

Social Navigation in Digital Libraries by Bookmarking*

Fiftarina Puspitasari, Ee-Peng Lim, Dion Hoe-Lian Goh, Chew-Hung Chang,
Jun Zhang, Aixin Sun, Yin-Leng Theng, Kalyani Chatterjea, and Yuanyuan Li

Nanyang Technological University, Singapore 639798

Abstract. In the age of Web 2.0, users are increasingly familiar with social tagging or bookmarking where comments and ratings are added by users to objects on the web for public consumption. Such comments and ratings are represented in bookmarks which can be used for information or opinion sharing, user interest discovery, and content recommendation. In this paper, we investigate social bookmarking in digital libraries and derive the design requirements for digital library incorporating social bookmarking. Instead of implementing social bookmarking functions in digital library systems from ground zero, we have chosen to explore the possibilities of integrating pre-existing digital library systems with pre-existing social bookmarking systems, and to derive a feasible system architectural design. We also present a case study where G-PORTAL, a geography digital library system, is integrated with SCUTTLE, an open source social bookmarking system.

1 Introduction

Social bookmarking[5] is no stranger to web users. Many users employ social bookmarking (e.g., del.icio.us¹, CONNOTEA², SCUTTLE³) to tag web objects for future references, for sharing, as well as for expressing opinions on the bookmarked objects. With social bookmarking, users can more easily collaborate by tracing each other's activities recorded by the social bookmarks, and this is also known as *social navigation*[4]. Moreover, ratings given in bookmarks can be used to evaluate the quality of bookmarked objects[7].

Most existing digital library systems consist of some content and metadata repositories created by professional librarians. They address mainly the information search and browsing issues but not issues related to user collaboration. Hence, social bookmarking is so far not a core feature in digital libraries despite its popularity among Web 2.0 applications. There have also been very little study on system issues related to extending digital library systems with social bookmarking.

* This work is funded by the Centre for Research in Pedagogy and Practice, National Institute of Education, Singapore.

¹ <http://del.icio.us>

² <http://www.connotea.org>

³ <http://scuttle.org>

There are several obvious benefits introducing social bookmarking to digital libraries. Apart from giving users alternative means to search and browse content collections, social bookmarking creates a space for users to interact with one another forming online communities that are essential in knowledge creation, sharing and dissemination[9,12].

In this paper, we study the provision of social bookmarking in a digital library system from an architectural standpoint. To the best of our knowledge, there have not been much technical study on how social bookmarking systems can be extended to bookmark digital library objects, and how a digital library system can be extended to provide social bookmarking feature. We therefore adopt a fresh approach to determine the design requirements for social bookmarking in digital libraries. This leads us to an integration framework with a set of strategies to incorporate social bookmarking into a digital library system. We also present a case study where G-PORTAL[8], a digital library system for geography learning, is integrated with SCUTTLE, an open source social bookmarking system[11].

The remaining sections of this paper are organized as follows. Section 2 gives an overview of related research. In Section 3, the requirements of social bookmarking in a digital library system are derived, followed by our proposed integration framework in Section 4. Section 5 presents the integration between G-PORTAL and SCUTTLE as a case study. Finally, Section 6 gives our conclusions and outlines the future research directions.

2 Related Work

2.1 Digital Libraries with Social Bookmarking Support

Very few existing digital libraries provide bookmarking or more often called annotation services for users to mark useful content objects. DLESE, for example, introduces an annotation metadata format and encapsulates annotations within its metadata records[2]. Based on the annotation metadata format, DLESE could incorporate a few annotated sub-collections. The bookmarking functions of DLESE however have not been reported.

In the DiLAS project, a decentralized framework that manages social bookmarks or annotations independently from a digital library is proposed[1]. The main idea here is to provide social bookmarking functions to any digital libraries with a very loose coupling architecture. Instead of using an existing social bookmarking systems, DiLAS has custom-built a social bookmarking system.

Our work is unique is that it introduces a framework for integrating social bookmarking and digital library systems. The framework includes both a reference architecture and the common database integration strategy. We also present a example case study to augment an existing digital library system with social bookmarking features.

