A Study of Ontology Convergence in a Semantic Wiki

Chrysovalanto Kousetti
Electronics and Computer Science
University of Southampton, UK

ck1004@zepler.net

David E. Millard
Learning Societies Lab
Electronics and Computer Science
University of Southampton, UK

dem@ecs.soton.ac.uk

Yvonne Howard
Learning Societies Lab
Electronics and Computer Science
University of Southampton, UK

ymh@ecs.soton.ac.uk

ABSTRACT

Semantic Wikis propose a combination of both easy collaboration and semantic expressivity; characteristics of the WikiWikiWeb and the Semantic Web respectively. In this paper we look to define and analyse the Semantic Wiki method, in order to explore the effect of different Semantic Wiki characteristics on the quality of the semantic networks authored within them. We look at a number of different Semantic Wiki implementations, including their semantic expressivity and usability. We focus on support for ontology creation, and perform an evaluation on the effect of type suggestion tools on ontology convergence within a seeded and unseeded Wiki (using Semantic MediaWiki and our own MOCA extension). We find that seeding a Wiki with typed pages and links has a strong effect on the quality of the emerging structure and that convergence tools have the potential to replicate that effect with an unseeded Wiki, but that they have limited impact on the reuse of elements from the evolving ontology.

General Terms

Design, Experimentation, Human Factors

Keywords

Semantic Wiki, Semantic Web, Emergent Ontologies

1. INTRODUCTION

Hypertext pioneers have always thought of hypertext as being a way of storing knowledge that can be then retrieved by humans, making it essentially a human to human activity [21]. This vision is starting to be fulfilled with users becoming both readers and authors through Web 2.0 applications, which allow the adding of more content, rather than more websites. However, the web as a continually growing collection of documents readable only by people does not take full advantage of the stored knowledge. The Semantic web is hoping to structure this knowledge, making it easier for machines to retrieve, manipulate and extend.

Semantic Wikis combine the notions of collaboration and structure by using wiki features for developing not only wiki pages but also formal structures in the form of ontologies.

Providing simple interfaces for authoring formal structures is not an easy task, yet the power of authoring using natural language in wikis seems to open an alternative way to achieving the Semantic Web vision. Although traditional wikis can themselves be viewed as folksonomies by allowing categorisation of wiki articles (which already improves the usability of wikis [31]), Semantic Wikis propose a more elaborate way of structuring the knowledge (as a fully typed hypertext network) that may prove to be more powerful.

In this paper we look at the way in which a number of Semantic Wikis deal with formalisation, ontology definition and expression. We then undertake a study using Semantic MediaWiki (SMW) into ontology convergence, the process of the types within a Semantic Wiki evolving over time into an ontology.

The paper is structured as follows: Section 2 begins by explaining the way in which Wikis and the Semantic Web have come together in the form of Semantic Wikis. Section 3 then looks at the requirements of a Semantic Wiki in terms of authoring support, knowledge representation and presentation, and Section 4 presents a brief survey of existing systems and their approaches. In Section 5 we present a definition of Ontology Convergence and describe our MOCA extension to SMW that uses *type suggestions* to encourage convergence. Section 6 and Section 7 describe our study of ontologies evolved within a seeded and unseeded copy of SMW, and with and without our MOCA extension (a total of four different configurations). We analyse the impact of each dimension on ontology convergence. Section 8 concludes the paper.

2. BACKGROUND

2.1 What is a wiki?

The first wiki system was developed by Ward Cunningham in 1994, under the name WikiWikiWeb[20]. In Hawaiian the word "wikiwiki" means "quick". The original definition of a wiki describes it as "The simplest online database that could possibly work" [9]. Several definitions [6,10,20,24,30] exist attempting to define what a wiki is today, some 14 years later. From the definitions, the design principles of the original wiki [8] together with some identified elements of wiki essence [13], we can extract some important key characteristics:

1. **Easy** – Easy implies that a wiki system aims to be used by the everyday user in a fast and simple fashion However, authoring may require limited expertise as wikis allow authoring through the use of natural language with the addition of special wiki notation to allow use of links and formatting [24].

2. **Create pages** – Allowing the creation of new pages.

3. **Edit pages** – Allowing the editing of an existing page.

4. **Link pages** – Linking to other websites but more importantly linking to other pages within the wiki in a straightforward manner.

5. **Collaboration** – Although collaboration is not a necessary ingredient. for example, Wikis can be used as a Personal Information Management system [3,4], it is still more common to allow several users to contribute to the growth of content. Usually such systems are accompanied by a version control system and different levels of user authority [24].

What we can identify from the above list is that requirement1, ease of use, impacts on the other more functional requirements. What distinguishes a wiki system is that a non-technical user (i.e. without any HTML knowledge) can create, edit and link pages in a fast and easy way [11].

2.2 What is the Semantic Web?

The first official article describing the Semantic Web was published in the Scientific American [5] describing how this extension of the Web would give data much more power, extending the functionality of the web to "bring structure to the meaningful content of Web pages, creating an environment where software agents roaming from page to page can readily carry out sophisticated tasks for users." [5].

Semantic Web knowledge structures are essentially triples of the form Object-Predicate-Subject, where each part of the triple is a globally unique name expressed as a URI. Triples can therefore be used to represent knowledge statements, for example Paper1 *has_author* Person1, and many triples from a knowledge graph (Person1 *works_in* Southampton, Southampton *is_in* UK).

These knowledge graphs can then be exchanged between machines using the RDF[1] format. What makes this approach powerful is that machine-readable schemas written in RDFS[2] or OWL[3] can constrain the types of relationships allowed between objects and subjects (for example, enforcing that a Person can only *work in* something that *is a* Place). In addition the knowledge graph can be used with inference rules written using casual, temporal and probabilistic logic. Inference allows the further growth of the available knowledge and the ability to reason about it [27].

