Proceedings of the Workshop on Semantic Web User Interaction at CHI 2008, Florence, Italy, April 5th, 2008.

User-Centered Design and Evaluation of Interface Enhancements to the Semantic MediaWiki

Frederik Pfisterer and *Markus Nitsche*
ontoprise GmbH
Karlsruhe, Germany

Anthony Jameson
DFKI, German Research Center for Artificial Intelligence
Saarbrücken, Germany

Catalin Barbu
International University in Germany
Bruchsal, Germany *

Abstract

Semantic wikis are a lightweight form of semantic technology that can enable more effective and active use of systems developed within the already popular wiki paradigm. But the goal of making the use and authoring of semantic wikis easy and attractive for a broad range of users raises several interaction design challenges: How can users unfamiliar with semantic technologies be helped and motivated to specify the many necessary typed links in a way that results in a useful web of annotations? And how can visitors of the resulting pages be helped to make maximal use of the added semantic information? With this case study, we discuss work conducted since the beginning of 2007 on the user-centered design and testing of interface enhancements for the Semantic MediaWiki developed at AIFB, focusing on their application to scientific pages adopted from Wikipedia. We discuss, with reference to specific interface examples, how usability challenges were addressed in the initial conceptualization of the interface enhancements, how user requirements were anticipated, and how the interface enhancements were iteratively tested and refined. We then describe an intermediate evaluation involving 42 participants and sketch further work currently in progress. We conclude with comments on some general lessons learned from this work concerning the application of user-centered design in the semantic web area.

1 Introduction

Semantic wikis are a lightweight form of semantic technology that holds the promise of being able to enhance the already popular wiki paradigm by supporting the addition and exploitation of additional semantic metadata. [4] describes a leading deployment of the semantic wiki paradigm—the Semantic MediaWiki—, explains the potential added value of semantics, and summarizes experience to date with semantic wikis. The ultimate goal of these authors is the creation of a Semantic Wikipedia that blends the advantages of semantic wikis into the already extremely successful normal Wikipedia.

Along with technical challenges such as scalability issues, a major challenge for those interested in promoting semantic wikis is that of usability: As the examples in the next section will show, a semantic wiki by definition includes numerous annotations that are both syntactically and conceptually more complex than the hyperlinks found in a normal wiki. And some of the ways in which these annotations can be exploited require more complex operations on the part of users than those required by normal wikis.

One of the goals of the work described in this case study is to loosen the usability bottleneck of semantic wikis so as to bring closer to reality the vision of a Semantic Wikipedia—at least for certain domains, such as those of the traditional science subjects of biology, chemistry, physics.[1]

For concreteness, we begin this case study with a scenario that illustrates some of the potential added value of semantic wikis while at the same time presenting some of the new user interfaces to the Semantic MediaWiki that we have developed.[2] The subsequent sections will describe the user-centered methods that we applied while designing and iteratively evaluating these interfaces, with the aim of bringing to light not only generally useful methods for the user-centered design of semantically based systems but also some problems that seem more or less typical in connection with this type of system.

2 Scenario

Although most of the interface design challenges raised by semantic wikis concern the process of introducting semantic annotations, our scenario will begin with a focus on the process of exploiting the results of such annotations. After all, it must be possible to do important things with a semantically annotated wiki that cannot be done (at least not equally well) with a normal wiki, if the extra overhead associated with a semantic wiki is to be justified.

Consider a high school student who is preparing a presenta-

*The research described here is being conducted in the context of the larger, multistage project HALO 2, which has been funded since 2004 by Vulcan, Inc. Many colleagues at ontoprise GmbH and AIFB contributed to the development of the interfaces described here.

[1] The larger project HALO 2 in which this work is being conducted has other goals as well, one of which will be mentioned at the end of the paper.

[2] The complete set of interfaces, along with the associated backend code, is available under an open-source licence from http://ontoworld.org/wiki/Halo_Extension.

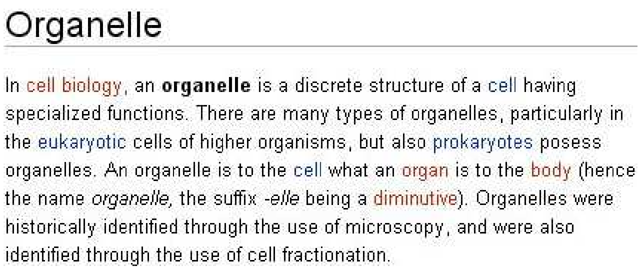

Figure 1. The first paragraph of a copy of Wikipedia's article on organelles in the Semantic MediaWiki.

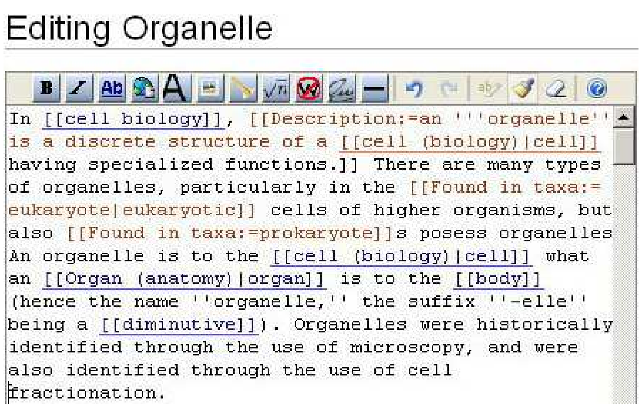

Figure 2. The source text of the same paragraph, showing the typed links (color-coded in red).

tion for his biology class on the topic of "organelles". Suppose further that, like many students, he consults the English-language Wikipedia. As the beginning of the article on organelles shown in Figure 1 shows, the article contains many relevant facts, but what is the student supposed to do with them? If he wants to collect information about various types of organelle and their properties, he will have to create a new document of his own and start typing or copying and pasting text on the basis of the Wikipedia text.

