AUTOMATYKA • 2009 • Tom 13 • Zeszyt 2

Grzegorz J. Nalepa*

Analysis of Selected Knowledge Representation Methods in Semantic Wikis**

1. Introduction

In last years the most important development of the Internet concerned not the lower network network layers, but the higher application or service layers related to the World Wide Web technology, or the Web. This is mainly due to the fact, that while the speed and storage capacities of the Web increased by orders of magnitude, its search and processing capabilities remained almost unchanged on the conceptual level.

This phenomena led, almost a decade ago, to the proposal of the Semantic Web [1]. In this architecture a number of higher level semantic facilities built on top of the Web would allow not just to search data but to reason with knowledge. In fact, this was the point where the focus of the Web development moved from content (data) to knowledge (in a broad sense). A decade later, number of semantic technologies are available and widely used, starting from the data structuring XML, to meta-data annotations with RDF and ontologies with RDFS and OWL (observe the well-known Semantic Web layer cake).

While these technologies provided knowledge encoding and representation solutions, the challenge remains to provide an efficient knowledge processing and reasoning with rules on the Web. This is in fact the point, where most of the current Semantic Web research focuses. Recent standards include RIF and SWRL. Another important issues include the quality of the knowledge stored on the Web.

Besides knowledge representation and reasoning, a sensible knowledge engineering solution for the Web is another important challenge. While the Semantic Web initiative targets mainly representation aspects, it does not directly address the specific problems stemming from the massively parallel and collaborative nature of the Web. Social networks, that provide specific services on top of the Web and the Semantic Web, try to cope with

* Department of Automatics, AGH University of Science and Technology, Krakow

** This work has been partially performed in the framework of the EU ICT Project INDECT (FP7-218086).

these problems. Recently the technology of the wiki systems has gained importance with respect to the collaborative knowledge acquisition and engineering.

The development of semantic wikis [2–4] allowed to use the Semantic Web methods and tools on top of the existing content-centered wikis. Existing semantic wikis allow introducing semantic information (e.g. meta-data, ontologies) into a wiki. In fact, they often allow to build a wiki around an ontology, which improves its conceptual coherence. Most of the semantic wikis reached a stage where the reasoning capabilities have to be added. This is where some limitations of existing solution become exposed.

The rest of the paper is organized as follows: In Sect. 2 wikis as web-based knowledge engineering systems (a special case of social networks are discussed). Then in Sect. 3 the development of semantic and knowledge wikis is considered.Semantic wikis pose some new challanges for knowledge evaluation in the wiki as discussed in Sect. 4. The possible applicability of the Semantic Wikis technology in the context of the INDECT Projec is given in Sect. 5 Finally a short evaluation of the technology and concluding remarks are given in the final section.

2. Wiki systems

Wiki systems appeared in the mid 90s. According to Wikipedia the first system called „wiki" (WikiWikiWeb) was established 15 years ago. Tho goal of these systems was to provide a conceptually simple tool for massively collaborative knowledge sharing and social communication. Wikis were meant to help build certain communities interested in given topics. Clearly some of them grew large and general, such as the Wikipedia.

A wiki system is a community-driven collaboration tool. It allows users to build content in the form of the so-called wiki pages, as well as uploaded media files. Wikipages are plain text documents containing special wiki markup (e.g. for structuring content) thus creating the so-called wikitext. The wikitext is simplistic and human readable, making it a much more accessible tool than HTML/XML. Pages are identified by a unique keyword (name) and usually grouped within the so-called namespaces. Pages are linked to each other and to external websites thus creating a hyperwikitext structure. An important feature of wikis is the integrated version control functionality, very helpful in a collaborative environment. It allows registering all subsequent versions of every page, thus allowing to see introduced content differences. All wiki edits may be identified by user names and time stamps, so it is possible to recreate any previous state of the wiki at any given time. From the technical point of view a wiki has a regular web-based client-server architecture. It is run on the webserver and accessed by a regular browser. Wikis introduce a range of access control mechanisms from simple ones, to full-fledged ACL (Access Control Lists) solutions. On the server side wikis require different runtime environment (e.g. PHP, JSP), possibly with a relational database system (RDBMS). A comprehensive comparison of different wiki systems can be found on www.wikimatrix.org.

