Using semantic technology in intelligence analysis

Christian Mårtenson
FOI - Swedish Defence Research Agency
SE 164 90 Stockholm, Sweden
Email: cmart@foi.se

Andreas Horndahl
FOI - Swedish Defence Research Agency
SE 164 90 Stockholm, Sweden
Email: andreash@foi.se

***Abstract*—In this paper, we discuss how modern intelligence information management can be enhanced by using wiki and semantic web technology concepts, and how the fusion of these concepts, the semantic wiki, could be beneficial in helping decision-makers achieve situational awareness. The Semantic MilWiki, a semantic wiki prototype developed at the Swedish Defence Research Agency, is described, as well as initial findings regarding its usability in intelligence contexts. The Semantic MilWiki has so far been tested in a workshop devoted to high-level intelligence processing and in an experiment on information processing for low-level commanders.**

I. Introduction

The intelligence community, in Sweden and elsewhere, is currently facing a number of new challenges. The most apparent is the new kinds of threats posed by global terrorism and by the Swedish Armed Forces increasing engagement in international operations other than war (OOTW), such as peace-keeping or peace-enforcing. These threats make it essential to broaden the scope of intelligence gathering in terms of geographical as well as cultural coverage. In combination with comprehensive civil-military operational approaches this leads to an increased demand for information sharing, within the intelligence organization as well as with coalition partners and civil agencies. In short, there is an overwhelming amount of information to be managed and interpreted in order to maintain acceptable situation awareness. The magnitude of this task requires new tools for efficient information management and also new methodologies for how to handle information.

In modern warfare, which mostly deals with OOTW, the physical locations of platforms and people still play a role for achieving situation awareness, but other parameters are necessary to complete the picture. The most important things are no longer the positions of threatening platforms. Instead, it is vital to keep track of more soft information, such as the mood of the people and the allegiances and loyalties of various militant groups that are operating in the area. While traditional systems mostly rely on information from sensors, it is also becoming more and more critical to be able to fuse textual information, and display, for example, relevant summaries of what is being reported in an area.

Such alternative situation pictures can fruitfully be displayed in a wiki-like setting. A wiki is a website with special functionality dedicated to supporting online collaborative information management using a standard web browser [1]. Most users are used to navigating the web and searching for information, which means that they will find the wiki interface appealing and immediately useful. Wiki pages also allow for integration of different kinds of information (text, images, sound and video files).

A *semantic* wiki has the potential to be an even more powerful tool to create and display alternative situation pictures in. It offers an appealing way to combine structured and unstructured information in a single collaborative framework. Natural language texts and hyperlinks can be annotated collaboratively using formal semantics. The semantics can be queried and the results can be inserted as dynamic content into wiki pages mixed with plain text. Implicit information can be inferred by a semantic reasoner and the consistency of the information can be verified against the underlying ontology. Our hypothesis is that intelligence analysts will benefit from a semantic wiki through:

- Improved search capabilities
- Information objects with dynamic content
- Enhanced context navigation
- Inference of new facts
- Consistency checking
- Shared semantic tagging

As a non trivial side-effect the adherence to a standardized ontology will give immediate interoperability advantages when connecting to other systems [2].

Results from semantic queries can be visualized as different kinds of dynamic information objects, e.g. tables, timelines or social networks. These are easily integrated in wiki articles, which provide rich possibilities to add context in form of text and hyperlinks, plain as well as semantic. In such a system many useful possibilities emerge for analysts managing intelligence. Decision makers will start performing their situation and impact assessment by simply surfing the relevant semantic wiki articles.

II. The Semantic MilWiki

In 2005, the Swedish Defence Research Agency decided to develop a prototype of a military command and control wiki knowledge base, called MilWiki KB (the Military Wiki Knowledge Base) [3], built on MediaWiki, the open source wiki platform used by Wikipedia. The purpose of the MilWiki KB was to study how a wiki system could be used for knowledge management in everyday use across hierarchies and boundaries in a military organization. The prototype system was used in Demo 06 Spring, an experimental exercise

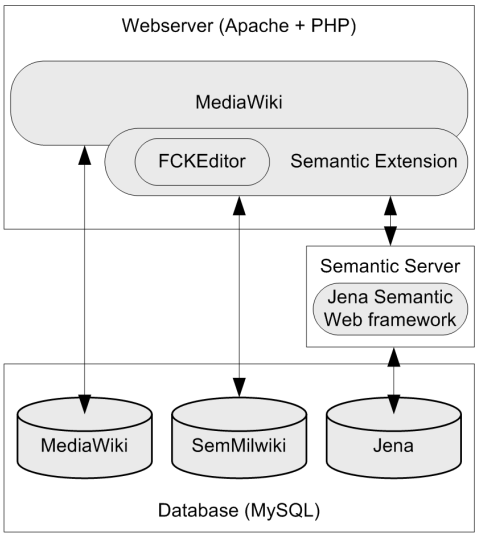

Fig. 1. Overview of the the Semantic MilWiki arcitechture.

organized by the Concept Development Centre of the Swedish Armed Forces.