Consider now what is possible if this page has been transformed into a semantic wiki through the addition of *typed links*, each of which captures a single fact about organelles. (As can be seen in the source text of the beginning of the organelles article, shown in Figure 2, these facts are represented with annotations that are somewhat more complex than the usual hyperlinks found in a wiki pages.)

One resulting enhancement that is a standard part of Semantic MediaWiki is a *fact box* at the end of the page (see Figure 3) that summarizes all of the facts in the article that have been captured by typed links. In addition to offering a concise overview of the facts in the article, the fact box makes it easier for the user to navigate systematically to related pages, which are ordered in the box according to the nature of the relationship.

A more powerful way of exploiting the additional semantics is by querying the Semantic Wikipedia as if it were not only a web of hypertext but also an intertwined database. Con-

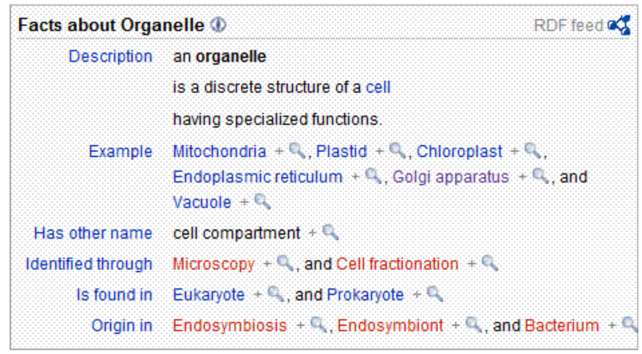

Figure 3. The fact box at the bottom of the article on organelles.

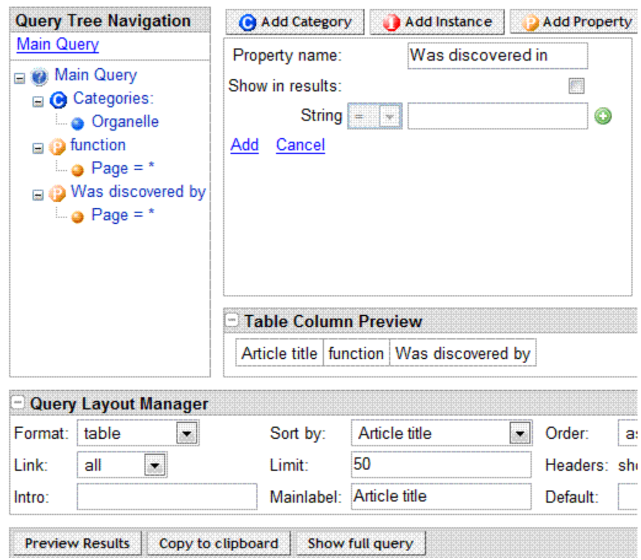

Figure 4. The Query Interface of the enhanced Semantic MediaWiki just as the user is specifying that the property "Was discovered in" should be included in the result table.

cretely, suppose that our student would like to start his research by creating a table that lists organelles along with their functions, the persons who discovered them, and the year in which they were discovered. He can then invoke the Query Interface shown in Figure 4 to specify the relevant category ("organelle"), and the three relevant properties ("Function", "Was discovered by", and "Was discovered in"), as well as various parameters that will determine the exact form of the table. Since the student probably does not know in advance the name of each of these properties (for example, the third one might also be called "Year of discovery"), the system includes an *autocompletion* feature: As the student types characters into the field for the desired property, the system displays all property names that include the substring that he has typed (see Figure 5).

The student can quickly preview the table to check that it looks right, and when satisfied he can copy the code generated for the query into a wiki page that he is building up for

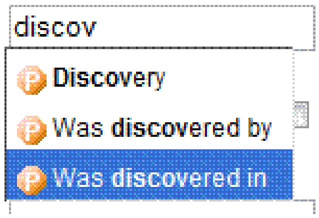

Figure 5. The autocompletion pop-up menu showing possible completions in the context of the previous screen shot.

Article title	Was discovered by	Was discovered in	Function
Peroxisome	Christian de Duve	1965	
Lysosome	Christian de Duve	1949	Digestion of macromolecules Digestion of bacteria Recycle organic material Digestion of viruses Digestion of disabled organelles Important for cell maintenance
Endoplasmic reticulum	Ernest F. Fullam Albert Claude Keith R. Porter	1945	Transport of proteins Folding proteins Protein biosynthesis
Golgi apparatus	Camillo Golgi	1898	Creation of lysosomes Transport of lipids Modifies proteins Disperse proteins and lipids with the cell Process and package macromolecules
Centrosome			
Chloroplast			Photosynthesis

Figure 6. A table created by the example user in our scenario via the query formulated in the previous screen shots.

himself (Figure 6). He can now click on the hyperlinks in this table to find further pages of interest, or he can consider various ways of expanding the table by adding additional fields or conditions. If at some point he thinks that his table could be of interest to other Semantic Wikipedia readers, he can insert it into the article on organelles so that others can access it without having to formulate a query themselves.

The Semantic MediaWiki also supports timelines and other visualizations that are appropriate for particular types of facts.

If we grant that benefits such as those mentioned so far make the upgrading of Wikipedia pages to Semantic Wikipedia pages seem like a worthwhile goal, the question arises of how all of the necessary annotations are going to be added. To some extent automatic methods are applicable (cf. [4]), but they cannot in general capture all of the potentially annotatable facts that are expressed in the text—let alone facts that are not present but that an author might want to add, for example after noticing that a given fact is apparently missing from the Wikipedia pages. Accordingly, our new interfaces offer support for manual annotation, although such annotation will presumably often be combined with automatic annotation.

