ABSTRACT

Ontologies are complex artifacts. They should seek consensus on the use of a set of modeled concepts. Some authors propose that these devices would be beneficial if they were built collaboratively. This article aims to address the use of a semantic wiki as an alternative to the collaborative construction of ontologies, and describes its ontological structure. Wikis are known as tools for collaborative construction of content. The semantic wiki is a research effort to integrate the concepts of wikis with the semantic web. The case study presented shows an implementation in Semantic MediaWiki: the best known and most used semantic wiki features by the academic community and the organizational environment.

Keywords: Semantic Web, Semantic Wikis, Ontology, Collaboration.

1. INTRODUCTION

Everyone agrees that user interaction building social networks is the cornerstone of "Web 2.0". Applications offer dynamic contents and rich interfaces, providing the means to add or edit contents. Blogs, wikis and photo / video / text sharing sites have increased user sharing and participation.

Wikipedia, a successful example of web technology, helps the sharing of knowledge between people, leaving individuals free to create and modify its contents. But, Wikipedia is designed to be used by people - software does not understand or can not automatically handle the contents of Wikipedia. In parallel, the "semantic web", which is a set of technologies that help knowledge sharing across the web among different applications, is starting to gain weight. Recently, the concept of "semantic wiki" has emerged, integrating the advantages of the wiki with semantic web technologies.

A semantic wiki is one that has an underlying knowledge model described on its pages. Classic or syntactic wikis are made up of text and untyped hyperlinks. Semantic Wikis, on the other hand, allow its users to identify information about the data described
on the pages, and relations between pages, so that it may be inspected or exported as a database.

In contrast, the contents of classical wikis are organized in a narrative and non-structural form. It is known that the structured information assists the work of information processing software, for example, eliminating ambiguities. The semantic web reaches maturity using ontologies to guide the semantic markup and decreases the gap of understanding between the machine and humans.

The semantic wikis were proposed in the early 2000’s, and began to be seriously implemented by 2005. In 2010, the best known wiki software seems to be the Semantic MediaWiki, whilst the best known semantic wiki is the Freebase\(^1\).

There is a wide variety of application scenarios for semantic wikis. To name a few: the engineering of ontologies, knowledge management and educational environments. This paper presents a semantic wiki and illustrates how such a wiki can be used for the collaborative engineering of ontologies. Semantic Wikis allow the sharing of formal knowledge models and formally structured ones so that software can process them properly.

Semantic Wikis are presented and analyzed in a case study.

2. METHODOLOGICAL PROCEDURES

As stated earlier, the objective of this paper is to show how a semantic wiki can be used for collaboration in the engineering of ontologies.

In order to meet that end, we did exploratory and qualitative research, where the improvement of ideas occurred from the selection and reading of seminal authored books, to support the main constructs elucidated below. With the aim of filling the gap between the content of the books used and the state of the art in semantic wikis, several major journals were accessed, resulting in an extensive reference list of technical and scientific papers, referenced and listed at the end of this study.

However, for the appropriate depth of this paper, we present a practical case study of an implementation of a semantic wiki. In chapter 5, will be reporting the experience and observed situations and how they reflect on the semantic web.

3. EXPLANATION OF CONCEPTS

For the theoretical foundation, an explanation of some concepts will be given, in light of the literature of the semantic web, collaboration and cooperation. Then, an analysis of semantic wiki will be presented as well as its ontological structure.

3.1 Collaboration and Cooperation on the Web

\(^1\) http://wiki.freebase.com/wiki/Main_Page
According to Choo (1997), the tacit and explicit knowledge of the members of organizations, collectively, is shared collaboratively. In order to encourage such organizations, through social and economic needs, knowledge is created, stored and disseminated.

Piaget (1995) suggests that this group interaction stimulates human beings to share knowledge restricted to individuals, making it collective and expanding knowledge. In the processes of learning, collaboration and cooperation are crucial to a community. To cooperate is to work with someone, to operate together with them, and build something with others. According to Zaidan and Bax (2010), there is the possibility of creating a new relationship between information and its users, whose technological devices stimulate individual interaction with information.

