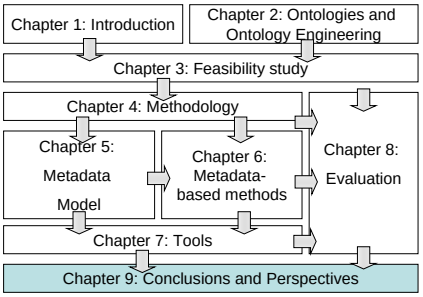
9 Conclusions and Perspectives



In this chapter we summarize the results of our work and outline its implications for the Semantic Web community. Section 9.1 pinpoints the main contributions of this thesis and outlines the planned future work. Section 9.2 puts the proposed solution in a broader context, scrutinizing further directions of research for ensuring the feasibility of ontology reuse at Web scale.

9.1 Summary and Future Work

Compliant to the general understanding of reuse in adjacent engineering disciplines *ontology reuse* can be defined as *the process in which existing ontological knowledge is used as input to generate new ontologies*. It is widely acknowledged that an efficient and effective operation of the reuse process is a pre-condition for the large scale take-up of semantic technologies. However, the challenges associated with achieving this objective—in Semantic Web context or beyond—are also well-known, given the inherent limitations of realizing highly reusable, commonly agreed knowledge conceptualizations. The current state of the art in the ontology engineering field states the need for additional instruments to increase the reusability of existing ontological sources in new application settings, and to aid humans in carrying out this process. This thesis is targeted at the achievement of these goals.

In order to spot the bottlenecks of current ontology reuse processes we performed an extended feasibility study comprising both self-conducted case studies and an in-depth review of recent literature in the area. The feasibility study was concluded by an analysis and a specification of requirements for methodologies, methods and tools. It fundamentally pointed out the role of context for the success of a reuse endeavor, both with respect to the process performance, and its outcomes.

Building upon the results of this field inquiry we designed a fine-grained methodology for the reuse of Web ontologies. The methodology provides a detailed description of the process, emphasizing application-oriented best practices and guidelines, as well as methods and tools which are likely to be useful to partially automatize it.

We created a metadata model for ontologies, which contributes to an European initiative for standardization in this field and is already an integral part of several ontology reuse tools. The metadata model formed a viable basis for the development of a method for ontology evaluation. This demonstrated with promising results how to operationalize the principles introduced in our reuse methodology using a semantics-aware extended information retrieval

model. Further on, we explained how the proposed metadata can be utilized to inject a context-dependent behavior into ontology merging and integration. In a joint effort with other researchers in our institute we elaborated on a rule-based approach to the non-trivial question of how to select the appropriate ontology merging strategy in a particular application setting.

The core theoretical considerations of our research have been prototypically implemented within the PROMI project. PROMI is a Java-based platform aiming to provide methodological and technological support to participants in a reuse process. It has been developed at the Free University of Berlin and has been recently published as an open source project. A further tool, OntoMeta, has investigated heuristics for an automatic generation of metadata information about ontology with positive results.

Our research has been evaluated using a combination of quantitative and qualitative evaluation approaches. Besides professional reviews and goal-free evaluation, we applied the case study validation methodology in order to compare the operation of reuse processes with and without the proposed methodology in a real-world case study in ontology engineering. The associated methods have been evaluated in two in situ user studies. The results of the overall procedure demonstrated that for a class of application which is considered to be representative for the success of Semantic Web technologies—information retrieval and semantic annotation—our approach improves ontology reuse as regards the invested efforts, the user-perceived process operation efficiency and the fitness of use of the reuse outcomes in the target application setting. The PROMI implementation was tested using common techniques in software engineering.

We intend to continue the research conducted in this thesis in several directions:

Methodology: at methodological level the process description could benefit from the consideration of further application scenarios which are rapidly emerging with the recent uptake of semantic technologies in industrial setting, and from a more detailed analysis of the existing ones. Further on, one could use the proposed ontology reuse methodology and the case studies carried out in order to derive *ontology reuse patterns* which would provide a proved and tested basis for the development of high-quality ontology-based applications.

