

# An Interactive Working Tool for Qualitative Text Analysis

Dong Han\*, Kilian Stoffel\*

\*Information Management Institute, Pierre-à-Mazel 7, CH-2000 Neuchâtel, Switzerland  
dong.han, kilian.stoffel@unine.ch,  
<http://www2.unine.ch/imi>

**Abstract.** In order to address the increasing needs of qualitative studies in numerous domains, a working tool is proposed in this paper. The research was motivated by the deficiencies of semantic analysis and interoperability found in industrial products widely used. Our tool is designed to be dynamic and flexible and allows the annotation of original documents. It provides analytical mechanisms for acquired knowledge while offering reasonable interactivity, reusability and portability. Currently it is being used by the research teams working on environmental studies. Based on real studies we can conclude that it integrates well with business processes and leads to very satisfactory results.

## 1 Introduction

Based on the analysis of a large number of applications in different domains, we were able to discover the phenomenon that academia and enterprises are both confronted with in the same way – information is getting more complex not only because of the increasing amount of data, but also the format used to represent it. An increasing amount of information is getting to be provided in the form of qualitative descriptions. The traditional way in which analysts apply quantitative studies whilst domain experts deal with qualitative information is not appropriate any more. There is a clear need to provide an integration of them both from theoretical and applied points of view.

The contribution of this paper <sup>1</sup> is to present a working tool by making a proposition for how to overcome the deficiencies identified in a thorough analysis. The goal is to assist the experts, including researchers in specialized domains as well as practitioners in enterprises, in their qualitative analysis tasks. It aims to provide a user-friendly graphical interface, interactive functional operations, effective semantic mechanisms and flexible data persistence. These features contribute to the uniqueness and advantages of our system compared to its counterparts.

---

<sup>1</sup>This work is supported by Swiss National Science Foundation(SNSF) project "*Formal Modelling of Qualitative Case Studies: An Application in Environmental Management*" (Project No. CR2112\_132089/1).

## 2 State of the Art

Among all the branches of qualitative studies, grounded theory(GT) (Glaser and Strauss, 2008) is coming to play an important role. GT starts from data (Coleman and O’Connor, 2007) and ends with theories based on codes and concepts. Meanwhile, another concept, ontology (Gruber, 1993), has been used in different applications to represent knowledge. Presently, *ATLAS.ti* and *NVivo* are two representative products globally used for qualitative studies. Both of them provide operations and annotation on primary documents, but without further analysis per se. In Han and Stoffel (2011) we have proposed methodologies based on ontologies for qualitative case studies in the domains of sustainability providing the adequate analysis capacities. We concluded that we needed more appropriately designed systems to support diverse domains, not limited to those with some particular features. Furthermore we urge for a dynamic and flexible architecture. This paper addresses some of these elements while inheriting the utilization of GT and ontologies from Han and Stoffel (2011).

## 3 System Design and Implementation

The developed system, named *Qualogier*, is established with significant improvements over standard approaches. Figure 1 and 2 present respectively the architecture and screenshot of the system.

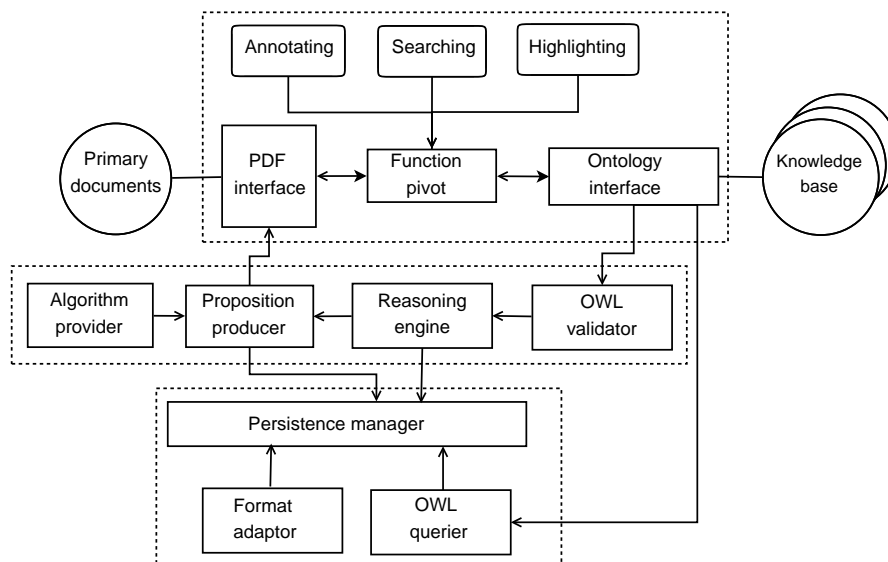


FIG. 1 – System architecture

Basic and annotational functionalities are provided for PDF documents based on ICEpdf (ICEpdf.org, 2011). In addition, we use the *OSGI* framework to register different bundles into a *function pivot* to support dynamic life cycles of these bundles instead of the static compositions. Ontologies are validated before they enter the semantic process for their conformity

to the OWL specification. We use *facade pattern* to manage all the algorithms involved to support effective semantic analysis and better extensibility. Since all the knowledge is stored in ontology files. Different formats such as HTML, XML, and CVS are plugged into our system as *adaptors* for persistence. For better user experience, modified information is saved automatically without user operations once the focus is relocated.