2.2 Social Bookmarking Systems

Social bookmarking systems have become popular with their attractive features on storing and organizing online bookmarks. Most of these systems focus on

sharing and visualizing bookmarks on web page URLs. Some of them are designed for certain target user groups. For example, CONNOTEA is designed to help scientists and researchers manage useful references to research literatures. Social bookmarking has also been incorporated into many different Web 2.0 applications, e.g., Flickr⁴, and YouTube⁵.

Despite their flexibility to bookmark any web content, social bookmarking systems has little control over the quality of bookmarked objects as well as the quality of bookmarks themselves. Digital library systems offer a solution to address the quality issues. With the digital library content carefully constructed and reviewed by library professionals, the quality of content and metadata objects and the consistency in controlled vocabulary classification in digital libraries are expected to be high. Hence, social bookmarking digital library content and metadata should be an attractive option when knowledge sharing is concerned.

3 Design Requirements for Digital Libraries with Social Bookmarking

In this section, we discuss digital library objects that can be good candidates for objects of bookmarking and how the bookmarks can be represented in the digital library context. Subsequently, we suggest some bookmarking features desired in a digital library. We divide the design requirement of a digital library with social bookmarking capabilities into several sub-areas as described below.

Objects for Bookmarking

The objects to be bookmarked can be divided into two aspects, namely *object type* and *granularity*. The object type refers to the kind of objects that can be bookmarked. Digital library systems usually maintain two main classes of objects, (i) content and (ii) metadata objects. For digital libraries that use content objects residing at public websites, the metadata objects can serve as surrogates and be used for bookmarking.

From the granularity standpoint, we examine three possible groupings of objects for bookmarking. One may want to bookmark: (a) a single object, (b) a collection of objects or (c) a fragment of an object. The granularity chosen will depend on the amount of information users want to bookmark. For example, a user may wish to bookmark a collection of news reports related to United States Presidential Election (type (b)), or a biography of some music composer appearing in a music history article (type (c)). Most existing social bookmarking systems support bookmarking of single objects only.

Bookmark Representation

Within a digital library, bookmarks can be treated as content or metadata objects and be represented in some record structure. There are mainly two possible record structures, namely:

⁴ <http://www.flickr.com>

⁵ <http://www.youtube.com>

- *Single fixed record structure*: A fixed set of attributes are defined to store a bookmark. This approach has been adopted by most existing social bookmarking systems. The attributes may include owner, reference to bookmarked object, date, keywords, ratings and comments.
- *Multiple record structures*: This bookmark representation allows a customized set of attributes to be used for representing a bookmark. This is necessary for bookmarks with special purposes. To allow bookmarks to be processed uniformly, these record structures will have to share a common set of mandatory attributes (e.g., owner, date, keywords, etc.).

Bookmark Access

This set of design requirements concerns the access to bookmarks. In the existing social bookmarking systems, bookmarks are either “owner accessible only” or “publicly accessible”. In a digital library, there is however additional access control over the digital library objects that may affect the access to the bookmarks. For example, the bookmarks of a bookmarked object should not be accessible by all public users if the latter is only accessible to a small group of users.

Bookmark Query

Social bookmarks are essentially structured information that can be queried based on their attributes, including keywords, references to bookmarked objects, owners, and others. In the digital library context, querying bookmarks can be associated with querying digital library objects. For example, one can search content or metadata by querying their bookmarks and vice versa.

Bookmark Organization

Bookmark organization allows bookmarks to be navigated based on bookmark attributes as well as bookmark categories. We further divide the latter classification approach into the following schemes.

- *Folksonomy-based classification*. This scheme allows users to classify bookmarks by assigning one or more keywords to each bookmark. This keyword serves as an open category label. The set of category labels is open since any keyword can be used as opposed to a controlled set. Hence, a bookmark can be classified into multiple keyword categories. This kind of classification is popular among existing social bookmarking tools. When a user browses a category, s/he will not only view his/her bookmarks in the category, but also other users’ bookmarks under the same category.
- *Controlled-vocabulary based classification*. This scheme uses hierarchical structure to organize bookmarks. In this case, the categories are predefined by some digital library administrator(s) or designated experts. This classification scheme is similar to the one implemented in existing digital library systems. The main advantage of such a system lies in the quality of classification.
- *Mixed classification*. This scheme allows the bookmarks being classified based on both folksonomy and hierarchical category structures. One example of mixed classification is *tag-bundle*, which was introduced by del.icio.us. Tag

bundle is used to group keywords into some user-defined categories. This allows users to define tag bundles of keywords for navigating bookmarks.

