EnTag Project Cover Sheet

Project Information			
Project Acronym	EnTag		
Project Title	Enhanced Tagging for Discovery		
Start Date	1 September 2007	**End Date**	31 October 2008
Lead Institution	UKOLN		
Project Director	Michael Day		
Project Manager & contact details	Koraljka Golub UKOLN, University of Bath, Bath, BA2 7AY tel: +44 (0) 1225 383619 email: k.golub@ukoln.ac.uk		
Partner Institutions	University of Glamorgan; STFC; Intute (MIMAS The University of Manchester) Non-funded supporting partners: OCLC Office of Research, USA; Danish Royal School of Library and Information Science		
Project Web URL	http://www.ukoln.ac.uk/projects/entag/		
Programme Name (and number)	*Repositories and preservation programme*		
Programme Manager	Balviar Notay		

Document Name			
Document Title	*Final report*		
Reporting Period	*n/a*		
Author(s) & project role	Koraljka Golub, Catherine Jones, Marianne Lykke Nielsen, Brian Matthews, Jim Moon, Douglas Tudhope		
Date	14 November 2008	**Filename**	EnTag-D1.6-Final-report.doc
URL	*n/a*		
Access	☑ Project and JISC internal		

Document History		
Version	**Date**	**Comments**
Final	14 November 2008	Submitted to JISC
Final revised	23 February 2009	Revised after feedback

Enhanced Tagging for Discovery (EnTag)

Final report

Authors: Koraljka Golub, Catherine Jones, Marianne Lykke Nielsen, Brian Matthews, Jim Moon, Douglas Tudhope

Contact person: Koraljka Golub

Date: 14 November 2008

Table of Contents

Acknowledgements

This project was funded by the Joint Information Systems Committee (JISC) of the Higher and Further Education, under the Repositories and preservation programme. Thanks to Debra Hiom of Intute for help with recruiting participants, the Intute study participants, and the STFC authors who participated in the evaluation process.

Executive Summary

The EnTag project explored the combination and comparison of controlled and folksonomy approaches to semantic interoperability in the context of repositories and digital collections. The aim was to investigate the effect on indexing and retrieval when using only social tagging versus when using social tagging in combination with a knowledge organization system. Two different contexts were explored: tagging by readers (Intute) and tagging by authors (Science and Technology Facilities Council (STFC)). The major development was that of Intute.

For each of these a separate demonstrator was developed, one operating on data extracted from Intute (Intute 2008), and the other operating over STFC's repository (STFC ePublication Archive 2008) in which tagging was conducted by authors submitting papers to the repository. A user study was conducted for each demonstrator, which allowed a general comparison of a repository versus digital collection context, a different knowledge organization system, interface and user community.

Three major methods to collect user data were log analysis, questionnaires, and interviews. The evaluation of the Intute demonstrator involved comparing basic and advanced system for indexing and retrieval implications. The test setting comprised 28 students in political science and 60 documents covering 4 topics of relevance for the students. Dewey Decimal Classification was used. The STFC study involved 10 authors depositors. The ACM Computing Classification Scheme was used.

The results of the Intute study showed the importance of controlled vocabulary suggestions (to produce ideas of tags to use, to ensure consistency and retrieval, to make it easier to find focus for the tagging, etc.) Furthermore, the value and usefulness of the suggestions proved to be very dependent on the quality of the suggestions. The suggestions must be user-oriented as regards to level of specificity, perspective and currency. Most tags were added by typing them directly in, as common in social tagging applications; of the other features used, the most frequent one was DDC suggestions, and another tagger's cloud. That the participants appreciated the suggestions was also seen from their comments. Both simple tagging and enhanced tagging provided additional entry points (for retrieval) beyond the original indexing. There was some evidence that vocabulary-based suggestions, in particular, provided additional access points beyond the literal text. Most participants claimed that they would be willing to use similar tools in real life.

The results of the STFC study show that there is a general pervading sentiment amongst the depositors that choosing terms from a controlled vocabulary was a "good thing" and in fact better than own terms. The participants could overall see the point of the adding terms for information retrieval purposes, and could see the advantages of consistency of retrieval if the terms used were from an authoritative source. Most claimed that they would be willing to use a tool similar to the one provided, albeit with some reservations and suggestions about the interface. ACM classification was however not seen as good enough for the purposes of this group.

