Semantic Wikis for Knowledge Workers*

Eyal Oren
Digital Enterprise Research Institute
National University of Ireland, Galway
eyal.oren@deri.org

ABSTRACT

Knowledge is crucial in our society and fundamentally created by individuals. Knowledge workers require support for authoring, finding and reminding, knowledge reuse, and collaboration. We present personal semantic wikis as supporting technology for knowledge workers, outline an architecture, and introduce our prototype implementation.

1. INTRODUCTION

Managing and enabling knowledge is key to success in our economy and society [4]. Knowledge is fundamentally created by individuals [4, p. 59]. Supporting these knowledge workers in their *personal* knowledge management is therefore crucial.

Current tools for personal knowledge management are limited: traditional (physical) approaches are not automated, computerised approaches are restrictive, do not support ad hoc structures, and do not support associative work.

Based on the requirements of knowledge workers we present personal semantic wikis as a promising supporting technology. We introduce some background on knowledge workers, and present our architecture and implementation of a personal semantic wiki.

2. BACKGROUND

Knowledge is "justified true belief"[4, p. 21]: it is a personal belief justified by information. Individuals are continuously personally committed to knowledge creation [6]. This personal commitment relies on the intentions and autonomy of individuals [3]. Intention defines the understanding and actions of an individual, autonomy gives him the self-motivation and freedom to absorb and create knowledge.

There are two types of knowledge, tacit knowledge and explicit knowledge [7]. Explicit knowledge is transmittable and codified in a formal language; tacit knowledge is hard to formalise but is embedded in individual experience, insights and skills.

Knowledge is created through conversions between tacit and explicit knowledge [4, p. 62–73]: (a) different sources of explicit knowledge can be *combined* to form new knowledge; (b) tacit knowledge can be *externalised* into explicit knowledge, through metaphors and codification; (c) explicit knowledge is *internalised* into tacit knowledge by acting, doing, and learning; and (d) tacit knowledge can be transferred through socialisation – without language, but through observation, imitation, and practise.

Knowledge workers produce new information by combining an existing body of knowledge [1]. "Ideas are formed in the minds of individuals and are developed in social interactions" [3]. Supporting knowledge workers means to support individual internalisation (learning) and externalisation (writing), and to support their interaction, their socialisation, their sharing of knowledge. We derive the following requirements [9]:

1. **authoring**
2. **finding and reminding**
3. **knowledge reuse**
4. **collaboration**

Since in online collaboration true socialisation is not an option, knowledge has to be shared explicitly through externalisation and internalisation. To externalise knowledge one needs a simple and adequate authoring environment. To internalise knowledge one needs to find it and be reminded of it [10]. Finding involves recall directed search and recognition based scanning [2]. Searching requires metadata, scanning requires preserving spatial layout and physical marks [1]. Since people are passive in finding information [8], related information should be presented automatically for associative browsing.

3. SEMANTIC WIKI

Semantic wikis are enhanced classical wikis. Classical wikis are collaborative hypertext environments, focused on open access, ease-of-use, and modification. Semantic wikis allow users to make formal descriptions of resources by annotating the pages that represent those resources. Where a regular wiki enables users to describe resources in natural language, a Semantic wiki enables users to additionally describe resources in a formal language.

Using the annotations (descriptions) of resources semantic wikis offer additional features over regular wikis. Users can find information directly ("show me all authors") navigate the wiki associatively ("go to other books by John Grisham") or create views from such queries. and users can

*This material is based upon works supported by the Science Foundation Ireland under Grants No. SFI/02/CE1/I131 and SFI/04/BR/CS0694. Extended versions of this paper have been submitted to WWW2006 and ESWC2006.

reuse background knowledge in the system ("all poets are authors; show me all authors").

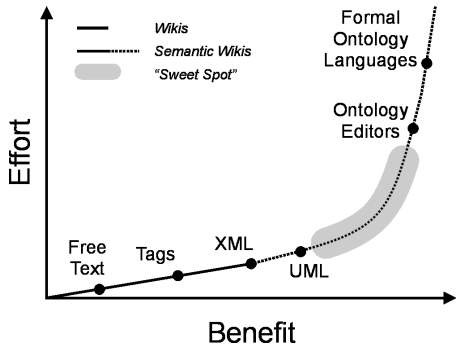

Figure 1: Effort and benefit in knowledge authoring

Semantic wikis support the identified requirements of knowledge workers: **authoring** of syntactical, structural, and semantical knowledge is possible through one uniform interface with gradual increase of benefit and effort (shown in figure 1); **finding** information is possible through querying and associative browsing; **knowledge reuse** is possible through views and rules for background knowledge; and **collaboration** is possible through shared storage or peer-to-peer connections.

