Tag Trails: Navigation with Context and History

Jacek Gwizdka
School of Communication,
Information and Library Studies
Rutgers University
New Brunswick, NJ USA 08901
chi2009@gwizdka.com

Philip Bakelaar
School of Communication,
Information and Library Studies
Rutgers University
bakelaar@rci.rutgers.edu

CHI 2009, April 4 – 9, 2009, Boston, MA, USA
ACM 978-1-60558-246-7/09/04.

Abstract
We describe a technique for preserving and presenting context and history while navigating web resources described by keywords. We use tagging and tag clouds as an application area for our technique. The technique is illustrated by employing it in a prototype that interfaces data from a social tagging website used to bookmark academic articles. The prototype displays a "tag trail" which can reveal contextual connections between web resources and the associated tags. We argue that the user's understanding of web resources is aided by making such connections explicit.

ACM Classification Keywords
H5.4. Information interfaces and presentation: Hypertext/Hypermedia: Navigation; H3.3. Information storage and retrieval: Information Search and Retrieval: Search process.

General Terms
Design

Keywords
Browsing, tag clouds, tagging

Figure 2. A sample tag cloud. The tag cloud represents topics covered in an HCI course taught by the first author.

Introduction

Tag clouds are typically associated with collaborative tagging and Web 2.0 websites such as delicious.com, flickr.com, technorati.com, and youtube.com. Such websites allow their users to associate words (a.k.a. tags) with their own and other users' content. The three sets of entities, users, web resources, and tags become associated with each other in the process of collaborative tagging (Figure 1) [9].

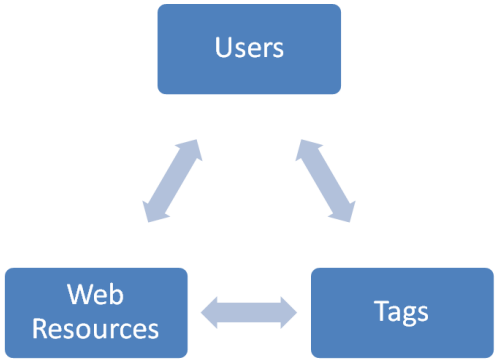

Figure 1. Relational diagram showing users, web resources, and tags. Some users interact with web resources and generate tags, while other users utilize those tags to navigate the web resources.

The associations between tag sets and other entities can then be used to facilitate social navigation, individual resource re-finding, defining communities of interest, and understanding the content of web resources [3], to name just a few. These activities are frequently supported by tag lists or tag clouds [10] (Figure 2) that have become one of most popular visual representations of tag sets.

A tag cloud may be shown next to a list of resources to represent all (or most frequent) words associated by a website user with these resources. A tag cloud may be shown next to a user name (or a group of users) to represent tags most frequently employed by this user (or the group). In both cases, the words in the tag cloud typically serve as links that can be followed to obtain a new view of the associations among the users-resources-tags. For example, a user can follow a tag-link to obtain a list of resources associated with this new tag, or a user can follow a username-link to explore resources tagged by another user. This kind of navigation has been called *pivot browsing*. Millen et al. [8] describe it as a lightweight navigation mechanism that enables users to reorient their view by clicking on tags or user names.

Navigating the user-resource-tag space by pivot browsing does not preserve the history of tag-space navigation. It currently works upon the assumption that each step in navigation is separate from other steps (previous and future), and, if there is any dependency/relationship, a user needs to remember the history of her exploration. Yet the search and navigation process is not a series of individual steps but rather an iterative process [7]. In this project, we focus on providing history of navigating the associations between tags and the web resources.

Related Work

Our concept of Tag Trails was influenced by the ideas from web navigation and information retrieval. In web navigation, breadcrumbs are a technique, which aids users in locating their current position and in navigating hierarchical web sites [5]. While tag-spaces are not hierarchical, the general idea of leaving a trail of

"crumbs" seems applicable. In information retrieval, clickable words can be considered as equivalent to queries that are sent to a system [4]. Information retrieval research has highlighted the usefulness of preserving query history [7]. Hence presenting a history of clicked tags seems applicable. These ideas have shaped the development of our techniques that are aimed at preservation of history and context in browsing tag-described information spaces.

