Visualizing Social Bookmarks

Joris Klerkx[1], Erik Duval[1]

[1] Departement Computerwetenschappen, Katholieke Universiteit Leuven,
Celestijnenlaan 200A, B-3000 Leuven, Belgium
{joris.klerkx.erik.duval}@cs.kuleuven.be

Abstract. Social bookmarking tools are very popular nowadays. In most tools, users tag the bookmarks to describe them. Therefore, it is often hard for users to discover implicit structures between tags, users and bookmarks. We think that this is essential for both end users to discover new bookmarks that could be of interest to them, and for researchers who want to study how people use social information retrieval tools. In this work, a cluster map visualization application is customized to enable users to explore social bookmarks in the del.icio.us system [5]. The design of the application aims to automatically identify tag and community structures, and visualizes these structures in order to increase the users' awareness of them.

Keywords: Information visualization techniques, social bookmarking tools, folksonomies, tags, cluster map.

1 Introduction

The ability to store or bookmark web page addresses has been one of the most important features of browsers since the beginning of the Web. Social bookmarking tools became possible when the process migrated from keeping the bookmarks on the client, to keeping them online on the Web, described by tags or terms. The purpose of social bookmarking tools is to tag the content of other users, mainly for the benefit of the tagger, although the bookmarks and tags are generally public, and users can establish networking opportunities [1]. The personal approach of tagging is an unstructured bottom-up approach of classifying content, in contrast to a top-down structured approach based on taxonomies, thesauri or ontologies. The semantic structures that result from a tagging approach are often referred to as a folksonomy.

The problem with tags is that they generally produce a flat namespace, rather than the hierarchical structures that taxonomies or other formal classification systems provide [1]. However, there can be rich implicit structures between tags, bookmarks and users. We designed a system that attempts to visualize these structures, so that end users can explore the social bookmarks in a playful, efficient and flexible way.

There are a number of social bookmarking tools, like CiteULike [2], Furl [3], BlogMarks [4], etc. Del.icio.us [5] is probably the most well-known social bookmarking tool, designed to store and share bookmarks on the web instead of saving in the browser. We chose del.icio.us as a source of data for our visualization

tool as it has many users, lots of data and a very easy API to access this data. Visualizing data of other social bookmarking tools can be done easily in the future.

We start this paper with a detailed overview of the design of our visualization in section 2. Section 3 describes the underlying technology. A short overview of related work is given in section 4 and we conclude the paper in section 5.
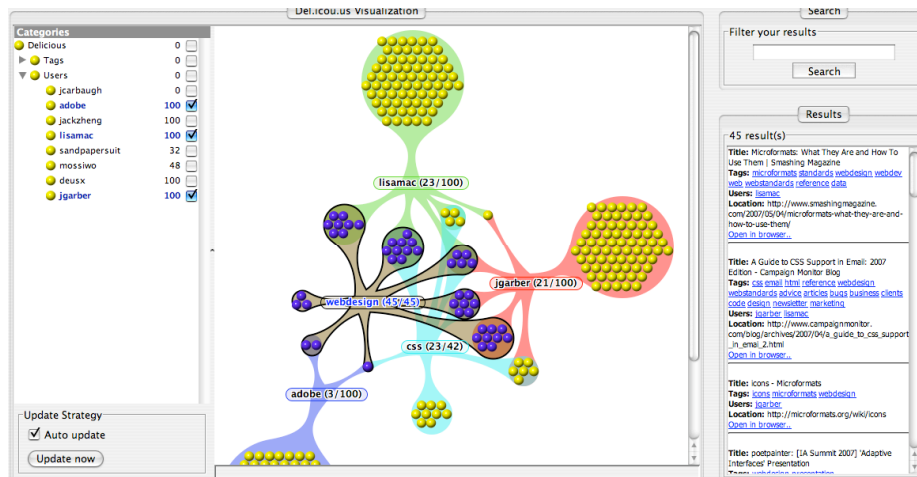

Fig. 1. A visualization of the del.icio.us social bookmarking tool.