2.3 What is a Semantic wiki?

Wikis promote collaboration and easy authoring of pages [11]. The Semantic Web aims to extend the current Web and structure it in order to make it more functional. A Semantic wiki is a system that allows collaborative authoring, editing and linking of pages, but also authoring and adding semantics to the data on the wiki itself. Figure 1, proposed by Tolksdorf and Simperl [29], positions

Semantic wikis as a solution for both creating and managing pages but also allowing knowledge representation.

Authoring semantics and the creation of ontologies has mainly been in the hands of "ontologists" and knowledge management experts. In order to make the Semantic Web vision reality, everyone needs to have the chance to contribute to it. The ease of authoring in wikis is the motivation for integrating Semantic Web technologies with wiki systems. Wiki systems will also benefit from the integration, with semantic searching and better browsing being only a part of the possible potential.
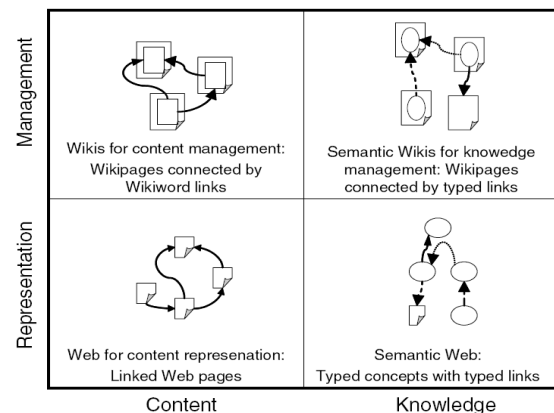

Figure 1.Hypermedia Structures under Consideration.[29] y-axis and x-axis represent the Web and the Semantic Web.

3. SEMANTIC WIKI OPERATIONS

In order to define more clearly what a Semantic wiki is, we need to identify the requirements for constructing a Semantic wiki system. In general the operation of a Semantic wiki can be divided into three parts: authoring, knowledge representation and presentation. These three dimensions can be used to categorise a wiki. Figure 2 shows how we have broken these three dimensions down.
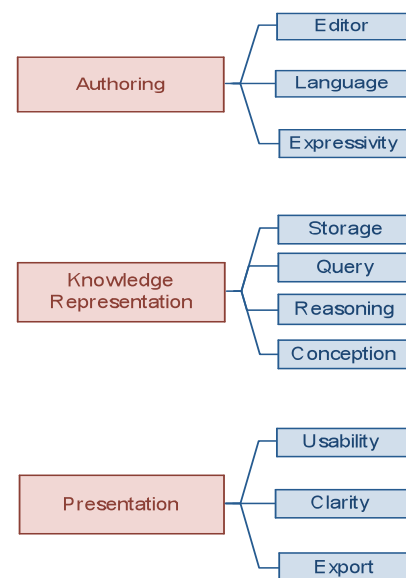
Figure 2.Operations of a Semantic Wiki

[1] RDF W3C recommendation: http://www.w3.org/TR/2004/REC-rdf-primer-20040210/, [accessed 2008-04-26]

[2] RDFS W3C recommendation: http://www.w3.org/TR/rdf-schema/, [accessed 2008-04-26]

[3] OWL W3C recommendation: http://www.w3.org/TR/owl-ref/, [accessed 2008-04-26]

3.1 Authoring

The authoring interface of wikis is what makes wikis such an exceptionally easy system as it avoids the necessity of technical knowledge in order to edit a page [11]. The same ease that also allows fast authoring must be transferred to a Semantic wiki system. Proposed implementations of Semantic wikis take different approaches in order to achieve this, with most trying to maintain the traditional wiki authoring interface which uses natural language and some special wiki notation.

The **editor** in a wiki is typically simple with some basic functionality. In a Semantic wiki, an editor becomes of greater importance since the wiki no longer only expects the provision of plain text and formatting, but also the inclusion of semantics. By tradition all wikis typically expect the **language** used to be natural language. Although this is true also in Semantic wikis, some of the implementations try to constrain how the user structures their text. Since we are trying to achieve the creation of ontologies through the use of the wiki authoring interface it is important for this interface to be functional enough to provide **expressivity** of the knowledge the user wants to input in order to allow the creation of a correct, rich and useful ontological structure.

3.2 Knowledge Representation

Traditional wikis typically store the different wiki pages either in plain text files or a database. The role of storing the wiki information becomes even more important in a Semantic wiki as representing the knowledge is an essential part of the Semantic Web. Typically wikis have some means of structuring the pages in order to allow easier browsing. The structures created are taxonomies usually in the form of hierarchies or even folksonomies [32]. In order to achieve the requirement of creating a Semantic wiki we need to structure the semantics in an ontology. How this ontology is stored can vary but being able to export the structure in the form of an acceptable standard, such as OWL, and the data represented in a standard format, such as RDF, should be common across Semantic wikis.

Storage is perhaps the central aspect of knowledge representation any wiki needs to deal with. Choosing a correct storage methodology requires trade-offs between different benefits including expressivity and efficiency. From the aspect of the Semantic Web, this storage needs to allow the creation of ontologies that can be extracted and used for both improving the wiki's usability but also by other Semantic Web applications. In evaluating the storage method, we need to evaluate the value and correctness of the created ontology according to these characteristics [15]:

- Consistency: no sentence can be contradicted by inference from other definitions and axioms.

- Completeness: anything that needs to be in the ontology is explicitly defined or it can be inferred from other defined definitions and axioms.

- Conciseness: does not contain unnecessary definitions and if no redundancies exist or can be inferred between definitions.

- Expandability: you can easily add more knowledge without requiring to make major changes to the existing structure.

- Sensitiveness: the ontology is more sensitive if small changes can alter easily how well-defined a definition is.