To continue with our scenario, suppose that our student notices in the table in Figure 6 that there seem to exist no annotations concerning the discovery of the centrosome: By clicking on the link "Centrosome" in the table, he visits the article on centrosomes in order to fill the gap. In the original Semantic MediaWiki, he would have to edit the text shown in the editing window on the left hand side of Figure 7. The figure shows how the source text around "Theodor Boveri" needs to be annotated. The HALO tools for the Semantic MediaWiki offer, in addition to some color coding for hyperlinks and typed links, a *Semantic Toolbar* that makes it possible to add or change an annotation without directly editing the source text.[3] In the example shown in the figure, the user has just selected the year "1888" in the text and clicked on a link in the "Properties" section on the right-hand panel, invoking the dialog box shown in figure, in which "1888" is already filled in. The user has also entered the name of the property (again via autocompletion). When he clicks on the link "Add", the appropriate annotation for the year of discovery will be added directly in the source text.[4]

Finally, suppose that the student would now like to find some additional properties that he could add to his table. Since he doesn't have a particular property in mind, autocompletion would be of no help; so the student visits the *Ontology Browser* (Figure 8) to get an overview of the various properties that organelles can have. He types the word "Organelle" into the filtering field and clicks on "Organelle" in the Category Tree, thereby narrowing the information shown in the browser to instances of the category "Organelle" and the properties of these instances. By clicking on "Nucleolus" in the middle section, he can see all of the properties that have been annotated for that type of organelle, and he notices several properties that he might add to his table. If he would like to add the property "Surrounded by" but is not sure whether

[3] After the work described in the main part of this paper had been completed, a more advanced Annotation Mode was introduced—see Figure 10 and Section 5.1.

[4] The user can also use the Semantic Toolbar to add annotations that are not associated with any particular part of the Wikipedia text—for example, facts that seem worth having available for the purpose of answering queries but that are not interesting enough to be worth mentioning in the text. When an annotated fact is mentioned in the text, however, there are advantages to associating the annotation with the text: 1. Any change or correction made later will affect both the text and the underlying formal representation. 2. It is easier for users to find the point in the text where a particular fact is mentioned, so that they can read about closely related information. More generally speaking, the value of the type of annotation illustrated by our scenario appears to be greatest when the reader is interested not only in text or only in raw data but rather in an interweaved combination of text and data.

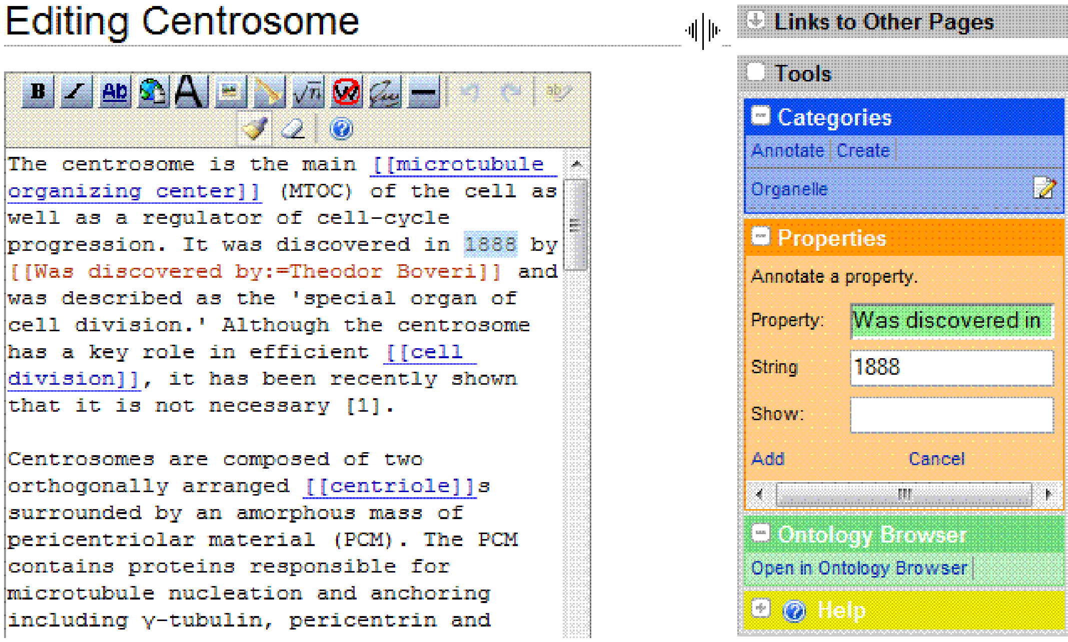

Figure 7. The user annotates the article on centrosomes with the help of the Semantic Toolbar (on the right).