2   A Cluster Map of Del.icio.us

We designed a visualization application for the social bookmarks in del.icio.us. First we discuss the design requirements in section 2.1. In section 2.2, we look at the data we use for the application. The customized cluster map visualization is discussed in section 2.3 and we conclude with a typical exploration use case in section 2.4.

2.1   Requirements

Our application for browsing social bookmarks was designed with a number of requirements in mind. First of all, we wanted to use and evaluate a novel paradigm for browsing social bookmarks, compared to the hypertext paradigm that is used on the del.icio.us website. Secondly, we wanted the system to be able to automatically identify tag and community structures. Such a structure is formed when two or more tags or users respectively describe or share common bookmarks. After the identification of these structures, they must be visualized for the users, so that they can become aware of them. We think that this is essential for both end users to discover new bookmarks that could be of interest to them, and for researchers who want to study how people use social information retrieval tools. On top of this, users

should be able to explore the social bookmarks in a playful manner in a fun and engaging space.

2.2 Data

In order to visualize the content of del.icio.us, we have to collect the data first. We start from one or more del.icio.us user names and collect their bookmarks and the tags that describe them. Del.icio.us enables users to add other users to their network, so that they can access the bookmarks of these users. In the second step, we collect the tags and bookmarks for all the users that are in the network of the users identified in step 1. In a third step, the same data is collected for all the fans of the chosen user(s) in step 1. User A 'is a fan of' user B if user A has added user B to his network. This relation is not necessarily bi-directional.

During the exploration of this initial set of data, end users can ask the system to expand the data with the tags, bookmarks and network of users of their choice.

2.3 Cluster Map Visualization

Fig. 1 shows the user interface of our del.icio.us visualization: it consists of three panes – a tree structure that shows the different users and tags in the currently loaded data, a cluster map visualization [7] and a search pane with integrated results list. All three panes are synchronized with each other. We discuss these panes in detail in the next paragraphs.
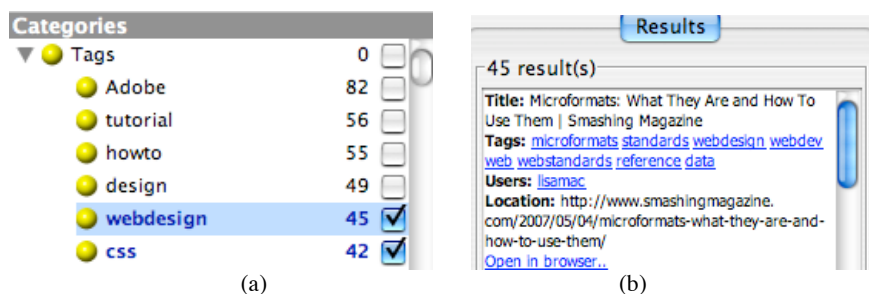

Fig. 2(a) Tree view of the tags that describe the bookmarks.
(b) Result view showing detailed metadata about a bookmark.

Tree Structure. This structure presents an overview of the tags and the users that are currently loaded in the system. Fig. 2a shows that each tag is presented as a node in the tree and the number of bookmarks, that are described with the tag, is indicated. In the design of this tool, we choose to follow the philosophy of "start with what you know, then grow" [6]. This means that, by default, nothing is visualized in the cluster map. After the checkbox in the tree structure is selected, the corresponding tag or user is visualized together with its bookmarks. In this way, the initial visualization carries less perceptual and computational burden to start with.
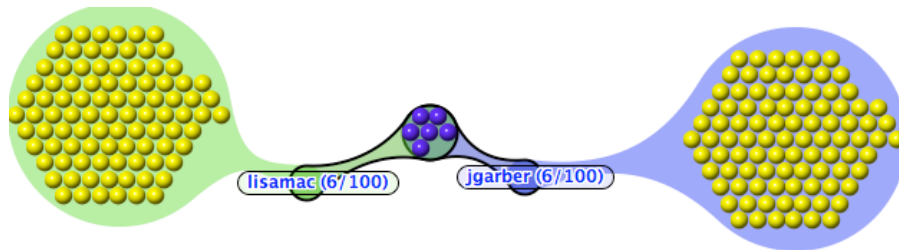
Fig. 3. Cluster Map showing 2 users with 100 bookmarks and 6 of them in common.