Furthermore, the structure must allow the retrieval of information by using some form of **query** language allowing for the key power of the Semantic Web to become available. The choice of query language will depend on the storage approach chosen. If RDF is used for example, a possible query language could be SPARQL[4]. Although **reasoning** is not of the highest priority in a Semantic wiki, providing facilities to reason about the structure created and generating more knowledge, can improve its usefulness. Finally, **conceptualisation** refers to the general conceptualisation approach taken by the wiki of how resources and information are represented within the wiki.
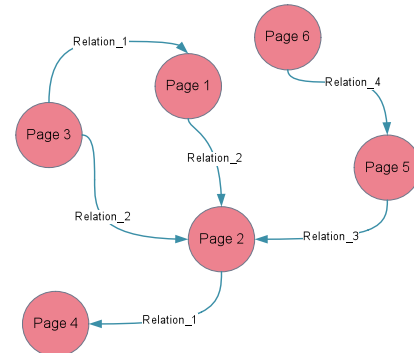

Figure 3. Conceptualisation: Every wiki page is a concept. Pages and therefore concepts are related with each other.
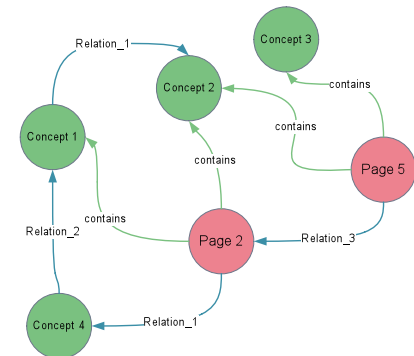

Figure 4. Conceptualisation: Every wiki page contains many concepts which can be linked to any page or each other.

Figure 3 and Figure 4 show two alternative conceptualisations. In Figure 3 pages are synonymous with concepts, tightly tying the hypertext to the knowledge model. In Figure 4, concepts are separate entities, to which pages are related. This second approach is a more sophisticated knowledge model (and matches the approach typically taken in RDF through the rdf:about relation) but the first is easier to author, as it doesn't make sophisticated distinctions between what is being written, and what is being formally expressed.

3.3 Presentation

The way the data is presented affects whether people will find what they are looking for in the wiki and therefore its usefulness. By making correct use of the knowledge structure created,

---

Semantic wikis can create a powerful presentation of the data, making it easier to browse and find more accurately what you are looking for. Semantic searches and displaying of relevant wiki pages are only some of the major presentation changes that can be achieved in a Semantic wiki.

One of the reasons wikis present the semantics to users is to make the wiki more usable. Therefore **usability** of the semantics needs to be addressed making them understandable, usable and useful. Furthermore the information presented needs to provide **clarity** for what is being described and conversely must not confuse the user. Many of the users accessing the wiki may want to extract the pure semantics and ontology behind the system so the wiki should provide an **export** facility. This facility not only refers to physical users but also intelligent agents.

4. SEMANTIC WIKI IMPLEMENTATIONS

In order to understand better what can be called a "Semantic wiki" an exploration of available Semantic wiki implementation was conducted. Through exploration, different characteristics, patterns, similarities and differences were identified. A list of available "state of the art" Semantic wikis [26] was retrieved from semanticweb.org[5]. Due to the large number of implementations, not all systems could be explored in this paper, so a selection of ten popular wikis was made.

4.1 Existing Systems

AceWiki[6] [19] is a slightly unusual wiki in that it is a new kind of Semantic wiki which uses the controlled natural language ACE [14]. Formal statements are the main content of the wiki itself. In this way, it tries to integrate ontology, rules and query language into one. It becomes quite limiting and, although it is written using natural language, what you are allowed to author is not so natural. The editor allows the user to either directly type in the statements or use a guided form of selecting from the existing ontology. AceWiki conceptualizes each wiki page as a concept and produces an OWL output of the underlying ontology.

IkeWiki[7] [24,25] aims to create instance data based on an existing ontology but also to be a (limited) tool for creating and editing ontologies. IkeWiki tries to keep the look and feel of popular wiki systems, such as Mediawiki. It makes use of the recommended Semantic Web technologies RDF and OWL. The presentation of the information makes use of the formal knowledge structure to improve navigation and provide recommendations while editing.

KawaWiki[16] aims to provide a complete formal structure of the data with proper use of RDF and RDFS. The architecture breaks down into three main layers:

- *RDFS* – The RDF Schema defines the underlying ontology and is used to validate the RDF Templates. The RDF Schemas are developed by the "Ontologists" group of expert users.

- *RDF Templates* – Defines the type of wiki pages that can be created and authored by the end user. The templates are created by Expert users. Templates hide the RDF complexity from the end user and also validate the user input.

- *Wiki Content* – The wiki content is what the end user gets to edit. The end user is limited in using the pre-constructed RDF Templates and all that is required is the filling of the template fields.

Makna[8][10] is an extension of the JSPWiki adding semantic functionality to it. Each page is a concept in the controlled ontology. "Wiki users are able to create semantic content (in form of RDF statements referencing pre-configured ontologies) in the classical Wiki manner" [10]. If any changes made to a page cause inconsistencies to the semantic model they are rejected. It also uses the JENA reasoning engine which allows the execution of complex queries. The authoring interface attempts usability improvement by using AJAX to provide suggestions.

Platypuswiki[9] [7] was the first wiki to be officially developed as a Semantic wiki solution and was first made public in 2004. This Java based implementation conceptualises pages as resources. Platypus wiki stores semantic statements in RDF and keeps them separate from the main content of the wiki. It requires the user to have technical knowledge to edit the RDF and consequently the semantics behind each wiki page.

OntoWiki[10] [2] is more of a collaborative ontology editor rather than a Semantic wiki. It does not have the familiar wiki interface of entering natural language text informally to represent a concept but supports several collaborative features and also allows the installation of plug-ins. The OntoWiki implementation conceptualises each page as a resource, storing triple statements in an RDF store. The basis of the implementation is the Powl [1] framework.