Tags as Navigation Context and History

Tags that were created by others can be generally viewed as serving two main purposes, i) promoting understanding of the associated set of web resources, and ii) facilitating navigation of the collection of web resources. At each step of the navigation process, a tag set serves to describe the set of resources, but it also serves as a starting point for further navigation.

Two complimentary uses of tags as navigation and search mechanism are currently found, 1) narrowing down search results and 2) the earlier mentioned pivot browsing. In the first approach, clicking a tag serves to progressively narrow down a set of web resources and hence an associated tag cloud. An example of this approach is delicious.com, which has recently updated their interface to include a "tag crumbs" component.

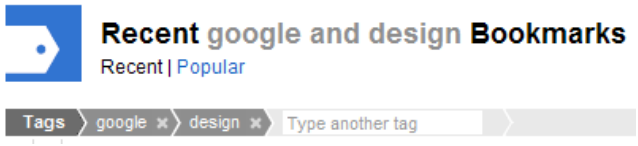

Figure 3. Delicious.com searching bookmarks, retrieved Jan 2 2008.

In the second approach, the tag cloud content is switched at each navigation step to a new one that contains tags associated with the new set of resources. The new set of tags describes these new resources and provides new points for further exploration. However, at the same the display is completely re-oriented around the tag or link clicked. Furthermore, the history of navigating tag-space is not preserved and no information is provided about relationships between tags (and tag clouds). This can be considered potentially confusing and disorienting to the user. Yet the authors have found this approach to be quite common across many sites that employ tags and tag clouds. Our work presented in this paper addresses this second approach and aims to improve the current state.

Data Set

In our exploration of preserving and displaying the navigation history, we use data from CiteULike, a social bookmarking web site that is focused on scholarly work. The data was downloaded on July 14th 2008 and loaded to a MySQL database. For the experimental system, only the records from January 1 to July 14 2008 were used. The data was sanitized to reduce the number of not useful tags (e.g., numbers, words that clearly referred to a particular browser, such as firefox, etc.). Approximately 800,000 records remained. Each record represents a single tagging instance that was submitted by a single user and assigned to a single web resource. For example, if one user tagged a single resource with 10 tags, there would be 10 records entered into the database. The data set contains tags entered by approximately 9,100 unique users and describing 224,000 articles.

Experimental Prototype

The Tag Trails prototype was developed in an iterative fashion using PHP and MySQL (Figure 4). The prototype has two main components, the results area on the left and the tag trail area on the right. The results page is created by an agent that gathers results from the CiteULike.org site based on the tag selected in the Tag Trails interface. The tag trail area is generated from the CiteULike data set that was loaded to the MySQL server. It displays the current tag cloud as well as three most recently viewed clouds. The main focus of the Tag Trails project was development of the tag trail area. The results area is incorporated to demonstrate how tag trails can be used to complement search and navigation of an existing set of resources.

Tag Trails uses three techniques to preserve context and history. First, Tag Trails stores the entire recently seen clouds so that a cloud history could be presented (Figure 4). The history is shown by displaying multiple tag clouds. As a user clicks through various tag clouds, the most recent are kept and the highest-frequency tags are displayed in alphabetical order. The older the history cloud the darker its background color. We chose to maintain three "history clouds" in addition to the current cloud. With this base display, we are able to apply the two following techniques to show contextual relationships and highlight possible connections between the clouds, of which the user may not have been aware otherwise.

Second, the relationships between the tag clouds are shown by using colors. The background color of the title bar of each tag cloud is used to indicate when that tag appears in another cloud; when it does, its background color is the same as in the title bar. Highlighting all tags using this "color context" technique allows the user to partially determine the level of continuity or relationship between the two tags clouds.

Third, the similarities and differences between the current cloud and the history clouds are displayed. We pair-wise compare tag sets associated with the current cloud and each history cloud. We show the result visually by highlighting the different tags in a distinct color (red). Using a distinct font color helps to keep this aspect from visually dominating the cloud. In the prototype, this feature is user-controlled, and can be set to "Off", "Highlight similarities", or "Highlight differences". Figure 4 shows the "highlight differences" mode, which we believe is the most useful. At one glance, the user can group either the black tags (similar) or the red tags (different) and determine which tags co-occur in the current cloud. We also provide a summary about the number of different tags and total tags in each history cloud in comparison with the current cloud. This information helps the user assess the extent of differences between the clouds.

