

SMW Based VRE for Addressing Multi-Layered Data Analysis – The Use Case of Classroom Interaction Interpretation

Cornelia Veja

German Institute for
International Educational
Research (DIPF)
Frankfurt am Main,
Germany
veja@dipf.de

Kendra Sticht

University of Münster
Münster, Germany
kendra.sticht@uni-
muenster.de

Christoph Schindler

German Institute for
International Educational
Research (DIPF)
Frankfurt am Main,
Germany
schindler@dipf.de
Open Education

Helge Kminek

Goethe University
Frankfurt, Frankfurt am
Main, Germany
kminek@em.uni-
frankfurt.de

ABSTRACT

Virtual research environments (VREs) based on Semantic MediaWiki (SMW) provide researchers with more effective means of collaborative knowledge creation. The main challenge is capturing the multi-layered and iterative research process by participatory and evolutionary design using an agile framework development. The paper focuses on the development of a VRE to enhance scholarly research practices with the specific goal to enable dialogic transcript data analysis in Social Sciences and Humanities, using digital and semantic technologies. To offer a user-friendly framework with a low technological barrier, the authors used existing SMW extensions and developed a novel extension for multi-layered dialogic analysis by adopting the researcher's theoretical model of Objective Hermeneutic.

As use case of addressing dialogic qualitative data analysis, the paper presents the VRE for classroom interaction which is based on two ontological models, the Objective Hermeneutic core ontology and the classroom interaction ontology, as a result of authors' work.

Author Keywords

Semantic MediaWiki, Open Collaboration Research, Semantic Web, Dialogical Theories, Conversation Analysis, Objective Hermeneutic, Virtual Research Environment, Qualitative Data Analysis, Humanities, Social Science,

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from Permissions@acm.org.

OpenSym '17, August 23–25, 2017, Galway, Ireland
© 2017 Copyright is held by the owner/author(s). Publication rights licensed to ACM.
ACM ISBN 978-1-4503-5187-4/17/08...\$15.00
<https://doi.org/10.1145/3125433.3125457>

ACM Classification Keywords

H.5.2. User Interfaces: User-centered design; K.4.3. Organizational Impacts: Computer-supported collaborative work; K.3.2 Computer and Information Science Education: Information systems education

1. INTRODUCTION

Humanities, Social Science and Educational Research are mostly seen as fuzzy or fluid, characterized by low levels of technological expertise in stakeholders and iterative research process [3]. The rise of Digital Humanities accomplishes these issues and creates the technological context for researchers, new computing resources and tools. It produces and uses new applications and techniques, which opens new possibilities for scientific research and takes Humanities research into the new digital era.

We argue that the indeterminacy of entities and properties in the qualitative data research process of Humanities, Educational and Social Science emphasizes the affordance of a VRE, enabling researchers to intra-link the corpus.

A Virtual Research Environment¹ is a web-based information system, which helps researchers collaborate across disciplines and has the potential to assist the researcher throughout the entire 'research life cycle'. This means giving researchers the continuous opportunity to create, modify and rearrange entities, properties, and relations. Nevertheless, it is necessary to articulate the potentials of the Semantic Web for qualitative data research appropriately, i.e.: a multi-faceted relationship among entities.

We argue also that SMWs partially meet the basic requirements (functional, structural and serving as user interface) for a VRE addressing small and medium research projects.

¹ <https://www.jisc.ac.uk/full-guide/implementing-a-virtual-research-environment-vre>

MediaWiki² is a form of Social Software web based platform, which enables collaboratively digital content creation, maintenance and information retrieval. MediaWiki is the wiki software behind Wikipedia³ and Wikimedia Commons⁴.

Semantic MediaWiki⁵ is an extension of MediaWiki. It uses semantic technology and allows annotating semantic data within wiki pages.

The work described in this paper aims to provide researchers an easy way to annotate and to structure specific research object in wiki without requiring the learning of formal ontology languages like RDF or OWL.

We model the sequential analysis of dialogic interaction that is based on conversation analysis (CA) [6], [9], and the interpretative act of Objective Hermeneutic (OH) as specific model for researchers. The model is implemented using collaborative VRE facilities, semantic technologies and bespoke tools.

The research questions investigated in this paper are, firstly, how to guide users in annotating semantic wiki contents according to their cognitive models, and secondly, how to build domain ontologies in a non-obtrusive way, taking unstructured knowledge and heuristic practices as the source of concepts and relations. The proposed solution for the first question is the development of VRE, based on Semantic MediaWiki (SMW) [10]. The latter solution is addressed through the interaction between various research fields: collaborative ontology modeling, Information Retrieval (IR) and specific research methodologies, Objective Hermeneutic. The goal of our work is to provide a wiki-based VRE for assisting researchers to easily apply Objective Hermeneutic method in qualitative data analysis of the dialogic transcript e.g. classroom interaction. Our main contributions in this work are four-fold:

- A new collection of semantic templates and forms to Semantic MediaWiki to capture the metadata of the transcript.
- Two new ontologies, the Objective Hermeneutic core ontology and the classroom interaction ontology.
- A new SMW extension: Objective Hermeneutic Interpreter (OHI) tool to support researchers in the interpretation process.
- Demonstration of the use of this extension in the classroom interaction data analysis. This approach allowed authors in defining dynamic domain ontologies, software supported research methodology and to build the learning knowledge bases for students.

The remainder of the paper is organized as follows: Section 2 discusses the context for the research and related work,

Section 3 outlines the design approaches and methodology, Section 4 discusses the overall architecture and describes components. Section 5 discusses our use case. Finally, Section 6 and Section 7 discuss future work and conclusion of our work respectively.

2. CONTEXT FOR THIS RESEARCH

There has been a recent move in social science towards multi-method approaches and a stage of qualitative research is often a precursor for quantitative data analysis, as Schmied [19] notes. Multi-methods require a semantic multi-disciplinary approach and this is reflected in recent research in this area.

The analysis of social interaction is a broad field in qualitative data research of Social Sciences, Education and Humanities, whereby a broad range of methodological approaches exist (e.g. ethnomethodology, conversational analysis, objective hermeneutics, dialogic approaches).

These kinds of research approaches rely mostly on protocols of interaction, which describes sequence by sequence the interactions or dialogues of different actors.

Studying this complex social phenomenon of interaction in Social Science is not a trivial task and accomplishes also an interpretation research process. Interpretive research tends to rely heavily on qualitative data [13]. One of the important classes of research use cases in Social Science is the human dialogic interaction.

For instance, in educational process, the teacher-student interactions at times appear to be ‘messy’ at first sight. In fact, the participants display awareness of differences in point of view and they attempt to acknowledge and to understand the other’s perspective. Also, classroom dialogue explicitly seeks to chain exchanges into meaningful sequence. It turns to be natural to use as research method the sequential analysis of the teacher-student talk that is based on conversation analysis methodology (CA) [6], [9]. The CA related analytic approach shows that each of the sequence items ‘is interdependent with its position in the sequence’ [11] ‘thus visualizing dialogic sequentially in its ongoing dynamics’. The sequential analysis is enriched by interpretation process, as part of qualitative data analysis.