Rhizome[11][28] is built using a set of Rx technologies which provide alternatives to traditional standards used by wikis and the Semantic Web: running on the Raccoon application server:ZML (WikiML alternative), RxML (an RDF product), RxPath (a language for querying an RDF model), RxSLT (syntactically identical to XSLT), RxUpdate (a language for updating the RDF model). In Rhizome, RDF is used to also express the wiki information and the main focus "is rather on management of knowledge than on advanced reasoning."[19]

SMW[12] [17,18,23,30] (Semantic MediaWiki) is an extension to Mediawiki[13], enabling it to become a Semantic wiki. It enhances WikiML annotations to allow users to include relations and properties to wiki pages. Several projects already use SMW, for example [22]. Semantic Mediawiki conceptualises pages as concepts [18] and stores the semantic data in Mediawiki's MySQL database, but it can also be exported as RDF. The immediate benefit of enabling the SMW extension in Mediawiki is the presentation of the semantic annotations of each page at the bottom, giving a summary and also improving browsing. Further benefits include semantic search and inline queries.

[5] http://semanticweb.org/, [accessed 2008-04-26]

[6] http://attempto.ifi.uzh.ch/acewiki/, [accessed 2008-04-26]

[7] http://ikewiki.salzburgresearch.at/, [accessed 2008-04-26]

[8] http://www.apps.ag-nbi.de/makna/, [accessed 2008-04-26]

[9] http://platypuswiki.sourceforge.net/, [accessed 2008-04-26]

[10] http://demo.ontowiki.net/, [accessed 2008-04-26]

[11] http://rx4rdf.liminalzone.org/Rhizome, [accessed 2008-04-26]

[12] http://semantic-mediawiki.org/wiki/Semantic_MediaWiki, [accessed 2008-04-26]

[13] http://www.mediawiki.org/wiki/MediaWiki, [acc' 2008-04-26]

SweetWiki[14] [6] identifies itself as a system that uses an ontology for the wiki rather than using a wiki for an ontology. The wiki relies on web standards, such as XHTML, XML and JSPSX. Furthermore, the Semantic web technologies used for the formal structure and its manipulation include RDF, RDFS and OWL, and SPARQL as the query language. The same interface is used for editing both the metadata and the content. The wiki makes use of two ontologies; one for the wiki structure and one for the content and what distinguishes SweetWiki is that many concepts are annotated on the same page.

WikSAR[15][3,4] not only tries to combine the notions of Wiki systems and the Semantic Web but also tries to add features to integrate the Semantic wiki onto the user's desktop. This enables it to be used as a better PIM (Personal Information Management) system. It allows the wiki to be integrated with the user's desktop and allows linking to the local machine (similar to Google Desktop[16]). One of its interesting features is "query chaining" which allows one query to be fed to another query allowing the creation of more interesting and useful queries.

4.2 Comparison

In order to put what each Semantic wiki has to offer into perspective two dimensions were devised to compare the Semantic and Wiki aspects:

Required Knowledge (x axis): This dimension refers to how much technical skill the user needs to have in order to use the wiki and contribute to the ontology. It divides the scale into a number of broad categories of users:

- *Everyday user* – A user familiar with the use of specific applications, but with no administrative, modelling or programming knowledge.

- *Power user* – A user familiar with the use of a wide range of applications and administration of their own computer, but with no modelling or programming knowledge.

- *Professional user* – An advanced user familiar with applications, administration, and with modelling and programming knowledge but no knowledge of Semantic Web technologies (such as OWL and RDF).

- *Ontologist* – Ontology expert user who knows how to create ontologies, how to manage knowledge; they know about the Semantic web and how it works and also how to use Semantic Web technologies.

Expressivity (y axis): This dimension refers to how expressive the final ontology is. The expressivity scale is divided into broad levels of expressivity and formalism.

- *Simple taxonomy* – A taxonomic classification of the wiki pages with no added semantics about relations between pages and concepts.

- *Relations between concepts* – Concepts within the wiki are linked to each other using semantic annotations.

---

[14] http://argentera.inria.fr:8080/wiki/data/Main/MainHome.jsp, [accessed 2008-04-26]

[15] http://wiki.navigable.info/, [accessed 2008-04-26]

[16] http://desktop.google.com/, [accessed 2008-04-26]

- *OWL Lite level* – Better formality than simple relations with some restrictions similar to the functionality offered by OWL Lite.

- *OWL DL, OWL Full level* – Very good formality of a rich ontology that has functionality similar to that of OWL DL and OWL Full, allowing for more restrictions and expressive relations with property characteristics such as transitive, functional, disjoint, etc.

Each Semantic wiki was approximately positioned in the graph of the two dimensions explained above, as shown in **Figure 5**, using the available information and demos found online for each one. A traditional wiki, Mediawiki, was also included on the diagram in order to allow comparisons to the Semantic wikis.

It is not our intention to claim that any position on the graph is better than the others, but it is useful to see the assumptions about knowledge expertise taken by the Semantic Wikis, and to see them in context with a more traditional Wiki.
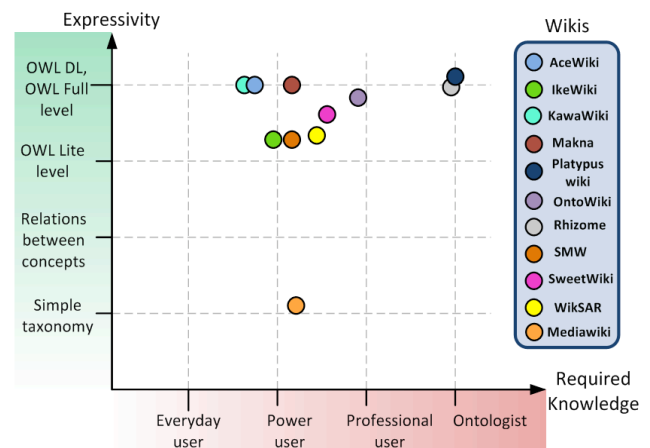

Figure 5. Graph to categorise Semantic wiki implementations.

Wikis on the left side of the graph restrict user input and require little user knowledge (such as AceWiki and KawaWiki). However they require users with ontology expertise to setup, monitor and help grow the wiki.