We constrain the tag trails area to approximately 33% of total web browser screen area to illustrate how it can augment rather than dominate the interface. When combined, these three techniques, "cloud history", "color context", and "similarity/difference" present several dimensions of information in a fairly compact space, allowing the user to make higher-level judgments and inferences about their search or navigation process and helping them to navigate the tag space to satisfy their information needs.

Figure 4. Tag Trails interface consist of two main areas: the result list on the left and the tag clouds on the right. Tag clouds on the right represent the tag trail. The interface uses "cloud history", "color context" and "similarity/difference" to show the history and context of user navigation or search process.

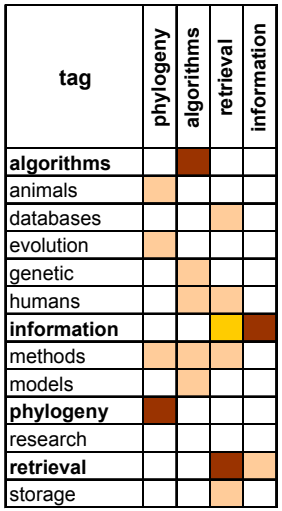

Figure 5. Heatmap that represents top tag co-occurrences for the same tag trail as shown in Figure 4. A similar visual representation has recently been demonstrated to be more effective in supporting explorartion of Google search results [5].

Acknowledgements

We thank several users who reviewed the prototype at various stages and offered feedback.

Summary and Future Work

Systems that currently support tag clouds and pivot-browsing often treat tags as discrete, independent entities. However, user-generated tags are part of a social and mental process in which relationships can be present. For instance, multiple tags may be entered at the same time, or common tags may be suggested prior to a user entering their own. So, we propose that tags should not be treated functionally as independent. The idea behind Tag Trails is to incorporate these tag dependencies and relationships and make them explicit to the user, through techniques such as "cloud history", "color context", and "similarity/difference". These techniques reveal the connections between tags as a user navigates through tag-space.

We argue that preserving navigation and exploration history in tag-space is useful to a user and can be implemented in a way that preserves the light-weight character of pivot-navigation, and we propose to improve navigation in systems and websites that utilize tag clouds and pivot-browsing by providing a history of tag space navigation and context. Our prototype demonstrates several techniques that have a potential to aid user in inferring this higher-level information. While our particular combination of techniques and visual construction may be applicable to a specific kind of social tagging system, the ideas can be modified to suit similar user tasks in other systems.

This research is still in its experimental phase. In the next step we plan to conduct a user study to determine the impact that our techniques have on the user's search and navigation process. We also plan to explore alternative visual representations, such as the heat map shown in Figure 5. This will hopefully lead to further refinement of the interface, and particularly to a better understanding of the design choices required to visually represent multiple dimensions of data using only text and color.

References

[1] Aery, Sean C. Breadcrumb Navigation Deployment in Retail Web Sites. A Master's Paper for the M.S. in I.S. Degree. July, 2007.

[2] Dubinko, M., Kumar, R., Magnani, J., Novak, J., Raghavan, P., Tomkins, A. Visualizing tags over time.

[3] Golder, S.A., Huberman, B.A. Usage patterns of collaborative tagging systems. *Journal of Information Science 32*, 2 (2006), 198-208.

[4] Golovchinsky, G. Queries? Links? Is there a difference? CHI 1997, p 407-414.

[5] Hoeber, O., Yang, X.D. HotMap: Supporting Visual Exploration of Web Search Results. *Journal of the American Society for Information Science and Technology*, 60, 1, 90-110, 2008

[6] Komlodi, A., Soergel, D., Marchionini, G. Search histories for user support in user interfaces, *Journal of the American Society for Information Science and Technology*, 57, 6, 803-807, 2006

[7] Kuhlthau, Carol C. Inside the search process: Information seeking from the user's perspective. *Journal of the American Society for Information Science*, 42, 5, 361-371. 1991.

[8] Millen, D. R., & Feinberg, J. (2006). Using social tagging to improve social navigation. Workshop on the Social Navigation & Community based Adaptation Technologies. (in conjunction with AH'06).

[9] Moulaison, H.L. Social tagging in the web 2.0 environment: author vs. user tagging. *Journal of Library Metadata 8*, 2 (2008), 101-111.

[10] Sinclair, J., Cardew-Hall, M. The folksonomy tag cloud: when is it useful? *Journal of Information Science 34,* 1 (2008), 15-29.