In 2007, the MilWiki KB became the Semantic MilWiki as it was extended through a plug-in developed at FOI, connecting it to the open source semantic repository and reasoner JENA. In addition, the open-source WYSIWYG-editor FCKeditor was added and extended with the capability to create semantic content. The content of a wiki article can be classified as a certain type (class), e.g., a Person, Organization or an Event. Then the user can add semantic statements that specify attribute values or relationships to other instances. The set of available attributes and relationships are filtered so that only the ones relevant to the chosen type are shown. The filtering speeds up the tagging process and helps keep information consistent.

The goal of the Semantic MilWiki concept is that it should be easy to use, even for a non technical user. From this perspective, creating a semantic query offers a real challenge. The query must match the structure of the ontology and be formulated in a formal language, in our case SPARQL. The SPARQL syntax is not optimal for fast manipulation by a human user and hence we have developed a SPARQL-wizard to alleviate query formulation. In the wizard, the user can browse the ontology hierarchies to choose the desired types and relationships. The same filtering technique is utilized as in the tagging process, which helps ensure query validity. This reduces complexity and saves time for the user. Figure 2 shows an example of the SPARQL-wizard GUI. This particular query will generate a list of all political persons that have been involved in an event of type Attack and also are supportive of the African Union (figure 3).

Once a query is formulated, the result is presented as a list of entities fulfilling the specifications in the query. The query can then be inserted in a wiki-article producing dynamic content at certain time intervals when the query result is automatically

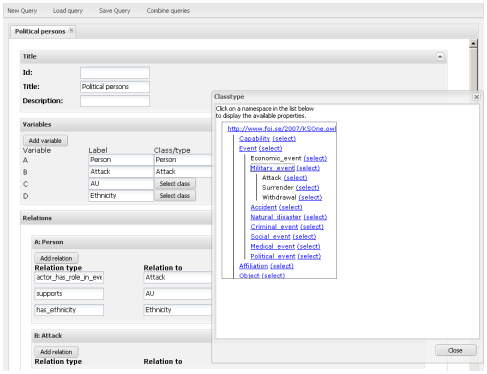

Fig. 2. The screenshot shows the SPARQL-Wizard. The type and relationships of the variables are chosen from pop-up hierarchies.

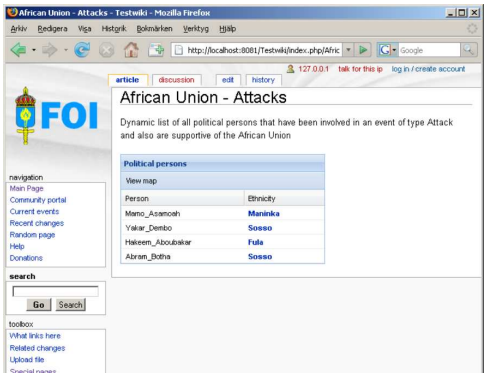

Fig. 3. The screenshot shows the result of the query in figure 2, inserted as a dynamic list in a wikipage.

updated. If new content matching the query has been entered somewhere in the wiki, it will show up on the dynamic page. Such dynamic queries can thus be used to get an immediate update on what is being entered into the knowledge base regarding items of interest.

For example, a user who is interested in all the clan leaders of an area can have a dynamic query that asks for information objects stored in the wiki and classified as a leader of a clan. When information is updated on one of these pages, it is then immediately visible on the dynamic query page, giving the user a situation picture of what is happening in the area.

III. Applications

We have identified two concepts under development at the Joint Concepts, Development and Experimentation Centre of the Swedish Armed Forces where semantic technologies and the Semantic MilWiki potentially can make contributions. In the following subsections the two concepts are briefly described and the motives for introducing a semantic wiki are explained.

A. Knowledge Support

Knowledge Support (KS) is a method and technology concept aimed at improving knowledge management of military operational intelligence. One approach for producing new

knowledge in KS is systems analysis. In systems analysis an area of interest is decomposed into entities and relationships. The possible interactions of the entities are then analyzed in order to predict the system's reaction to various activities. This analysis can yield new insights of the system as a whole and serve as decision support for the commander.

Semantically structured information can improve the information retrieval of basic facts supporting the construction of the system models. Semantic queries can generate precise lists of entities fulfilling the criteria for being sufficiently influential to be taken into account. If there is need to alter the number of entities feeding the analysis, the ontology structure can help to narrow or widen the query scope. The dynamic query feature of the Semantic MilWiki makes it possible to automatically detect changes in the basic facts influencing the system model.