In conclusion, we recommend that social tagging be allowed in the JISC context (e.g., repositories), enhanced with suggestions from a controlled vocabulary. More findings are needed so it is important to further analyze, experiment and pilot test tools derivative from both Intute and STFC demonstrators. It was shown that further developments and improvements are needed in the following major aspects: automated suggestions, controlled suggestions, tag input features such as auto-complete and spelling checking, controlled vocabulary presentation, other controlled vocabularies and the user interface. Detailed recommendations are discussed in Deliverable 5.1: Recommendations briefing paper.

Background

Knowledge organization systems have been used as tools for information discovery and retrieval in libraries and abstracting and indexing services, some for more than a century. Their benefits for improved information retrieval in the digital environment have been well acknowledged and recognized. They have devices to reduce the ambiguity of natural language when describing and retrieving items, and to allow access via browsing and navigation. However, there are costs associated with the use of knowledge organization systems – manual indexing or classification are a significant resource, especially when performed by trained indexers.

Social tagging applications, such as Flickr (2008) and Del.icio.us (2008) with their community-based user interfaces encouraging social tagging activity, currently attract much attention and are seen as key elements of new Web 2.0 services. They hold the promise of reducing indexing costs by drawing end-users into contributing, adding value as part of their interaction with information services. However, social tagging is less concerned with consistency than with making it easier for end-users to describe information items and to have access to other users' descriptions. Existing social tagging applications have not been designed with information discovery and retrieval in mind. The resulting folksonomies (collections of tags) are completely uncontrolled, lacking even basic control of word forms such as spelling variants, synonyms and disambiguation of homonyms (cf. Spiteri 2007; Guy & Tonkin 2006). Many users use tags only to organize own documents, and not to help the community (cf. Tonkin et al. 2008). On the other hand, natural language tags could cover aspects that are not available in a knowledge organization system, especially when it comes to new concepts; as such, they could help update the knowledge organization system.

The need for knowledge organization systems in relation to folksonomies has been reported in the literature. Weller (2007) compares ontologies and folksonomies, suggesting that they are not to be seen as rivals but complement each other. Noruzi (2007) provides seven arguments for why a folksonomy-based system should use a thesaurus, emphasizing that there is no way to maintain consistency over time or across folksonomy users without a thesaurus. For Connotea (2008), a service for organizing references, recently an add-on tool has been developed that allows taggers to select terms from a knowledge organization system (Entity Describer 2007). Smith (2007) explores the connection between folksonomies and Library of Congress Subject Headings (LCSH) and describes advantages and disadvantages of each. She suggests that their product (called LibraryThing for Libraries) may "provide a compromise between the constraints of controlled vocabulary and the relative wilderness of the folksonomy". Hayman (2007) argues for combining the best of the two worlds and describes its application on Australian collection of education resources (education.au). In their collection, knowledge organization systems are used for metadata creation and searching, and in order to keep pace with user needs, folksonomies are being explored. Users can tag resources by choosing from an established taxonomy or by entering their own terms. Users' own terms will be used later to feed back into the taxonomy to improve its quality. The Library of Congress is collaborating with Flickr (2008), in order to enhance bibliographic records for its images by end-user tags (Raymond 2008).

Aims and Objectives

Aims
The project was to investigate the combination and comparison of controlled and folksonomy approaches to semantic interoperability in the context of repositories and digital collections.

Objectives
The project was to:

- Investigate indexing aspects when using only social tagging versus when using social tagging in combination with a controlled vocabulary;
- Investigate above in two different contexts: tagging by readers and tagging by authors; and,
- Investigate influence of only social tagging versus social tagging with a controlled vocabulary on retrieval.

Methodology

The main focus of investigation was the effect of an enhanced tagging system. The enhanced system, with the capability of offering suggestions via a knowledge organization system, was compared against free social tagging. Two different contexts were explored: tagging by readers (Intute) and tagging by authors (Science and Technology Facilities Council (STFC)). For each of these a separate demonstrator was developed, one operating on data extracted from Intute (Intute 2008), and the other operating over STFC's repository (STFC ePublication Archive 2008) in which tagging was conducted by authors submitting papers to the repository.