Vygotsky (1998), though, explains that collaboration between people (peers) is an essential action for the learning process, because it shows the heterogeneity of the groups, helping to preserve the virtue of the cognitive process implicit in interactions and communications. In each member of a community, sources that stimulate group life are found, regulating individual actions. The emphasis lies in the essential condition of stimulations of the group life which, however, permeate its controls (Piaget, 1995).

We must introduce the most contemporary concept of mass collaboration, discussed in Tapscott and Williams (2006). According to the authors, due to fundamental changes in technology, demographics, business, economics, and in the world, we are in a new era when people participate in the economy like never before. In the past, collaboration existed on a small scale, and occurred only among relatives, friends, associates and homes, communities and workplaces.

3.2. Brief History of the Evolution of the Web

The web is creating new spaces and contexts for the construction of knowledge. Its first generation was characterized by the huge amount of information made available. But, its regular users, mere readers, could not change or edit the content themselves (Recuero, 2009). Web 2.0 adds principles that characterize it as a platform not only for consumption, but also for production of information shared by its users. Content can be published by users in a simple and straightforward way, increasing collaboration. Also notable is the gratuity present in most systems available, usually supported by advertising.

According to Zaidan and Bax (2010, p. 9), "technology is not neutral and with the advent of Web 2.0, society undergoes several changes, and among them, some may be considered fundamental." Perhaps the most significant is the possibility of computer-mediated socialization.

Web 3.0 can be characterized as a semantic social web. Furthermore, Web 4.0 is in the future, where artificial intelligence will be put in place (SPIVACK, 2007).
3.3. Wikis as collaborative tools

With the use of IT, society suffers significant impacts. In this research the focus is placed on the possibility of socialization of knowledge, the use of communication tools. Indeed, collaborative tools are able to provide the web users, in addition to cooperation and collaboration, knowledge construction in an interactive manner (Recuero, 2009).

The hypertext on the Internet emerged as a paradigm of social construction, as users negotiate and reconstruct and share knowledge (Majchrzak; Wagner; Yates, 2006).

In the context of enterprise collaboration, according to Tapscott and Williams (2006), we see that the world is moving towards dispersed and decentralized knowledge. As it is known, the term wiki is originally from the Hawaiian language, coined by Ward Cunningham who originally created the WikiWikiWeb in 1995. Wikis are used for different purposes which are:

- Maintenance of knowledge networks;
- Construction of knowledge in communities;
- Cooperation in the construction of knowledge;
- Knowledge management systems.

Wikis as tools of knowledge construction allow collaboration in the knowledge society (Majchrzak; Wagner; Yates, 2006, Ramalho, Vidotti, Fujita, 2007; Tapscott, Willians, 2006). They are not only on the Internet, but also on the intranets of organizations.

3.4. Semantic Web

It is widely known that the semantic web is the result of applying the technologies of knowledge representation to distributed systems in general, to fill the communication gap existing between humans and machines. Indeed, today the knowledge is implicit in web pages, and so it is difficult to be extracted and treated by the machine (Breitman, 2005; Mika, 2007).

In a classic article, "The semantic web" (2001), the semantic web is described as an extension of the current web in order to develop means to ensure that machinery can serve humans more efficiently. However, it is necessary to build instruments in order to provide a logical sense (syntactic) and semantic to computers (Berners-lee, 2001). Different fields of computational linguistics, database, knowledge representation, knowledge-based systems and service-oriented computing collaborate in the building of the semantic web.

3.4.1 Metadata

Allowing tagging of data and information is used for various purposes, such as content identification, description, location, etc. The realization of the semantic web is dependent on the joint construction of a worldwide network of metadata; in this context, they are also referred to as controlled vocabularies.
3.4.2 Taxonomy

Taxonomy is a scientific classification, proposed by a group of officials, experts in a particular field of knowledge. In IT, it is simply a hierarchical classification of entities as established relationships in the real world.

