

# Lightweight Community-Driven Ontology Evolution

Katharina Siorpaes

Digital Enterprise Research Institute (DERI), University of Innsbruck, Austria  
katharina.siorpaes@deri.at

**Abstract.** Only few well-maintained domain ontologies can be found on the Web. The likely reasons for the lack of useful domain ontologies include that (1) informal means to convey intended meaning more efficiently are used for ontology specification only to a very limited extent, (2) many relevant domains of discourse show a substantial degree of conceptual dynamics, (3) ontology representation languages are hard to understand for the majority of (potential) ontology users and domain experts, and (4) the community does not have control over the ontology evolution. In this thesis, we propose to (1) ground a methodology for community-grounded ontology building on the culture and philosophy of wikis by giving users who have no or little expertise in ontology engineering the opportunity to contribute in all stages of the ontology lifecycle and (2) exploit the combination of human and computational intelligence to discover and resolve inconsistencies and align lightweight domain ontologies. The contribution of this thesis is a methodology and prototype for community-grounded building and evolution of lightweight domain ontologies.

## 1 Problem Statement

Even though one can observe highly active research in the areas of ontologies and the Semantic Web, there is a lack of up-to-date domain ontologies for many fields. The complexity of reality makes building ontologies difficult in several ways: (1) Ontology engineering methodologies tend to focus on formal means for specifying ontologies. However, in order to describe the *intended* meaning of ontology elements, informal means are also needed. (2) Many relevant domains of discourse, such as e-commerce, show a high degree of conceptual dynamics, i.e. it is hard to keep up with the pace of change in reality [1]. (3) A prerequisite for using an ontology and thus committing to its view of the world is to be able to understand the meaning of concepts and relations [1]. This is problematic for many users, since they cannot easily figure out what they would be committing to when using a particular ontology file from the Web, as the discourse that lead to a certain specification is not available. (4) Most ontologies are built by a group of engineers and the user community does not have direct control over the evolution of the ontology.

## 2 Motivation

A community-oriented approach has several advantages as compared to an isolated engineering-oriented approach, where a small group of engineers carefully constructs the ontology and releases it to the user community at a later point of time: (1) We believe that ontology building is inherently a collaborative task, for two reasons: first, ontologies are supposed to be community contracts [2, 3]; second, the combination of required domain expertise and modeling skills is more likely to be found in a group than in a single individual. (2) Timeliness of a domain ontology is important in rapidly changing domains. A community can keep up with the pace reality is changing more easily. (3) The user community is more likely to agree on a view of the world that is represented by the ontology. Therefore, it is likely that this community will also actually use and further develop the ontology as it is not a subjective conceptualization based on an outdated state of the world. (4) The burden of creating, maintaining the ontology can be shared by those benefiting from the ontology. Similarly, inconsistencies can be discovered and resolved faster. In [4], we provide evidence for the assumption that a community can be trusted with the task of building lightweight ontologies by showing that the conceptual reliability of Wikipedia articles is very high (93% of URIs convey a stable ontological meaning, even though Wikipedia is maintained by a highly diverse group of actors).

## 3 Approach and Design Principles

In the following we outline the design principles of this approach: **Tapping the “wisdom of crowds”**: We believe that a diverse community consisting of domain experts as well as knowledge engineers is smarter and more agile than only a small ontology development team. Instead of forcing one view of the world onto the user communities, we envision co-existence and interoperability of conflicting views and to support the community in achieving consensus. **Openness and multimedia richness**: Traditional ontology building environments usually impose quite high entrance barriers on a user while wikis allow many users to contribute easily with only basic Web-editing skills. The culture of Wikis is the underlying paradigm of this work. Additionally, we propose the use of multimedia elements to better convey the informal part of the intended meaning of a concept. **Integration of external data**: with the increasing popularity of “Web 2.0”, a wealth of data is available, which can be exploited for improving the process of ontology construction. **Combination of human and computational intelligence**: Building ontologies is a task that depends on human intelligence, both as a source of domain expertise and for producing a consensual conceptualization as well as aligning ontologies and resolving inconsistencies. We aim at developing functionality that combines human and computational power and thus supports users in achieving several ontology building tasks.