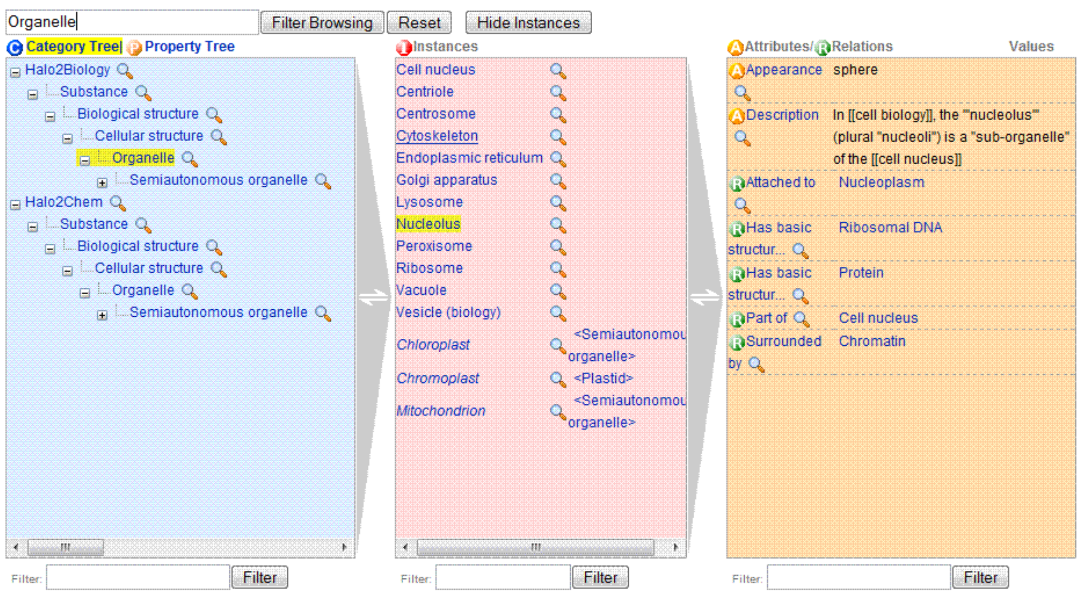

Figure 8. The user looks for properties of organelles with the help of the Ontology Browser (explanation in text).

it is applicable to a large enough proportion of the organelles to be worthwhile, he can click on the link "Surrounded by" and then on the big gray arrow between the second and third columns to move the "flow" in the opposite direction: The system will now show in the middle column all instances of "Organelle" that have a property "Surrounded by".

Although the functionality and interfaces presented in this scenario are only a fraction of those developed in our work, the scenario can serve as a background for a discussion of our experience in designing and testing these interfaces.

3 Requirements Gathering and Iterative Design and Testing

For reasons that will be explained below, there was no sharp division between the analysis of users' requirements, the design of an initial set of user interfaces, and the iterative testing and refinement of these interfaces.

3.1 Considerations Concerning the Target User Group

One overall goal of this work is to help realize the vision of the Semantic Wikipedia—in particular, for scientific Wikipedia pages.

In contrast to a common situation in user-centered design, there is no predefined target user group, such as the members of a given company or profession, for whom the new system must meet particular requirements. Instead, there must be *some* set of users who together can and will create a useful scientific Semantic Wikipedia; and there must be some (presumably overlapping) set of users that can benefit from the scientific Semantic Wikipedia. Whether, for example, the former group comprises high school students, retired persons, or (more likely) the union of a number of types of users, each with their own usage patterns, is not very important in judging the success of the enterprise. Accordingly, instead of picking out one or more user groups, analyzing their specific characteristics and requirements, defining personas, etc., we focused on meeting requirements that appear to apply to a broad range of potential users with certain general characteristics: an interest in helping to build up and/or in exploiting a common knowledge source such as a scientific Semantic Wikipedia, coupled with a lack of familiarity with semantic technologies.

In view of experience with the normal Wikipedia (see, e.g., [4]), we never expected to be able to make all Semantic Wikipedia users roughly equally capable of performing such tasks as annotation and querying. Instead, it is more realistic to expect that many users will do a lot of writing and only minimal annotation, while a relatively small number of users will handle a large proportion of the work of adding metadata and ensuring its consistency and correctness.

It soon turned out to be inevitable to assume that there would be a class of users, called *gardeners*, who would take on the most sophisticated tasks, such as merging two independently introduced relations that evidently have the same intended meaning. This gardener role has some similarities to the administrator role in Wikipedia: Wikipedia administrators likewise have access to technical features that help with the maintenance of articles. But our specialized gardening tools presuppose that gardeners have at least a basic understanding of knowledge bases and semantic technologies, whereas the role of a Wikipedia administrator can in principle be taken by anyone who has been an active contributor.

3.2 Analysis of Use Cases

One important method for formulating both usability requirements and more technical requirements was the collection and analysis of use cases. Some use cases were those that were already covered by the existing Semantic MediaWiki implementation. A number of new use cases were contributed by team members from AIFB on the basis of their deep understanding of wikis and their experience in developing and testing the original Semantic MediaWiki. In addition to the central, frequently occurring use cases for the end user, such as the ones illustrated in Section 2, this set included some less obvious use cases such as “Find pages that have not yet been annotated”, as well as use cases for system administrators and gardeners. For the usability members of the team, this set of 60 use cases served as a reminder of all of the functions that might have to be served by the new user interfaces, though it was accepted in advance that not all use cases in this set would actually ultimately be supported.

3.3 Analysis of Existing Systems, Requirements, and Guidelines

As can be seen in Section 2, many aspects of the envisioned enhancements of the Semantic MediaWiki could already be found in some form in other existing systems based on semantic technologies.

The most obvious design goal was that of minimizing the extent to which users should have to write or edit the source text of typed links and especially queries, which is a time-consuming and error prone activity for users who are not accustomed to this type of notation.

A second general principle was that users cannot be expected to be familiar with the ontology underlying the semantic Wikipedia pages; the autocompletion feature can be seen as a lightweight way of making users aware of existing categories and properties. In situations in which users want to see a more explicit representation of the ontology, the Ontology Browser was intended to serve this purpose in a relatively intuitive way.

As regards the Query Interface more specifically, there exist a large number of interfaces that allow users to formulate queries using elements of an ontology as building blocks. Some general recommendations on the basis of experience with such systems, together with more general user interface design principles, have been formulated in [3] and (more recently, after the end of our design phase) in [8]. For example, both of these works emphasize the importance of enabling users to formulate queries quickly and to refine them easily on the basis of the results obtained—thereby taking into account the fact that users do not in general have a good mental model of the system’s ontology and therefore cannot in general plan carefully in advance in order to formulate an optimal query. The Query Interface shown in Figure 4 supports iterative and exploratory querying by offering quick previews of the results of a query, by making it easy to navigate to the pages that appear in the query results so as to better understand the details of the results, and by making it possible to modify the query until the results appear satisfactory.