3.4.3 Ontologies and languages

In information systems, an ontology is a formal conceptualization and consensus (for a group) of domain knowledge. They are conceptual models that capture in an explicit way the shared vocabulary used by applications. It is a guarantor of a communication with control of ambiguities. Ontologies make up a sort of lingua franca of semantic web.

Building an ontology implies the making of a conceptual model of a domain in a formal language, enabling inferences by computers. This modeling activity is a formal description of some aspects of the physical world (real), aiming at representation (Breitman, 2005; Mika, 2007).

3.4.5 XML, RDF e OWL

Created in 1998, XML (eXtensible Markup Language) has constituted the serialization syntax for various other formalities. RDF (Resource Description Framework) is a standard recommended by W3C in 2004, whose history began in 1995, proposed in 1999. RDF represents metadata in the form of statements about properties and relationships between resources on the web (Breitman, 2005). It provides simplified semantics with good representation for the treatment of metadata, but it does not provide necessary inputs required for the expression of an ontology (Mika, 2007).

OWL (Web Ontology Language) is more comprehensive and expressive than RDF. A comparison is proposed for teaching purposes by Mika (2007) between the formalisms: entity-relationship diagram (E/R), UML (Unified Modeling Language), XML and RDF / OWL.

Table 1: Comparison between concepts and semantic languages and non-semantics

<table>
<thead>
<tr>
<th>Origin</th>
<th>Primitive</th>
<th>Expressivity</th>
<th>Distributed representation</th>
<th>Formal semantics</th>
</tr>
</thead>
<tbody>
<tr>
<td>E/R</td>
<td>1976</td>
<td>Relation</td>
<td>•</td>
<td>No</td>
</tr>
<tr>
<td>UML</td>
<td>1995</td>
<td>Objects</td>
<td>**</td>
<td>No</td>
</tr>
<tr>
<td>XML</td>
<td>1998</td>
<td>Entity</td>
<td>**</td>
<td>Yes</td>
</tr>
<tr>
<td>RDF / OWL</td>
<td>2004</td>
<td>Resource</td>
<td>****</td>
<td>Yes</td>
</tr>
</tbody>
</table>


² With UML 2.0.
3.5 Semantic Wikis

Semantic wikis extend the traditional wikis allowing the annotation of content explicitly, making them more readable for the machines. Some limitations of wikis can be solved, namely (Krötzsch; Vrandečić; Volkel, 2005; Krötzsch, et al., 2007):

- Content consistency: often in traditional wikis, the same information can appear on several pages. The semantic markup allows for greater information consistency, avoiding the ambiguity at the time that users are making the content insertion.

- Content access: big wikis have many pages, making it a challenge to search and to compare information. Using search with syntax closest to SQL\(^3\), semantic wikis allow you to return to the desired content.

- Content reuse: the motivation of wikis is to provide information. The content of unstructured traditional wikis allows only reading by browsers.

Semantic wikis enable semantic enrichment of the content that is still easily manipulated by the user. They allow:

- Classification and annotation for the links;
- Dynamic presentation of contents;
- Richer navigation;
- Metadata;
- Data in triples;
- Semantic search;
- Embedded queries;

3.5.1 Examples of semantic wikis:

Nowadays, there are some semantic wikis available, which can also be used as free software. Among them: AceWiki\(^4\), Kiwi\(^5\), Knoodl\(^6\), OntoWiki\(^7\) and Semantic MediaWiki (SMW)\(^8\).

The Semantic MediaWiki (SMW) is an extension, also MediaWiki\(^9\) is free. The research was begun in 2005 in Germany AIFB (Informatics Institute of the Faculty of Economics and Business Engineering\(^10\)) in cooperation with KIT (Forschungszentrum...
Karlsruhe GmbH and Universität Karlsruhe\textsuperscript{11}). The responsible researchers are Krötzsch Markus, Denny Vrandečić and Max Volkel.