Cluster Map Visualization. For the visual presentation of the social bookmarks in del.icio.us, we customized an existing visualization technique [7]. A bookmark is represented as a small circle in the visualization. Each bookmark belongs to the collection of one or more users. In Fig. 3, two users are shown, each having hundred bookmarks in their collection. Those two users have six bookmarks in common. This is represented in the visualization by the smaller common cluster of bookmarks in the middle. By using the tree structure pane that was described in the previous paragraph, users can select which users and tags are drawn on the cluster map.
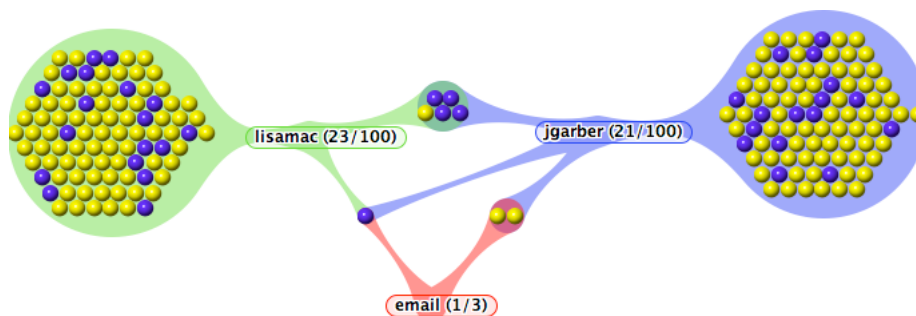

Fig. 4. Cluster Map Visualization, showing 200 bookmarks with 2 users and 1 tag 'email'.

Bookmarks can be clustered by the users that have them in their collection, or by the tags that describe them. This can be seen in Fig. 4, where the tag 'email' is shown. Users 'lisamac' and 'jgarber' have one bookmark in common which is tagged by 'email'. The user 'jgarber' has 2 extra bookmarks tagged with 'email' that are not in the collection of the other user. In Fig. 4., two colors are used to represent bookmarks – yellow and blue. The blue color means that the bookmark belongs to a selected set of bookmarks. Such a set can be created by clicking on a user or a tag in the tree pane, or by performing a keyword query in the search box. In Fig. 4, the blue bookmarks are also tagged with the tag 'webdesign' which is used to create a selection by clicking on this tag in the tree pane (Fig. 2a). The blue bookmark between 'lisamac' and 'email' is therefore in the collection of the users 'lisamac' and 'jgarber' and is described by the tags 'email' and 'webdesign'.

Result list. This list, shown in Fig. 2b shows detailed metadata about the bookmarks that are either selected in the cluster map, or bookmarks that match the search terms when a query was performed. The metadata cover the title, the location, user(s) that added the bookmark and tag(s) that describe the bookmark. A user can interact with the detailed information by clicking on e.g. a tag that describes a bookmark. If this tag is not already drawn on the cluster map, the visualization automatically updates itself and the tag classification is shown.

2.4 Typical Use: Exploring Social Bookmarks

As an alternative to exploring social bookmarks on the del.icio.us webpage, we offer a novel access paradigm that enables a user to explore social bookmarks in a playful manner in a fun and engaging space. There are a number of ways that users can start the exploration of the social bookmark space. First of all, a user can start from an egocentric point-of-view, with a visualization of his bookmarks, much along the lines of the philosophy of "start with what you know, then grow" [6]. The user can select a number of bookmarks in the visualization or perform a keyword query after which detailed metadata about the resulting bookmarks is shown in the result list (Fig. 2b). The metadata contain all the tags that describe the bookmarks, i.e. not only the tags of the user itself. The user can click on those tags and by doing this add them to the cluster map, where the layout of the bookmarks updates itself to represent the new sub-clusters, like in the example of Fig 4.