## 4 Methodology and Contribution

The envisioned research methodology is as follows: (1) Analysis of a trade-off between expressivity and agility of an ontology meta-model suitable for a broad audience. (2) Deployment of a suitable model for ontology evolution. (3) Integration of external resources to support lexical and multimedia-based enrichment: tagging, i.e. describing objects with keywords to enhance content retrieval, avoids the limitations known in ontology building as they allow the instant maintenance of tags. Statistical analysis of the large amount of data produced in tagging application can be used for ontology construction [5, 6]. Additionally, Web resources, such as Google, Wikipedia, or Freebase, and terminological resources, such as Wordnet, will be exploited for ontology construction and background enrichment. (4) Specification of user roles and processes supporting the community in achieving consensus. (5) Combination of human and computational intelligence: this includes discovering and resolving inconsistencies and aligning ontologies by combining matching algorithms with human intelligence, mapping Wikipedia special pages to an ontology building environment, and achieving stable releases of ontologies. (6) Application of various techniques for visualization of ontologies and user interfaces to foster comprehensibility. (7) Investigating user incentives for contributing to ontology building in a community-driven environment. (8) Implementation of a community portal based on the methodology and evaluation: the portal will be released to a community who will create and maintain several domain ontologies. The data will be logged and analyzed, in order to find out about the domain coverage, stability of concept definitions, cost, user participation, and precision of usage.

## 5 Related Work

The work related to this thesis can be divided into the following areas: **Traditional ontology engineering** methodologies are described and compared in [3]. **Collaborative and community-driven ontology engineering:** OntoSaurus [7] is a web-based tool for editing and browsing ontologies; however, it has no explicit support for collaborative or community-based work. Domginue [8] describes Tadzebao and WebOnto: Tadzebao supports asynchronous and synchronous discussions on ontologies. WebOnto complements Tadzebao by supporting collaborative browsing, creation and editing of ontologies. Vrandecic and colleagues [9] describe the DILIGENT knowledge process where ontology evolution and collaborative concept mapping are applied to deal with conceptual dynamics of domains. OntoEdit [10] is a collaborative ontology editing environment. Kotis and Vouros describe the HCOME methodology in [11], which is a “human-centered” approach: ontology development is defined as a dynamic process with a focus on ontology evolution. HCOME is a decentralized engineering model where everyone first formalizes her own ontology and shares it in a further step within the community. This is different to our approach as the ontology is not open to the community at all times. Braun and colleagues [12] present an ontology maturing process consisting of four steps: emergence of ideas, consolidation in communities, formalization,

axiomatization. They regard the evolution of an ontology as maturing from tags to formal ontologies via the community. Gibson and colleagues [13] distinguish between ontological and meta-ontological data and propose that more attention should be dedicated to the meta-ontological data and to allow the community to discuss ontologies at an early stage of development with a Web 2.0 user interface, where no formal model is required. They emphasize the integration of natural language. **Semantic Wikis:** [14] describe Makna, a Wiki engine that was extended with generic ontology-driven components that allow collaborative authoring, querying, and browsing Semantic Web information. IkeWiki [15] allows annotating links, typing of pages, and context dependent content adaptation. [16] have the objective to make the knowledge within Wikipedia, the online encyclopedia, machine-accessible by adding semantic information. Platypus Wiki [17] focus on the creation of RDF (instance) data, Platypus Wiki aims at augmenting a wiki with semantics. WikiFactory [18] is a framework that allows the automatic generation of domain specific wikis. The main difference to our work is that existing approaches aim at augmenting existing wiki content with semantics instead of using a Wiki-like infrastructure as an environment for collaboratively building ontologies.

## 6 Expected Impact

The final goal of the thesis is to develop a deployable prototype for wiki-based ontology construction by large, unsupervised audiences. The process of community-driven ontology construction and maintenance will be supported by background intelligence in the tool that helps the community by suggestions based on statistically significant patterns or external lexical or structural resources. Eventually, thus should hand back the control over the evolution of ontologies to the community in order to overcome the current shortage of domain ontologies.

**Acknowledgments.** I would like to thank Martin Hepp for supervising my thesis and his continuous support. The work presented in this paper has been funded by the Austrian BMVIT/FFG under the FIT-IT project myOntology (grant no. 812515/9284).