The main features of SMW are:

- MediaWiki extension;
- Nothing is overwritten in MediaWiki;
- The functions are called when necessary;
- Simple semantic marks;
- You do not need a great knowledge of ontologies.

Technologies for installation are the same for* the aforementioned MediaWiki (Krötzsch; Vrandečić; Volkel, 2005). (***mesmos para = the same for/ mesmo que = the same as)

4. SEMANTIC WIKI ANALYSIS AS A COLLABORATIVE BASIS FOR ONTOLOGY CONSTRUCTION

The collaboration aspects were embryonic and some explored the semantic web community in the first half of the 2000s. Tolksdorf and Simperl (2006) report that research efforts were focused on issues related to knowledge representation. Correndo and Alani (2007) explain that ontology collaborative constructions were little supported by ontology editors. In fact, some ontologies need to be agreed upon by user communities. To reach this agreement, it requires the support of tools and methodologies that will allow users to express and write down, collaboratively, their points of view. To address these requirements, many tools have evolved primarily from Web 2.0 and Web 3.0.

In the state-of-the-art semantic approaches, ontologies have been sustained as the key to the most advanced technologies to support knowledge workers. Still, we see that is not given due importance in conscious elaboration to ontology projects, and they are made without the minimum requirements for construction. In full operation of the semantic web systems, the concept of collaboration seen in the previous section, is strongly committed to this deficiency. Braun, Schmidt and Zacharias (2007) show that the main challenge is to construct implicit and informal ontologies for the explicit formal models needed for semantic approaches, including semantic wikis.

These authors said that to integrate the work of ontology constructions with the natural appearance (implied) of vocabularies, the followig is needed:

- Ontology construction integrated with the usual tasks of users, for example, annotation or navigation;
- It should not be assumed that modeling users are fully expert, and the complexity of ontology task editions should be reduced to a minimum. This also means balancing the expressiveness of ontologies with the usability of the editing program, for example, reducing buildings to taxonomic structures;

\footnote{http://www.kit.edu/kit/english/index.php}
• Finally, the strong collaboration, not only because ontologies share concepts, but also because it is necessary to disseminate them to various communities.

Kousetti, Millard and Howard (2008) corroborate and add that semantic wikis are the perfect combination of collaboration and semantic expressiveness. Attempting to analyze semantic wikis as a basis for collaborative ontology construction, these authors explain that the creation of ontologies has been in the hands of experts, as well as knowledge management with specialists in this area. In order to realize the semantic web, all participants in the process have the opportunity to contribute and, therefore, to collaborate. The ease of authorship in wiki pages is a great motivator for integration and interoperability of the semantic web. The knowledge users, not technical in ontologies, require a minimal understanding of how to operate in semantic wikis, without the need for deepening the ontological concepts.

Concluding this analysis of semantic wikis for ontology collaborative construction, it is necessary to resort to Lim and Ko (2009), who show that the ontology construction is the main part of semantic web applications. These authors propose a semantic wiki that experts in a field, not technicians in ontology, can easily and collaboratively organize, evaluate and refine the content semantically marked. When a new wiki page is generated, the models with semantic marking on the pages are also generated, and then the collaboratively created ontology is automatically generated.

In this same sense, Kasisopha and Wongthongtham (2009) demonstrate the ontologically-based evolution on semantic wikis. They start with the assumption that the maintenance of an ontology and the introduction of new versions to users is time-consuming and costly; also, running the risk of new versions being built before the previous ontology is put into production. The semantic wikis provide the management of ontologies in that, they are characterized by the ease, creation and editing of the semantic search, the authentication and quality management.

The semantic wiki chosen in this study comply with these aspects to the extent that users add semantic structures to the text in a collaborative environment on the Internet or Intranet. The same can not be confirmed if the user is building and editing their ontology in a stand alone editor to be imported later in the application of the semantic web. The next topic will confirm these concepts through a practical construction of a case study.

