

# **An ontology-based approach and software for creating and maintaining semantic wikis**

**Alvaro Ortiz-Troncoso<sup>1</sup>**

<sup>1</sup>**Museum für Naturkunde - Leibniz Institute for Research on Evolution and Biodiversity**

Corresponding author:

Alvaro Ortiz-Troncoso<sup>1</sup>

Email address: alvaro.OrtizTroncoso@mfn.berlin

## **ABSTRACT**

Semantic wikis support research at the Museum of Natural History Berlin by providing services for developing and documenting collection management and imaging procedures or annotating archive material. On the other hand, some wikis are available on the Web, as publication medium for conference presentations or providing additional material for exhibitions. A total of 16 semantic wikis are currently online on the museum's servers. Semantic wikis combine the metadata and conformance aspects required by knowledge management, with the possibility to access and build-upon a corpus of collaboratively collected knowledge. On the down side, maintaining a semantic wiki is challenging. Coping with increasing complexity of the wikis led to the development of a methodical approach for simplifying the creation and maintenance of semantic wikis: modelling the semantic relationships underlying the information in the wiki as an ontology, and then automatically creating the wiki from the ontology constructs. Benefits gained: a methodical approach for creating and maintaining wikis greatly simplifies these tasks. Furthermore, reusing vocabularies and taxonomies throughout several projects becomes much more manageable.

## **1 INTRODUCTION**

Wikis are web platforms allowing for collaborative authoring of information. With certainty, the best known wiki is Wikipedia, but wikis are also being used to setup corporate intranets, private collaboration environments or knowledge bases. A semantic wiki is a wiki whose functionality has been extended to encode information in a structured and machine-readable way: information encoded semantically forms a "consistent logical web of data" (Berners-Lee et al., 1998), that can be used for enriching information with metadata, checking the conformance of information against an agreed upon schema, or accessing and combining data from different sources (Oren et al., 2006). Semantic wikis therefore combine the metadata and conformance aspects required by knowledge management, with the possibility to access and build-upon a corpus of collaboratively collected knowledge.

Semantic wikis have been used at the Museum für Naturkunde Berlin (Museum of Natural History Berlin) since 2015. A total of 16 semantic wikis are currently online on the museum's server. Most of these wikis are part of the museum's intranet, and support research by providing services for developing and documenting collection management and imaging procedures or annotating archive material. On the other hand, some wikis are available on the Web, as publication medium for conference presentations or providing additional material for exhibitions. An example of a publicly accessible semantic wiki running at the Museum of Natural History Berlin is the Wiki supplementing the Panda exhibition. The construction and provisioning of wikis has been the subject of two research projects at the Museum (Patzschke et al., 2016). The experience gained demonstrates that creating and maintaining a semantic wiki is challenging, a major difficulty being that the complexity of the information stored in the wiki tends to increase as the user base grows, and new ideas are integrated into the wiki's semantic structure (Kiniti and Standing, 2013).

## 2 METHODOLOGICAL APPROACH

Coping with this increasing complexity led to the development of a methodical approach for simplifying the creation and maintenance of semantic wikis. One way to create a semantic wiki is to model the semantic relationships underlying the information in the wiki as an ontology, and then to customize a wiki to reflect the constructs described by the ontology (Di Iorio et al., 2006).

At this point it might be necessary to clarify the concept of ontology. A basic definition of what an ontology is in the realm of knowledge engineering was given by Gruber and Borst (cited in (Corcho et al., 2003)): an ontology is made of clear-cut and formal definitions of the concepts in a domain, and these definitions are accepted by the community of practitioners of that domain. Another way to put Gruber and Borst's definition: an ontology is a collaboratively constructed, formalized mind map. Another definition of ontology can be gained from information theory. Pieterse and Kourie (2014) differentiate between five types of fundamental information structures: linear lists, hierarchical taxonomies, poly-hierarchical lattices, thesauri combining hierarchies with various kinds of relations between concepts, and ontologies, which extend thesauri with a set of rules for inferring new knowledge. This definition emphasizes the role of logical inference rules, and paves the way for the role ontologies play in machine learning (see Zhou (2007) for a road map).

Ontology engineering is a vast topic by itself, yet a simple methodology could follow these basic steps (Noy and McGuinness, 2001): obtaining domain information by consulting experts and gathering specific vocabularies (e.g. in the case of Museum of Natural History Berlin, zoological taxonomy, stratigraphical vocabulary and others); subsequently, modelling the ontology in an external tool such as Protégé (Musen, 2015), and storing the result as an RDF/XML file (Gandon and Schreiber, 2014). Software implemented at the Museum automates the conversion of the ontology file into web pages and forms in the wiki. This software is distributed under an Open Source License, in the hope that it can be useful for computer scientists researching computer-mediated collaboration and ontology engineering.

