

Creating Dynamic Wiki Pages with Section-Tagging

Denis Helic
IWM, TU-Graz
Inffeldgasse 16c
8010 Graz, Austria
dhelic@iicm.edu

Anwar Us Saeed
IWM, TU-Graz
Inffeldgasse 21A
8010 Graz, Austria
anwar.ussaeed@student.tugraz.at

Christoph Trattner
IICM, TU-Graz
Inffeldgasse 16c
8010 Graz, Austria
ctrattner@iicm.edu

ABSTRACT

Authoring and editing processes in wiki systems are often tedious. Sheer amount of information makes it difficult for authors to organize the related information in a way that is easily accessible and retrievable for future reference. Social bookmarking systems provide possibilities to tag and organize related resources that can be later retrieved by navigating in so-called tag clouds. Usually, tagging systems do not offer a possibility to tag sections of resources but only a resource as a whole. However, authors of new wiki pages are typically interested only in certain parts of other wiki pages that are related to their current editing process. This paper describes a new approach applied in a wiki-based online encyclopedia that allows authors to tag interesting wiki pages sections. The tags are then used to dynamically create new wiki pages out of tagged sections for further editing.

Categories and Subject Descriptors

I.7.1 [Document and Text Processing]: Document and Text Editing – *Document management, Version control.*

General Terms

Design, Experimentation

Keywords

Section Tagging, Wiki systems, Austria-Forum

1. INTRODUCTION

The popularity of social software has brought up new user generated content and metadata resources in the form of wikis, blogs, social tagging and bookmarking applications. These new systems have emerged as a major force driving the reshaping of information spaces on the World Wide Web to better serve both collaborative and personalized information needs of users. In social software applications Web has drifted towards users' content creation instead of the commercial content as a major contributing factor to Web resources.

For instance, wikis are used for sharing, management, and organization of knowledge. Wikipedia is a user-created encyclopedia and a well known example of a wiki system. Wiki systems are asynchronous, collaborative authoring and content versioning systems where any user can add and edit content. A new version of the page is stored in the system after each editing operation [2].

In wiki systems, user's content-creation/authoring processes involve laborious tasks like information selection from diverse

resources, restructuring, modification, and adaptation of information object according to the perceived context [6]. The reuse of existing content in the form of copy-paste mechanisms in order to restructure and create new documents is applied by authors frequently. For example, a typical editing workflow in wiki systems involves investigating volumes of information where in fact only small part of that information is relevant to the current user need. Thus, the user has to browse all the resources again and again to review the related pieces of information from their relevant or selected resources. This typically requires a lot of effort and time.

On the other hand resource organization with tagging and bookmarking services like Delicious, Citeulike or bibsonomy have received community focus due to ease of use and information discovery mechanisms. In social tagging and bookmarking applications users assign free form keywords and annotations to the addresses (URLs) of an information resource (e.g., a web page) [8]. These keywords relate the current user context to the content of a tagged resource. The weighted set of keywords (tags) assigned to a resource by all users within a system is called the tag cloud. Tag cloud is a visual representation of tag terms in which their font is scaled according to their frequency weights.

As [3] suggests the user motivation to tag a resource might be organizational or communicational on one hand, and on the other hand the users tag resources for their personal use and/or to share them with others. For example, users who tag resources for their personal use in an organizational sense use social tagging applications to organize interesting, important, and related resources according to their current needs. The tags are applied as a support for later search and retrieval of tagged resources via search or navigating the tag cloud. Typically, the tag cloud provides an overview of defined tags showing only the tags themselves but not the actual content of the tagged resources. The resources are represented via navigable links. Another motivation of using tags is to share them with other users and in such a scenario tags are typically used in a communicational sense to send signals to other users about resources that might be of interest in a more general case.

Regardless of users tagging scope- personal resource organization or sharing it with others- they have to tag the whole resource. This, however, does not always fulfill the users need. For example, users are often viewing content and are interested only in one part of the whole content. For future use users tag and bookmark it with a keyword that would be helpful later to retrieve the content. In this case users tag the whole content with a navigational keyword useless to represent the context of resource

but a useful one for them to reach the content section of their interest. This unrelated navigational tag in tag cloud will create noise. But users have no option to tag a particular interesting section within the whole resource. Such an option of tagging a part of resource may increase the user efficiency for later content retrieving, as well as help reducing noise from document tag cloud and providing a separate content-focused section tag cloud.

To overcome above mentioned problems we present a novel modified social tagging approach. The benefit of such an approach has been illustrated in a wiki system on the example of simplifying the editing process. We call this new approach section-tagging as it supports users to assign keywords and annotate sections of a wiki page.

