Browsing Information Tags Space

Andrea ŠTEŇOVÁ*

Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova, 842 16 Bratislava, Slovakia
andrea.stenova@gmail.com

Software systems process information whose amount is growing exponentially thanks to the advent of the Internet. We have found out recently, that it is no longer enough for us to store the data, but we need something more. To be able to organize and simply and effectively browse stored information, we need additional information about these data. Therefore we use metadata (data about data) that can provide information about what data are, describe their structure, or give us additional information about the user who created them and how they were created. Semantic web was created thanks to fact, that web was enriched by metadata [1].

Adding metadata to our content has various advantages. Their importance lies in ability to store, organize and provide information about the data whom they are assigned to. This gives us simpler and faster searches in metadata. Metadata can also allow a better understanding of the data and display the data change over the time. They provide us information about the users, who read these data, worked with them or created them.

One type of metadata is information tags, which contain structured information associated with a particular piece of content, such as the number of clicks on the link on the page, keyword characterizing paragraph in the article or the number of lines of method's source code.

Metadata are usually generated and processed by machine, and their amount makes it difficult for people to effectively read and understand them. There are various problems with comprehensibility of data by users. When there is too much of data, we cannot display them at once, because they would be not easily legible. However, if we don't display all of the data, we need to ensure that the user is able to get the information he is looking for. When searching in such data, there might be a problem with lack of knowledge of a particular dictionary. The user does not have to know exactly what he is looking for and which sources are most relevant for him [2].

However, since the metadata can contain valuable information about the content, as well as the users working with it, it is important to enable the understanding and

* Supervisor: Karol Rástočný, Institute of Informatics and Software Engineering

Spring 2013 PeWe Workshop, April 5, 2013, pp. 41–42.

analysis of this information to the people. To ensure readability and understandability we can use navigation of user through the data with the help of their visualization.

Its importance depends on the importance of metadata it displays. Currently, it helps us to understand the huge amount of data and allows us to more easily navigate within them. Research methods are therefore dealing with how to handle metadata and how to navigate users through them.

To understand navigation we often use visualization of data. The most commonly used visualization methods use graphs or tables. Graphs can be two, three and multidimensional. When displaying a graph to a user, it is important to focus on readability and actuality of the information, or display him also change of information in time.

Navigation can be resolved by using graphs, namely using zoom techniques, displaying only the selected subgraph, clustering of graph nodes, or other techniques to clarify content without losing context [3]. Furthermore, we can use the tables to navigate users, faceted browsers, etc. Each of these techniques has its advantages and disadvantages, and its use is appropriate for solving various problems.

In our work we would like to propose method that would help users browsing information tags space. We want to focus on structured metadata about user activity, as well as the structure and content of data. Problem with metadata is their change in time. Only updated data, or their combination with its changes in time, can be presented to users.

Another problem is computational complexity of algorithms to display data. If we do not display to user all of the data and calculation of its representation is challenging at same the time, we need some way to predict user's actions to optimize the computational time.

Our solution must be understandable and easily readable by users. To support that, we will use some of the existing visualization solutions. We plan to verify our solution in domain of project PerConIK.

Acknowledgement. This contribution is the partial result of the Research & Development Operational Programme for the project Research of methods for acquisition, analysis and personalized conveying of information and knowledge, ITMS 26240220039, co-funded by the ERDF.

References

[1] Rástočný, K., Bieliková, M.: Maintenance of Human and Machine Metadata over the Web Content. In *Current Trends in Web Engineering*. 2012. p. 216–220.

[2] Katifori, A. et al.: Ontology visualization methods - a survey. In *ACM Computing Surveys*. 2007. Vol. 39, no. 4.

[3] Herman, I. et al: Graph visualization and navigation in information visualization: A survey. In *IEEE Transactions on Visualization and Computer Graphics*. 2000. Vol. 6, no. 1, p. 24–43.