The ontological approach for ‘Dialogic classroom talk’ is open to different mutually crossing perspectives [1]. We adopted also the ontological perspective which is more compatible with dialogic interaction and has the potential to improve the practice of dialogue in education [21].

In German-speaking countries, one of the main methodologies in Social Science used for qualitative data analysis of human dialogic interaction [5], [17], is Objective Hermeneutics.

² <http://www.mediawiki.org/>

³ <https://www.wikipedia.org/>

⁴ https://commons.wikimedia.org/wiki/Main_Page

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

Objective Hermeneutics was developed in the 70s mainly to record events in familiar interactions [14]. It allows the identification of the typical and characteristic structures of an interaction and brings out the principles behind the interaction.

According to Oevermann [15] and Wernet [22], the inner psychic reality is only methodically understandable and interpretable in forms of expression like texts. These texts find their appreciable manifestation in so called 'protocols of interactions'. These protocols are the research object of Objective Hermeneutic.

This methodology generates multi-layered deep-structure analyses of cases by reconstructing actions and meanings.

Just like CA, which has more international sound, OH regards meaning structures as the outcome of interaction, but it distinguishes between different dimensions of meaning in interaction. With regard to the case-elements, one can say that OH puts more emphasis on the actor element, while CA is mostly focused on actions [12].

In following the course of interaction, the actual situational context is gradually taken into account and the 'inner context' of the sequence is analyzed [23]. The hypothesis is the result of cumulative work about the selective pattern that is characteristic for the interaction structure.

Unlike CA, OH emphasizes on strong hypotheses from the outset of the analysis, hypotheses that can evolve in the course of sequential analysis, or can even be rejected.

The collaborative act of interpretation [16] is realized in a common space of imagination, wherein multiple layers of interpretations interfere and makes use of falsification and abduction [5]. In this respect, annotations are crucial in interpretations. Either they are note taking or created via a software tool, the specific of dialogic interpretation requires a high degree of annotations customization.

Semantic technologies and annotations

Different software-mediated frameworks for qualitative data research, like tools for the flexible coding system, are allowing researchers to annotate research resources according to a classification system that may evolve over time. However, these tools have been criticized for the limited metadata interoperability of resources and limited research findings. There are certain limitations in creating intra-links of the research object, which is crucial in the iterative act of interpretation. [18].

Annotations can contain interpretations, reviews and additional information about the resources they belong to. They reflect a variety of opinions about a resource, which establishes an interesting context, exploitable for information retrieval. Layers of annotation (private layer, collective layer, public layer) can coexist and reflect different views of the information resources.

A social platform solution, like semantic wikis, are promising environments, providing the users with a low level of technological expertise, an easy way to manage machine-processable knowledge. Semantic wikis yield the creation of added-value services based on the semantics of web pages. They support metadata insertions through semantic annotations and easy-link relations between wiki pages.

Semantic wikis provide the collaborative environment for users to contribute semantic markup (annotations). Annotations are required to refer to an ontological model defining concepts and properties that can be associated to pieces of wiki contents. Generally, those annotations are designed to create instances of domain ontologies and their related properties, whereas other wikis use semantic annotations to provide advanced metadata regarding wiki pages.

A semantic wiki has an underlying model of the knowledge described in its pages, allowing capturing or identifying further information about the pages and their relations. The formalization of the knowledge model should be available as RDFS or OWL, so that machines can process and reason on it. An article in wiki will have a part of semantic data about the page itself and another part about the page's subject.

Semantic MediaWiki (SMW) is a semantically enhanced Wiki engine that enables users to annotate wiki's content with explicit, machine-readable information. It supports adding structured and semantically annotated information into wikis using a specific syntax. SMW is based on a simple and unobtrusive mechanism for semantic annotation.

From the functional point of view, SMW may use and extend the MediaWiki template mechanism by offering to process semantic properties as parameters of the template. On this basis, semantic forms facilitate user input by hiding the complex syntax of the semantic templates.

The reason we use SMW as support for VRE:

- Low technological barrier and easy semantic annotation
- Steep learning curve
- Easy inter-linking
- Extensive use of templates and semantic forms for non-technical users (Data Aggregation with Templates and Semantic Forms)
- From the knowledge expressivity perspective SMW represents a balanced solution between expressivity and required knowledge.
- Easy customization and modularity
- Standardized development procedure for additional functionalities
- Query language for information retrieval
- XML/RDF export to other semantic applications

SMW based VREs

VREs take advantage of multiple perspectives, expertise, varying research methodologies, different personalities and attitudes and thus increase the likelihood of discoveries and inventions [20]. VREs based on semantic technologies leverage these features by shared-ness, extra knowledge base and possibility to exchange knowledge to external tools.

A SMW based VRE gives the opportunity to use a mature already existing framework, extensible and under an open license. SMW offers a modular semantical platform with various extensions, which offers a fundamental affordance of qualitative research for small/medium Social Science research projects.

SMW based VRE in Social Science enables distributed and collaborative qualitative analysis, easily and not burdening researchers with additional tasks. It also serves a community of practice and represents an attempt of formalization in this respect. In Social Science, researchers usually expect lower technical barrier in using a tool and SMW based VRE helps them to collaboratively create a common digital corpus of data and common ontologies.

In order to provide the community of practice with the whole array of commodities needed to accomplish the community's goal, SMW, together with our bespoke tool make the research process transparent (Open Science). VRE based on SMW offers services for semantic annotations, eases formalization of hypotheses, connectivity and collaborative space, along with the possibility of retracing back the data.

Due to the flexibility of SMW, customized functions and tools can easily be added pertinent to specific research project goals. Also, SMW empowers users in terms of fast knowledge accumulation and enables to participate in lightweight collaborative ontology design without explicitly stated as such.

Furthermore, we employ SMW as support of argumentation for more complete structure of the case debate. In this way, our VRE uses one of the most important features in MediaWiki, discussion wiki page, and offers technological support for consensus choice which will be reached by discussions, controversies and agreements.

Another requirement was the ability to manage users and their behaviors to ensure compliance with certain policies, as is the case of Objective Hermeneutic Protocol. This task is heavily based on SMW services, and only a minor part was directly handled by OHI extension.

As Candela states [18], 'it is also important to promote the systems and make the scholars aware of how the systems and tools can simplify their workflows and more widely disseminate their work'. In this respect, an important added value of our VRE is the availability for learning/teaching

Objective Hermeneutics or Conversation Analysis to students.

Related works

Semantic wikis are wide spread technologies, and several semantic wikis based on SMW extension exist in various domains. To the best of the authors' knowledge a similar semantic wiki or VRE to facilitate the research in conversation analysis domain does not exist.