In order to get an initial understanding of the usability of the Semantic MilWiki we participated in a workshop conducted by the KS team in November 2007. The Semantic MilWiki was used stand-alone with the purpose of answering questions that arose during the systems analysis and support the generated system model with facts, such as a list of non-official key persons of the country including interesting attributes such as their ethnicity, religion and organizational belonging.

B. Non-hierarchical Information Management

The Non-hierarchical Information Management concept (NHIM) aims at enhancing the capabilities of decision-making through methods and technologies exploiting non-hierarchical information paths in the military organization. The hypothesis is that the quality and speed of low-level decision making will improve if information is allowed to flow directly between peer nodes, without first being interpreted and filtered by higher command levels. One risk with such an approach is that low-level commanders will be overloaded with information irrelevant to their task. To avoid this, semantic querying can be used to specify the exact information needs of each individual commander. It is important to note that the non-hierarchical information flow is not meant to replace the standard flow, instead it will complement it. It will always be necessary to feed information from low levels to intelligence processing cells at high levels who will produce official intelligence reports and recognized situation pictures.

In May 2008 an explorative experiment was conducted by us in collaboration with the NHIM team. Eight commanders from the Swedish Armed Forces participated in exercises and discussions to investigate how to describe and formalize information needs given a certain role, task and mission environment. In the exercises, a game was played where the participants acted as platoon commanders performing an escort task in a fictitious conflict area. The Semantic MilWiki was used as technical platform. The information needs expressed by the participants were entered as subscription queries by technical personnel using the SPARQL Wizard. The results were then presented on wiki-pages as dynamic lists and on special pages with a map display.

IV. Lessons learned

Both the KS workshop and the NHIM experiment indicated that the types of queries possible in a semantic wiki are of potential use for the type of tasks presented. However, there are a number of issues that have to be solved before introducing a semantic system in the operational environment.

In KS, the sources of information available to the analysts during the systems analysis exercise consisted of a number of documents describing the region of interest and a Subject Matter Expert (SME). Some of the content had been manually tagged by us in advance, but there was not time for a complete mark up. Hence, many queries could not be fully answered. This leads to the important but hardly surprising observation that the critical issue regarding use of semantic technology in KS is the tagging. Intelligent computer support is needed to help users do this.

A lesson learned from NHIM was that expressing information needs as subscription queries is time consuming. Therefore it is important to be able to reuse subscriptions as templates when facing similar needs. The templates can be tailored to new needs by tuning a number of parameters. In a time critical situation the default values of the template can be used to speed up the process. Parameters can also be imported from other applications, such as planning and Geographical Information System (GIS) tools.

In NHIM reports were tagged with involved actors, observed materiel (consumables and/or equipment type), the type of event that had occurred and states that this event might affect. Actors, materiel and event types are *objective* features of the situation that the report describes. Affected states on the other hand are based on the *subjective* conclusions drawn by the reporter.

There are obvious pitfalls when using subjective tags to mark and search information. There is a risk that the conclusions drawn are wrong and that not everybody agrees that the reported event really affects the states that the reporter claims. However, an experience from NHIM was that the states were of interest to the subscribers and intuitive to use when formalizing the information needs to subscription queries. Examples of popular states were ”Road condition” and ”Threat level”.

V. Future work

The initial testing of the Semantic MilWiki has revealed a number of interesting issues for further studying. As mentioned above, the semantic tagging would benefit greatly from automatic tool support. There are a number of academic as well as commercial systems for information extraction that could be integrated into the concept and used for this. With such a component the amount of semantic information could grow quickly, which emphasizes the need for the system to be scalable. In particular, we would need to look at alternative architectures for semantic storing and querying.

The development of the Semantic MilWiki is part of an FOI project dealing with situation and impact assessment [4]. The Semantic MilWiki will be integrated with the other decision

support tools developed by that project and used to store all the information that is processed by these tools as well as the results of the automatic and semi-automatic analysis performed by them.

REFERENCES

[1] B. Leuf and W. Cunningham, *The Wiki way: quick collaboration on the Web*. Addison-Wesley Longman Publishing Co., Inc. Boston, MA, USA (2001).

[2] M. Eklöf and C. Mårtenson, *Ontological Interoperablilty*. Technical Report, FOI-R1943SE, FOI, Sweden (2006)

[3] M. Brännström and C. Mårtenson, *Enhancing situational awareness by expoiting wiki technology*. CIMI (2006)

[4] P. Svenson et al., *Using the impact matrix for predictive situational awareness*. Proceedings of the 10th Int. Conf. on Information Fusion (2007)