The wikis on the right of the graph (Rhizome, Platypus wiki and to some degree OntoWiki) although quite expressive are hard to use by non-expert users, as they require knowledge of Semantic Web technologies or elaborate special syntax.

Wikis in the middle of the graph balance out expressivity and required knowledge, such as SMW, IkeWiki, WikiSAR, and SweetWiki.

5. ONTOLOGY CONVERGENCE

Many Semantic Wikis take an instance first approach to creating ontologies that means that the ontology is expressed implicitly through all the instances of pages and links. This is lightweight in that Semantics are created on-demand as opposed to a requirement where a full schema or ontology is declared up front. The disadvantage is that because there is no ontology constraining the types of pages and links that can be created, the implicit ontology evolves in a chaotic manner, possibly resulting in a structure where multiple terms are used to represent the same semantic relationships.

We use the term Ontology Convergence to mean the process of the implicit ontology evolving to a single model. A system with good ontology convergence would end up with a coherent, consistent and well structured implicit ontology, while a system with bad convergence would be incoherent, inconsistent and badly structured, with multiple overlapping and synonymous types.

5.1 MOCA for Semantic MediaWiki

We wanted to study ontology convergence in a Semantic Wiki. We chose to look at a Semantic MediaWiki (SMW) because MediaWiki is one of the most popular wikis (with the largest wiki online already using it: Wikipedia[17]) and its extension, SMW, is perhaps the most mature Semantic wiki implementation [19].

The SMW tries to maintain the WikiML simple approach of the traditional Mediawiki by adding Semantic capabilities to its syntax[18], some of which are shown in Table 1.

Table 1. SMW added syntax

Name	Syntax	Comment
Properties/ Relations	`[[propertyname: :value]]`	For defining properties and relations.
Properties (alternative)	`[[property name:=value]]`	An alternative way to define properties.
Property type	`[[hastype::numb er]]`	Defining the property type of a property concept.

The syntax is not only similar to that of a traditional Mediawiki but it also looks like a triple with the subject being the page being edited itself, the predicate being the property name and value either the linked page or literal value. Although the syntax is simple, especially if the user is familiar with the wiki interface, the user also needs to have an understanding of the background ontology in order to maintain a concise ontology.

Support systems have been developed to try to assist wiki authoring and contribution to the background ontology by providing suggestions and help to the authors. In order to assist the convergence for an existing project, FREMA[19], using SMW, we developed the Mediawiki extension MOCA (MediaWiki Ontology Convergence Assistant). MOCA was developed as an attempt to help the FREMA authors, who were not necessarily technical people and might be classed as everyday users, to be able to use the FREMA wiki[20] and use it to contribute to the background ontology.

Since we developed MOCA other convergence tools have emerged, in particular the Halo extension developed by ontoprise.de[21], which also tries to support Semantic Mediawiki authors to add semantics but also provides other elaborate features for browsing and searching.

---

[17] http://www.wikipedia.org, [accessed 2008-04-26]

[18] SMW annotations manual can be found at: http://semantic-mediawiki.org/wiki/Help:Annotation, [accessed 2008-04-26]

[19] http://www.frema.ecs.soton.ac.uk/, [accessed 2008-04-26]

[20] http://frema.ecs.soton.ac.uk/wiki, [accessed 2008-04-26]

[21] HALO extension page [accessed 2008-04-26]: http://www.mediawiki.org/wiki/Extension:Halo_Extension,

For our study we concentrated on MOCA for simplicity reasons (as it only focuses on authoring assistance). MOCA uses AJAX to provide assistance in the edit page of the wiki, giving help with recommendations for types using the existing background ontology and insertion of annotations without knowledge of the syntax (see Figure 6). MOCA also provides feedback, via a helper wizard while authoring, on what is missing from the page and how it can be corrected (see Figure 7).
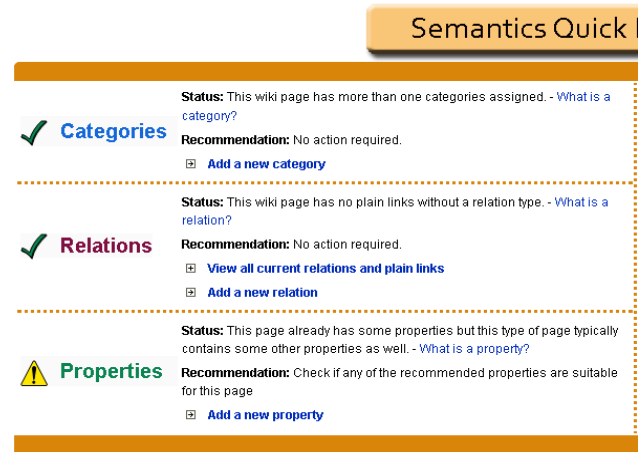

Figure 6.MOCA Semantic wizard.

Figure 7.MOCA Semantics Quick Fix Panel.

6. CONVERGENCE STUDY

In order to explore how ontologies emerged in a semantic wiki, and what effect tools like MOCA might have, we undertook a small-scale study with variably skilled users constructing an ontology through a Semantic wiki.

Setup: To perform this evaluation four different wikis, under the name of TVwiki (shown in Figure 8),were installed in order to look at four different dimensions:

- *SMW1 - Unseeded Semantic Mediawiki:*Standard SMW installation with no initial pages.

- *SMW2 - Seeded Semantic Mediawiki:*Standard SMW installation with some initial pages.

- *MOCA1 - Unseeded Semantic Mediawiki with MOCA:* SMW installation with MOCA and no initial pages.

- *MOCA2 - Seeded Semantic Mediawiki with MOCA:* SMW installation with MOCA and some initial pages.

Sample: The sample was chosen to approximately reflect the four types of users explained in Section 4.2. As this was a small-scale evaluation, we took one user from each user category for each one of the wiki setups giving us a total of 16 participants in 4 groups.