One of the closest wikis to the one introduced in this paper is that of WiWork [8]. This paper presents WiWork, which was designed with the goal to provide documentations and manuals to domain experts and common users interested in CRISP activities and topics.

Another interesting work is WikiKnowledge [2] that uses SMW for the materials science community to create and curate consolidated vocabularies. It contains knowledge about public services, focuses on the labor market, on health, and on education domains.

Scratchpads⁶ are a VRE framework that has been constructed to support the needs of specialists interested in natural history. It is based on Content Management System Drupal. This framework was criticized for low usage of outbound connectivity.

Closer to our work, as domain of discourse, is VIVO project⁷, dedicated to fostering team work and scholarly communication across disciplinary and institutional boundaries. It enables the discovery of research and scholarship worldwide. As a research subject, it is much related with education, but the focus is not on conversational analysis, but it is mostly an institutional approach.

TEXTUS⁸ is an open source platform that helps researchers and students to collaborate around and work with collections of digitized public cultural and historical texts.

myExperiment⁹ is one of VREs that have started to use Semantic Web technologies to reach a higher level of interoperability. MyExperiment aims to create a platform to collect and semantically annotate scientific workflows. However, one of the main challenges of e-research - linking research data with further artifacts and offering them for re-use - is only captured in a limited way.

Similarities with the research presented in this paper is authors' previous work, Semantic MediaWiki for Collaborative Corpora Analysis (SMW-CorA) [18] which aims to adjust Semantic MediaWiki (SMW) for research:

⁶ <http://scratchpads.eu/>

⁷ <http://vivoweb.org/>

⁸ <http://okfnlabs.org/textus/>

⁹ <https://www.myexperiment.org/>

enabling researchers to create a research corpus, enrich, annotate and analyze digital objects.

Our work differs from each of these methods in research subject, objectives and methodology. We combine VRE requirements with Semantic MediaWiki technology and implementation of Objective Hermeneutic methodology for interpretive analysis. As a use case, this methodology is applied to classroom interaction analysis.

3. DESGN APPROACH AND METHODOLOGY

We use an empirical approach, as the result of the study in a multi-domain analytical experts group and made use of existing theory of Objective Hermeneutic. The researchers' methodology is driving the framework.

We used a participatory and evolutionary design approach with an agile development framework.

A group of researchers distributed across Germany actively participated to VRE design and analyzed classroom interactions for around one year. A multi-disciplinary team worked collaboratively to analyze requirements, established communication structure and build capacities. Every two weeks we organized workshops where the team discussed the evolution of the VRE, as part of a participatory approach.

This approach led to rapid prototyping, while the researchers continuously tested the environment. After this evolutionary development cycle was completed, the VRE has been used in seminars at the Goethe University of Frankfurt am Main, where several groups of students

intensively used the framework for learning the OH method and analyzed classroom interaction transcripts.

The implementation of VRE combines the advantages of a collaborative space for researchers, semantic technologies and Objective Hermeneutic Protocol, as a natural methodology for researchers. We extend SMW framework with a specific extension in order to support the intensive and multi-layered space of annotations.

We combined the experience gained in previous VRE development with domain experts' experience in applying OH research method, as we described in Section 2. To summarize the step-wise methodology:

- Case structure hypothesis is developed, based on former hypotheses.
- The next step is to find out what general types of contexts the stories' collection represents, and in doing so gradually develop the structural features (hypotheses) of these types that are relevant for the interaction.
- The next step emphasizes the contrast with the actual situational context. In this respect, it considers the general and the concrete situational context of an utterance. Also, it allows for collecting and analyzing multiple perspectives of interpretations.

The requirements formulated by the domain experts group found the successful expression by integrating their research and VRE practice.

The SMW gives the opportunity to develop this VRE as stepwise process and straightened out for early prototyping.

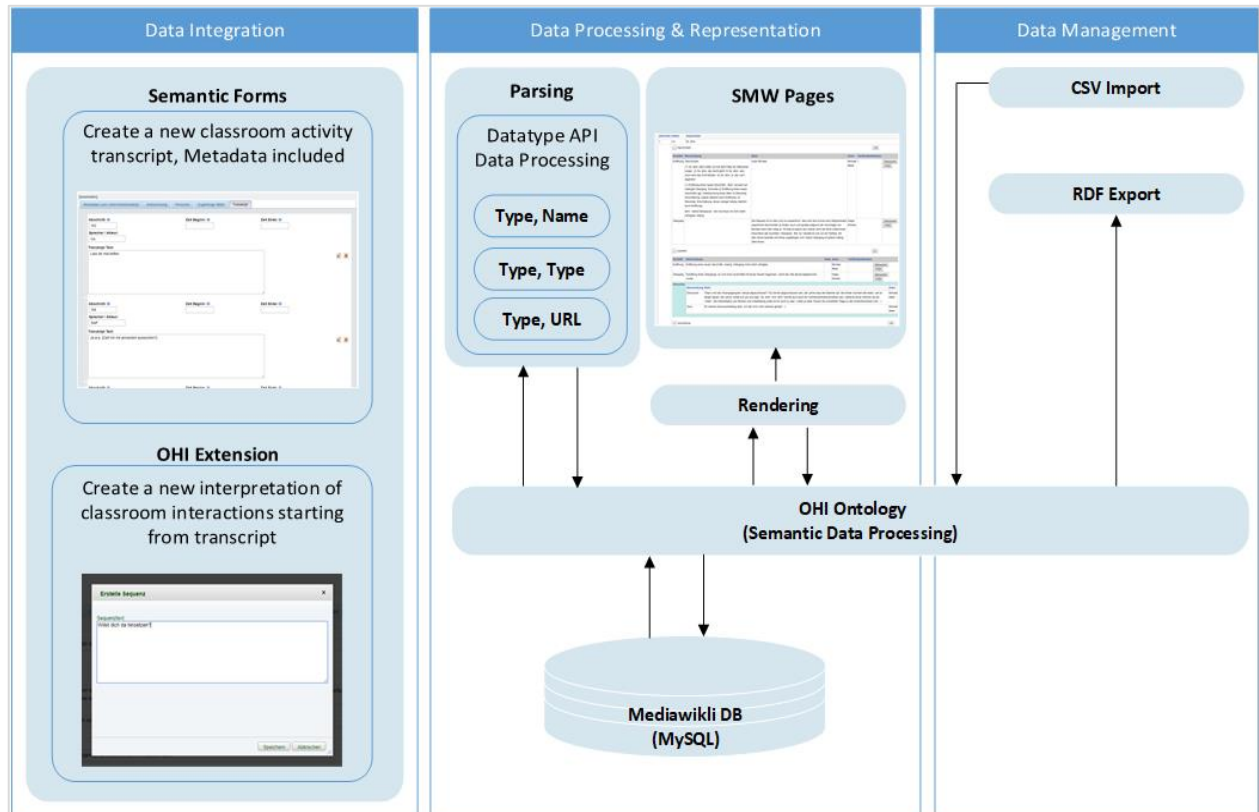


Figure 1: Overall Architecture

The early usage of the VRE prototype has pointed out two important aspects:

- The increased participation of the end-user as a result of the incremental awareness and gained experience was converted into knowledge. In turn, the knowledge has been shared within the community and after a while, formalized and externalized.
- The learning curve of system users started from very low and went up in time. After a while, the semantic technologies supported by Semantic MediaWiki were very well understood by researchers.