3.1 Architecture Overview

A semantic wiki consists (at least) of the following components: a user interface, a parser, a data analyser, and a data store, as shown in figure 2.

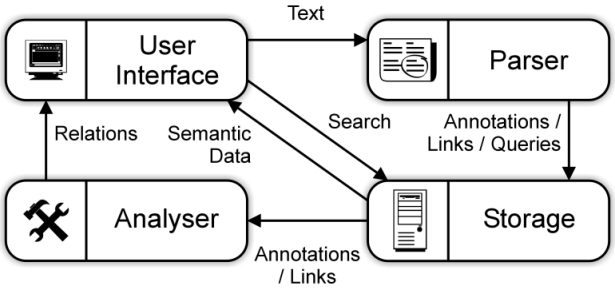

Figure 2: Architecture of a semantic wiki

Users can browse, edit, and query pages via the user interface. When users edit a page, the user interface notifies the parser. The parser analyses the text, and extracts annotations and links. All data (text, annotations, etc.) are stored in the semantic storage. From the data in the storage, the analyser computes sets of pages that are related to the current page, which are displayed by the user interface. Queries are posed to the storage, and the results are displayed by the user interface. All these operations should happen unobtrusively in the background, as to provide the user a responsive application.

3.2 Annotation language

The for the user of a semantic wiki most visible change compared to conventional wikis is the modified *annotation language*. For semantic wikis the annotation language is not only responsible for change in text style and for creating links, but also for the semantic annotation of wiki pages and for writing embedded queries in a page.

3.2.1 Annotation primitives

Internal links are written in CamelCase or by enclosing them in brackets; external links are written as full absolute URIs, or are abbreviated using namespace abbreviations. Internal wiki links are expanded to absolute URIs using the usually configurable wiki base namespace, namespace abbreviations are expanded using configurable namespace definitions.

Semantic annotations are written on a separate line, and, following RDF conventions, consist of a predicate followed by an object. Predicates can only be resources, objects can be either resources or literals. Annotations are expanded to triples using *the resource that the page is describing* as the subject of the triple. To annotate the current wiki page itself, instead of the resource that the page is describing, annotations have to be pre-pended with an exclamation mark (since annotations of resources are more common than annotations of pages).

An example page is displayed in figure 3. It describes John Grisham, an author published by Random House. Information about John Grisham is mixed between natural text in English and formal annotations.

```
Page:  JohnGrisham
John Grisham is an author and retired lawyer.

rdf:type foaf:Person
dc:publisher RandomHouse
```

Figure 3: Example page

3.2.2 Advanced annotations

In case an author of a wiki page does not have naming authority over the resource it is necessary to explicitly specify the URI of the described resource. For example, figure 4 shows a page that describes the research institute DERI. The page uses the "semper:about" predicate to relate the page to the resource (DERI, identified by `urn://deri.ie`) that it is describing. The annotations state, using the Semantic Web Research Community ontology, that DERI is a research institute, founded in June 2003, and located in Galway. The last annotation, prepended with an exclamation mark, refers to the page instead of the resource; it states that Eyal Oren is the creator of that page.

```
Page:  DERI Galway
DERI Galway is one location of the Digital
Enterprise Research Institute, researching
Semantic Web technology; our main page is at
http://deri.org.

semper:about urn://deri.ie
rdf:type swrc:Organization
swrc:location "Galway"
swrc:created "2003-06-01"
!dc:creator EyalOren
```

Figure 4: Annotating real-world resources

3.2.3 Embedded queries

Embedded queries (generating views) are written using triple patterns, sequences of subject, predicate, object, that can contain variables (names that start with a question mark). A triple pattern is interpreted as a query: triples matching the pattern are returned. Patterns can be combined to form joins.