A second way of exploring is browsing the tree structure to find interesting tags to visualize. Users can order the tag tree alphabetically or by the number of bookmarks they describe. By adding tags to the visualization, the corresponding bookmarks, possibly tagged by different users, can be explored in the cluster map.

Users can also interact with the visualization itself by clicking on nodes and expanding these nodes and the visualization with new information. Upon clicking a bookmark, the user can choose to show all associated tags. Depending on the choice, the network and the fan data of all the users that saved this bookmark are loaded into the system so that they become available for exploring. The reason that we do not add the data of all users that saved a bookmark automatically from the start is the usage throttling and abuse monitoring software at the del.icio.us website [8].

A last way of exploring del.icio.us social bookmarks with our tool is to enter one or more user names in a dialog box. All the data of those users is then loaded in the system and ready for exploring.

3 Underlying Technology

The tool for visually searching and analyzing del.icio.us social bookmarks is created with the open and extensible information visualization framework that we created as part of our research on the use of information visualization techniques for flexible and efficient access to learning repositories [9]. We want to make it easy to add new visualization techniques as well as new data sets, possibly delivered in various

formats and structured according to various metadata schemes. This was a requirement because we wanted to reuse this framework in different contexts. Visualization techniques like tree-maps, hyperbolic trees, node-graphs, fisheye-views, etc. are supported by the framework by plugging in existing visualization components into our framework.

For the visualization of the del.icio.us bookmarks, we plugged the Aduna Cluster Map software [10] into our framework. The Aduna Cluster Map software is a library that contains functionality for creating visualizations of collections of hierarchically classified objects. By integrating this library in our framework, the Aduna visualization technique is available for the application, described in this paper, but also for other case studies developed with our framework. Moreover, this integration demonstrates the open and extensible nature of our framework.

With our framework, we can now also visualize the bookmarks as e.g. a tree-map where the bookmarks are classified per user (Fig. 5), in a flexible and efficient manner, or with any of the other techniques that our framework provides.
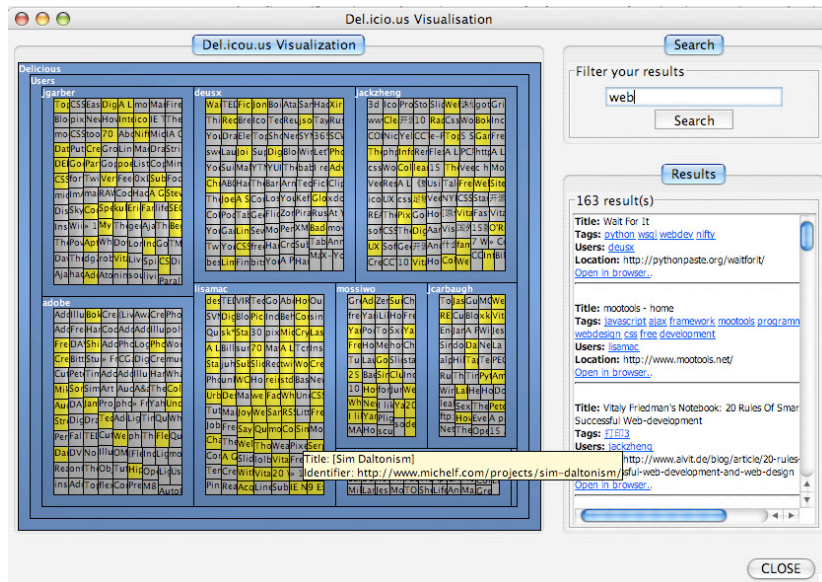

Fig. 5. Tree Map Visualization: blue rectangles represent the users, yellow rectangles represent the bookmarks that match the selection made by the keyword 'web' in the search box and gray rectangles are bookmarks that are not selected.