Stages: The evaluation was in general divided into three phases:

- *Phase1:* Accept participation and answer five basic user profiling questions. During this phase users were also

given the option to view a "crash course" video and visit the SMW annotations manual.

- *Phase2:* Receive a 5 day period during which participants would perform given tasks on a TODO list using an assigned wiki setup (from the four mentioned above).

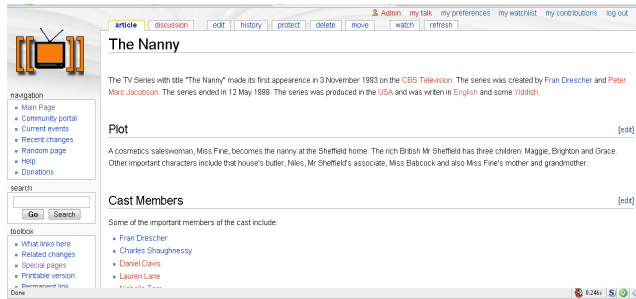- *Phase3:* A twenty minute debriefing interview to learn about the user experience.

Figure 8. Screenshot of TVwiki page setup for evaluation.

Although the sample size is too small for detailed quantitative analysis, the resulting wiki structures are complex enough for us to draw some initial conclusions about ontology convergence according to the 2 dimensions (seeding and assisted authoring).

6.1 Wiki use

The use of the wiki was driven by how motivated or available the participants were. We made a number of observations:

- Ontologists were very keen to create annotated links in general but also to define types in properties.

- Most participants used existing wiki pages to copy and paste the structure.

- Half of the users in both installations with MOCA made use of the MOCA extension at some point during the evaluation.

100% of the users were familiar with the Mediawiki interface (mainly through reading) as they all used Wikipedia.Figure 9 shows the percent of pages, in each of the wiki setups, which were given at least one category:

$$[\% \text{ of categoried pages}] = \frac{[\text{number of categorised pages}]}{[\text{number of pages}]}$$

Figure 10 shows the percent of links, in each of the wiki setups, which were annotated:

$$[\% \text{ of annotated links}] = \frac{[\text{number of annotated links}]}{[\text{number of links}]}$$

We can observe that the seeded wikis and wikis with MOCA have a higher proportion of annotated links.

6.2 Emergent Structure

The RDF for each one of the wiki setups was extracted after the end of the evaluation. By using the RDF Gravity Visualisation

Tool[22], graphs were constructed of the emergent structure (including the seed pages). Most information was removed (including literal values, URIs and categories) from the graphs leaving only the pages and their relations, in order to provide a clearer visualisation. The diagrams generated for SMW1, SMW2, MOCA1 and MOCA 2 are shown respectively in Figure 11, Figure 12, Figure 13 and Figure 14.
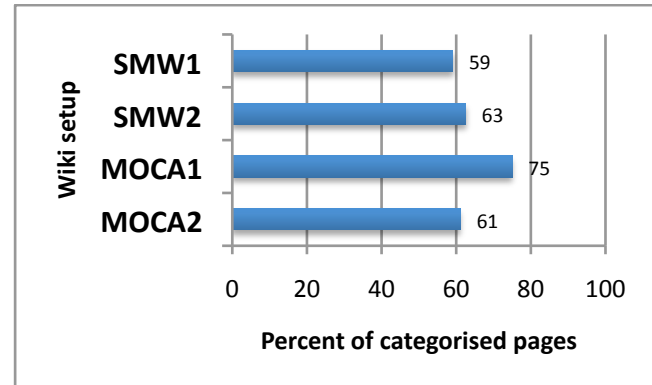
Figure 9. Graph showing percent of categorised pages)
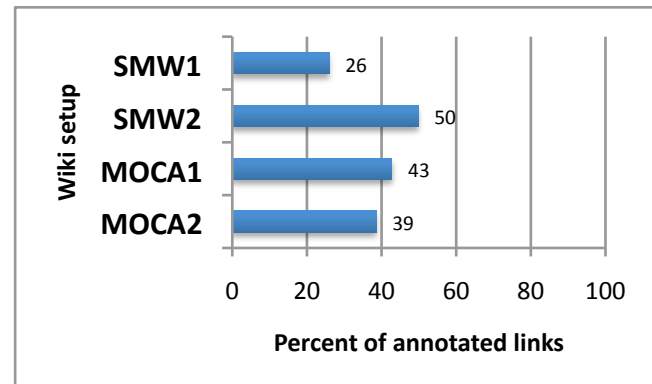
Figure 10. Graph showing percent of annotated links.

It is very clear from the visualisations that the structure created in SMW1 is the least complex, with fewer nodes and relations than wikis with either seeding or that were using the MOCA extension. By comparing SMW1 and MOCA1 we can say that MOCA seems to have helped an unseeded wiki to flourish, encouraged the creation of more nodes and relations, and seems to compensate for the lack of seeding.

6.3 Emergent Ontology

We can also examine the use of page types (classes) and relation types within the wiki structure in order to evaluate the quality of the implicit ontology. Rather than make a qualitative judgement we have used two metrics, based on the reuse of classes and relation types. Our assumption is that greater reuse (i.e. a higher value of instances per class and relations per relation type) means a smaller, tighter ontology.

---

[22] http://semweb.salzburgresearch.at/apps/rdf-gravity/index.html, [accessed 2008-04-26]

Figure 11. SMW1 emergent structure.

Figure 13. MOCA1 emergent structure.

Figure 12. SMW2 emergent structure.

Figure 14. MOCA2 emergent structure.

Figure 15 shows the number of instances of each class that were created in each wiki, and the total instances per class. **Figure 16** shows the total relations per relation type (the data set is to large to show).

The metrics reveal that seeding has a significant effect on the reuse of types within the ontology (in both charts we see a step up from SMW1 to SMW2 and MOCA1 to MOCA2). But the metrics also show that the ontology convergence tool does not seem to have the same effect (no significant difference from SMW1 to MOCA1 and SMW2 to MOCA2), and in fact there is even a minor decrease in reuse.

