Personalized Search and Exploration with MyTag

Max Braun, Klaas Dellschaft, Thomas Franz, Dominik Hering, Peter Jungen, Hagen Metzler, Eugen Müller, Alexander Rostilov, Carsten Saathoff
ISWeb Working Group, University of Koblenz-Landau
Koblenz, Germany
{maxbraun, klaasd, franz, hdomra, peju, haegarm, vikam, alero, saathoff}@uni-koblenz.de

Categories and Subject Descriptors:
H.5.4 Hypertext/Hypermedia: Navigation

General Terms: Design.

Keywords: Personomy, Searching, Tagging.

1. INTRODUCTION

Nowadays Web 2.0 platforms like YouTube [6], Flickr [3] and del.icio.us [2] provide large amounts of resources such as videos, photographs and social bookmarks. Common to the platforms is the classification by so called tags that can be used for organization and retrieval.

A current limitation of tagging platforms is their confinement to a single media type. Furthermore, a magnitude of platforms exists for each media type. Thus, in both cases of searching resources of either the same or of different media type, a user has to search multiple platforms [7]. For example, a user needs to search on del.icio.us, RawSugar [4] and Bibsonomy [1] to find bookmarks or on Flickr and YouTube to find media related to e.g. an artist.

Another limitation results from the ranking of resources as implemented by platforms such as YouTube, Flickr, and del.icio.us. Usually, the overall popularity of a resource is used for ranking search results. A personalized search is currently missing that takes the interests of a user into account.

MyTag[1] aims at solving the previously described limitations of current tagging platforms by enabling cross-media search across images, video, and social bookmarks. It offers transparent access to different single-media platforms currently including Flickr, YouTube, and del.icio.us (cf. Section 2). Furthermore, it introduces personalization features such as a personalized ranking of search results (cf. Section 3). The architecture of MyTag ensures its extensibility towards further tagging platforms (cf. Section 4).

2. CROSS-MEDIA SEARCH AND EXPLORATION

On its start page, MyTag welcomes its users with a single input field where the search tags can be entered. By default, all platforms currently integrated into MyTag are then searched for the tags. The search results for each media type are presented as separate results sets in different columns that can be sorted by popularity and creation date (cf. Fig. 2).

For every resource in the result sets, its title, a preview, and the creation date are shown. For bookmarks, the preview is a snapshot of the website that is provided by a mashup with the service of Snap.com. The snapshot is shown by moving the mouse over the icon next to the title of the bookmark. Furthermore, for each resource its associated tags are shown. Clicking on the preview of a resource opens a new window that offers additional details about the resource and links to its occurrences on the original platform. Above the results, a tag-cloud is shown that summarizes the most frequent tags in the result set. The font size of each tag is proportional to its occurrence frequency.

In MyTag, clicking on a tag refines the previous search query by adding the tag. This helps in disambiguating the search and giving it the intended direction, and thus in dealing with the problem of homonyms. The user can easily explore different directions in adding or removing tags. This paradigm is in contrast to platforms like Flickr and del.icio.us, where selecting a tag results in a new search with that tag only.

3. PERSONALIZATION

Two personalization features are provided for search: First, a search can be restricted to resources uploaded by the user. This feature requires that a user enters her external account names for Flickr, del.icio.us, and/or YouTube into her profile. Searching own resources is implemented by using the corresponding feature from the integrated tagging platforms.

The second personalization feature allows for ranking search results based on the user's personomy. The personomy is automatically built based on the resources the user picks from the result set. It is modeled by a vector $\mathbf{p}$ of tag frequencies representing the previous search interests of the user. As it is based on the implicit feedback given by selecting from the search results, no additional user effort is required to gain personalization. Using implicit user feedback is a very promising approach to personalizing search results or web browsing in general (cf. [10] and [9]). This feature adds an advantage compared to systems such as Flickr and del.icio.us, where personalization requires adding resources to the system, i. e. the explicit feedback of users.

[1] http://mytag.uni-koblenz.de. This work has been partially supported by the European project Semiotic Dynamics in Online Social Communities (TAGora, FP6-2005-34721).

WWW 2008, April 21–25, 2008, Beijing, China.
ACM 978-1-60558-085-2/08/04.