The early prototyping method has proved to be successful and opened new insights from testing data and other opportunities for further research (see Figure 7).

4. ARCHITECTURE

The following section describes the architecture of our VRE (see Figure 1). The VRE development process is stepwise, we explain first how the data is collected via the semantic forms and how specific templates and forms are integrated into the SMW. Secondly, we describe how each entity is processed and how each entity is represented on a wiki page.

4.1. Data collection and integration

SMW use the content templating mechanisms inherited from MediaWiki, under the name of seeding wiki pages. Using templates is useful for structured input, allowing non-technical users to add annotations to the wiki without

too much effort. SMW use aggregation of properties results into templates, and then an aggregation of templates results into a semantic form (see Figure 2).

Semantic forms give an intuitive interface to templates:

- Allow the creation of templates using forms for adding and editing pages
- Are defined using a mini-scripting language
- Are created on-the-fly, based on existing data, form definition and the templates that the form outputs



Figure 2: SMW templates and forms

Our VRE aims to foster collaboration and facilitates domain experts to edit data easily using our extension. Domain experts with low level of technical expertise can easily edit and create new data on the wiki using semantic forms. For this purpose, SMW also provides an “Edit with Form” option, which allows users to edit each page via user-friendly forms as depicted in Figure 3 .

Our system uses two kind of templates aggregated into one semantic form. The first template allows collection of general description metadata of the transcript: descriptive metadata of actors, contextual description metadata and administrative metadata. The second template collects the sequences which represent the body of transcript and the target of researcher’s interest.

Other templates are used in the process of interpretation. Each of the annotation types has its own underlying

The screenshot shows a web interface for editing transcript data. At the top, there is a navigation bar with tabs: "Metadaten zum Unterrichtstranskript", "Aufzeichnung", "Personen", "Zugehörige Bilder", and "Transkript". Below the tabs, there are two main form sections, one for section 153 and one for section 154. Each section contains the following fields: "Abschnitt:" (with a dropdown arrow), "Sprecher / Akteur:" (with a text input), "Zeit Beginn:" (with a time input), "Zeit Ende:" (with a time input), and "Transkript Text:" (with a large text area). The text area for section 153 contains "Lass dir mal helfen." and for section 154 contains "Ja-a-a. {Darf ich mir jemanden aussuchen?}". There are also small icons for adding and deleting entries.

Figure 3: Edit with Forms

template definition, allowing users to annotate the research object in a layered manner.

4.1.1 Objective Hermeneutic Interpreter Extension

In order to support the endeavor of researchers, we designed and developed a new SMW extension: Objective Hermeneutic Interpreter Extension (OHI). This extension allows users to follow the natural workflow they are accustomed to on paper-based environments.

The OHI extension might support the CA method also and states this in OH Protocol. OHI extension uses the OH Protocol that enforces certain rules (functional and for management). OH Protocol is a wiki page in a special namespace (MediaWiki).

In the first step, the research object is the transcript, represented as a wiki page. In this step, the templates are setup via OH Protocol. Tool supports automatically templates creation into wiki. OHI uses in the background semantic templates and creates additional entities in the wiki.

The next steps are related with the interpretation process and fully supported by our extension via semantic annotations. OHI extension supports annotations management at many levels:

- First level: annotations of transcript sequences, as base for interpretation. The result is an interrelated collection of wiki pages and a user interface for interpretation.
- Second level: first level of annotations allows to be annotated again, using 4 types of annotation, in conformity with Objective Hermeneutic Protocol.
- Third level: the second level of annotations refines the second level of annotations, and opens an additional space for users' collaboration and debates.

The OHI extension is responsible and also gives the opportunity for creation of *hypothesis of the structure of the case, related to the theoretical framework*. Depending on the research method, OHI allows semi-automatic creation of hypotheses wiki pages at the setup of the case (OH) or not (CA). These wiki pages will serve as base for research reports and scientific paper writing.

The OHI tool is responsible with the user interface. Using semantic queries and built in functions, the extension presents different data views to researchers: transcript view, interpretation data management, personal view, filtered open a semantic dimension view, hypotheses view. Inheriting MediaWiki facilities, a friendly multi-lingual user interface is very easy to be implemented in VRE. In this paper we present only the German interface.

4.2. Data processing and representation

Each of the transcripts has been mapped to a wiki page. The translation of the template instance to a wiki page source is accomplished by the parsing phase. Data collected from users via semantic forms are transformed within the wiki for appropriate representation (see Figure 3).

The rendering phase is responsible of how each entity is represented in a wiki page. The rendering phase is accomplished by the corroborative actions of MediaWiki template engine, which translate template calls into wiki text and SMW engine which translate template calls parameters into semantic properties and typed entities.

Par example:

```
{{Transcript
|SequenceNumber=2
|Actor1=Sm18
|Actor2=Lm
|Beginning Time= 5 October 2006 08:30:00
|Ending Time = 5 October 2006 08:31:00
|TranscriptText=Keiner.
}}
```

This is a general template call, using mandatory and optional parameters. Depending on the use-case, parameters like: Acteur2, Beginning Time and Ending Time may be omitted. These parameters are rendered to semantic properties and linked to the transcript itself. OHI tool and SMW services model domain ontology in wiki.

4.2.1 Overall Domain Ontology

Ontologies are designed to define concepts and relationships related to the knowledge base. The methodology described in this paper helped designing and structuring two different OWL ontologies. The first is the OH Core ontology representing the conceptual model defining the main topics, research fields, activities, communities, users, and technologies managed. The second is the classroom interaction ontology, being the conceptual model of one specific VRE topic, used to model the wiki pages about OH knowledge.

In SMW, semantics are formally defined using the OWL DL based upper level ontology: *Semantic Wiki Vocabulary and Terminology (SwiVT)*¹⁰. This ontology gives the possibility to reuse and aggregate knowledge between different semantic wikis. This ontology includes the most basic terms involved in the markup metadata model used by Semantic MediaWiki.

