

Visual Knowledge Discovery on Semantic Wikis

Tomislav Štritof, Markus Schatten

Faculty of Organization and Informatics

University of Zagreb

Pavlinska 2, 42000 Varaždin, Croatia

{tštritof,markus.schatten}@foi.hr

Abstract. *A different approach to visualizing formalized knowledge on social semantic web applications is proposed and implemented in to the TaOPis system. A knowledge visualizer is build on top of the semantic wiki subsystem, using OWL based ontologies which are generated from user supplied meta-data. The application displays visual connections between classes and instances, as well as provides access to object properties and allows for visual search of related tags. Various use-cases are presented and analyzed which show that visual search has major advantages to formal niKlas-based querying.*

Keywords. TaOPis, knowledge visualization, semantic wiki, OWL

1 Introduction and Related Work

The visualization of knowledge bases and ontologies is not a new field [9, 1, 2, 5] and there have been various tools and techniques proposed [10]. On the other hand visualisation the structure and dynamics of wikis is subject to intensive research and development in recent years [16, 4, 8, 17]. The visualization of semantic wikis and its use in knowledge management, however, has got rather small attention of the scientific community. There were some interesting applications like Fact Boxes and Ontology browser for the Semantic Media Wiki engine [7], ontology graph visualization in WikiSAR [3], or visualization of wiki semantics using topic maps [6].

Herein we will present a different approach to se-

mantic wiki visualization with an accent on knowledge discovery. A visualization module for the TaOPis system has been developed that is directly connected to the user defined semantics of a given wiki. For TaOPis, a community management Web 3.0 application, already having a formal querying mechanism (the niKlas language) this is an excellent addition since it fosters ease of use. On the other hand, visual knowledge discovery isn't as powerful as niKlas since it doesn't make use of the social network surrounding the wiki.

2 TaOPis, niKlas and OWL

TaOPis is a social semantic web application for self-organizing communities providing suitable tools like semantic wiki systems, forums, blogs, ranking mechanisms, content filtering, tagging etc [12, 13, 15]. The system has been used for almost three years for various purposes, but most projects deal with knowledge management due to the fact that the system is used in a knowledge management course at the Faculty of organization and informatics. Especially its semantic wiki subsystem is of special interest since it accumulates meta-information which is of great importance to finding relevant information.

TaOPis uses the so called niKlas wiki language which has build in querying mechanisms [14]. The niKlas syntax, which is similar to phpBB code¹, defines beside different text formatting commands, a query command based on frame logic [11] and especially the FLORA-2 reasoning engine [18] with the following syntax:

¹phpBB is a free open-source forum and messaging board system available at <http://www.phpbb.net>

```
[query flora2_query. ]
  {[amalgamate project_list ]}
  {[probability constraint]}
  {[header]header_formatting[/header]}
  answer_formatting
[/query]
```

whereby *flora2_query* is a normal (restricted) *FLORA-2* query with defined return variables, *project_list* is a list of projects (each having its own knowledge base) which should be amalgamated, *constraint* is a probability constraint making use of social network analysis, *header_formatting* is the optional header (possibly formatted using **niKlas** code),² and *answer_formatting* is a **TAOPis** formatting that can contain variables used in the *FLORA-2* query. The *answer_formatting* is repeated for any answer returned by the *FLORA-2* reasoning engine by using the generated *FLORA-2* ontology of the semantic wiki as a knowledge base.

As an illustrative example the following query would generate a list of users.

```
[query ?_:user[ name->?n ],
      sort(?n,asc). ]
[header][b]Users[/b][[/header]
name: ?n
[/query]
```

The *FLORA-2* query *?_:user[name->?n, surname->?s]* is issued against the dynamic knowledgebase of the system. The obtained results are then replicated in the answer formatting; each answer prints out one answer formatting. On the other hand the header section will be printed only once. Thus the result of this query would be similar to:

```
[b]Users[/b]
name: Markus
name: Tomislav
```

In the second phase, after the query generated a formatting, the rest of the **niKlas** code is translated to HTML but other target languages could be implemented. In this case, the HTML encoded answer would be:

```
<b>Users</b>
name: Markus<br />
```

²In the current version subqueries are not allowed.

```
name: Tomislav<br />
```

TAOPis implements a dynamic ontology export facility, allowing *FLORA-2* and Web Ontology Language (OWL) export. The ontologies are dynamically constructed depending on the user-supplied meta data. In this way **TAOPis** wikis can be used to implement intelligent reasoning agents or an ontology visualization like it is done in this paper.