With regard to the annotation facilities, one general requirement (studied in special depth in [5] and [6]) is that users who take the trouble to add metadata should receive some more or less immediate reward for their efforts. The Semantic Toolbar takes a step toward meeting this requirement by quickly reflecting each new annotation in the list of annotations contained directly in the Semantic Toolbar (not visible in Figure 7 because of being obscured by the active dialog box)—as well as in the Ontology Browser. Probably more important in this respect is the Query Interface, which allows the user to make immediate use of any new annotations to improve the results of queries that the user wants to exploit for himself, to add a table or other visualization to a public Semantic MediaWiki page, or perhaps both.

3.4 Participants

In keeping with the goal of developing a scientific Semantic Wikipedia, we recruited 7 subject matter experts (SMEs), aged between 23 and 29, each of whom studied in an area related to physics (2), chemistry (3) or biology (2). All of these participants, of whom 3 where female, had been working with computers for 5 to 10 years; currently working with them an average of 15 hours per week. All of them used wikis (including Wikipedia) on a regular basis (more than twice a week), but none of them had ever written or edited a wiki article before participating in our studies. None of the participants knew much about the semantic web; only 2 of them had even heard the term.

The participants took part in the design and development phase of SMW for over 7 months, the overall goal being to add semantic annotations to scientific Wikipedia pages that had been copied from the English-language Wikipedia. Each participant did this work for several weeks. At first, they often visited the offices of the development team; as time went on, more of their work was done at home, and their tasks were less specifically defined.

The participants regularly filled in questionnaires, most of the questions being multiple-choice but allowing an opportunity for the participant to write comments. During the early phase of testing, they were often observed and interviewed by members of the usability team; as time went on, their performance was assessed more in terms of annotations that they made and the queries that they formulated.

The reasons for having a small number of participants work over a period of months—as opposed to studying a number of different groups of participants–were as follows:

1. It would have been impractical repeatedly to recruit new participants and give them the necessary background information and instruction; by reusing the same participants we were able to elicit a great deal of information with a reasonable amount of effort.

2. In the real-life use of a Semantic Wikipedia, we expect many users to work with the system over an extended period of time. Feedback from experienced users is therefore of interest, as are larger patterns involving more than one user that arise when a group of users works with the system over a long period of time.

This approach does, course, have some drawbacks, which will be discussed below.

3.5 Examples of Changes Based on User Testing

As would be expected, the regular feedback from these users produced a continuous stream of ideas for improving the design of the interfaces. Two examples can serve as illustrations:

Improvements to the Autocompletion Facility

There are many ways of making more sophisticated an autocompletion facility such as the one shown in Figure 5. Since each such improvement typically brings some performance and/or implementation cost, the selection of the improvements to realize should be driven by user feedback. One improvement that the SMEs insisted on was the display, for each element in an autocompletion list, of the *type* of the corresponding ontology element (e.g., category or property)—which is important in contexts in which elements of different types can appear within a single list. Less obviously, they suggested that autocompletions beginning with the same substring as the one typed by the user should appear at the beginning of the list.

Improvements to the Ontology Browser

The Ontology Browser represents a special interface design challenge in that it not only aims to make accessible to users something that they are not familiar with outside of the Semantic MediaWiki—an ontology—but also aims to allow them use it in a flexible and efficient way. Our SMEs found the Ontology Browser basically appealing right from the start, but not surprisingly they initially had difficulty understanding some aspects of it. Over a period of four months, they made many suggestions about the details of navigation and layout within the Ontology Browser which together led to a useful and reasonably usable (though still relatively challenging) interface.

Evolution of System Usability Scale Scores

In accordance with the goal of continually increasing the overall usability of the Semantic MediaWiki over time, we administered the System Usability Scale (SUS, [1]), starting just two weeks after the SMEs began their work, to obtain a baseline for comparison with later results. The 10 questions of SUS yield an overall usability score on a scale from 0 to 100, where scores between about 60 and 70 have come to be regarded as reflecting normal, acceptable usability.

The initial administration, which concerned essentially the unenhanced Semantic MediaWiki, yielded an average SUS score of 42, with the scores spread over a broad range from 15 to 67.5.

12 weeks later, the SUS scale was administered again, to the set of SMEs that were working that time (including 5 of the same SMEs as with the previous administration plus 3 replacements for the 2 who had since left the project). This time, the average score was 73, and there was no longer large variation among the scores.

At face value, this jump from a clearly unsatisfactory average score to a good one would seem to confirm the success of the overall effort to enhance the usability of the Semantic MediaWiki. But some caveats need to be borne in mind, aside from the obvious point that the two sets of SMEs only partly overlapped:

The SMEs for the second administration of the scale had by and large much more experience in working with the Semantic MediaWiki than the original SMEs had had. Some of the statements in the SUS scale would be expected to be answered more positively by more experienced users, even if the system itself were the same (e.g., “I felt very confident using the semantic functions of the wiki”).

Even with regard to feedback other than the SUS results, we suspected that our regular SMEs had become quite different from persons who would use the enhanced Semantic MediaWiki for the first time. Therefore, we decided to test the core aspects of the system in September 2007 with a group of users who not only had no relevant experience with semantic technologies but who also had less motivation then our regular SMEs to respond favorably (or even patiently), to the enhanced Semantic MediaWiki: A group of students in an introductory course in human-computer interaction who participated as part of a required laboratory session.