## 3 IMPLEMENTATION DETAILS

The process by which the conversion of the ontology into a wiki is automated is implemented as follows: the RDF/XML file representing the ontology is parsed into an in-memory semantic model. Then, for each object in the model, a data access object (DAO) is created using a factory class. A connection is established with the wiki's web service (API). Finally, each DAO object saves itself to the wiki, using the wiki's API. With an ontology as input, the software knows two commands: "import" creates the necessary wiki pages, "delete" removes the corresponding wiki pages. The layout of the wiki is customizable through templating: templates can be used to specify the order in which elements should appear, as well as to filter out or hide some elements. The software is written in Python, templates are in XSLT. The resulting wiki pages are compatible with the semantic extension to the popular MediaWiki software.

Ontology constructs are mapped to the wiki artefacts according to these rules: ontology classes are mapped to forms, templates and page categories; data attributes are mapped to property pages of the corresponding data type; relations between objects are rendered as links between pages. The cardinality of attributes is enforced through appropriate form input fields. Class inheritance is supported. Furthermore, navigation elements are added to the wiki, to enable creating, listing and editing content following the structure defined by the ontology. The information stored in the wiki using these forms and templates is thus structured and consistent, and can be accessed using semantic queries.

## 4 CONCLUSIONS AND FUTURE WORK

The method presented here and the associated software support research at the Museum of Natural History Berlin by providing a way to create collaboration environments for working on complex scientific projects, while enforcing structure and consistency. Conceivably, the method and software could be applied to other research institutions and even to areas of industry that face similar knowledge management challenges. Furthermore, as ontologies can be reused and extended (Noy and McGuinness, 2001), the methodical approach facilitates knowledge transfer between projects, as taxonomies and vocabularies found useful in one project can be used to construct a wiki for a second project. As discussed above, the implemented software supports the conversion of semantic classes, attributes and relations, class inheritance and cardinality. This is a subset of the constructs defined by the RDF/XML specification

94 (Gandon and Schreiber, 2014). More work is needed to extend the possibilities of the software so as to  
95 pave the way for more expressive collaboration environments.

## 96 **5 ACKNOWLEDGEMENTS**

97 Software development at Museum für Naturkunde Berlin was sustained in part by the project “Knowledge  
98 Transfer Concept for Research Contents, Methods and Competences in Research Museums”, funded by  
99 the German Federal Ministry of Education and Research (BMBF), grant no. 01IO1632.

## 100 **REFERENCES**

- 101 Berners-Lee, T. et al. (1998). Semantic web road map.
- 102 Corcho, O., Fernández-López, M., and Gómez-Pérez, A. (2003). Methodologies, tools and languages for  
103 building ontologies. where is their meeting point? *Data & knowledge engineering*, 46(1):41–64.
- 104 Di Iorio, A., Fabbri, M., Presutti, V., and Vitali, F. (2006). Automatic deployment of semantic wikis: a  
105 prototype. In *SemWiki*.
- 106 Gandon, F. and Schreiber, G. (2014). Rdf 1.1 xml syntax. [Online; accessed 1-July-2019].
- 107 Kiniti, S. and Standing, C. (2013). Wikis as knowledge management systems: issues and challenges.  
108 *Journal of Systems and Information Technology*, 15(2):189–201.
- 109 Musen, M. A. (2015). The protégé project: a look back and a look forward. *AI Matters*, 1(4):4–12.
- 110 Noy, N. and McGuinness, D. (2001). Ontology development 101: A guide to creating your first ontology.
- 111 Oren, E., Möller, K., Scerri, S., Handschuh, S., and Sintek, M. (2006). What are semantic annotations.  
112 *Relatório técnico. DERI Galway*, 9:62.
- 113 Patzschke, E., Ortiz-Troncoso, A., and Abele, A. (2016). Wiki-ansatz und kollaboratives arbeiten im  
114 forschungsmuseum: Schlussbericht.
- 115 Pieterse, V. and Kourie, D. G. (2014). Lists, taxonomies, lattices, thesauri and ontologies: paving a  
116 pathway through a terminological jungle. *KO KNOWLEDGE ORGANIZATION*, 41(3):217–229.
- 117 Zhou, L. (2007). Ontology learning: state of the art and open issues. *Information Technology and*  
118 *Management*, 8(3):241–252.