In the mentioned export facility, any wiki page, forum post, blog entry or comment, user or project is considered to be an instance of a proper class³. For example an instance of a wiki page might look like the following in an exported *FLORA-2* ontology:

```
object_id_5352:wiki_page[
  'poglavlje' -> {'4'},
  'author' -> {'markus.schatten'},
  'url' -> 'http://autopoiesis.foi.hr/
           wiki.php?name=KM+FOI&
           parent=NULL&page=semmodtao',
  'title' -> 'semmodtao',
  'naslov' -> {'TaOPis - FAQ'},
  'type' -> 'wiki'].
```

The same instance would look as follows in OWL:

```
<WikiPage rdf:ID="object_id_5352">
<poglavlje>4</poglavlje>
<author
  rdf:resource="#markus.schatten"/>
<url>
http://autopoiesis.foi.hr/
  wiki.php?name=KM+FOI&
  parent=NULL&page=semmodtao
</url>
<title>semmodtao</title>
<naslov>TaOPis - FAQ</naslov>
</WikiPage>
```

As one can see, each instance has some predefined attributes from which the url attribute is of special importance for this study. The url attribute always points to the absolute URL of a given object and thus provides us with the possibility to find it on the wiki system.

³There is a small predefined vocabulary in **TAOPis** but new classes, attributes, relations (*FLORA-2* and OWL) and rules (only for *FLORA-2*) can be added at will.

3 Semantic Wiki Visualization Implementation

All data needed for visualization is gathered from an OWL file which is generated dynamically for each wiki. The OWL export is essentially an XML file with defined class hierarchy where classes have supplemented metadata. OWL language is proposed and recommended by W3C as standard a format for data integration (information and metadata) and one of the fundamental building blocks of the Semantic Web.

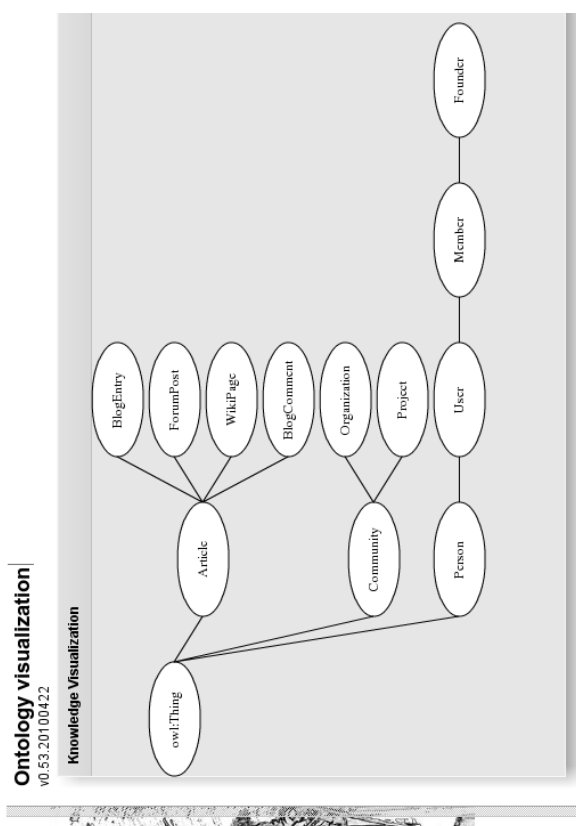


Figure 1: Sample visualization of a semantic wiki ontology

The implementation of Wiki Visualization is written in the MXML/ActionScript3 programming language and requires Adobe Flash Player VM on the end user side to be run. The application itself is open source and can be compiled with free available Adobe Flash Builder 4 SDK. First part in the process of visualizing begins with parsing

the OWL/RDF XML file. As the parsing goes, the data from the file is populated in tree data structure. Each node in the tree represents a class from the ontology. Other information about a specific class is also bound to the tree node, such as object information (attributes) and class instances (members). The second part of the application deals with visualizing the data from the gathered tree data structure which consists of drawing ellipsis (representing classes), tweaking position of ellipsis to look like a tree, drawing connections between ellipsis and lastly, making them interactive (setting mouse event handlers and functions). Figure 1 shows a sample visualization.