4 The Intermediate Evaluation

4.1 Goals

The purposes of this study were to

1. obtain a global overall assessment of the usability of the enhanced Semantic MediaWiki for users with no relevant specialized knowledge or experience and with no special reason to be interested in using the system;
2. obtain qualitative feedback on the most important areas for improvement seen by this group of users;
3. Perform a failure analysis of the problems experienced by these users, which could be especially informative because of the relatively large number of participants.

4.2 Method

Participants

A total of 42 students in two sections of an introductory course in human-computer interaction participated during their first lab session (i.e., at a point in time at which they had virtually no knowledge of human-computer interaction). Their sole incentive, aside from the fact that their participation was a required part of the course, was the promise that they would learn a lot and that their feedback would be taken seriously in the further development of the system.

Materials

A large number of Wikipedia pages concerning well-known scientists was made available within a test environment of the Semantic MediaWiki.

Design

Since the total available time of 90 minutes was not long enough to allow a meaningful comparison of two different systems, the participants worked only with the then-current version of the Semantic MediaWiki interface enhancements. (The screenshots in Section 2 are based on a more recent version, which benefited in part from feedback from this study.)

Procedure

After a brief introduction to the Semantic MediaWiki and some instruction and practice concerning annotation of Wikipedia pages with the Semantic Toolbar, each subject was randomly assigned one of the pages about a scientist and asked to annotate, within a fixed period of 20 minutes, as many as possible of the answers to a number of questions. These questions were formulated in a way that did not reveal the exact content of the annotation(s) to be made; for example:

1. When was the scientist born?
2. What was his field?
3. What did he discover?

In the second half of the study, participants were given some instruction and training concerning the formulation of queries with the Query Interface (cf. Figure 4). They were then given a number of queries of which they were to formulate as many as possible within 20 minutes, storing the results in a separate page of their own (as in the scenario described in Section 2). Again, the formulations of these tasks did not reveal which attributes or relations should be used in the formulation of the query.

After finishing this second phase, the participants filled out the SUS scale, having been instructed to answer the questions with regard to the combination of the annotation tools and the Query Interface.

Dependent Variables

In addition to the responses to the SUS scale, the participants answered questions about what they liked most and least about each of the two parts of the system with which they had worked: the Query Interface and the Semantic Toolbar.

The objective data that resulted from the study were the annotations and the queries that the participants had created.

4.3 Results

SUS Scores

The average score on the 100-point SUS scale was 54.8 (with a standard deviation of 17.2).

Participants' Comments

In answering the open questions about annotation with the Semantic Toolbar, the autocompletion feature was at once the most liked and the most disliked feature: The participants recognized its importance in helping them to identify the relevant property or attribute name in each case; but they noted that even with autocompletion it was not always possible to find the relevant term (i.e., when the term does not share a substring with the string typed by the user). The participants thus showed an awareness of one of the most pervasive problems with this type of annotation; note that, if they had had time to learn to use the Ontology Browser, they would have had additional means of finding the relevant term to use. More important for the evaluation in terms of the SUS scores were the many complaints about the slow response time of the autocompletion feature, which was apparently due to the fact that it was being used by more than 20 persons at the same time, an unusual situation for this test version of the Semantic MediaWiki. This particular problem is not theoretically interesting, but it does illustrate the need to look for the reasons underlying global scores such as those for SUS and to pay attention to scalability issues when designing evaluations.

In answering the open questions about the Query Interface, a number of participants pointed out that it enables fast formulation of queries once you have figured it out but that a number of aspects of the interface are not self-explanatory on a first encounter. In response, some enhancements to details of the interface have since been made. For example, the main menu bar was moved from the very top of the interface to the very bottom (as is shown in Figure 4), in accordance with the standard position of interface components in MediaWiki. Immediately accessible online help has also been added.

We are still working on ways to make the interface more self-explanatory, so that users need to consult the on-line help as little as possible. Further planned improvements include the elimination of the explicit distinction in the interface between categories, instances, and properties. Additionally, a stylized English representation of each (partial) query will be shown even while it is being constructed, so that the user can more easily check whether the meaning of the query corresponds with his or her intention.

Analysis of Annotations

One convenient aspect of a wiki from point of view of evaluation is that a lot of information about the users' behavior is captured in the form of annotations and stored queries, reducing the need for instrumentation of the system for logging purposes. An analysis of the annotations in our intermediate evaluation illustrates the kind of lesson that can be learned from this sort of study.

Each student made 4.2 annotations on the average in their first 20 minutes of annotating with the Semantic Toolbar, and only 50.3% of these annotations were fully correct. The low total number may be due to the slow response time and the unfamiliarity of the system, but the number of fully or partly incorrect annotations requires a closer look. Figure 9 shows a breakdown of the 73 incorrect annotations in terms of the nature of problem. The two categories on the right, "Unrecognized date format" and "Unrecognized characters", which account for 53% of all incorrect annotations, simply reflect the limited ability of the system to recognize what are basically correct annotations (e.g., specification of a date of birth in a date format that the system does not recognize, or the inclusion of extraneous brackets within an annotation). These phenomena are now being treated as the type of system bug that comes to light when a system is tested with novice users. More fundamentally problematic are the cases where the user either chose an incorrect property name or introduced a new, redundant property even though an applicable one already existed. Examination of the cases where these problems arise helps us to judge what improvements to the autocompletion feature and the Ontology Browser are

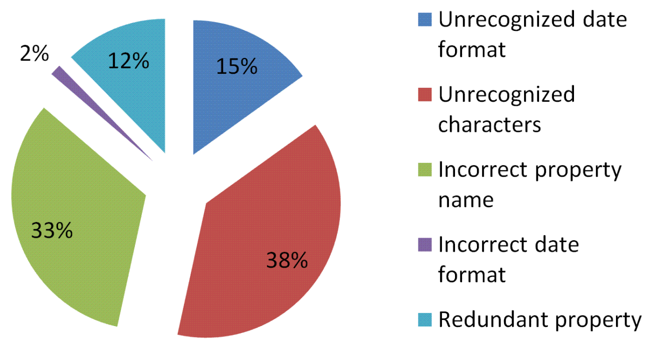

Figure 9. Categorization of the 73 (partly or fully) incorrect annotations in the intermediate evaluation in terms of the nature of the error.

most likely to lead to improvements.