Figure 5 shows the earlier example page about John Grisham, including an embedded query at the bottom of the page. The query returns all books written by JohnGrisham; it creates a view on the data that is displayed below the page text.

```
Page:  JohnGrisham
-----------------------------------------
John Grisham is an author and retired lawyer.

rdf:type foaf:Person
dc:publisher RandomHouse

this query shows all his books:
?book dc:creator JohnGrisham
-----------------------------------------
TheFirm dc:creator JohnGrisham
TheJury dc:creator JohnGrisham
ThePelicanBrief dc:creator JohnGrisham
```

Figure 5: Page showing embedded query

3.3 Information access

Information access is offered through structured navigation and various querying facilities.

A semantic wiki provides the metadata necessary to navigate the information in a structured way. For example, knowing that John Grisham is an author, we can automatically show all other authors in the system, and offer navigation to them.

One approach for viewing and navigating related items is faceted browsing [11], in which the information space is partitioned in conceptual dimensions that constrain the currently visible elements of the information space. For example, a collection of art works can consist of facets such as type of work, time periods, artist names, geographical locations, etc. To implement faceted browsing one normally needs a schema that defines the facets, but in browsing arbitrary RDF we cannot rely on having a schema definition of the data. We therefore use **faceted triple browsing** [9], a technique for partitioning, displaying, and navigating arbitrary schemaless RDF data. The facets of related items are shown in the sidebar of the wiki.

4. SEMANTIC WIKI: IMPLEMENTATION

We have implemented the architecture in a prototype called SemperWiki [5]. SemperWiki is a desktop application but the architecture is equally applicable to Web systems.

SemperWiki is a semantic personal wiki that can be used used for personal knowledge management. The main advantages compared to a normal wiki are intelligent navigation, semantic search, and embedded queries. All information in SemperWiki can be annotated semantically, and all information can be exported and shared on the Semantic Web.

SemperWiki is implemented in Ruby[1], using the GTK[2] windowing toolkit for the graphical programming. It is open source, consists of around 1500 lines of code, and can be downloaded at `http://semperwiki.org`.

5. RELATED WORK

Compared to other tools for personal knowledge management, semantic wikis fill a clear gap in flexible authoring [9]. Compared to other semantic wikis, SemperWiki offers high usability and supports integration with existing Semantic Web knowledge [5].

6. CONCLUSION

We have presented semantic wikis as supporting technology for knowledge workers. We introduced requirements for knowledge workers and shown how semantic wikis address these. We have described our architecture for semantic wikis and discussed our prototype briefly.

Acknowledgements. Thanks to Max Völkel, Stefan Decker, and John Breslin for joint work that lead to this paper.

7. REFERENCES

[1] A. Kidd. The marks are on the knowledge worker. In *CHI '94: Proc. of the SIGCHI conf. on Human factors in computing systems*, pp. 186–191. ACM Press, 1994.

[2] M. Lansdale. The psychology of personal information management. *Applied Ergonomics*, 19(1):55–66, 1988.

[3] I. Nonaka. A dynamic theory of organizational knowledge creation. *Organization Science*, 5(1):14–37, Feb. 1994.

[4] I. Nonaka and H. Takeuchi. *The Knowledge-Creating Company*. Oxford University Press, New York, 1995.

[5] E. Oren. SemperWiki: a semantic personal wiki. In S. Decker *et al.*, (eds.) *Proceedings of the Semantic Desktop Workshop*. Galway, Ireland, 2005.

[6] M. Polanyi. *Personal Knowledge: Towards a Post-Critical Philosophy*. Routledge & Kegan Paul Ltd, London, 1958.

[7] M. Polanyi. *Tacit Dimension*. Routledge & Kegan Paul Ltd, London, 1966.

[8] J. Teevan, C. Alvarado, M. S. Ackerman, and D. R. Karger. The perfect search engine is not enough: a study of orienteering behavior in directed search. In *CHI '04: Proc. of the SIGCHI conf. on Human factors in computing systems*, pp. 415–422. ACM Press, 2004.

[9] M. Völkel and E. Oren. Personal knowledge management with semantic wikis, Dec. 2005. *Submitted*.

[10] S. Whittaker and C. Sidner. Email overload: exploring personal information management of email. In *CHI '96: Proc. of the SIGCHI conf. on Human factors in computing systems*, pp. 276–283. ACM Press, 1996.

[11] K.-P. Yee, K. Swearingen, K. Li, and M. Hearst. Faceted metadata for image search and browsing. In *CHI '03: Proc. of the SIGCHI conf. on Human factors in computing systems*, pp. 401–408. ACM Press, 2003.

[1] `http://ruby-lang.org/`.
[2] `http://gtk.org`.