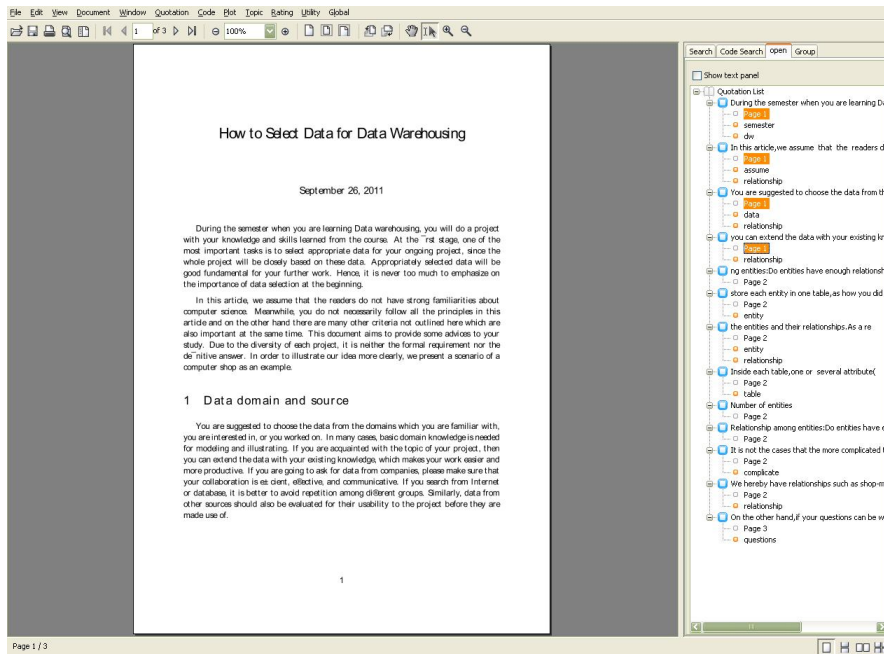


FIG. 2 – Main interface

## 4 System Evaluation and Scenario Applications

To characterize the features and advantages of Qualogier compared to ATLAS.ti, we propose a generic framework, not limited to these two software packages, to evaluate qualitative software by stressing the underlying criteria: (1) *Interactivity*: The operations of Qualogier are more consistent with other PDF viewers to ensure facility and seamless transition for the users. (2) *Interoperability*: In contrast to ATLAS.ti, our system is able to integrate with other projects, domain experts and a number of software packages because of its standard oriented approach. (3) *Portability*: Qualogier runs on different platforms while ATLAS.ti is limited on MS Windows.

Qualogier is currently used in our cooperative research teams for analyzing the impact of environmental factors on financial firms. More than 30 people with different operating systems, languages and background are working for the projects using this system. As shown in figure 3, we refined some simple but typical use cases with *BPMN* based on their daily work as a potential demonstration.

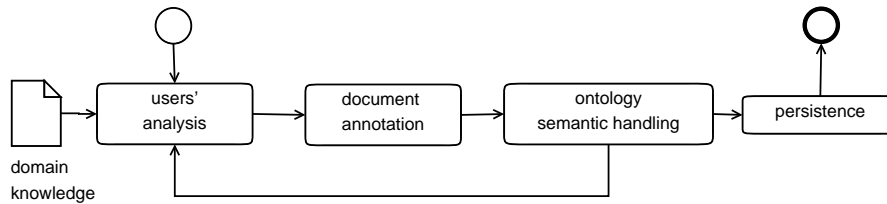


FIG. 3 – A use case process with Qualogier

## 5 Conclusion and Future Work

In this paper, an ontology based analytical tool is presented for qualitative studies in different domains. As the further development plan, we will use Latent Dirichlet Allocation to optimize the proposition mechanism to find the trade-off between accuracy and efficiency. Another task is to formalize and modelize the system to set up the scenario from the users view.

## References

- Coleman, G. and R. O'Connor (2007). Using grounded theory to understand software process improvement: A study of irish software product companies. *Information and Software Technology* 49, 654–667.
- Glaser, B. and A. Strauss (2008). *The Discovery of Grounded Theory: Strategies for Qualitative Research*. Transaction Publishers.
- Gruber, T. R. (1993). A translation approach to portable ontology specifications. *Knowledge Acquisition* 5, 199–220.
- Han, D. and K. Stoffel (2011). Ontology based qualitative case studies for sustainability research. In *Proceedings of the AI for an Intelligent Planet*.
- ICEpdf.org (2011). Homepage of icepdf. <http://www.icepdf.org>.

## Résumé

Pour répondre aux besoins croissants dans plusieurs domaines des études qualitatives, un outil a été proposé dans ce papier. Ce développement a été motivé par les lacunes présentes dans l'analyse sémantique et l'interopérabilité des produits industriels utilisés qui y est omniprésente. Cet outil est conçu sur une architecture dynamique et flexible et permet l'annotation de documents originaux. Il fournit des procédés d'analyse des connaissances acquises, tout en offrant une interactivité raisonnable, la réutilisabilité et la portabilité. Actuellement elle est utilisée par les équipes de recherche pour les études environnementales. Sur la base des projets pratiques, nous concluons qu'il s'intègre bien avec les processus métier et conduit à des commentaires satisfaisants.