4.4 Discussion

The overall quantitative results from this evaluation, such as the average SUS scores and the rates of success at annotation, are disappointing at first glance; but closer examination shows that they are due in large part not to genuine usability problems but to technical limitations of the current version of the system, such as the inability to serve more than 20 simultaneous users with short response times and the inability to recognize all of the formats that novice users employ when entering annotations. So in part this study illustrates the difficulty of conducting a quantitatively meaningful intermediate evaluation, in the style of a summative evaluation, in the middle of a long series of iterative testing and redesigning: It may not make sense overall in this context to invest a lot of resources in the removal of this type of technical bottleneck while the system is still being developed rapidly. An extensive final evaluation of the Semantic MediaWiki enhancements, lasting several weeks and preceded by a dry run and a quality assurance phase, is planned for the second half of 2008, and in this context it should be possible to obtain quantitatively more meaningful results.

At the same time, this study did yield a good deal of qualitative information, which was especially valuable because the participants had quite different properties from those of the SMEs with whom we had been working regularly—properties that may be more typical of the sort of user who tries out the Semantic MediaWiki in a realistic situation without special preparation or motivation.

5 Current Work Based on the Results of Previous Tests

The current phase of the project is being devoted not only to the further detailed improvements of the user interface enhancements but also to the development of new user interfaces on the basis of the experience of the first few months of the project. These current developments will be described briefly here just to indicate how they were inspired by our experience with users.

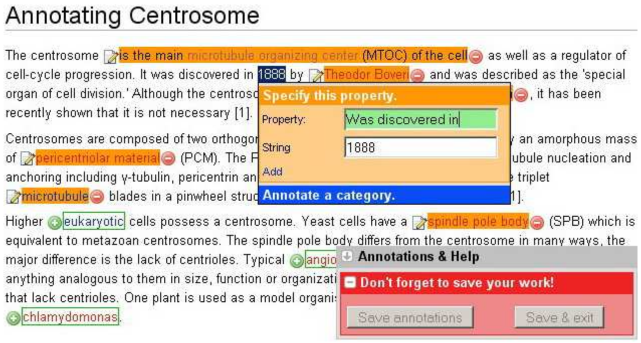

Figure 10. Example screen shot of the new Annotation Mode that lets users introduce new annotations directly in the rendered version of the article without needing to edit the source text of the page.

(Each part of the text highlighted in a solid orange color corresponds to an existing annotation, which the user can edit by clicking on the associated icon. A part highlighted with a box corresponds to an existing normal hyperlink, which the user may want to make into a semantic annotation by clicking on the "+" icon.)

5.1 A Separate Annotation Mode

Although the Semantic Toolbar (Figure 7) makes it unnecessary for a user to add typed links directly to wiki source text, it still does expose users to the source text, and especially beginning users have remarked that they find a source text rather daunting and confusing. Accordingly, a dedicated *Annotation Mode*, in which the user can add or change annotations without seeing the source text, was part of the design of the HALO tools from the start. But since the detailed design and especially the efficient implementation of such an Annotation Mode was known in advance to be challenging and time-consuming, we postponed the attempt until relatively late in the project, using the Semantic Toolbar (Figure 7) as a temporary surrogate that was just good enough to allow SMEs to work with the entire set of tools. In this way, it was possible to do a great deal of iterative user testing of other parts of the set of tools while in parallel testing low-fidelity prototypes of the new Annotation Mode, which were viewed by our SMEs as representing a considerable improvement. Figure 10 shows the recently integrated Annotation Mode, which is now part of the publicly available set of HALO tools.

5.2 Intelligent Task Routing

One complaint often heard from our users is that they would often like to do some useful annotation but do not know where to start. The scenario presented in Section 2 shows how a user can stumble upon some work that needs to be done (e.g., in the form of a gap in a table), but a different type of support is needed for users who simply feel like doing something useful.

One approach that has been pursued with Wikipedia is that of *intelligent task routing*: The system gives to each individual author hints about specific tasks that need to be done and which seem especially appropriate for that author, given their knowledge and interests. For example, if the user in our example scenario (Section 2) made a number of annotations like the one illustrated in Figure 7 and later requested recommendations of further tasks, the system might list several other cell biology pages that had few or no annotations.

A successful implementation of this idea in the context of Wikipedia was described in [2] under the name SUGGESTBOT.

Because of the additional semantic information included in a semantic wiki, this type of intelligent task routing appears to be at least as usable and useful within a semantic wiki.

5.3 Serving as a Knowledge Source for More Complex Modeling

The original motivation for the enhancement of the Semantic MediaWiki within the HALO 2 project was actually not the creation of a Semantic Wikipedia itself but rather the creation of a large knowledge base that would be suitable for use as a knowledge source by scientists engaged in more complex modeling activities than those typical of a Semantic Wikipedia. For example, the knowledge capture system AURA, developed as part of the HALO 2 project in a team led by SRI International, is used by experienced scientists to encode the knowledge contained in advanced high school science text books into a sufficiently rich form to allow the answering of questions from the American Advanced Placement exams. In some cases, it is possible for users of AURA to import from a suitable scientific semantic wiki considerable amounts of relevant knowledge, so as not to have to enter it all themselves (an obvious example being the information contained in the periodic table of the elements in chemistry). We are working on a mapping and merging tool that will enable users who have little familiarity with ontologies to find useful knowledge within a Semantic Wikipedia and import it into a rich modeling system. Our conception of how best to deal with this challenge— and even of the types of knowledge that it can be worthwhile to exchange in this way— has changed considerably on the basis of our experience with users of the Semantic MediaWiki enhancements.