4 Integration Framework

Having identified the social bookmarking requirements of digital libraries, we now present a framework for integrating a digital library system with a social bookmarking system. The framework consists of a **reference architecture** to describe the generic architectural elements of a digital library system with social bookmarking functions, and some **integration strategy** to outline the possible integration approach. Due to space constraint, we will only present the common database integration strategy in this paper.

4.1 Reference Architecture

Figure 1 depicts the reference architecture with both digital library and social bookmarking system elements. Similar to digital library architectures adopted by National Science Digital Library (NSDL)[6] and DELOS[3], we divide the system into four levels, namely *user*, *user interface*, *service* and *storage*.

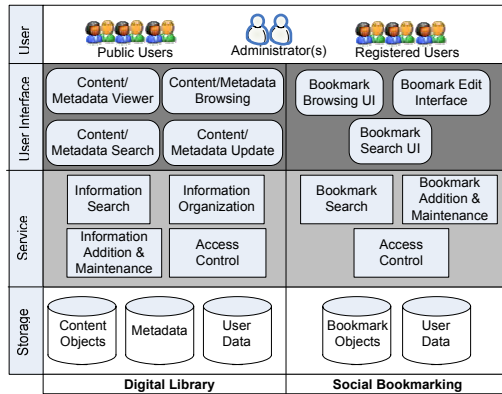


Fig. 1. Reference Architecture to Incorporate Social Bookmarking in DL

- **User Level.** A digital library system or a social bookmarking system serves a population of public users and registered users. In cases where the system is designed for internal use, only registered users may be supported. Administrators in a digital library or social bookmarking system are usually responsible for managing user accounts. For digital libraries, administrators may include librarians who maintain the digital collections and perform other library tasks.
- **Storage Level.** At the storage level, a digital library provides a repository for content objects, metadata and user data. While the object collections in

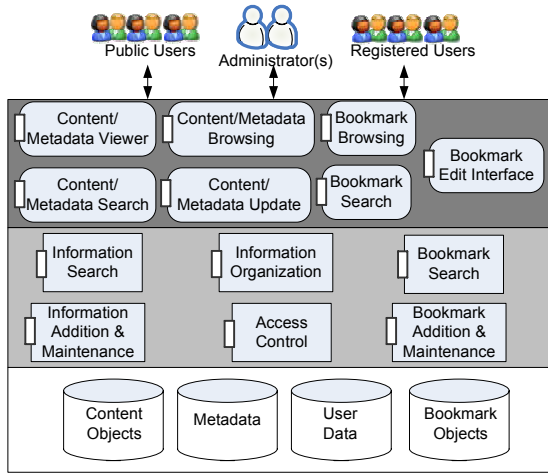


Fig. 2. Integration Strategy (Common Database)

most digital libraries are to be shared by all users, there are cases where digital library services are designed to maintain sub-collections for individuals or user groups[10,8]. For them, only the authorized users can access the sub-collections. A social bookmarking system, in contrast, maintains a collection of bookmarks. Usually, these bookmarks are open for public access except when they are specifically created for personal consumption by the owners.

- **Service level.** At the service level, the digital library and bookmarking systems provide similar services (i.e., search, addition and maintenance) over their objects. Information organization service is less important in bookmarking system due to the popular approach of using keywords to organize bookmarks. Hence, Figure 1 shows this service adopted only by digital libraries. Access control service exists in both kinds of systems to support user authentication and user access right maintenance (i.e., who owns what objects) and determines which users can perform what operations (e.g., view, update, delete) on which objects.
- **User interface level.** The digital library and social bookmarking systems adopt some user interface (UI) modules. Content/metadata viewer, one of these UI modules, is always present in a digital library system since each content/metadata object carries richer amount of information compared to a bookmark. The social bookmarking system, on the other hand, has more emphasis on social navigation and therefore has a more comprehensive bookmark browsing/search interface that allows users to find bookmark objects easily.

4.2 Common Database Integration Strategy

The common database strategy is suitable for an existing digital library system that needs to be extended with bookmarking features. The strategy involves the

construction of a single system with both the digital library and bookmarking services adapted to the common database design as shown in Figure 2. The integrated system requires only one access control service. To minimize development efforts, the modules from the digital library and social bookmarking systems still have to provide APIs for them to interoperate with each other.