One of the most interesting wiki systems for developers is DokuWiki (wiki.splitbrain.org/wiki:dokuwiki). It is designed to be portable, easy to use and set up. Like number of other solutions DokuWiki is based on PHP. However, it does not require any relational database back-end. It allows for image embedding, and file upload and download. Pages can be arranged into namespaces which act as a tree-like hierarchy similar to directory structure. It provides syntax highlighting for in-page embedded code for programming languages using the GeSHi parser. Furthermore, it supports extensive user authentication and authorization mechanisms including ACL. Its modularized architecture allows the user to extend DokuWiki with plugins which provide additional syntax and functionality. A large number of plugins extending the core functionality is available. The mechanism of templates provides an easy way to change the presentation layer of the wiki.

All wiki systems provide an abstract representation of the content they store. They all provide standard searching capabilities. However, they lack facilities helping in expressing the semantics of the stored content (besides simple tagging mechanisms, that can later be used to create the so-called folksonomies). This is especially important in the case of collaborative systems, where number of users work together on the content. This is why wikis became one of the main applications and testing areas for the Semantic Web technologies.

3. Semantic wikis

A step in the direction of enriching standard wikis with the semantic information has been made by the introduction of the so-called semantic wikis, such as the IkeWiki [2], OntoWiki [9], SemanticMediaWiki (SMW) [3], or SweetWiki [4]. In such systems the standard wikitext is extended with the semantic annotations. These include relations (represented as RDF triples) and categories (here RDFS is needed). It is possible to query the semantic knowledge, thus providing dynamic wiki pages. Ultimately these extension can also allow for building an ontology of the domain (usually with OWL [5]) with which the content of the wiki is related. This extension introduces not just new content engineering possibilities, but also semantic search and analysis of the semantically enhanced content.

Number of semantic wiki systems are available, most of them in the development stage providing demo versions. In a recent paper [4] a comprehensive overview of semantic wikis technology has been given, with number of systems described w.r.t to their features and implementation details. The development of semantic wikis is very dynamic. Since that paper has been written some of these systems are no longer supported. Several new systems are currently available, including Kiwi or KnowWe (and KnowWe2). A new FP7 project Kiwi (www.kiwi-project.eu) aims at providing a collaborative knowledge management based on semantic wikis (it is the continuation of IkeWiki effort). The website semanticweb.org/wiki/Semantic_Wiki_State_Of_The_Art provides an up-to-date list of systems and their features. Some important features of selected systems available up to recently are presented in the Table 1.

Table 1
Selected Semantic Wikis features comparison

Feature	*AceWiki*	*IkeWiki*	*Kiwi*	*KnowWE*	*SMW*	*SweetWiki*
Status	active	discontinued	active	active	active	discontinued
License	LGPL	GPL	CDDL	LGPL	GPL	LGPL
Platform	Java	Java	Java	Java	PHP	Java
Reuses engine	no	no	no	yes	yes	no
Wiki object model	no	no	yes	no	no	yes
Assisted annotations	yes	yes	yes	yes	no	yes
Social tagging	no	yes	yes	yes	no	yes
Edit ontology in wiki	yes	yes	yes	yes	yes	no
Representation	ACE	OWL	OWL	OWL	RDFS	RDFS
Queries	no	SPARQL	SPARQL	SPARQL	wikiml	SPARQL
Reasoning	no	Jena	misc	d3web	KAON	Corese
Versioning for wikipages	no	yes	yes	yes	yes	yes
Versioning for metadata	no	no	partial	partial	partial	partial
Discussion pages	no	yes	yes	yes	no	no

From the knowledge engineering point of view expressing basic semantics is not enough. In fact a knowledge-based system should provide effective knowledge representation and processing methods. In order to extend semantic wikis to knowledge-based systems, ideas to use a problem-solving knowledge have been introduced. An example of such a system is the KnowWE semantic wiki [10–12]. In such a system the semantic knowledge is extended with the problem-solving domain-specific knowledge. The system allows for introducing knowledge expressed with decision rules e.g. with SWRL [6] and trees related to the domain ontology. It could be said, that conceptually it is built on top of the simpler wikis, e.g. the SMW. Extended representation equivalent to the expressive power of Prolog [7] can be also considered.

4. Knowledge quality issues

Semantic wikis with both fact and rule representation can be considered knowledge-based systems. The distributed knowledge development process in a wiki poses new problems in knowledge engineering when compared to the classic development of monolithic knowledge bases. In case of most of the semantic wikis the focus of the current research is on the knowledge representation, integration and authoring. However, with the growing amount of knowledge contained in semantic wikis, the knowledge quality issues seem to be critical [14].

Formal verification of knowledge-based systems is a mature field, where a number of important results have been brought up in the last decades. The research in verification has been especially active in the field of rule-based expert systems [17]. A taxonomy of formal properties for the verification of such systems has been presented in [18], with some more recent follow-ups such as [19].