6 Some Lessons Learned

We think we have learned a lot of small and large lessons about the design issues that arise with interfaces like the ones that we have been working on. But results like these are best viewed as data points that contribute to a larger picture, which can best be grasped when a number of similar systems are analyzed together (as is done, for example, in [3] and [7]).

More relevant for this case study description are lessons about the process of applying the methods of user-centered design in the context of semantic technologies. Even though the application of these methods has long been commonplace with many types of interactive systems, it appears that there is still a need within the semantic web community to provide concrete examples and to comment on the conditions for their successful application.

6.1 The Importance of High-Level Support for User-Centered Design

The company that has funded research on the HALO project has organized and staffed the project in a way that ensures that usability specialists are involved in all phases of the design and development process, working closely with both the project management and the development team, starting from the initial requirements specification and finishing with a large-scale evaluation. By contrast, in all too many projects in the semantic web area, the consideration of users is still seen as a burdensome obligation that needs to be fulfilled, for example, to satisfy reviewers—an obligation that threatens to impede technical progress and, worse yet, to reveal that the system that looks so brilliant in a carefully choreographed demonstration is flatly rejected by real users. It is a great help for the user advocates in a development team if they do not have to engage in a drawn-out tug-of-war with those who run the project but rather are continually supported by them. Those who are contemplating embarking on a project in this area under less favorable circumstances may want to consider whether they can improve the organizational conditions from the start.

6.2 The Importance of Collocation

Most of the members of the development team for the Semantic Mediawiki enhancements work within a single large room, constantly communicating synchronously face-to-face or asynchronously via the project's own semantic wiki, which is seen as a successful example of "eating one's own dog food". This extremely broadband communication ensures that user-related issues are not lost by the wayside or handled in an ineffective way, as so often happens when the usability-related aspects of a project are outsourced to some remote group—which may soon come to feel like a barking dog chasing after a car as it tries to keep up with the creative and energetic developers in the project.

On the other hand, we have found that including one usability consultant who is not collocated with the rest of the team (although there are frequent visits and electronic communication) has the advantage of occasionally interjecting new perspectives and calling into question implicit assumptions.

6.3 The Absence of a Clearly Defined Target User Group

As was discussed in Section 3.1, the goal of the work described here is not to help a particular predefined user group to perform particular tasks but rather to support the creation of a large common artifact—a scientific Semantic Wikipedia. Which people end up creating it and which ones end up benefiting from it is not of central importance as long as both of these things happen and the benefits are substantial. This type of situation can easily arise with semantic technologies, where the general goal is to create a web of data that can be valuable in various ways to various people. This situation raises the question of what to do about customary user-centered design methods such as carefully specifying user groups and their properties and needs and consistently keeping the requirements of these particular groups in mind during the design process. Should we try to guess what types of user are going to be using our system and design for these users? Or should we work backward from the tasks that we want be performed, providing support for these tasks that seems likely to be useful for a broad range of potential users?

Although we have largely taken the latter course in the work discussed here, we welcome discussion of this issue.

References

1. John Brooke. SUS - a quick and dirty usability scale. In Patrick W. Jordan, Bruce Thomas, Bernard A. Weerdmeester, and Ian L. McClelland, editors, *Usability Evaluation in Industry*, pages 189–194. Taylor & Francis, London, 1996.
2. Dan Cosley, Dan Frankowski, Loren Terveen, and John Riedl. SuggestBot: Using intelligent task routing to find work in Wikipedia. In Tessa Lau and Angel R. Puerta, editors, *IUI 2007: International Conference on Intelligent User Interfaces*, pages 32–41. ACM, New York, 2007.
3. Anthony Jameson. Usability and the semantic web. In York Sure and John B. Domingue, editors, *The Semantic Web: Research and Applications: Proceedings of the Third European Semantic Web Conference, ESWC 2006*, page 3. Springer, Berlin, 2006. Abstract of a keynote address. Slides available from http://dfki.de/∼jameson/abs/Jameson06ESWC.html.
4. Markus Krötzsch, Denny Vrandecic, Max Völkel, Heiko Haller, and Rudi Studer. Semantic Wikipedia. In *Proceedings of the 15th International Conference on the World Wide Web* , pages 585–594, 2006.
5. Luke McDowell. *Meaning for the Masses: Theory and Applications for Semantic Web and Semantic Email Systems*. PhD thesis, Department of Computer Science and Engineering, University of Washington, 2004.
6. Luke McDowell, Oren Etzioni, Steven D. Gribble, Alon Halevy, Henry Levy, William Pentney, Deepak Verma, and Stani Vlasseva. Mangrove: Enticing ordinary people onto the semantic web via instant gratification. In *Proceedings of ISWC 2003*, pages 754–770, Sanibel Island, Florida, 2003.
7. Victoria Uren, Philipp Cimiano, José Iriac, Siegfried Handschuh, Maria Vargas-Vera, Enrico Motta, and Fabio Ciravegna. Semantic annotation for knowledge management: Requirements and a survey of the state of the art. *Web Semantics*, 2005.
8. Victoria Uren, Yuangui Lei, Vanessa Lopez, Haiming Liu, Enrico Motta, and Marina Giordanino. The usability of semantic search tools: A review. *Knowledge Engineering Review*, 22(4), 2007.