The Sediment Metaphor: Serendipity for Browsing Ontologies

Katharina Siorpaes, Martin Hepp, and Andreas Klotz
Digital Enterprise Research Institute (DERI), University of Innsbruck, Austria
katharina.siorpaes@deri.at, mhepp@computer.org, andreas.klotz@deri.at

Abstract. Navigating ontologies can be a cumbersome task: existing approaches to ontology visualization have several shortcomings including that elements are neither sorted by relevance nor do they adapt to changing user requirements. We present a novel technique for user interfaces of semantic systems based on an adaptive extension of the tag cloud paradigm.

1 Motivation

The navigation within large vocabularies is a difficult task, and the state of the art of techniques for supporting users in accessing relevant conceptual elements is unsatisfying. Existing approaches, such as lists or trees, have two main limitations: (1) Elements are typically sorted according to relations between them (mainly the taxonomic relations) or in alphabetical order, but the ordering does not reflect the relevance of the elements. (2) Just like ontologies change due to conceptual dynamics in many domains of discourse, the relevance of single ontological elements within vocabularies changes. Tag clouds, also known as weighted lists, are interface models for visual support for the information retrieval of tags. They have quickly gained popularity as a lightweight way for visualizing terms. In this paper, we propose a novel approach to ontology visualization, based on the extension of tag cloud visualization by four novel aspects: (1) Reflection of relevance of ontology elements. (2) Adaptation to user search behavior. (3) Concepts with decreasing significance lose prominence while the process of fading is sufficiently slow for elements to remain in the general focus. (4) Forgotten elements can reappear. We call our novel approach "sediment metaphor", since it imitates the ideal of objects slowly sinking to the ground, unless kept on top by some action.

Related Work: Aumüller and Rahm [1] describe Caravela, where a tag cloud is used for browsing authors. Additionally to the font size they also use different shades depending on the citation index of the author. Fluit et al. [2] provide an overview of existing approaches which focus on graph visualization. The difference of these approaches to our approach is that the sediment metaphor approach makes use of tag clouds, reflects the importance of elements, and adapts to changing user needs.

2 The Sediment Metaphor

The sediment metaphor (Fig. 1) is a novel user interface technique for navigating ontologies making use of the tag cloud as a visualization paradigm. It can be used for both single- and multi-user scenarios and relies on three design principles: **serendipity**, **adaptability**, and **slow percolation**.

```
FOR ALL x SET relevancyOfX=1; //indicates relevance of element
SET alpha=0.8; //factor for decay
SET dateOfLastDecay; //date of last
IF element X is clicked THEN
relevancyOfX + 1; //increase relevancy
relevancyOfX * alpha; //decay
IF dateOfLastDecay > 24hours THEN
FOR ALL x relevancyOfX * alpha; //decay for all elements
FILTER elements to list l WHERE relevancyOfX = 1;
CHOOSE one element y from l; //choose random element
SET relevancyOfY=biggestValueInList/2; //increase its relevancy
GENERATE tag cloud;
```

Listing 1. Pseudocode

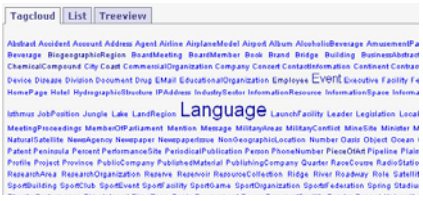

Fig. 1. Sediment Metaphor Tag Cloud

Listing 1 describes the sediment metaphor (Listing 1) using the Pseudo code Standard[1]. The sediment metaphor prototype is implemented in Java with a Sesame triple store [3] in the background within the myOntology[2] platform. For the background research on myOntology, see [4].

3 Evaluation

In the following, we describe our evaluation methodology, which tries to analyze whether retrieving ontology concepts in a big vocabulary is more efficient using (1) the sediment metaphor tag cloud as compared to (2) an alphabetically sorted list.

Methodology, Data, and Participants: We used the concepts of the Proton ontology [5], including the top, upper, and knowledge management module, for the experiment comprising 258 concepts. The concepts were available both in the sediment metaphor tag cloud as well as in an alphabetically sorted list. We chose five arbitrary concepts and reflected a decreasing distribution of importance by creating 100, 70, 49, 35, 21, and 14 cards of the concepts. This creates a set of 144 cards, which we brought into a random order. Five participants who have experience using computers and the Web in general were recruited. Each participant was asked to do a reaction time test[3] using the same mouse before carrying out the experiment, so that we could later evaluate whether differences in the time for accomplishing the task were not mainly caused by differences in cognitive or motoric abilities. Then, each participant had to go through the stack of cards in order to find and click each concept once in the tag cloud (1st round) and once in a list (2nd round). When looking for elements from the list, participants were allowed to use the find functionality provided by the browser Firefox[4]. Additional to taking the time, participants were interviewed about their experience after the experiment.

Results: The following table shows the results of our evaluation:

ID	Gender and education	Average Reaction Time	Time Sediment Metaphor	Time List	Experiences
A	Female, Master in computer science	0.22	09:59:02	18:14:04	Intuitive to use, very handy
B	Male, Bachelor in computer science	0.21	08:38:19	16:25:05	Much faster
C	Female, Student of medicine	0.23	10:49:01	20:13:27	First time to use a tag cloud, better than list
D	Female, PhD in history	0.21	07:56:34	16:57:07	Much more efficient
E	Male, Student	0.28	09:18:09	18:23:14	Good, even for very small elements in cloud
		Ø 0.23	**Ø 9:20:13**	**Ø 18:02:35**	

4 Discussion and Outlook

The results of the evaluation show that a visualization paradigm adapting to user or community behavior is a promising approach to ease the task of navigating large vocabularies. Users find it easy to use tag clouds. We plan to evaluate the sediment metaphor in the multi-user mode as well. Additionally, it is likely that we will use clustering techniques representing proximity of elements (e.g. concepts that are connected through relations).

Acknowledgments. The authors would like to thank Michael Waltl for his support in preparing the prototype. This work has been funded by the Austrian BMVIT/FFG under the FIT-IT project myOntology (grant no. 812515/9284).

References

1. Aumueller, D. and E. Rahm. Caravela: Semantic Content Management with Automatic Information Integration and Categorization. ESWC 2007. Innsbruck, Austria: Springer.
2. Fluit, C., M. Sabou, and F. Van Harmelen, Supporting User Tasks through Visualization of Lightweight Ontologies, in Handbook on Ontologies in Information Systems, S. Staab and R. Studer, Editors. 2003, Springer.
3. Broekstra, J., A. Kampman, and F. Van Harmelen, Sesame: A Generic Architecture for Storing and Querying rdf and rdf schema, International Semantic Web Conference ISWC 2002.
4. Siorpaes, K. and M. Hepp, myOntology: The Marriage of Collective Intelligence and Ontology Engineering, Workshop Bridging the Gap between Semantic Web and Web 2.0 at the ESWC 2007.
5. Terziev, I., A. Kiryakov, and D. Manov, BULO Guidance, in SEKT Deliverable 1.8.1. 2005.

[1] http://www.csc.calpoly.edu/~jdalbey/SWE/pdl_std.html
[2] http://myontology.org
[3] http://getyourwebsitehere.com/jswb/rttest01.html
[4] http://www.mozilla-europe.org/en/products/firefox/