According to [19] three most important groups of properties in the verification of knowledge-based systems include the following:

– Consistency of the knowledge base means that no contradictory conclusions can be inferred from valid facts.
– Completeness of knowledge means – in a vague sense – that no information is missing.
– Conciseness means, that no redundant (unnecessary) knowledge can be found in the knowledge base.

Let us consider how the three classic rule verification criteria may be applied with respect to the distributed knowledge base in a wiki system.

Wikis are composed of wiki pages. So, a wiki can be described as a distributed knowledge-based system, where a number of knowledge bases exist. It is a distributed system, because everyone can work on his own knowledge base. Pages are usually grouped within namespaces related to their common semantics, which can be explicitly marked in a semantic wiki ontology. Pages in different namespaces can be interconnected, as well as pages can reference pages in other wikis (interwiki connections).

Considering the general knowledge wiki architecture several verification scopes need to be considered:

– single page scope – where the given property is analyzed only in a single wiki page, and all links are ignored.
– namespace scope – where every page in a group is considered to be a component of a single namespace-wide knowledge base, so the given formal property must hold with respect to the whole group. This means that all the links to the pages in the namespace have to be considered, an external links ignored.
– wiki scope – this global wiki scope treats the whole wiki as a single knowledge base, interwiki wiki links are ignored.
– interwiki scope – the most complex case interwiki links should be analyzed.

With respect to the above scopes the properties may be interpreted as follows.:

Wiki Consistency means that no contradictory information is contained in the unit. For the practical verification the given wiki unit (page, namespace) needs to be analyzed to detect contradictory facts or rules. Inconsistency is likely to appear in a distributed environment such as wiki, where a number of independent authors extend the knowledge base. It should be detected on-line, during the wiki editing session. However, it is worth noting, that considering the evolutionary nature of the wiki knowledge, such an inconsistency between two versions of a given page could be in fact a hint for knowledge refinement, so it is not obvious which of the above contradictory facts is "correct". considering the changed page contents.

Wiki Completeness means that no information with respect to the given ontology is missing in the unit. This should be considered with respect to all the pages and knowledge bases in the wiki. In a general case completeness verification is hard. It is possible to conduct such a procedure in cases where the domain of a certain property is given. This is possible in case of wikis designed according to a domain ontology.

Wiki Conciseness can be interpreted as a state where no redundant information is contained in the unit. In general cases it means that no identical facts or rules are inserted. In more specific cases it could also mean that new facts are more general, or that new rules subsume the older versions. In case of multi-page scope (namespace,wiki) practical implementations of the verification algorithms should consider comparing pairs of pages as units for properties verification.

5. Semantic wikis for the indect gis system

In the context of the INDECT project semantic wikis can play an important role. Within the WP4 a GIS system supporting the collaboration between citiziens and the authorities, e.g. the Police, w.r.t. to presence of threats is to be implemented. In the system a web interface for gathering citizen-provided feedback has to be provided. The interface supports a distributed knowledge acquisition facility. It helps building a knowledge base on threats, basing in information from citizens and the Police.

Semantic wikis provide knowledge representation methods for semantic annotations and metadata encoding. Thus, a hirerchical architecture for the GIS system is considered, where a semantic wiki supports the geographic core of the system w.r.t. to semantic metadata encoding. As a part of annottaions an ontolgy suporting thretas categorization would be develpoed. This would support the citizens in entering the information which is semanticaly coherent. The wiki builtin facilities for information versioning would help to track different reports on threats. The rule-based inference mechanisms provided by some semantic wikis allow for basic reasoning with knowledge contained in the wiki, e.g. infering new inormation, basing on the provided one.

6. Concluding remarks

Wikis are an example of basic yet very successfull technology, enhancing user productivity in distributed content authoring. The impose very low requirements on the user runtime, which makes them very accessible and easy to implement. Semantic wikis extend the core wiki technology with knowledge representation including semantic annotations, concept ontologies and inference. Thanks to the rapid development of semantic wikis, new systems provide advanced semantic web features.

It is expected, that the semantic wiki technology would play an important rle for develepoling a hierarchical GIS system for the INDECT Project, where the main geogrpahic compeonent would be supplemented with a knowledge base layer implemented by such a wiki. In this case, the distributed knowledge authoring facilities provided by wikis are suitable for gathering information from multiple sources, including both individual citiziens and authorities.

As a part of future implementation of the wiki, evaluation of the possible use of the XTT [15, 16] rule representation for rules is consiered. The framework would allow for a unified knowledge representation and inference.