4   Related Work

There a quite a few social bookmarking initiatives, like CiteULike [2], Connotea [11], del.icio.us [5], Furl [3], BlogMarks [4], etc. A thorough general review on a number of these initiatives can be found in [1]. Tag Clouds are normally used to visualize the tag structures of one or more users. Many visualizations of tag clouds were created for del.icio.us. Notable such visualizations include HubLog [12] that

enables graphical browsing of del.icio.us tags in a mind mapping way. Extisp.icio.us Text [13] provides a random textual scattering of user tags, sized according to the number of times that they've been used. Revealicious [14] is a set of 2D graphic visualizations that enables the user to browse, search and select tags and bookmarks. Vizster [6] is a tool that is designed for visualizing the online social network Friendster, as a browseable network of social relations. Vizster is very useful for sociological research but does not take tag structures into account.

All these initiatives are based on either visualizing the tag structures or the community structures where in our design both structures are taken into account for the visualization, so as to make more apparent implicit community and tag structures.

5 Conclusion

In this paper, we presented our design of a visualization application for social bookmarks. Our design tries to visualize implicit structures between tags, users and bookmarks. We customized a cluster map visualization technique for this purpose. The social bookmark tool del.icio.us was chosen as the source for the data to be visualized.

In the future, we will extend the system with data of other social bookmarking tools. In this way, users can explore social bookmarks from multiple tools in our application. Other possible data sources are folksonomies like Flickr [15] where photographs are described by tags. Validating the usefulness and the effectiveness of our design for visualizing del.icio.us is very important. Therefore, we will proceed by gathering user data of the application. We will gather this data by integrating the Contextualized Attention Metadata (CAM) framework [16] into our information visualization framework. CAM can be used to capture the attention a user spends on content in an application. We will analyze this data, as well as the data of usability tests we plan to do, to validate the effectiveness and usefulness of our design.

References

1. Hammond, T., Hannay, T., Lund, B, Scott, J. Social Bookmarking Tools (I): A General Review. D-Lib Magazine, 11(4), April 2005, http://dx.doi.org/10.1045/april2005-hammond
2. CiteULike. http://www.citeulike.org
3. Furl, http://www.furl.net
4. BlogMarks, http://www.blogmarks.net

5. Del.icio.us, http://del.icio.us

6. Heer, J. and Boyd, D. 2005. Vizster: Visualizing Online Social Networks. In Proceedings of the Proceedings of the 2005 IEEE Symposium on information Visualization (October 23 - 25, 2005). INFOVIS. IEEE Computer Society, Washington, DC, 5.

7. Fluit, C., van Harmelen, F., Sabou, M., Ontology-based Information Visualization: Towards Semantic Web Applications, in Visualising the Semantic Web (2nd edition), 2005, Springer Verlag.

8. Del.icio.us abuse monitor: http://del.icio.us/help/json/url

9. Klerkx, J., Meire, M., Ternier, S., Verbert, K., Duval, E., "Information Visualization: Towards an Extensible Framework for Accessing Learning Object Repositories", Proc. ED-Media'05, AACE, Montreal, Canada (2005), pp. 4281-4287

10. Aduna ClusterMap Library: http://www.aduna-software.org/projects/display/CLUSTERMAP/

11. Connotea: http://www.connotea.org/

12. HubLog: http://hublog.hubmed.org/archives/001049.html

13. Extisp.icio.us: http://kevan.org/extispicious

14. Revealicious: http://www.ivy.fr/revealicious/

15. Flickr: http://www.flickr.com

16. Wolpers, M., Martin, G., Najjar, J. and Duval, E., Attention Metadata in Knowledge and Learning Management, Proc. IKNOW'06, Graz, Austria.