6.4 User Perceptions

After the study period we ran structured interviews in order to gain some insight into the user experience of each Wiki. We noted a number of interesting comments:

- A perception many of the users had, especially in seeded wikis (or unseeded wikis for which other users had already contributed some information), was that there was only a defined set of annotations they could use. User quote: "I didn't know you could add categories other than 'Series' "

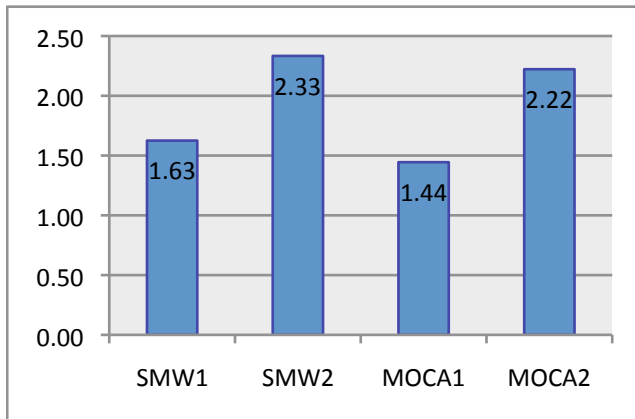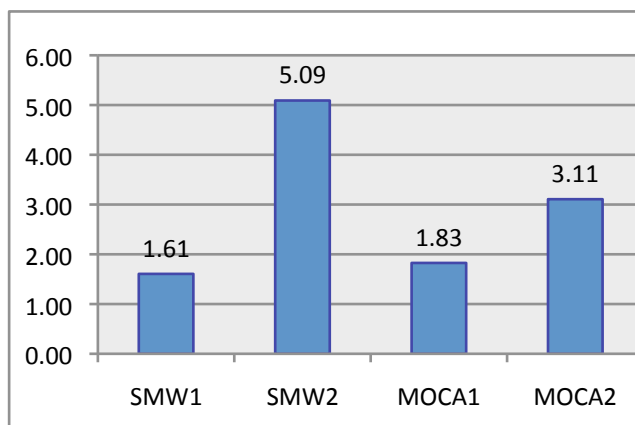| Category (Class)	SMW1	SMW2	MOCA1	MOCA2
Series	4	6	4	4
Actor	1	5	2	4
Actress	1	0	0	0
Channel	2	1	1	2
Character	2	1	1	0
Country	1	1	1	1
Episode	1	0	0	0
Language	1	0	0	0
Comedy	0	3	1	3
Director	0	3	0	2
Producer	0	3	0	1
Romance	0	2	1	1
Science_Fiction	0	1	0	0
TV_Network	0	1	0	0
Writer	0	1	0	2
Drama	0	0	1	0
Mystery	0	0	1	0
No. Instances	13	28	13	20
No. Classes	8	12	9	9
Instances per Class	**1.63**	**2.33**	**1.44**	**2.22**

Figure 15: Instances Per Class

Figure 16: Relations Per RelationType

- Three of the users found MOCA to be useful especially by using it as a checker to help them refine the page to better suit what already had been done in similar pages. User quote: "I looked at the warnings and added stuff it recommended before saving".

- All participants (with the exception of one participant already familiar with the SMW) found the crash course to be very useful and could probably not have started using the system properly without it.

- Ontologists made observations about the Semantic web side of functionality available on SMW, with suggestions such us: allowing the use of other ontologies (i.e. FOAF[23]), some basic inference (i.e. if an actor plays in a series then the actor's page should automatically be categorised to be an actor).

- Participants disliked the need to resave a page in order to get it to refresh its inline queries. User quote: "My series wouldn't show on the homepage so every time I had to go to edit and then save".

- Most ontologists and professional users tried using the inline queries.

- Ontologists suggested there should be help with adding property datatypes. User quote: "It would be nice if there was something helping you adding a datatype".

7. Analysis

Our study shows that Type Suggestion Tools like MOCA do encourage users to add link types in a Semantic Wiki, and also encourage them to create more structure. However in both cases these tools appear to have only the same impact as seeding the Wiki with a clean ontology and initial structure.

We have also shown that seeding the Wiki with a good implicit ontology does significantly increases the reuse of types, but that a tool like MOCA does not increase the reuse of types beyond this. This may seem counter intuitive, after all MOCA is a type suggestion tool, but it may be because the act of suggesting a class or type actually prompts users to think more deeply about classification, and therefore to extend the implicit ontology with more specific types rather than reuse existing ones.

Our conversations with users also suggest that a tool like MOCA fails to impact the ontology because it falls between two types of user; it is too complex for novice users to feel comfortable with, but expert users will just as happily use the wiki mark-up itself. The success of seeding, and our observations of user behaviour, show that for these novice users it is *copy and paste* that dominates both their normal editing and semantic interactions, and that some form of cut and paste palette may well be more useful than a wizard interface.

The small number of participants in our study has made it difficult to distinguish between failure of reuse and refinement of the ontology. We are looking at doing an extended evaluation to study the impact of type suggestion tools over a longer period of time, and intend to further explore the quality of the implicit ontology, looking at factors such as consistency and coherence, in addition to reuse and coverage.

8. CONCLUSION

Although Semantic wikis are relatively new applications they open a fascinating path to the achievement of a Semantic Web created by people, for people. Semantic wikis could generate useful and significant amounts of semantic information, which

---

[23] http://xmlns.com/foaf/spec/, [accessed 2008-04-26]

would alternatively take a long time and a lot of effort by knowledge experts to gather. The greatest penetration of Semantic wikis will most likely occur by implementing the Semantic Wikipedia [18,30]. However if the emerging semantic structures are to be coherent, consistent and well structured it will be necessary to support ontology evolution and convergence. Our study has shown that type suggestion tools may play a part in this, but that it may be more useful to support novice users with reuse through a cut and paste metaphor, and leave expert users to evolve the ontology itself using more advanced tools and the original mark-up. We are currently looking at other visualisations of structure within a semantic wiki and also the possibility of using cut and paste suggestions to encourage the creation of structure and reuse of types.