Having a single system clearly reduces confusions to the end users. The common database strategy however may also pose some inconsistencies in UI design. However, the use of common databases also makes it possible for some of the UI modules to be modified to suit the user needs. Such modifications will be further elaborated in our case study (see Section 5).

5 Case Study: G-PORTAL and SCUTTLE

5.1 Overview of G-PORTAL

As mentioned earlier, G-PORTAL is developed for learning purpose. It provides a repository of resources surrogated by metadata records and shared among G-PORTAL users. To support learning, G-PORTAL introduces a concept of *project*, which is a set of metadata records identified and assembled for a specific learning task. Users participate in the learning task by contributing/browsing metadata records within the same project space as well as organizing metadata based on user-defined classification schemes. A project can be configured to be accessible by all users, selected users, or only the project owner. When a project is assigned shared access to multiple users, these users can be grouped into one user group. A user group can be granted to access several different projects since it represents users sharing common interest and learning tasks.

5.2 Social Bookmarking to G-PORTAL

To support better learning, we have decided to augment G-PORTAL with bookmarking capability for metadata records in G-PORTAL using the common database strategy. This strategy reduces the amount of development efforts involved as we can keep most G-PORTAL modules intact while modifying some for the integrated system. The integration makes use of SCUTTLE, a open source social bookmarking system implemented using PHP[11]. The bookmarking services offered by SCUTTLE include bookmark navigation, adding bookmark, updating and deleting bookmarks.

After integration, the combined system contains modules from G-PORTAL and SCUTTLE share a common set of databases. The user and access control data of bookmark objects and digital library metadata objects are combined at the storage level. The service level of the combined system consists of a common access control module, metadata-related modules from the original G-PORTAL system and bookmarking modules from SCUTTLE. All these modules are extended with APIs to allow inter-module calls. At the UI level, similar merger of modules also takes place. In particular, the metadata viewer of G-PORTAL is combined with the bookmark edit interface of SCUTTLE so as to facilitate bookmarking when viewing metadata of G-PORTAL.

5.3 Bookmark Representation

We define G-PORTAL's bookmark representation to contain mandatory attributes and optional attributes, as shown below (with attribute type indicated in brackets):

- **Bookmarked object name and URL (mandatory)** together identify a bookmarked object. In G-Portal, a metadata object name refers to its resource name, and the metadata object's URL is the web reference to the metadata object. As the same metadata object may appear in different projects, it is important to keep the project information in the bookmark. Therefore, the URL contains the metadata object id and the id of project where the metadata object is bookmarked.
- **Bookmark owner(mandatory)** attribute stores the user id of the bookmark contributor.
- **Date of creation(mandatory)** records when the bookmark is created.
- **Description(optional)** stores users' comment or description on a bookmarked object.
- **Rating(optional)** is a value between 1 and 5 inclusive, where 1 represents 'least useful' and 5 represents 'most useful' respectively. This attribute can be used by users to rate the usefulness of metadata records.
- **Media object(optional)** attribute allows users to associate one or more media objects (e.g., images, videos, etc.) with the bookmark. These media objects have to be made available on the web and be identified by their URLs.
- **Bookmark keywords(optional)** is a set of keywords associated to the bookmark. Like other social bookmarking systems, the list of keywords is left open so that a folksonomy based classification of bookmarks can be supported.

5.4 Bookmark Access

The bookmark access control in the combined system is designed to be consistent with that of G-PORTAL. G-PORTAL enforces access control over its metadata objects by designating for each project a coordinator who specifies the group(s) of users allowed to access the project. Hence, the access control module only allows a user to add bookmarks to metadata records in a project when the user is authorized to access the project. Furthermore, an authorized user of a project is also allowed to view bookmarks on any metadata records in the project.

5.5 Bookmark Browsing and Search

Bookmark browsing and search interface is one single module in the integrated system. A user can browse and navigate bookmarks both in a bookmark list and a tag cloud as indicated in Figure 3. The bookmark list shows the result of a bookmark query, whereas a tag cloud shows a visual summary of keywords such that more frequently used keywords are shown in larger font sizes and rarely used ones in smaller font sizes. Hence, one can quickly identify recently used and popular keywords from the tag cloud.