Our OH data model can be mapped to the vocabularies like: Wf4Ever Research Object Model¹¹ (ro), Object Reuse and

¹⁰ <http://semantic-mediawiki.org/swiwt/>

¹¹ <http://wf4ever.github.io/ro/>

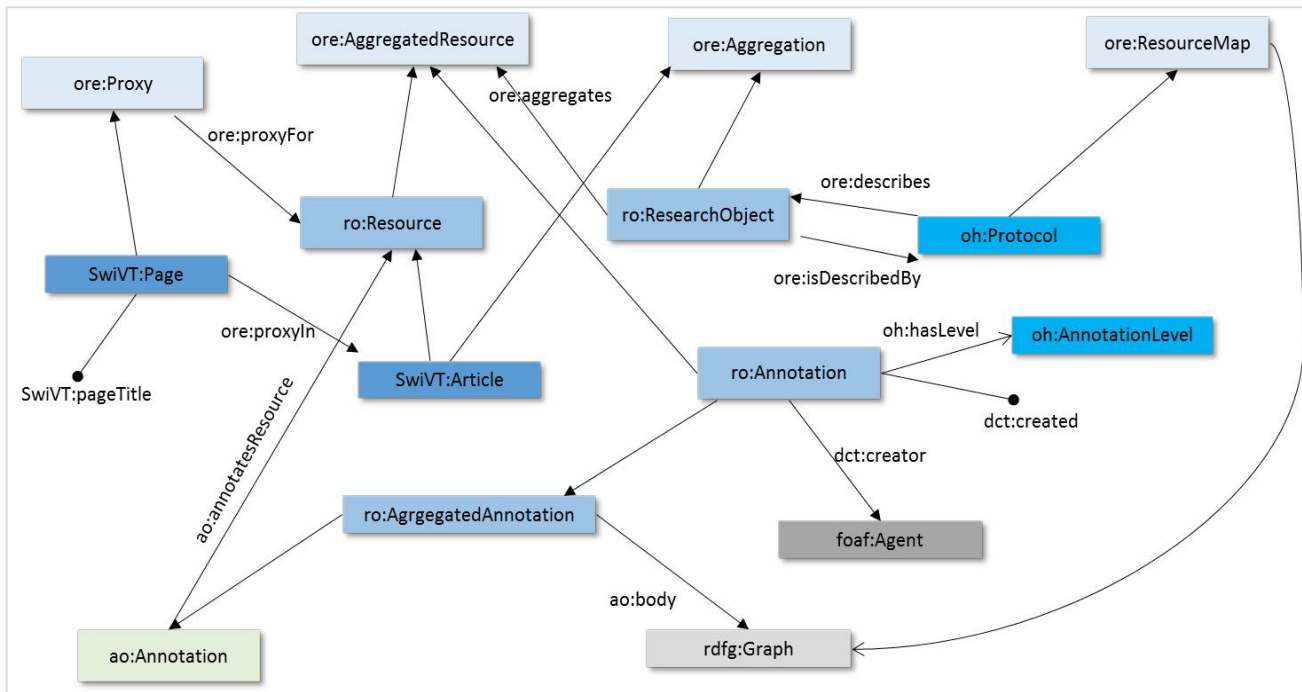


Figure 4: Overall OH Ontology

Exchange¹² (ore), Named Graphs¹³ (rdfg), Web Annotation Data Model¹⁴ (ao), Friend of a Friend¹⁵ (foaf) and Semantic Wiki Vocabulary and Terminology (SwiVT).

We extended above-mentioned ontologies and created the domain ontology, depicted in Figure 4 (oh).

The central entity in our ontology is the Research Object. The Research Object is an aggregation of resources, enriched with aggregated annotation. A research object aggregates a few resources expressed as wiki pages, that are used and/or produced in a given dialogic transcript interpretation using OH research method (OH) and is described in a oh:Protocol (OH model).

The RO model uses two existing vocabularies to provide aggregation and annotation functionality. Object Reuse and Exchange (ORE) is used for specifying aggregation of resources and the Annotation Ontology (AO) is used to represent annotations. OH model extends and refine the annotation introducing oh:AnnotationLevel, necessary to express OH interpretative act. As a subclass of ro:SemanticAnnotation the ao:body must point to an rdfg:Graph which contains the actual annotation.

During the annotation process, the research object structure becomes an aggregation of interlinked entities. As result, the content granularity changes and becomes finer, while content typologies become wider.

4.3. Data management

The transcript data can be automatically collected via a CSV import step also and translated into wiki syntax script. Using the SMW facilities, collected data may be exported in RDF format to other external application or other semantic wikis.

5. USE CASE FOR CLASSROOM INTERACTION QUALITATIVE DATA ANALYSIS

A group of researchers distributed across Germany and interested in classroom interactions used the VRE in practice over several months. The group has been involved in the design and testing processes.

An initiative to create a digital archive of transcripts has resulted in a database and related pdf files. Insufficient quality metadata were collected and made available.

Prior of using VREs, researchers used paper based transcriptions and all the process of interpretation was manually carried out. The entire process was tedious, requiring face-to-face meetings and had resulted in a several numbers of unstructured papers along with other separated and unstructured digital sources, like media files.

The VREs equipped with OHI tool give the opportunity to use a unique framework, to structure information, to collect and inter-link the data along with semantic annotation. All the process is collaborative and fully supported by SMW facilities and OHI tool.