5. SUPPORT OF THE ONTOLOGY COLLABORATIVE CONSTRUCTION

Let us illustrate, with a practical example, how a semantic wiki can serve to support the ontology engineering process.

In the engineering of ontologies, domain experts and knowledge engineers work together to create a formal ontology. Semantic Wikis can support this process:

• Domain experts (non-technical) have an easy way to explain their knowledge;

---

12 A term that means that the application can be run and controlled by an operator independently.
Semantic Wikis and the Collaborative Construction of Ontologies: case study

- Domain experts and knowledge engineers can work together, each bringing their personal experience, and
- Knowledge can be incrementally formalized in an evolutionary process, starting with informal texts.

At first, it is observed that some semantic wikis, such as a Knoodl and Ontowiki, which are not part of this case study, allow the import of ontology already built and in this sense are oriented not exactly towards the support of its engineering, but especially towards its use. In such cases, ontology provides the basis for content annotations included in the wiki.

The SMW, the semantic wiki chosen in the study does not allow the natural importing of a previously built ontology without installing extensions (Vrandečić; Krötzsch, 2006). It can be considered this as a limitation of the semantic wiki (Auer, 2006; Dietzold, et al., 2010); however, it should be noted that the scenario of the SMW is one on which the semantic structure would be built collaboratively and not merely imported (Krötzsch; Vrandečić; Völkel, 2005).

Regarding the SMW, its simplicity and collaborative ease of use, inherent in its nature as a wiki tool, allows to see the interest of its use in certain projects. Especially, we will see it in the support of collaborative engineering of ontologies.

As an example, consider a researcher who wants to create a semantic portal in a given Congress (FIG. 1). It begins to fill the semantic of wiki with texts, images, and maybe even audio content, available through an easy-to-use editor. It can also describe simple relationships between content.

At some point, his experience in technology is no longer enough. Then, a knowledge engineer and his experience in the Semantic Web come into play. The knowledge engineer, without understanding specifically the topic (Congress), shall include the simplest relationships to form an ontology. At the same time, the researcher can continue to fill the system with more content.

The semantic structure of the SMW includes:

- Categories and subcategories (classes for meta-classification);
- Properties (annotated links);
- Data Types (to design values);
- Instances (their own wiki pages);
- URI (Uniform Resource Identifier - character string to identify a web resource).

Figure 1 illustrates the class hierarchy that will be used in the case study. When inserted into SMW, such classes transform themselves into categories and subcategories, the basis for ontology development. In the class hierarchy of Figure 1, the root class is "Academic".
It is noteworthy that, as in any engineering project, the use of a support tool obviously does not exempt a previous study from the project requirements, as well as planning and preparation of its ontological structure, at least initially, before leaving to use the semantic wiki.

In the figures that follow are the representations of the ontological structure of the SMW, ending with a comparison with a typical triple RDF (subject - predicate - object), the basis for all ontologies.

We will begin the presentation of case study showing, in Figure 2, a wiki page called "Contecsi 2010," developed in a traditional wiki. We can see the text-only form and the unstructured content form.

**Figure 2. Traditional wiki page.**

Source: Research data.

Figure 3 reveals a piece of code on this page, the option of editing the wiki code. It is observed that there are two links:

- Internal link: syntax [[City of São Paulo (city)]], which refers to an internal wiki page;
- External Link: syntax [http://...Contecsi 2010].
Using only links in the format of traditional wikis does not allow the formalization of content structure for programmatic use of the tool itself. So the software can do nothing to support users with regards to the ease of navigation, retrieval content form, collaboration, reuse, and other resources of manipulating content. These and other limitations are the ones that the semantic wikis could be reduced to (Schaffert, et al. 2009; BAO, et al., 2010).

Table 2 presents the design of a semantic structure greatly simplified for instructional purposes. It uses only one instance (or wiki page), which presents the Contecsi/2010 event page. One category: Congress and its properties: city, rating, event date and publication type.