A user study was conducted for each demonstrator, which allowed a general comparison of a repository versus digital collection context, a different knowledge organization system, interface and user community. In the EnTag Intute demonstrator, the Dewey Decimal Classification (DDC) forms the knowledge organization system, and in STFC the ACM Computing Classification Scheme.

The **Intute** study involved 28 participants who completed the study. They were all politics students at British universities, with one from the European University Institute. A call for participation was put together. They were recruited mainly from a written call for participation throughout the country. Each participant was given 4 tasks, and in each task 15 documents were to be tagged – 60 in total. Each task covered one topic of relevance to the politics student. Two tasks were controlled and two tasks free. In order to reduce the learning influence, tasks were rotated. A hypothetical group project scenario was outlined as a rationale and motivation for the tagging activity (by users as reader/searchers).

In each task the participant was to first search for documents and then tag 15 of them. In controlled task, they were told to choose the top 15 documents, while in free tasks they could choose any documents they found relevant. In the case that a URL had become unavailable, the instruction was to move on to the following document. Tagging instructions specified that tagging each document should on average take between 5 and 10 minutes. They were to describe as many aspects and topics they thought appropriate for the task. They were also reminded to open the URL, but need not follow further internal links within a Web site. In case of very long documents, they were to focus on its abstract, introduction, conclusion, headings and table of contents. Additional instruction was added to tasks for Enhanced Tagger, to try to consider the suggestions from the controlled vocabulary.

Topics for the controlled tasks were suggested by a subject expert, PhD student in politics, who also evaluated whether there were at least 20 documents in the database relevant to the topics. The controlled task for Simple Tagger was on the topic of European integration. The controlled task for Enhanced Tagger was on the topic of peacekeeping.

After signing the participation form and completing a pre-study questionnaire, the Instructions document was sent out. It was the main document that each participant was given. It introduced the study and described each step the participant was supposed to do. Major steps comprised the following:

1) Technical requirements for using the system (with reference to the Settings document)
2) Learning the system (reference to the Training document)
3) Task 1
4) Task 2
5) Task 3
6) Task 4
7) Final questionnaire
8) Email

Before starting the study itself, each participant was given a Training document through which she was to learn the system and try out tagging. The Settings document described how to enable scripting in Internet Explorer and Firefox browsers, and how to zoom the screen display for better viewing in Firefox.

The main method of data collection was logging the steps the participants conducted in the demonstrator. In order to help contextualize and explain the results better, questionnaires were also used. Apart from the pre-study questionnaire for collecting background information about the participants, the participant was to complete a post-task questionnaire, after every task, and a post-study questionnaire, after finishing all the tasks.

The **STFC** study involved 10 participants who are scientifically active and will deposit papers (or records of papers) within the STFC repository. They all worked within one discipline for consistency and ease of comparison, each have published refereed papers, and have deposited a number (> 10) papers within the ePubs repository, and can thus be considered a regular author and depositor. The sample set of authors was by necessity small and thus the results are qualitative rather than statistical. The discipline was that of computer science and information technology, for which there was an easily available controlled vocabulary: the ACM Computing Classification Scheme.

The exercise sought to address a number of questions on the improvement of tagging from an author's perspective:

- What do they feel the purpose of tags is, and how would they like them to be used?
- Whether they consider using controlled vocabularies a worthwhile exercise over and above free tagging. Does the more accurate and consistent tagging achieved merit the overhead of browsing and selecting from a controlled vocabulary? Or does the freedom which free tags offer to allow them to pick their own terms give a more satisfactory result in their terms? (accuracy and efficiency)
- Do they find that the controlled vocabulary makes them consider tags which they would not have otherwise considered? (exhaustivity)
- Do they choose terms at a deep level in the hierarchy? (specificity)
- Is the controlled vocabulary structure (hierarchy, related terms) intuitive and easy to use?
- Is the user interface intuitive and easy to use?
- User satisfaction (is this something which they would like to see provided as a feature of the input system to ePubs?)
- Are there any further tools or uses for tags (controlled or free) which the author would like to see?