¹² <https://www.openarchives.org/ore/1.0/datamodel>

¹³ <https://www.w3.org/2009/07/NamedGraph.html>

¹⁴ <https://www.w3.org/TR/annotation-model/>

¹⁵ <http://xmlns.com/foaf/spec/>

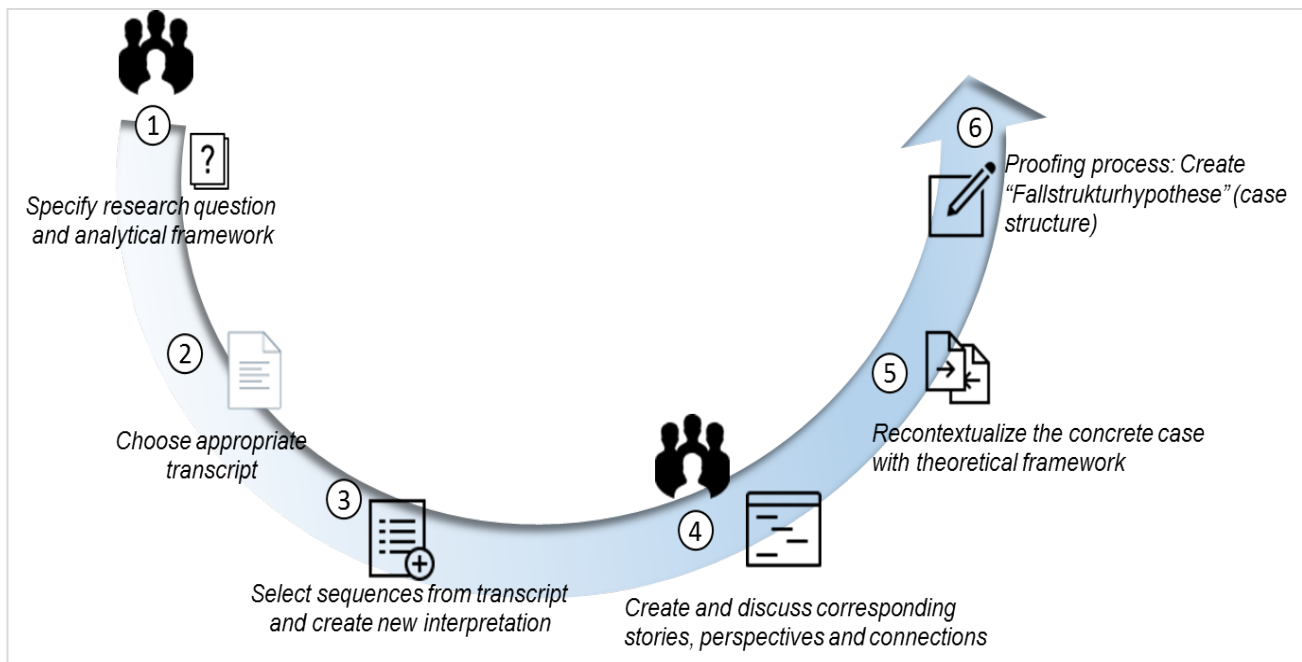


Figure 5: The Objective Hermeneutic Protocol workflow

The classroom interaction ontology, as an instance for OH Core ontology is depicted in Figure 6. The annotations space is layered and connected with the use-case. Each classroom activity was transcribed verbatim by a researcher or trained research assistant and carefully checked by either the junior researcher or the lead researcher.

Starting from these pdf documents, the transcripts were subsequently digitized in SMW based VRE using semantic class instances as wiki pages and properties of the domain ontology. The wiki page transcript contains general descriptive metadata of the teacher, students, classroom and subject of the lesson. Sensitive metadata, like personal data has been anonymized. The transcript contains also administrative metadata of the context in which the record has been transcribed (date of record, the person responsible with transcription, person who checked and who corrected the transcript). Also, the transcript might contain media files like: images of the classroom, images of the blackboard and sound files used in tuition. These files are uploaded along with metadata and linked to the transcript wiki page.

The classroom dialogue representation in wiki consists in creation of sequences; therefore, every single action of an interaction will be defined as a sequence.

As it is specific for the classroom dialogic interaction, the emphasis is on the first 'actor', aka the person who is speaking. The TranscriptText parameter means the action in which the 'actor' is involved.

Each of the sequences are parsed and translated as a template call that links the main part of the transcript (descriptive metadata) with the dialogic part (metadata of sequences). At the data level, each of sequences is represented by a set of semantic properties. The whole transcript now consists of calls of different SMW templates. Starting from this workflow step, the OHI extension supports the qualitative analysis of the transcript, as well as the metadata filtering and visualization.

The central procedure of Objective Hermeneutic method is the sequential analysis of the interaction written down in a protocol, named Objective Hermeneutic Protocol. These sequences then will then be investigated by the same rules to reproduce and examine a case structure [14]. In a subsequent step, researchers rely on the OHI tool to analyze the transcript in conformity with Objective Hermeneutic Protocol. Simultaneously they create the hypotheses. The lead researcher was fully involved in both types of analyses in order to allow for comparison between the results from these different analytical approaches.

The interpretation process is described in detail as follows (see Figure 5): After specifying the research question and choosing an appropriate interaction protocol (in this case, the transcript of classroom teacher-student interactions), the researchers start the sequential analysis of the interaction by selecting individual sequences. For every individual sequence of interest researchers formulate context stories.

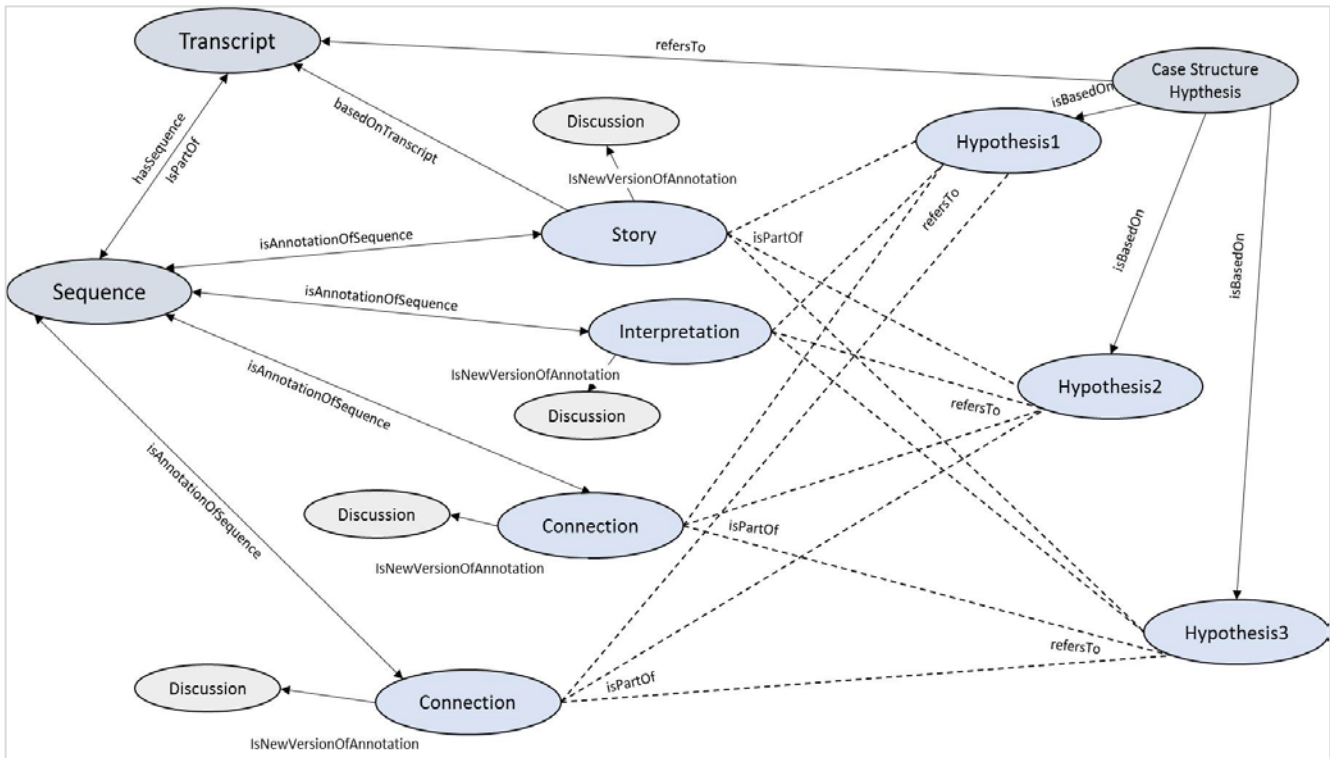


Figure 6: The classroom interaction ontology

Further, analyzing differences or commonalities of these context stories it may develop perspectives and connections. At the data level, they create multi-level and typed instances of semantic annotations, automatically linked with the sequence of the transcript. OHI automatically creates additional entities of type as Story, Connection, Contextualization and Discussion.