## References

1. Hepp, M., Possible Ontologies: How Reality Constrains the Development of Relevant Ontologies. *IEEE Internet Computing*, 2007. **11**(7): pp. 96-102.
2. Hepp, M., D. Bachlechner, and K. Siorpaes. OntoWiki: Community-driven Ontology Engineering and Ontology Usage based on Wikis. In *Proceedings of the 2005 International Symposium on Wikis (WikiSym 2005)*. 2005. San Diego, California, USA.
3. Gómez-Pérez, A., M. Fernández-López, and O. Corcho, *Ontological Engineering*. 2004: Springer.
4. Hepp, M., K. Siorpaes, and D. Bachlechner, Harvesting Wiki Consensus: Using Wikipedia Entries as Vocabulary for Knowledge Management. *IEEE Internet Computing*, Vol. 11, No. 5, pp. 54-65, Sept-Oct 2007.

Katharina Siorpaes

5. Van Damme, C., M. Hepp, and K. Siorpaes, *FolksOntology: An Integrated Approach for Turning Folksonomies into Ontologies*. In *Proceedings of the Workshop Bridging the Gap between Semantic Web and Web 2.0 at the ESWC 2007*. 2007, Innsbruck, Austria.
6. Specia, L. and E. Motta. *Integrating Folksonomies with the Semantic Web*. In *Proceedings of the European Semantic Web Conference (ESWC 2007)*. Springer LNCS: 2007. Innsbruck, Austria.
7. Swartout, B., et al., *Ontosaurus: a tool for browsing and editing ontologies*. In *Proceedings of the 9th Banff Knowledge Acquisition for Knowledge-based systems Workshop*. 1996. Banff, Canada.
8. Domingue, J. Tadzebao and *WebOnto: Discussing, Browsing, and Editing Ontologies on the Web*. In *Proceedings of the 11th Knowledge Acquisition for Knowledge-Based Systems Workshop*. 1998. Banff, Canada.
9. Vrandečić, D., et al., *The DILIGENT knowledge process*. *Journal of Knowledge Management*, 2005. **9**(5): pp. 85-96.
10. Sure, Y., et al., *OntoEdit: Collaborative Ontology Engineering for the Semantic Web*. In *International Semantic Web Conference 2002 (ISWC 2002)*. 2002. Sardinia, Italia.
11. Kotis, K. and G.A. Vouros, *Human-centered ontology engineering: The HCOME Methodology*. *Knowledge and Information Systems*, 2005. **10**(1): pp. 109-131.
12. Braun, S., et al., *Ontology Maturing: a Collaborative Web 2.0 Approach to Ontology Engineering*. In *Proceedings of the Workshop on Social and Collaborative Construction of Structured Knowledge at the 16th International World Wide Web Conference (WWW2007)*. 2007: Banff, Canada.
13. Gibson, A., K. Wolstencroft, and R. Stevens, *Promotion of Ontological Comprehension: Exposing Terms and Metadata with Web 2.0*. In *Proceedings of the Workshop on Social and Collaborative Construction of Structured Knowledge at 16th International World Wide Web Conference (WWW2007)*. 2007: Banff, Canada.
14. Dello, C., E. Paslaru Bontas Simperl, and R. Tolksdorf, *Creating and using semantic content with Makna*. In *Proceedings of the Wiki meets Semantics workshop at the ESWC2006*. 2006. Budva, Montenegro.
15. Schaffert, S., *IkeWiki: A Semantic Wiki for Collaborative Knowledge Management*. in *1st international workshop on Semantic Technologies in Collaborative Applications STICA06*. 2006. Manchester, UK.
16. Völkel, M., et al. *Semantic Wikipedia*. In *Proceedings of the 15th International Conference on World Wide Web (WWW2006)*. 2006. Edinburgh, Scotland.
17. Campanini, S.E., P. Castagna, and R. Tazzoli, *Platypus Wiki: a Semantic Wiki Wiki Web*. In *Proceedings of the 1st Italian Semantic Web Workshop Semantic Web Applications and Perspectives (SWAP)*. 2004. Ancona, Italy.
18. Di Iorio, A., V. Presutti, and F. Vitali, *WikiFactory: An Ontology-Based Application for Creating Domain-Oriented Wikis*. In *Proceedings of the 3rd European Semantic Web Conference, ESWC 2006*. 2006, Springer LNCS: Budva, Montenegro.