The evaluation used the following methodology:

1. An author will be invited to a session of approximately 40 minutes.
2. A worksheet will be provided for the author, with notes and guidance on the task. A copy of this worksheet is provided in Appendix 2.
3. The author will be invited to use the tagging interface to select tags for a number of papers which he or she has authored or co-authored (6 would be a reasonable number). This should take no more than about 20 minutes.
4. The author will be free to select appropriate descriptive tags from any of the three approaches offered (free text, controlled vocabulary, tag cloud) as their preference.
5. At a later point in the trial, the observers may suggest that the authors try a different way of tagging using a different component of the tool in order to gauge their response to a different approach to their normal preference.
6. The tagging system will collect a number of statistics on the session.
    - Tags collected on the paper
    - Length of time to tag a paper.
    - Proportion of free to controlled tags
    - Number of tags selected
    - Depth of hierarchy of tags selected
    - Tags deleted.
7. The session will be observed by a member of the EnTag team, who has a observer guidance sheet (see Appendix 3 for a copy of the notes for observers).
8. The member of the EnTag team will then discuss the exercise with the author to record impressions of the exercise:

- Is a tagging interface something they would use when depositing papers?
- Do they prefer to enter free text tags, or use the controlled tags?
- Is the controlled vocabulary a reflection of how they would categorise their work?
- Is the controlled vocabulary hierarchy intuitive and easy to navigate?
- Is the user interface an appropriate one for tagging?
- Are there any improvements they would like to see (e.g. more use of the tag cloud? Keyword suggestions from abstract?)

The main method of data collection was logging the steps the participants conducted in the demonstrator. The Tagger tool collected information as follows.

- Tags collected on the paper
- Length of time to tag a paper.
- Proportion of free to controlled tags
- Number of tags selected
- Depth of hierarchy of tags selected
- Tags deleted.

The observers looked for user behaviour, whether the papers already had keywords assigned and general observations.

Implementation

For each part of the project, separate demonstrators were implemented. Since the Intute demonstrator was the major development, it is discussed here in more detail. For user studies and their implementation, see the previous section.

1) Intute demonstrator

The EnTag simple social tagging system is based upon ideas used in other Social Tagging Websites. Intute provided part of their Social Sciences database which the users can search to find abstracts of web resources that might interest them. The URL to the web resource is provided so that the user can view the actual document. The EnTag software permits a user to tag Intute resources. The EnTag Enhanced social tagging system provides the same capabilities as the Simple system with the additional feature of automatic tagging suggestions based on the Dewey Decimal Classification (DDC). Visual Studio, C# and ASP .NET were used for the development.

Object Oriented (OO) Analysis was employed for the original architectural design but this was not found very helpful for this software application. OO analysis derives classes from real world objects whereas web applications are largely based around building web pages. Furthermore, web applications are not really OO in nature; very little state information is held and many of the classes are really just containers for functionally decomposed solutions. In addition the state information is for the most part held in Session variables which have global scope and destroys the OO principle of encapsulation.

The application design is expressed in UML. The Use Case model was a valuable aid to determining the user requirements and sequence and activity diagrams have been used to design the algorithms employed. The database design was achieved using conventional E-R modelling but expressed in UML. The data used was extracted from the formats that were provided by our partners at Intute and OCLC.

The prototype proved extremely robust; no errors were reported in the EnTag software for the duration of the user study. Some users did experience a few problems with the dated hardware and software configurations they were using. Each of the reported problems was resolved within a few hours, although, from one or two user comments, it appears that some configuration problems were not reported. A modern social tagging system requires quite a high bandwidth and the relatively complex interface of the EnTag system needs a reasonably high screen resolution.

Please refer to a separate document “Software Report EnTag system” for more details.

2) STFC demonstrator

The STFC demonstrator was an Apache Cocoon application using combined Java and XML techniques; the underlying database was Oracle. This linked dynamically to the ePubs intuitional repository so that the user did a search in the repository and in a specially adapted edit mode, there was an option to enter the tagger system. The tagger application was then started and the title and abstract of the work was transferred together with any existing free tags. The controlled vocabularly was imported into SKOS and then into the tool. There is the potential to have more than one controlled vocabulary within the tool.

Outputs and Results

1) Intute demonstrator (D.3.1)
Link to Intute demonstrator will be on EnTag website.