References

[1] Berners-Lee T., Hendler J., Lassila O., *The semantic web*. Scientic American, May 2001.

[2] Schaert S.: *Ikewiki, A semantic wiki for collaborative knowledge management*. [w:] WETICE '06: Proc. of the 15th IEEE International Workshops on Enabling Technologies: Infrastructure for Collaborative Enterprises, Washington, DC, USA, IEEE Computer Society, 2006, 388–396.

[3] Krötzsch M., Vrandecic D., Völkel M., Haller H., Studer R., *Semantic wikipedia*. Web Semantics 5, 2007, 251261.

[4] Buffa M., Gandon F., Ereteo G., Sander P., Faron C., *Sweetwiki: A semantic wiki*. Web Semantics: Science, Services and Agents on the World Wide Web, 2008, in press.

[5] Mcguinness D.L., Van Harmelen F., *Owl web ontology language overview*. w3c recommendation 10 february 2004. Technical report, W3C, 2004.

[6] Horrocks I., Patel-Schneider P.F., Boley H., Tabet S., Grosof B., Dean M.: Swrl, *A semantic web rule language combining owl and ruleml*. w3c member submission 21 May 2004, Technical report, W3C, 2004.

[7] Bratko I., *Prolog Programming for Articial Intelligence*. 3rd edn. Addison Wesley, 2000.

[8] Nalepa G.J., Wojnicki I., *Proposal for a prolog-based knowledge wiki*. G.J., Baumeister, J. (eds.), 4th Workshop on Knowledge Engineering and Software Engineering (KESE2008) at the 32st German conference on Articial Intelligence: September 23, 2008, Kaiserslautern, Germany, Kaiserslautern, Germany, 2008, 43–54.

[9] Auer S., Dietzold S., Riechert T., *Ontowiki – a tool for social, semantic collaboration.* [w:] Cruz I.F., Decker S., Allemang D., Preist C., Schwabe D., Mika P., Uschold M., Aroyo L. (eds.), International Semantic Web Conference. Volume 4273 of Lecture Notes in Computer Science., Springer, 2006, 736–749.

[10] Baumeister J., Reutelshoefer J., Puppe F., *KnowWe: community-based knowledge capture with knowledge wikis*. [w:] K-CAP '07: Proc. of the 4th international conference on Knowledge capture, New York, NY, USA, ACM, 2007, 189–190.

[11] Baumeister J., Puppe F., *Web-based knowledge engineering using knowledge wikis*. [w:] Proc. of the AAAI 2008 Spring Symposium on Symbiotic Relationships between Semantic Web and Knowledge Engineering, Stanford University, USA, 2008, 113.

[12] Reutelshoefer J., Baumeister J., Puppe F., *Ad-hoc knowledge engineering with semantic knowledge wikis*. [w:] Proc. of SemWiki 2008 – The Wiki Way of Semantics, Workshop co-located with the 5th European Semantic Web Conference, Tenerife, Spain, 2008.

[13] Ben-ari M., *Mathematical Logic for Computer Science*. Springer-Verlag, London 2001.

[14] Baumeister J., Nalepa G.J., *Verication of distributed knowledge in semantic knowledge wikis.* [w:] FLAIRS-22: Proc. of the twenty-second international Florida Articial Intelligence Research Society conference: 1921 May 2009, Sanibel Island, Florida, USA, 2009, to be published.

[15] Nalepa G.J., Ligęza A., *A graphical tabular model for rule-based logic programming and verication.* Systems Science 31(2), 2005, 89–95.

[16] Nalepa G.J., Ligęza A., *Xtt+ rule design using the alsv(fd).* [w:] Giurca A., Analyti A., Wagner G., (eds.), ECAI 2008: 18th European Conference on Articial Intelligence: 2nd East European Workshop on Rule-based applications, RuleApps2008: Patras, 22 July 2008, Patras, University of Patras, 2008, 11–15.

[17] Giarratano J.C., Riley G.D., *Expert Systems*. Thomson, 2005.

[18] Vermesan A.I., Coenen F. (eds.), *Validation and Verication of Knowledge Based Systems. Theory, Tools and Practice*. Kluwer Academic Publisher, Boston, 1999.

[19] Ligęza A., *Logical Foundations for Rule-Based Systems*. Springer-Verlag, Berlin, Heidelberg, 2006.

[20] Intelligent information system supporting observation, searching and detection for security of citizens in urban environment. FP7, 2009–2013, http://www.indect-project.eu/.