Methods: a first direction of research and development relates to the future of the ontology metadata model. This is currently refined in collaboration with several European research institutions towards a fully-fledged schema for the ontology engineering field. These efforts are planned to be concluded by a submission to the W3C consortium. The merging and integration method could be extended with an alternative approach for processing dependency rules and with a more detailed description of the involved matching algorithms. Similarly, we will investigate the benefits of putting more emphasis on the formal description of the application systems and scenarios using ontologies in the context of the ontology evaluation task.

Tools: the PROMI framework necessitates further testing, documentation and functionality extensions in order to increase its usability and applicability in a wide range of situations. A concrete R & D project has already been approached in conjunction with the MOMA framework. Further on, a re-engineering of the task-oriented components in

the direction of Web services is likely to contribute to the modularization of the system, thus increasing its extendability. A last direction of future work is situated in the area of ontology learning. Such services could be linked to the PROMI framework in order to increase the number of input types accepted by the merging component.

9.2 Outlook

The last section of this thesis outlines a research agenda for ontology engineering compiled from the experiences and lessons learned in our PhD research.

Provision of ontological content : The question of how to automatically build ontologies from less structured forms of organizing knowledge is being intensively researched in the Semantic Web community. In particular we experience a multitude of methods and tools aiding humans in extracting ontological knowledge from textual resources. While these achievements provide the fundamental building blocks necessary for extracting Semantic Web information from the conventional Web, more research is needed in order to integrate ontology learning (and knowledge acquisition in general) into ontology engineering. Further on, our experiences indicated that the question of creating ontologies from semi-structured resources, be that XML, databases or classification systems in many domains of interest, is only marginally investigated by the community. A holistic approach to this vital issue for the success of the Semantic Web is required.

Ontology discovery : Even after fifteen years of intensive research and development, the issue of how and where to find ontological resources is solely partially solved. There is a need for fully-fledged ontology search engines and for repositories, which cover a representative amount of ontologies and offer advanced capabilities for searching, rating and describing them.

Ontology engineering methodologies : Prior to answering the question whether an existing ontology is reusable in a new context, ontology engineers require means to reliably perform a feasibility study indicating that ontology reuse is a viable ontology engineering strategy at all. Ontologies can be built using various approaches. The most common classification differentiates between building by scratch, ontology reuse and ontology learning [80]. It is likely that fragments of the same application ontology could be constructed using several of these alternatives. In order to efficiently carry out such a complex process we need quantifiable means to analyze alternative engineering strategies from a cost/benefit perspective. Orthogonal to this issue, there is a need for fine-detailed methodologies explicitly coping with this diversity.

The uptake of ontology-driven technologies : A final critical aspect for the industrial uptake of ontologies and ontology-driven applications is the availability of stable, easy-to-use tools enabling this process. This implies on one hand established ontologies modelling domains of interest which are likely to be relevant for a wide range of business sectors. On the other hand, it postulates best practices and guidelines explicitly targeted at industry practitioners, which aid them in developing and deploying such

resources. Finally, there is a strong need for mature tools supporting various stages of an ontology life cycle, which can be easily integrated into the existing environments.

Advocating the efficiency gain arisen by explicitly dealing with the context-sensitive nature of knowledge reusability, this thesis describes an ontology reuse integrated environment aiming to answer to this problem at methodological and technological level. As evidenced by our experimental results, the proposed methods and tools bring benefit to ontology developers and users in real-world application settings with respect to the user-perceived process operation efficiency, the relative development costs, and the fitness of use of the reuse outcomes. Feasibly reusing the vast amounts of domain knowledge increasingly available on the Web in form of ontologies, complemented by the enumerated research issues, can contribute in the next years to the prognosticated evolution of semantic technologies towards mainstream IT.