A description of the Intute demonstrator follows. There are three major interfaces: searching, basic tagging, and enhanced tagging. The searching interface offers a tag cloud with tags linked to documents that they index, names of taggers linked to documents they indexed, as well as a free-text search box where searching can be limited to tags, title and description fields. Once a document is selected from search results, a tagging interface appears. Two types of tagging interfaces are developed: the basic, with tagging features usual in popular tagging services, and the enhanced one, with options from the knowledge organization system. Both tagging interfaces have the following options from which to select tags: a global tag cloud (an alphabetical list of all tags in the system, with different font sizes relative to popularity); tags assigned by a specific user; and list of own tags. A user may also type in a tag. The enhanced interface (Fig. 1) involves suggestions from the knowledge organization system, presented in three frames. In the first frame, a list of DDC classes are automatically suggested based on matches with a tag entered by the user (by any means). If the user selects a class its narrower and broader classes are shown in the second frame, allowing interactive browsing of the hierarchical context. Simultaneously, in the third frame a tag-cloud like list of DDC captions, relative index and LCSH mapped terms is presented as a source of suggestions from which the user may optionally select. While the current Intute implementation is DDC specific, the techniques could be generalized to employ elements such as synonym sets and related concepts from, for example, any SKOS encoded vocabulary.

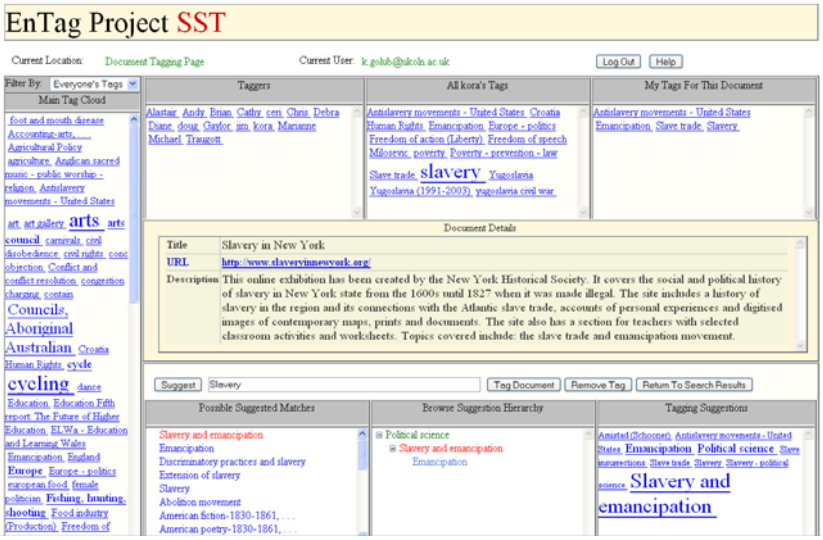

FIG. 1. Enhanced tagging interface.

An example of how enhanced tagging is implemented follows. After searching for the term “slavery” in all fields, a list of documents is returned. We choose the document with title “Slavery in New York”, an online exhibition on history of slavery in New York State (http://www.slaveryinnewyork.org/). We find the document useful and want to tag it. By clicking on “Tag” button, the enhanced tagging page opens. We enter tag “slavery” and click on “Suggest” button to get suggestions based on DDC terminology. The first frame (Possible Suggested Matches) then displays matching classes, which in our example includes:

Slavery and emancipation
Emancipation
Discriminatory practices and slavery
Extension of slavery
Slavery
Etc.

Since the document is about the slave trade and emancipation movement, we click on “Slavery and emancipation”, which results in the second frame (Browse Suggestion Hierarchy) as shown in Fig 1, with “Slavery and emancipation” highlighted, and its broader and narrower classes. The third frame lists optional Tagging Suggestions. The terms chosen in this scenario are:

Antislavery movements - United States
Emancipation
Slave trade
Slavery

Thus, from the initial “slavery” tag, three other tags are derived via the knowledge organization system.

2) STFC demonstrator (D.3.2)
The Tagger interface is supplied in conjunction with the ePubs metadata editing tool so that tags can be entered for a specific publication by its authors. The figure below (Fig. 2) shows a screen shot of a typical tagging screen.