To practically implement and test the idea we extended the functionality of an online encyclopedia called Austria-Forum with section tagging along with the conventional social tagging. The rest of the paper is organized as follows. The next section describes in more details the Austria-Forum system. The third section discusses the idea of section tagging in Austria-Forum and how it may be used to support content retrieval and to simplify a typical editing workflow. The fourth section describes the implementation of section-tagging idea within Austria-Forum. The last section provides conclusions and an overview of the future work.

2. AUSTRIA-FORUM

Austria-Forum [4] is a networked information system that manages a very large repository of information items, where new information items are easily published, edited, checked, assessed, and certified, and where the correctness and a high quality of each of these items is backed by a person that is accepted as an expert in a particular field. Consequently, each of the information items is citable as any other editorially checked content and might be used in education, scientific research, or journalism. The content of Austria-Forum is always related to Austria – as such Austria-Forum might be seen as an Austrian online encyclopedia.

In the first experimental phase of Austria-Forum the system had an editorial board of more than 20 editors and a growing community of users. The number of users who contributed with the content was more than 100. The number of unique users who have visited the site is around 4000 each month.

The current number of contributions is around 80000 (including pictures and videos as well as the content converted from the well-known Austrian cultural information system AEIOU [1]), out of which around 6000 are user-generated contributions – approximately 8% of all contributions. Most of these user contributions are pictures and photos, with a small number of blogs, discussion forum posts, and comments. Although these numbers are quite substantial for a site that has been online experimentally a more active community involvement is desired. Community tools and facilities are already present in the system. However, as a number of users suggested, usability and a better integration of different community tools with the main system needs to be improved.

Therefore, the original system that was technically based on an in-house developed content-management system has been replaced by an open-source wiki software called JSPWiki [5]. The idea here is that more users will be attracted to a well-known

collaborative authoring tool such as wiki. Moreover, the intention is to offer a number of community tools that will support users in retrieving information quickly and reduce the complexity of editing workflow. Among such tools is also the above presented section-tagging tool.

Even if the Austria-Forum wiki is still under development, it nearly offers ideal environment to test the concept because a huge amount of test data is available.

3. SECTION TAGGING AND PERSONALIZATION

Section tagging is a novel social tagging approach which allows users to annotate the content of interest within a resource using free form keywords.

The implemented approach differs from existing tagging and bookmarking services in the following way. First, it allows the tagging of subdocument level content. Second, tag retrieves not merely the set of links annotated by tag keyword but also the actual content of the tagged sections. Thus, when the user clicks on a tag all sections from wiki pages that have been tagged with the particular term by the specific user are dynamically loaded and presented to the user in the form of a standard wiki page.

The section of a wiki page is a self explaining piece of information about some topic of interest. Tagged content snippets in the case of section tagging have conceptual relationship to perceived structure of an information object that the user relates to the tag terms. Hence, the context of information snippet of user's interest is more relevant to the user perception of an information object in relation to the tag terms. The underlying idea of such an approach is based on personalized content aggregation from different wiki pages because the wiki system may not hold the required information in one page but typically in various pages.

Personalization in Austria-Forum refers to the content annotation and aggregation from different wiki pages according to users' intent. A typical personalization scenario involves users collecting, customizing, and modifying diverse text snippets from different wiki pages within an informational focus being described by the given tag keyword.

System offers two levels of personalization:

- Users can tag and annotate sections of wiki pages as well as full pages and hence personalize the content of interest.
- A dynamic personalized wiki page content view is created for a user by aggregating all sections tagged by him with a particular keyword. The aggregated sections are retrieved from the same versions of wiki pages which were used while tagging. The rank of a particular section within this aggregated set is determined by the frequency of same tag assigned by other users to this section.

The resulting dynamic personalized wiki page can further be collaboratively edited to create a logically complete information object reflecting the particular user context. The system facilitates further the personal/collaborative knowledge creation and management. Dynamic wiki pages created by collecting snippets of information from diverse wiki pages allow users to restructure and organize information on multiple axes of personalization.

Currently, the section tagging is primarily used for supporting editing workflow in the system. For example, suppose that an author is writing a new contribution on the Mozart's birth house. Before writing about the birth house the author wants to have an introductory section about Mozart that includes the basic biographical information, the list of Mozart symphonies and a picture of the Mozart monument in Vienna. The basic biographical information is included in the first section of the page on Mozart biography, the list of symphonies is described in the page on Mozart's work and the Mozart monument is depicted in the page that talks about monuments in Vienna. Thus, the author tags all the appropriate section in pages in question with a tag "Mozart". In the personal section-tag cloud the tag "Mozart" is now visible. When the author clicks on that tag a new dynamic wiki page including three tagged sections from three different wiki pages is created on the fly. The author chooses to save the dynamically created page in the system. Now, the author can access the new page as any other wiki page and edit it by restructuring sections and adding new sections about Mozart's birth house.