Fig. 3. Query Bookmarks by Project

Using the query box, a user can select the desired query option in a drop down list and enter query term(s). The query options supported include:

- *Query on all accessible bookmarks.* This option allows users to view all bookmarks belonging to projects accessible to the users.
- *Query on self created bookmarks.* Users can browse and query all bookmarks he or she has created.
- *Query by project.* This option caters to users who want to query bookmarks created for metadata records within a project. Figure 3 illustrates how a user searches bookmarks by project. In this case, the user first selects the *Search in my project* query option followed by selecting one of the accessible projects. In this example, the user selects the *BTNR Project*. The user finally provides the query term “*litter content*”. Upon submission, the query results are returned as a bookmark list.
- *Query by user group.* As mentioned in Section 5.1, a user can belong to multiple user groups. To query bookmarks created by group members that may be relevant to users’ need, users can choose *Search in my groups* option followed by specifying the group name and query term(s).

6 Conclusion

Social bookmarking is essential to digital libraries. This paper outlines the important design requirements for including social bookmarking in a digital library system. An integration framework consisting of a reference system architecture and a common database integration strategy has been developed. With this framework, a systematic approach to incorporate social bookmarking into digital library systems can be adopted. The paper also describes our experiences

in integrating G-PORTAL digital library system with SCUTTLE, an open source social bookmarking tool.

Looking ahead, one will expect more digital libraries to incorporate social bookmarking to enhance collaboration among their users. There will be case studies of using strategies other than common database to be reported in the future. As part of our future work, we plan to conduct user evaluation on the usability of the G-PORTAL's social bookmarking functions and its impact on learning.

References

1. Agosti, M., Ferro, N., Panizzi, E., Trinchese, R.: Annotation as a support to user interaction for content enhancement in digital libraries. In: Working Conference on Advanced Visual Interfaces (2006)
2. Arko, R., Ginger, K., Kastens, K., Weatherley, J.: Using Annotations to Add Value to a Digital Library for Education. *D-Lib Magazine* 12(5) (2006)
3. Candela, L., Castelli, D., Pagano, P., Thanos, C., Ioannidis, Y., Koutrika, G., Ross, S., Schek, H.-J., Schuldt, H.: Setting the Foundation of Digital Libraries: The DELOS Manifesto. *D-Lib Magazine* 13(3/4) (2007)
4. Dieberger, A., Dourish, P., Hook, K., Resnick, P., Wexelblat, A.: Social Navigation: Techniques for Building More Usable Systems. *Interactions* 7(6), 36–45 (2000)
5. Hammond, T., Hannay, T., Lund, B., Scott, J.: Social bookmarking tools (i), a general review. *D-Lib Magazine* 2(4) (2005)
6. Lagoze, C., Arms, W., Gan, S., Hillmann, D., Ingram, C., Krafft, D., Marisa, R., Phipps, J., Saylor, J., Terrizzi, C., Hoehn, W., Millman, D., Allan, J., Guzman-Lara, S., Kalt, T.: Core services in the architecture of the national science digital library (NSDL). In: Proceedings of the 2nd ACM/IEEE-CS Joint Conference on Digital Libraries, pp. 201–209 (2002)
7. Lauw, H.W., Lim, E.-P., Wang, K.: Summarizing review scores of unequal reviewers. In: SIAM International Conference on Data Mining (2007)
8. Lim, E.-P., Goh, D.H.-L., Liu, Z., Ng, W.-K., Khoo, C.S.-G., Higgins, S.E.: G-Portal: A Map-based Digital Library for Distributed Geospatial and Georeferenced Resources. In: ACM/IEEE-CS Joint Conference on Digital Libraries (2002)
9. Lomas, C.: Seven Things You Should Know About Social Bookmarking, Online (2005) (accessed January 3, 2007)
10. Shipman, F., Hsieh, H., Moore, J., Zacchi, A.: Supporting Personal Collections across Digital Libraries in Spatial Hypertext. In: ACM/IEEE-CS Joint Conference on Digital Libraries (2004)
11. SourceForge.net, <http://sourceforge.net/scuttle>
12. Yew, J., Gibson, F., Teasley, S.: Learning by tagging: group knowledge formation in a self-organizing learning community. In: International Conference on Learning Sciences (2006)