methods and the potential that new technologies can have for a certain research method.

## 5 Conclusion

In the four projects presented in the paper, it became clear that the creation of qualitative, widely branched networks of relationships can be a core competence of qualitative research. With our approach we allowed the researchers to articulate their needs and build infrastructure and tools that allow them to create browsable and explicit connections of sources and analysis in their research. Our research lead to a great acceptance of tools in the field, which can be seen in joint paper publications (e.g. [Ve17]) as well as that the tools are still in use, e.g. for teaching.

A central challenge, however, remains: The cooperation between researchers from the various disciplines with the aim of intensifying the understanding of the different methods. Another open issue is to understand the capacities of digital tools in order to use them fruitfully in qualitative research. Tool development in the Digital Humanities should not be self-referencing, but help researchers to better answer their research questions.

## 6 Acknowledgements

The projects were realized at the eHumanities center CEDIFOR funded by BMBF (01UG1416C and 01UG1816C) and by the Leibniz Senate Competition Committee 2016.

## References

- [Bo10] Borgman, Christine L: The digital future is now: A call to action for the humanities. 2010.
- [Dr12] Drucker, Johanna: Humanistic theory and digital scholarship. *Debates in the digital humanities*, pp. 85–95, 2012.

- [RS18] Rajan, Vinodh; Stiehl, H. Siegfried: From Eye-to-Eye to Hand-in-Hand: Collaborative Solution Building in Interdisciplinary Manuscript Research. In (Burghardt, Manuel; Müller-Birn, Claudia, eds): INF-DH-2018. Gesellschaft für Informatik e.V., Bonn, 2018.
- [SE16] Schindler, Christoph; Ell, Basil: Linking Qualitative and Quantitative Research by Thin Descriptions through Semantic Graphs. Experiments in Apparatus Design with Semantic CorA. In: Proceedings of Digital Humanities 2016, short papers. Digital Humanities Conference 2016, 2016.
- [St18] Steinhardt, Isabel: Open Science-Forschung und qualitative Methoden—fünf Ebenen der Reflexion. *MedienPädagogik: Zeitschrift für Theorie und Praxis der Medienbildung*, 32:122–138, 2018.
- [Ve17] Veja, Cornelia; Sticht, Kendra; Schindler, Christoph; Kminek, Helge: SMW Based VRE for Addressing Multi-Layered Data Analysis: The Use Case of Classroom Interaction Interpretation. In: Proceedings of the 13th International Symposium on Open Collaboration. ACM, p. 10, 2017.
- [Wa12] Warwick, Claire: Studying users in digital humanities. *Digital humanities in practice*, 17(5):1–21, 2012.
- [We13] Wernet, Andreas: Hermeneutics and Objective. *The SAGE handbook of qualitative data analysis*, p. 234, 2013.