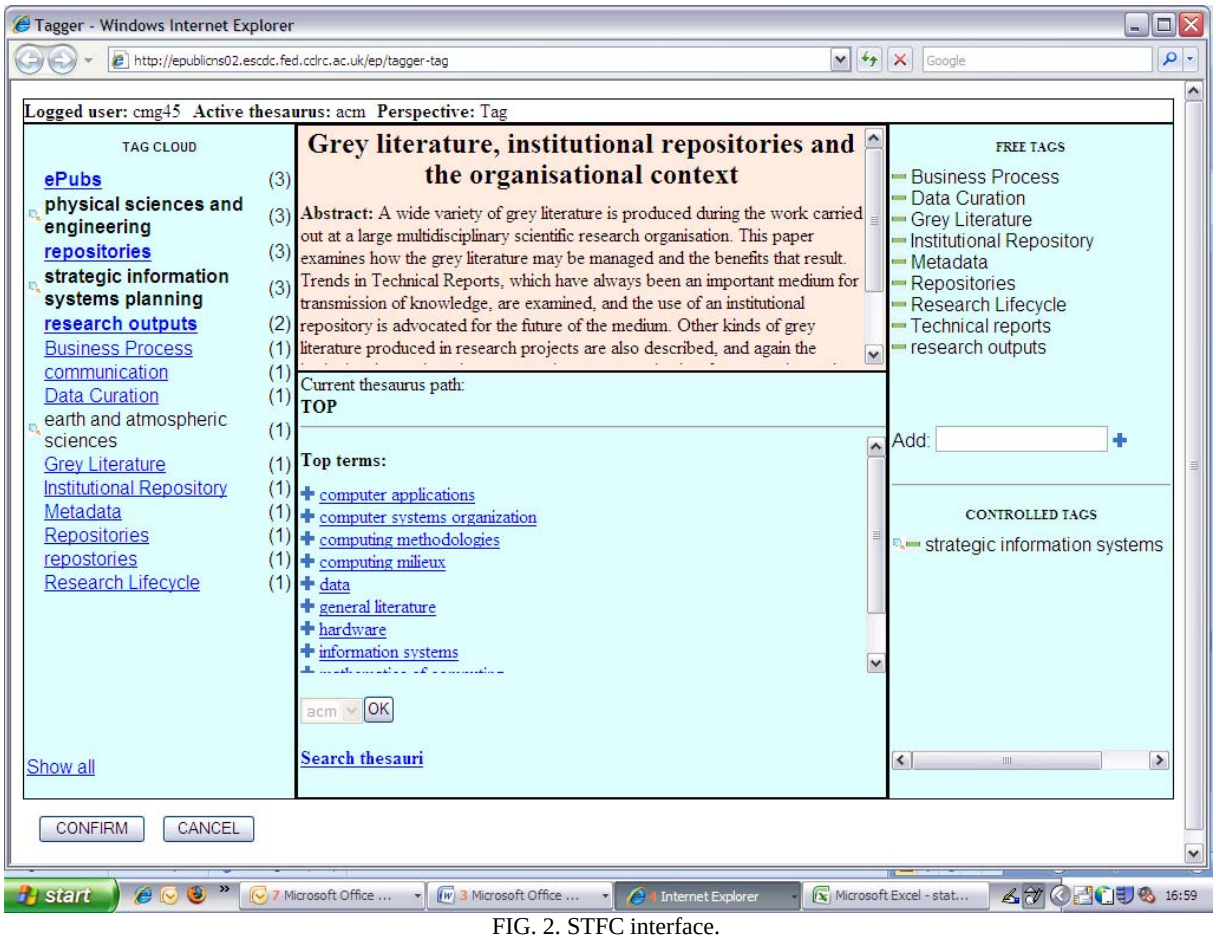

FIG. 2. STFC interface.

The screen is divided into four main areas.