4. IMPLEMENTATION ASPECTS

As described before, the core of the section-tagging mechanism is to allow users to tag not only a whole wiki page, but also to tag a particular section (identified with a heading). In this way users add semantic information to arbitrary sections of different wiki pages. In the next step, it is possible to extract sections referred by a particular tag and to create a new personalized wiki page out of tagged content snippets.

The implementation of the section concept is comprised of two functional modules, called Section-Tagging (ST) and Personalized-Content-Creation (PCC) module.

The JSPWiki system is based on a clean and extensible plug-in and filter architecture that allows easy addition and configuration of new modules.

The filter mechanism allows on the fly parsing and modifying of wiki pages before they are rendered.

On the other hand, the plug-in mechanism allows server-side code to be referenced from within a wiki page. This code dynamically produces wiki content that can be included in the wiki page that refers to the plug-in.

Thus, technically the ST module is a filter module as it inserts section-tagging functionality into already existing wiki pages by pre-processing them; the PCC module is a plug-in module that dynamically creates a new wiki page according to the selected tag and the tagged sections from various wiki pages.

4.1 Section-Tagging Module

ST module is a filter for pre-processing of rendered wiki pages. This unit is responsible for extending document object model (DOM) of a rendered wiki page via a JavaScript module called ST form module. As shown in Figure 1, this module supplies a simple to use pop-up form (red colored box in front of section) that visualizes particular semantic section information by an onmouseover effect and letting the user tag a section using the onclick event. Moreover the ST form module also supplies the database connector module with information about the currently tagged section number and page version.

The actual centerpiece of the ST module is a unit called ST plug-in. It loads and manipulates the data from the ST data storage backend module, extracts user data from the ST security module and handles data sent by the ST form module via XMLHttpRequest (see Figure 2).



Figure 1: ST form module

As a data storage module the open-source content-management system Scuttle [7] is deployed. The database itself is not accessed by the API which the systems offers but by the database connector module which extracts user data such as username and IP address directly from the JSPWiki user session module. This user data record is stored together with a special section URI to the Scuttle database by the plug-in module every time a section is tagged by the user, in order to guarantee an unambiguous relationship between user and tagged sections.

In order to have a clear relationship between page sections, page versions and corresponding tags and still offer a readable URI without changing the database structure itself, the well known (X)HTML method of creating links within a hypertext document was adopted in the following form:

http://<URI>#<section ID>_<version>

Thus a section of a wiki page can be easily addressed to a tag and vice versa by adding a fraction identifier holding information about the section ID (<section ID>) and page version (<version>).