<table>
<thead>
<tr>
<th>Instance</th>
<th>Classes or Categories</th>
<th>Attributes or Properties</th>
<th>Data types</th>
</tr>
</thead>
<tbody>
<tr>
<td>category</td>
<td>property</td>
<td>type</td>
<td></td>
</tr>
<tr>
<td>Contecsi 2010</td>
<td>Congresses</td>
<td>City</td>
<td>String</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Classification</td>
<td>String</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Event Date</td>
<td>Date</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Publication type</td>
<td>String</td>
</tr>
</tbody>
</table>

Table 2: The planning for semantic structure of the example.

To formalize the category, following the SMW syntax, the following line has to be added to the page: [[Category:Congresses]].

We have to create individual pages for the properties setting. In the figure below, the page illustrates the Classification property. Note that it is possible to make a list of accepted values. If they are not listed correctly ("Allows value" accepted only such values as Excellent, Good, Bad) at the moment of data entry, an error will appear.
In the code above, one has the data type for this property, described in SMW syntax as: \[[\text{has Type::Type::String}\]]

Finally, the whole semantic structure of the simplified study case is displayed on the Contecsi/2010 page, shown in Figure 5.

Figure 4: SMW code for the "Classification" property.
Source: Research data.

Figure 5: Semantic page for “Contecsi 2010”.
Source: Research data.

One draws attention to a feature of SMW called "Facts about" (Figure 5) which explicitly shows the semantic relationships present in the content provided to SMW. This feature allows the navigation through the semantic content. For example, clicking on Excellent, it displays all the contents of the "Congresses" type (Figure 6).
This return from a click on the properties has advantages for the end user. However, it is noteworthy that for the semantic wiki administrator, all the navigation on pages and data can be built from complex sentences, also called inline queries, or embedded queries. As an example, from the code `{{# ask: [[Category:Conference]]}}` embedded in a wiki page, any maintenance that is done, on pages whose category is Congress, will be automatically updated, resulting in what is shown in Figure 7.

It is interesting to note that a typical triple RDF (subject-predicate-object), corresponds to the "classification property" of the "Contecsi 2010", as shown in Figure 5, and could be expressed as follows:

\[ <\text{Contecsi2010} \text{ Classification} \text{ Excellent}> \]

6. FINAL CONSIDERATIONS

It is known that the construction of an ontology is a complex task that should be, preferably, performed by multi-disciplinary teams; however, the most common tools to build ontologies (e.g. Protégé\(^{13}\)) do not yet allow for remote collaborative teamwork.

After a brief presentation of teh Semantic Web key concepts, the article presented an instructional case study aimed at illustrating how a semantic wiki tool could be used as an interesting support option for the collaborative construction of ontologies. The ability for collaboration and cooperation becomes rich and effective with an easy and simple documentation process.

\(^{13}\) http://protege.stanford.edu/
It can be said that the article illustrated how what we call "the wiki philosophy," enhanced by methods and techniques of the semantic web approach which gives rise to a simple, but effective tool which seemed useful in certain phases of the ontology engineering, or even as a support throughout the process. It is noteworthy the support for ontology documentation, which, once represented in a wiki, already constitutes its own documentation.

The article makes explicit the semantic structure underlying the SMW tool and shows how it deals with some major problems of today's wikis, such as inconsistency of content, access and knowledge reuse. The usability of semantic wikis, Semantic MediaWiki in particular, was verified. Even people who are not specialists in logic or ontology could use it. However, it is important to point out the limitations of the textbook case study, with a very simplified knowledge domain, in representing all the SMW features.

As noted above, the SMW tool is not the only one. Future work could make more comparisons with others like Knoold and OntoWiki.

It is believed that the use of a semantic wiki as a tool to support remote collaborative work has advantages. One can cite, for example, the flexibility to absorb, in a simple way, proposals from other team members while preserving the formal structure that organizes the content. That is, it is believed that the use of semantic wikis to support the remote collaborative construction of ontologies is a good commitment between flexibility (necessary for humans) and formal (necessary to machines).

REFERENCES