4 Use-Cases

Figure 2 shows the use case diagram of the application. The user's role is to pick Visual Knowledge Discovery option on a project page on the TaOPis web application and then visualization starts. User can then explore class attributes or pick classes to open a list of wiki pages tagged with the adequate attribute-value tag.

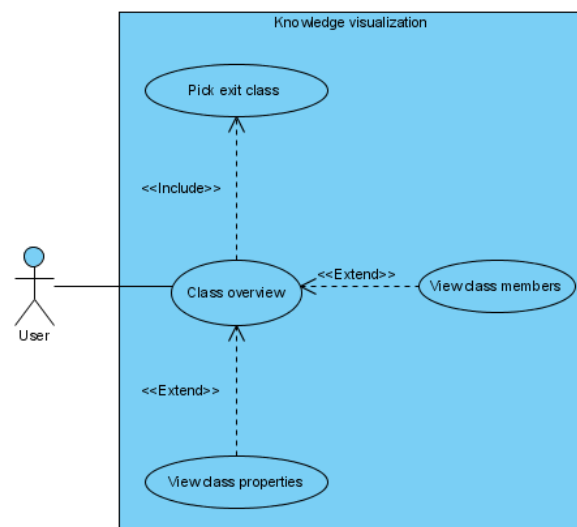


Figure 2: Knowledge visualization use-case diagram

For example if a user clicks on class definicija⁴ the output would be similar to the one on figure 3. As

⁴Definicija is Croatian for definition.

one can see the list contains all pages tagged with `class:definicija` all over the system (not only on the particular project the visualization was used on). These pages represent the instances of the selected class. By clicking on some instance the page shows up, together with its properties (attributes and values) which are again followable links. By clicking on a selected attribute-value pair a list of pages all over the system, containing this properties shows up. In this way users can discover similar pages and discover new knowledge.



Figure 3: Partial list of wiki pages tagged with `class:definicija`

In order to do a similar thing using the `niKlas` querying mechanism the user would have to issue a query of the following form:

```
[ query ?_ :class [
  attribute -> value ,
  title ->?p
  url ->?u ] . ]
[ amalgamate
  " project_1 "
  " project_2 "
  . . .
  " project_N " ]
[ link=?u>?p ]
[ /query ]
```

Where `class` is the class name the user wants to follow and `attribute` and `value` represent the selected property. Such a query requires relatively advanced knowledge of the user. In order to get results from multiple projects the user would have to use the amalgamation facility. Thus, ontology visualization yields better ease of use for the less conversant users.

5 Conclusion

We showed how a simple class hierarchy visualization system build on top of `TAOPIS` can be used to discover new and unknown knowledge on semantic wikis. The advantages over the build-in `niKlas` querying mechanism are a visual approach, intuitive user interface and interactivity which in the bottom line fosters ease of use. Still, the ontology visualization facility isn't as powerful as the `niKlas` querying mechanism since it only allows for a pre-defined set of queries (browse instances by class, browse instances by property). Thus a combination of the two approaches for knowledge discovery would be ideal.

We envision to build a visual querying mechanism (builder) by introducing logic variables into the visualization. In this way queries could be constructed by creating visual templates, which then in turn could be translated to `niKlas` queries. These and similar issues are subject to our future research.

References

- [1] ANDRIENKO, G., AND ANDRIENKO, N. Knowledge-based visualization to support spatial data mining. In *In Proceedings of the 3rd Symposium on Intelligent Data Analysis* (1999), SpringerVerlag, pp. 149–160.
- [2] ANNE STOREY, M., MUSEN, M., SILVA, J., BEST, C., ERNST, N., FERGERSON, R., AND NOY, N. Jambalaya: Interactive visualization to enhance ontology authoring and knowledge acquisition in protégé. In *in Protégé. Workshop on Interactive Tools for Knowledge Capture (K-CAP-2001)* (2001).
- [3] AUMUELLER, D., AND AUER, S. Towards a semantic wiki experience - desktop integration and interactivity in wiksar. In *1st Workshop on The Semantic Desktop, Next Generation Personal Information Management and Collaboration Infrastructure* (2005).
- [4] BIUK-AGHAI, R. P. Visualizing co-authorship networks in online wikipedia. In *International Symposium on Communications and Information Technologies, 2006. ISCIT '06* (September 2006), pp. 737–742.