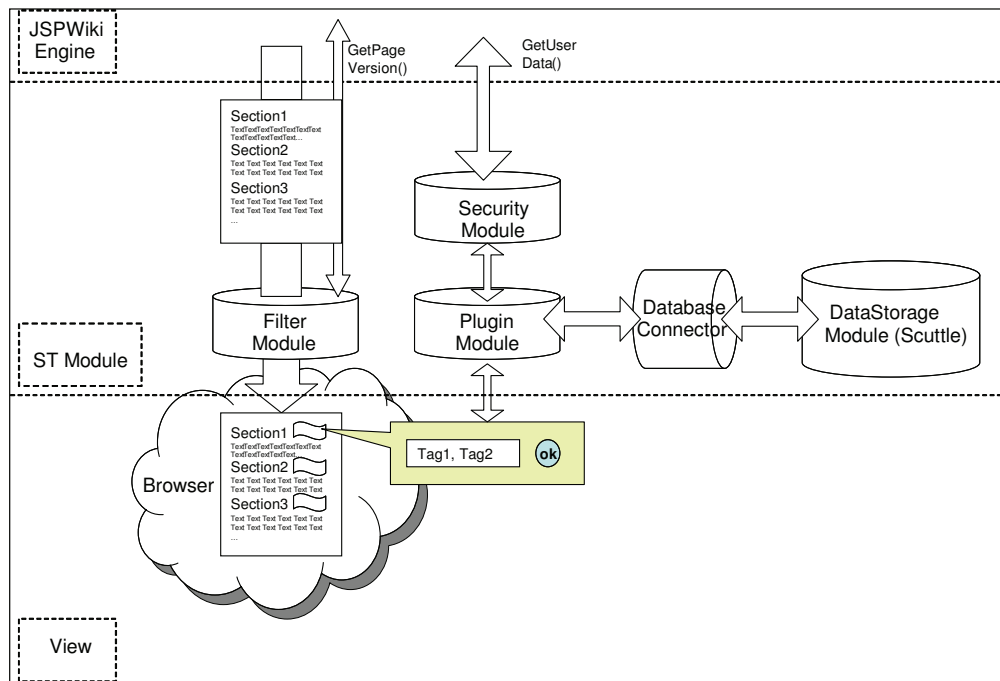


Figure 2: Architectural diagram of the ST module

4.2 Personalized-Content-Creation Module

The PCC module is implemented as a plug-in that can be included in any wiki page. Currently, this module is included in a personalized wiki page that is shown on the right-side of the user screen. It shows a standard tag cloud with tags assigned by a particular user to wiki page sections of interest. When a user clicks on a tag the PCC module retrieves all tagged sections using the appropriate wiki page versions. The sections are then dynamically combined into a wiki page that is shown to the user. The user has then the possibility to edit and modify this new wiki page using the standard wiki editor and to save the editing operations in a completely new wiki page for later retrieval. Moreover, the dynamic page can be still retrieved at all times by simply clicking on the appropriate tag. Note that the dynamic page is always created on the fly, thus whenever the user adds tags to sections of some other wiki pages this will be reflected in the dynamic page as the page will include the new sections.

5. CONCLUSIONS AND FUTURE WORK

A novel approach for tagging sections of wiki pages has been presented. This approach is able to personalize the users' content in an efficient way. This approach has reduced the manual effort required to author a wiki-page about a topic. Often, the wiki system may not have the required info in one page but typically in various pages. Therefore, a combination of the social tagging approach with the wiki concept in an innovative manner facilitates an easy retrieval of the relevant content in the form of a new dynamically created wiki page. Such dynamic wiki pages created by collecting snippets of information from diverse wiki pages allow users to restructure and organize information on multiple axes that best fit their current needs.

The future work includes:

- Testing and evaluating the section-tagging approach with a number of users during the experimental phase of Austria-Forum.

- Sharing of section tags between users, i.e. not only a personalized section-tag cloud should be generated but also a global one with tags from all users.
- Interesting aspects of global section-tag clouds will be the tag and section selection strategy in the case that there are numerous sections tagged by a particular tag. A collaborative filtering approach taking into the account the user profiles might be needed to limit the sections only to those that are most relevant.
- Extending the section-tagging approach to arbitrary Web resources. This can be implemented as browser plug-in in future which will gather the tagged content in a dynamic wiki system as a Web based service.

6. ACKNOWLEDGMENTS

We would like to thank Prof. Hermann Maurer and Mr. Muhammad Tanvir Afzal for providing supporting discussions, valued inputs, and comments. The research of this contribution is funded by Higher Education Commission of Pakistan and TUGraz.

7. REFERENCES

- [1] AEIOU - Annotierbare Elektronische Interaktive Oesterreichische Universal-Informationssystem, <http://aeiou.iicm.tugraz.at>. (Last visited: March 30, 2009).
- [2] Alain Désilets , Sébastien Paquet , Norman G. Vinson, Are wikis usable?, Proceedings of the 2005 international symposium on wikis, p.3-15, October 16-18, 2005, San Diego, California. DOI= <http://doi.acm.org/10.1145/1104973.1104974>.
- [3] Ames, M. and Naaman, M. 2007. Why we tag: motivations for annotation in mobile and online media. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (San Jose, California, USA, April 28 - May 03, 2007). CHI '07. ACM, New York, NY, 971-980. DOI= <http://doi.acm.org/10.1145/1240624.1240772>.
- [4] Austria-Forum, <http://www.austria-forum.org>. (Last visited: March 30, 2009).
- [5] JSPWiki, <http://www.jspwiki.org>. (Last visited: March 30, 2009).
- [6] Nelson, L., Smetters, D., and Churchill, E. F. 2008. Keyholes: selective sharing in close collaboration. In CHI '08 Extended Abstracts on Human Factors in Computing Systems (Florence, Italy, April 05 - 10, 2008). CHI '08. ACM, New York, NY, 2443-2452. DOI= <http://doi.acm.org/10.1145/1358628.1358701>.
- [7] Scuttle, <http://sourceforge.net/projects/scuttle>. (Last visited: March 30, 2009).
- [8] Tony Hammond, Timo Hannay, Ben Lund, and Joanna Scott. Social Bookmarking Tools (I): A General Review. D-Lib Magazine, 11(4), April 2005.