OHI tool allows editing, changing and deleting these annotations, under certain restrictions like ownership. The visualization of the interpretation process is based on the rewarding feature of SMW, as a result of semantic queries. The visualization of the annotations is enhanced by OHI tool that allows filtering upon one or multi-dimension properties and it is customizable.

Apart from genuine wiki page feature of versioning, OHI tool supports creation and management of an additional level of annotation versioning, with the purpose of refining the knowledge embedded in the annotation wiki page and gives the chance of controversies.

OHI tool also supports the concrete case reconstruction, that is based on these developments (annotations) and re-contextualization. After the reconstruction of the sequential analysis of the case structure completes, researchers start to create a case structure hypothesis.

Based on this interpretative act, a detailed case structure is created which describes the conflicting motivations, interests and interactions of the actors. The proofing

process ends with creation of case structure hypothesis and may specify, confirm or to reject the former hypotheses. Case structure hypotheses is a separate wiki page, automatically seeded by OHI tool with references to the wiki page annotations as result of the interpretative act. Research cycle ends with research paper writing based on case structure hypotheses.

6. DISCUSSION AND FUTURE WORK

Interpretive research relies heavily on qualitative data, but can sometimes benefit from including quantitative data as well. The work presented in this paper offers to researchers in Social Science, Educational Research and Humanities the appropriate framework to carry out their research. Lightweight semantic technologies introduced by SMW offers the possibility of a mixed-mode designs that combine qualitative and quantitative data which are often highly desirable. The joint use of qualitative and quantitative data may help generate unique insight into a complex social phenomenon that is not available from either type of data alone.

In a VRE based on SMW, users' participation is layered upon the level of technical skill the user needs in order to use the wiki and contribute to the ontology.

The use of SMW as technological support for our VRE enriches the possibilities for researchers to also collect quantitative data, as result of IR technology (semantic queries). This feature requires a higher (professional) usage of SMW facilities.

This possibility gives other valences to their research and opens opportunities to assess the collaboration process and

infers new insights from data. For instance, Figure 7 illustrates the connection between sequence number and interpretation items. The example diagram proves that the most interesting point for researchers was the sequence 'S5'.

Limitation of the work

At the time of writing, the VRE we presented here offers only a limited support for argumentation and editing the case hypothesis to researchers. The researchers may select and include in different type of hypotheses the annotations, but cannot select very accurately a certain piece of annotation when they are editing the case structure hypothesis.

Another limitation of the actual version of VRE is the low granularity of the structure of argumentation in the case discussion. Also, an assistant to help researchers in their platform learning it might be helpful for the future versions.

Future works

As such, the knowledge base created as result of our work, is used to develop a 'grounded' theory about golden professionalism and creates a foundation for developing *pedagogy of professionalism*. As ongoing work, students in pedagogy are using several clones of the prototype VRE for better understanding the concepts of the theoretical research methodology of OH. The students' early positive feedback confirmed the researchers' feedback in our approach evaluation.

Argumentation is part of interpretation analysis and encouraging disagreement help stakeholders to expose a richer set of possibilities. That helps in identifying, processing and understanding the context and relationships between different aspects of the transcript.

Our intention is to introduce support for the argumentation ontology that should focus on the argumentation-related ontological entities. Also, by refining structure of argumentation, we can evaluate the controversy aspects of argumentation and export semantic data for further research in Social Science.

Our VRE is not subject limited to classroom interaction analysis. The VRE may be very easily adapted for evidence in criminal cases interpretation, or any domain of discourse where the conversational analysis or objective hermeneutic analysis may apply.

7. CONCLUSION

This work presents a model of VRE that facilitates the sequential analysis of dialogic interaction. This approach is theoretically grounded on Objective Hermeneutic (OH) method as specific method for researchers.

The VRE is implemented using collaborative VRE facilities, semantic technologies and bespoke tools.

The paper showed also how to model domain ontologies, without explicitly states this, structuring the domain of classroom student-teacher interaction knowledge managed by the wiki. The VRE based on SMW showed several advantages: it facilitates the task of structuring the knowledge according to ontologies, it guides users in annotating texts in an easy way and it provides means to process and reuse machine-readable knowledge.

Moreover, a particular methodology was successfully modelled within the VRE to directly support users in their research task, and indirectly to model the domain ontology. The outcome of the multidisciplinary research team is a VRE prototype, which has proven to be successful.

At the time of writing, the knowledge base of the prototype VRE is being used as practical framework to train students in pedagogy to understand the theoretical OH method.

In the future, the authors plan to include into the VRE architecture reasoning features and more querying features and to experiment it with both structured data (e.g., user-generated annotations encoded in RDF) and unstructured ones (e.g., wiki pages about argumentation).

Also, the author's intention is to structure the argumentation level, therefore the researchers may evaluate the controversy aspects of argumentation and export data for further research in Social Science.