- [5] BUITELAAR, P., EIGNER, T., GULRAJANI, G., SCHUTZ, A., SIEGEL, M., WEBER, N., CIMIANO, P., LADWIG, G., MANTEL, M., AND ZHU, H. Generating and visualizing a soccer knowledge base. In *EACL '06: Proceedings of the Eleventh Conference of the European Chapter of the Association for Computational Linguistics: Posters & Demonstrations* (Morristown, NJ, USA, 2006), Association for Computational Linguistics, pp. 123–126.
- [6] ESPIRITU, C., STROULIA, E., AND TIRAPA, T. Enwic: Visualizing wiki semantics as topic maps. In *Proc. of 8th Int. Conf. on Enterprise Information Systems* (2006).
- [7] GEDIGIAN, M. Semantic wiki visualization. Available at <http://vis.berkeley.edu/courses/cs294-10-fa08/wiki/images/8/84/FA08-MattGedigian-Paper.pdf>, 2008.
- [8] HOLLOWAY, T., BOŽIČEVIĆ, M., AND BÖRNER, K. Analyzing and visualizing the semantic coverage of wikipedia and its authors. *Complexity, Special issue on Understanding Complex Systems* 12, 3 (2007), 30–40.
- [9] HÜSER, C., REICHENBERGER, K., ROSTEK, L., AND STREITZ, N. Knowledge-based editing and visualization for hypermedia encyclopedias. *Communications of the ACM* 38, 4 (April 1995), 49–51.
- [10] KATIFORI, A., HALATSIS, C., LEPOURAS, G., VASSILAKIS, C., AND GIANNOPOULOU, E. Ontology visualization methods—a survey. *ACM Comput. Surv.* 39, 4 (2007), 10.
- [11] KIFER, M., LAUSEN, G., AND WU, J. Logical foundations of object-oriented and frame-based languages. *Journal of the Association for Computing Machinery* 42 (May 1995), 741–843.
- [12] MALEKOVIĆ, M., AND SCHATTEN, M. Leadership in team based knowledge management - an autopoietic information system's perspective. In *19th Central European Conference on Information and Intelligent Systems – CECIIS2008 Conference Proceedings* (September 2008), B. Aurer and M. Bača, Eds., Faculty of Organization and Informatics, pp. 47–52.
- [13] SCHATTEN, M., ČUBRILO, M., AND ŠEVA, J. A semantic wiki system based on f-logic. In *19th Central European Conference on Information and Intelligent Systems – CECIIS2008 Conference Proceedings* (2008), B. Aurer and M. Bača, Eds., Faculty of Organization and Informatics, pp. 57–61.
- [14] SCHATTEN, M., ČUBRILO, M., AND ŠEVA, J. Dynamic queries in semantic wiki systems. In *Proceedings of the 20th Central European Conference on Information and Intelligent Systems* (Sept. 2009), B. Aurer and M. Bača, Eds., Faculty of Organization and Informatics, pp. 13–20.
- [15] SCHATTEN, M., AND ŽUGAJ, M. Organizing a fishnet structure. In *29th International Conference Information Technology Interfaces Proceedings* (Cavtat-Dubrovnik, Croatia, June 25 — 28 2007), pp. 81–86.
- [16] VIÉGAS, F. B., WATTENBERG, M., AND DAVE, K. Studying cooperation and conflict between authors with history flow visualizations. In *CHI '04: Proceedings of the SIGCHI conference on Human factors in computing systems* (New York, NY, USA, 2004), ACM, pp. 575–582.
- [17] WATTENBERG, M., VIÉGAS, F., AND HOLLENBACH, K. Visualizing activity on wikipedia with chromograms. 2007, pp. 272–287.
- [18] YANG, G., KIFER, M., AND ZHAO, C. Flora-2: A rule-based knowledge representation and inference infrastructure for the semantic web. In *Second International Conference on Ontologies, Databases and Applications of Semantics (ODBASE) Proceedings* (Catania, Sicily, Italy, November 2003).