9. ACKNOWLEDGMENTS

10. REFERENCES

[1] Auer, S. 2005. Powl – A Web Based Platform for Collaborative Semantic Web Development. University of Leipzig. Denmark

[2] Auer, S., Dietzold, S., and Riechert, T. 2006. OntoWiki - A Tool for Social, Semantic Collaboration. University of Pennsylvania, Department of Computer and Information Science, Philadelphia, USA and Universität Leipzig, Institut für Informatik, Leipzig, Germany.

[3] Aumueller, D. 2005. Semantic authoring and retrieval within a Wiki. Department of Computer Science, University of Leipzig. Germany

[4] Aumueller, D., and Auer, S. 2005. Towards a Semantic Wiki Experience - Desktop Integration and Interactivity in WikSAR. Department of Computer Sience, University of Leibzig. Leipzig, Germany

[5] Berners-Lee, T., Hendler, J., and Lassila, O. 2001. The Semantic Web. Scientific American.

[6] Buffa, M., and Gardon, F. 2006. SweetWiki: Semantic Web Enabled Technologies in Wiki. Mainline Group, I3S Laboratory, University of Nice, France - Acacia Group, INRIA laboratory. France

[7] Campanini, S. E., Castagna, P., and Tazzoli, R. 2004. Platypus Wiki: a Semantic Wiki Wiki Web.

[8] Cunningham, W., 2008, Wiki Design Principles, c2.com, http://c2.com/cgi/wiki?WikiDesignPrinciples, [Online; accessed 2008-04-26]

[9] Cunningham, W., and Leuf, B., 2002, What Is Wiki, c2.com, http://www.wiki.org/wiki.cgi?WhatIsWiki, [Online; accessed 2008-04-26]

[10] Dello, K., Simperl, E. P. B., and Tolksodorf, R. 2006. Creating and using Semantic information with Makna. Institut für Informatik, Freie Universität Berlin. Berlin, Germany

[11] Désilets, A., Paquet, S., and Vinson, N. G. 2004. Are Wikis Usable? National Research Council of Canada, Institute for Information Technology. Canada

[12] Dillon, A. 2001. Beyond Usability: Process, Outcome and Affect in human computer interactions. Faculty of Information Studies, University of Toronto. Toronto

[13] Elements Of Wiki Essence, 2008, c2.com, http://c2.com/cgi/wiki?ElementsOfWikiEssence, [Online; accessed 2008-04-26]

[14] Fuchs, N. E. 2007. Reasoning, Rules and Semantic Wikis. University of Zurich. Zurich

[15] Gómez-Pérez, A. 2001. Evaluation of Ontologies. Facultad de Informática, Universidad Politécnica de Madrid. Madrid, Spain

[16] Kawamoto, K., Kitamura, Y., and Tijerino, Y. 2006. KawaWiki: A Semantic Wiki Based on RDF Templates. Kwansei Gakuin Universwity. Japan

[17] Krötzsch, M., Vrandečić, D., and Völkel, M. 2006. Semantic MediaWiki. AIFB, Universität Karlsruhe, Germany - FZI Karlsruhe,Germany

[18] Krötzsch, M., Vrandečić, D., Völkel, M., Haller, H., and Studer, R. 2007. Semantic Wikipedia. Institut AIFB, Universit¨at Karlsruhe (TH), Germany - FZI. Karlsruhe, Germany

[19] Kuhn, T. 2008. AceWiki: A Natural and Expressive Semantic Wiki. Department of Informatics, University of Zurich.

[20] Leuf, B., and Cunningham, W. 2001. The Wiki Way Quick Collaboration on the Web.

[21] Millard, D. E., Gibbins, N. M., Michaelides, D. T., and Weal, M. J. 2005. Mind the Semantic Gap. IAM Group, University of Southampton. UK

[22] Millard, I., Jaffri, A., Glaser, H., and Rodriguez, B. 2006. Using a Semantic MediaWiki to Interact with a Knowledge-Based Infrastructure. School of Electronics and Computer Science, University of Southampton. UK

[23] Muljadi, H., and Takeda, H. 2005. Semantic Wiki as an Integrated Content and Metadata Management System. National Institute of Informatics, Research Center for Testbeds and Prototyping. Tokyo, Japan

[24] Schaffert, S. 2006. IkeWiki: A Semantic Wiki for Collaborative Knowledge Management. Salzburg Research Forschungsgesellschaft/Salzburg New Media Lab. Austria

[25] Schaffert, S., Gruber, A., and Westenthaler, R. 2005. A Semantic Wiki for Collaborative Knowledge Formation. Knowledge-based Information Systems Group, Salzburg Research. Austria

[26] Semantic Wiki State Of The Art, 2007, ontoworld.org, http://www.semanticweb.org/wiki/Semantic_Wiki_State_Of_The_Art, [Online; accessed 2008-04-26]

[27] Shadbolt, N., Hall, W., and Berners-Lee, T. 2006. The Semantic Web Revisited. University of Southampton and Massachusetts Institute of Technology.

[28] Souzis, A. 2005. Building a Semantic Wiki. Liminal Systems.

[29] Tolksdorf, R., and Simperl, E. P. B. 2006. Towards Wikis as Semantic Hypermedia. Institute of Computer Science, Networked Information Systems. Berlin, Germany

[30] Völkel, M., Krötzsch, M., Vrandečić, D., Haller, H., and Studer, R. 2006. Semantic Wikipedia. Institute AIFB, University of Karisruhe. Karlsruhe Germany

[31] Voss, J. 2006. Collaborative thesaurus tagging the Wikipedia way. Wikimedia Deutschland e.V.

[32] Yeung, C.-m. A., Gibbins, N., and Shadbolt, N. 2007. Understanding the Semantics of Ambiguous Tags in Folksonomies. School of Electronics and Computer Science, University of Southampton. Southampton, UK