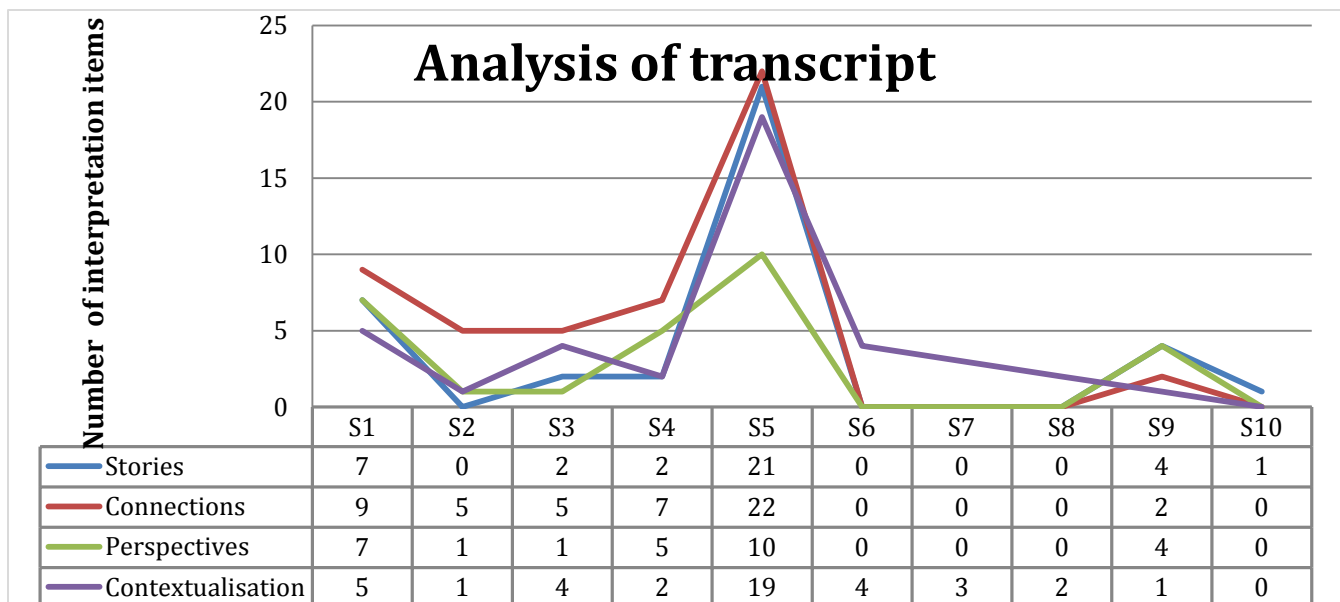


Figure 7: Connections between sequences and interpretation items

8. ACKNOWLEDGEMENTS

Initially, the realization of the research environment Semantic CorA was supported by the German Research Foundation (DFG); its further development for Objective Hermeneutics was funded in the context of the eHumanities Center CEDIFOR by the German Federal Ministry of Education and Research (BMBF) no. 01UG1416C.

9. REFERENCES

- [1] Arend, B., Sunnen, P. (2016). Dialogic Classroom Talk – Rethinking ‘Messy’ Classroom Interaction, EAPRIL Proceedings, ISSUE 2, ISSN 2406-4653
- [2] Boselli, R., Cesarini, M., Mercorio, F., & Mezzanzanica, M. (2013). Semantic Annotation of unstructured Wiki Knowledge according to Ontological Models. In WEB 2013, The First International Conference on Building and Exploring Web Based Environments. IARIA.
- [3] Candela, L., Castelli, D., & Pagano, P. (2013). Virtual Research Environments: An Overview and a Research Agenda. In *Data Science Journal*, Volume 12, 10 August 2013.
- [4] Drucker, J. (2013). Performative Materiality and Theoretical Approaches to Interface. In *digital humanities quarterly*. 7 (1). <http://www.digitalhumanities.org/dhq/vol/7/1/000143/00143.html>
- [5] Flick, U. (2005). Qualitative Research in Sociology in Germany and the US—State of the Art, Differences and Developments. In *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, 6(3). <http://www.qualitative-research.net/index.php/fqs/article/view/17>
- [6] Gülich, E., Mondada, L. (2008). *Konversationsanalyse: Eine Einführung am Beispiel des Französischen*. Tübingen: Niemeyer.
- [7] Houweling, L., Mayo, A. (2017). Educational Pedagogy of Professionalism, EAPRIL Proceedings, ISSUE 3, ISSN 2406-4653
- [8] Jaykumar, N.; Yallamelli, P.; Nguyen, V.; Lalithsena, S.; Thirunarayan, K. & Sheth, A. (2016) KnowledgeWiki: An OpenSource Tool for Creating Community-Curated Vocabulary, with a Use Case in Materials Science, World Wide Web 2016
- [9] Koole, T. (2015). Classroom Interaction. In K. Tracy (Ed.). *International Encyclopedia of Language and Social Interaction*. Wiley-Blackwell.
- [10] Krötzsch, M., Vrandečić, D. and Völkel, M. (2006): "Semantic Mediawiki," Proc. 5th International Semantic Web Conference (ISWC 06), vol. 4273, Springer, pp. 935-942.
- [11] Linell, P. (2009): *Rethinking Language, Mind, and World Dialogically: Interactional and Contextual Theories of Human Sense-making*, *Advances in cultural psychology: constructing human development* Advances in cultural psychology, IAP, ISBN: 9781593119959, pp. 72.
- [12] Maiwald, K. O. (2005). Competence and Praxis: Sequential Analysis in German Sociology. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, 6(3). Retrieved from <http://www.qualitative-research.net/index.php/fqs/article/view/21/46>
- [13] Martin, V. (2014): *Demystifying eResearch: A Primer for Librarians*, ABC-CLIO, ISBN 1610695216, 9781610695213, pp. 65-68.
- [14] Oevermann, U. (1981). Fallrekonstruktionen und Strukturgeneralisierung als Beitrag der objektiven Hermeneutik zur soziologisch-strukturtheoretischen Analyse. https://bildungsportal.sachsen.de/opal/auth/RepositoryEntry/10283810816/CourseNode/92942450281379/Oevermann_U._1981_Fallrekonstruktion_und_Strukturgeneralisierung.pdf (Last seen 04.16.2017).
- [15] Oevermann, U. (2000). Die Methode der Fallrekonstruktion in der Grundlagenforschung sowie der klinischen und pädagogischen Praxis. In K. KRAIMER (Hrsg.) *Die Fallrekonstruktion. Sinnverstehen in der sozialwissenschaftlichen Forschung*. Frankfurt am Main: Suhrkamp
- [16] Piez, W. (2010). Towards hermeneutic markup: an architectural outline. In *Digital Humanities 2010: Conference Abstracts*. <http://dh2010.cch.kcl.ac.uk/academicprogramme/abstracts/papers/pdf/ab-743.pdf>
- [17] Reichertz, J. (2004). Objective Hermeneutics and Hermeneutic Sociology of Knowledge. In U. Flick, E. v. Kardoff, I. Steinke (Ed.), *A Companion to Qualitative Research*. 290–295. London.
- [18] Schindler, C., Veja, C., Rittberger, M. and Vrandečić, D. (2011). How to Teach Digital Library Data to Swim into Research. In *Proceedings of i-Semantics conference*, September 7 - 9, 2011, Graz, Austria, ACM 978-1-4503-0621-8, pp. 142-149.
- [19] Schmied J. (1993) "Qualitative and quantitative research approaches to English relative constructions". In Souter C. & Atwell E. (eds.): 85-96
- [20] Voss, A. and Procter, R. (2009). Virtual research environments in scholarly work and communications. In: *Library Hi Tech*. 27, 2, 174-190. DOI = 10.1108/07378830910968146
- [21] Wegerif, R. (2008). Dialogic or dialectic? The significance of ontological assumptions in research on educational dialogue. *British Educational Research Journal*, 34, 3, pp. 347-361.
- [22] Wernet, A. (2000) *Einführung in die Interpretationstechnik der Objektiven Hermeneutik*. Opladen: Leske & Budrich
- [23] Wernet, A. (2014): *Hermeneutics and Objective Hermeneutics*. In: Flick, U. (Ed.) *The Sage Handbook of Qualitative Data Analysis*. London: Sage, pp. 234 - 246.