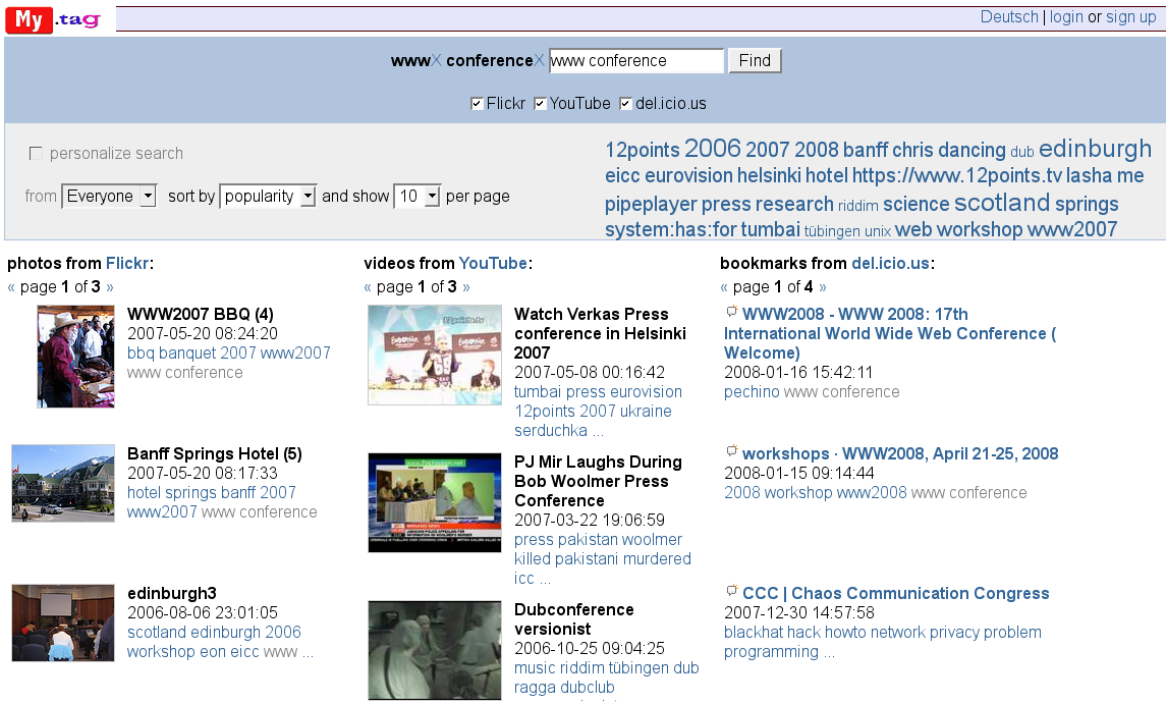

Figure 1: Screenshot of a MyTag Search Result

The current MyTag platform implements a ranking algorithm that combines information from the personomy and the tags assigned to resources of a result set. The tags of a resource are represented as a vector $\mathbf{v}$ of binary values indicating the presence of a tag. The rank r of a resource is then computed by the scalar product of the two vectors: $\mathbf{r} = \mathbf{v} \cdot \mathbf{p}$. It is then used for ordering the resources based on their rank value.

4. ARCHITECTURE

MyTag is implemented using the Ruby on Rails framework [5] as it supports efficient development of web-applications. The MyTag architecture realizes the model-view-controller paradigm (MVC). The three layers of the MyTag architecture are shown in Fig. 2. The view layer at the top is responsible for the interaction with the user while the control layer in the middle processes data from the model layer, e.g. by computing personalized rankings.

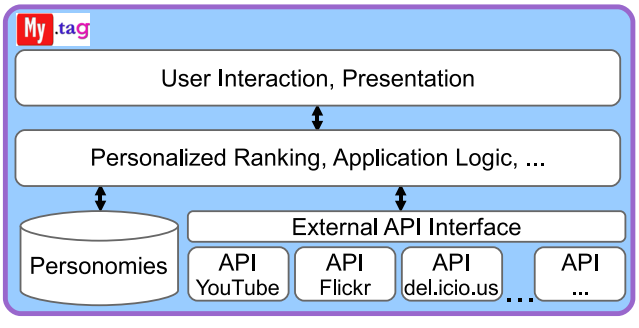

Figure 2: MyTag Architecture

The model layer consists of two core parts: First, the interface to the local database that contains the user profiles and personomies. Second, a generic interface that abstracts core functionality provided by the APIs of the external tagging platforms, ensuring MyTag's future extensibility.

5. CONCLUSIONS

MyTag provides intuitive cross-media search and exploration over multiple tagging platforms. Personalized rankings of search results are featured by incorporating implicit user feedback gathered during the search activities of a user. As future work, we plan to evaluate and enhance the personalized ranking algorithm and compare it with state-of-the-art approaches like FolkRank [8]. Furthermore, we will focus on algorithms for merging search results from different platforms into a single result set.

6. REFERENCES

[1] Bibsonomy website. http://bibsonomy.org.
[2] Del.icio.us website. http://del.icio.us.
[3] Flickr website. http://flickr.com.
[4] Rawsugar website. http://rawsugar.com.
[5] Ruby on rails website. http://www.rubyonrails.org/.
[6] Youtube website. http://youtube.com.
[7] G. Anadiotis, T. Franz, and S. Boll. W3C Multimedia Semantics Incubator Group: Tagging Use Case. http://www.w3.org/2005/Incubator/mmsem/wiki/Tagging_Use_Case, 2007.
[8] A. Hotho, R. Jäschke, C. Schmitz, and G. Stumme. Trend detection in folksonomies. In *Proc. of the International Conference on Semantic and Digital Media Technologies*, 2006.
[9] D. Mladenic. Web browsing using machine learning on text data. In P. Szczepaniak, J. Segovia, J. Kacprzyk, and L. Zadeh, editors, *Intelligent exploration of the web*. Physica-Verlag, 2002.
[10] K. Sugiyama, K. Hatano, and M. Yoshikawa. Adaptive web search based on user profile constructed without any effort from users. In *Proc. of the 13th WWW Conference*, 2004.