1. At the top-centre, the title and abstract of the publication selected for tagging is displayed.
2. At the bottom-centre, a browse interface for the thesaurus is shown. The chosen thesaurus can be changed by choosing from a drop down menu (only the ACM scheme is available for the trial). This has the following features
    a. The top-level terms are shown initially,
    b. The hierarchy can be expanded by clicking on a term (shows narrower terms and related terms)
    c. The current path to the top of the hierarchy is shown as a "breadcumb" trail along the top of the hierarchy. This can be used to backtrack.
    d. Terms can be selected for adding to the "controlled tags" list by clicking on the "+" symbol to the left of each term.
    e. The "search thesauri" link brings up a window where free-text terms can be entered which allow the thesuarus to be searched.
3. To the left a "tag-cloud" is displayed, ranked in order of frequency of use of the tag. These can be selected by either clicking on them (if they are free-text terms) or clicking on the spyglass symbol displated to their left (if the arise from the controlled vocabulary). This will enter them into the "free-text" add term box on the right of the screen, where they can be accepted as a tag for this paper. The tag cloud as a default shows only the terms used by this particular author. By clicking on "show all", the tag cloud of all authors can be displayed.
4. To the right, the current selected free-text (above) and controlled vocabulary terms (below) are shown. These can be deselected by clicking on the "-" sign to their left. In the centre of this panel, there is a free-text box, which the user can enter free-text terms and click on the "+" symbol to the right to enter them as tags for this paper. Multiple terms can be entered at once by separating them with commas.

Once a suitable selection of tags has been made using the tool, the user can accept them by clicking on the "Confirm" button at the bottom of the screen.

Currently the Demonstrator is behind the STFC firewall and is not publically available. The service itself runs via a Tomcat servlet. Although the code has some dependencies on the ePubs itself, it could be released as a separate package. STFC are considering whether to develop the system further for use in within the ePubs production system.

3) Intute evaluation analysis report (D.4.1)
Available as separate document.

4) STFC evaluation analysis report (D.4.2)
Available as separate document.

5) Recommendations briefing paper (D.5.1)
Available as separate document.

Outcomes

The project showed the importance of social tagging supported by controlled vocabulary suggestions, both at the time of tagging (indexing) and retrieval. We recommend that social tagging be allowed in the JISC context (e.g., repositories), supported by controlled vocabularies.

The results show the importance of controlled vocabulary suggestions (to produce ideas of tags to use, to ensure consistency and retrieval, to make it easier to find focus for the tagging, etc.).

Furthermore, the value and usefulness of the suggestions is very dependent on the quality of the suggestions.

Conclusions

The results of the Intute study show the importance of controlled vocabulary suggestions (to produce ideas of tags to use, to ensure consistency and retrieval, to make it easier to find focus for the tagging, etc.). Furthermore, the value and usefulness of the suggestions is very dependent on the quality of the suggestions. The suggestions must be user-oriented as regards levels of specificity, perspective and currency. Most tags were added by typing them directly in, as common in social tagging applications; of the other features used, the most frequent one was DDC suggestions, and a certain tagger's cloud. That the participants appreciated the suggestions was also seen from their comments. In general it was found that adding tags by users, either with or without suggestions from a controlled vocabulary, increases the potential for finding documents. Most participants claimed that they would be willing to use similar tools in real life.

The results of the STFC study show that there is a general pervading sentiment amongst the depositors that choosing terms from a controlled vocabulary was a "good thing" and in fact better than own terms. The subjects could overall see the point of adding the terms for information retrieval purposes, and could see the advantages of consistency of retrieval if the terms used were from an authoritative source. Most claimed that they would be willing to use a tool similar to the one provided, albeit with some reservations and suggestions on the interface. ACM classification was not seen as good enough for the purposes of this group.

Please refer to Intute study report (EnTag-D4.1-Intute-study), STFC study report (EnTag-D4.2-STFC-study), and recommendations (EnTag-D5.1-recommendations) for further details.

Implications

Both free tagging and vocabulary-based tagging can potentially serve to add access points. There is evidence from the Intute study that automatic suggestions of vocabulary-based tagging have potential to offer additional access points beyond the literal text and thus can enhance access compared to free text search engines.

Further work is needed related to the Intute study. A related point for further research would be to see whether the user facets are different from those assigned by librarians in the Intute database. Similarly, with the STFC study, further research is required to see whether subject librarians would chose tags differently to authors, and develop processes so that authors and subject librarians could assist each other in the selection of tags.

Most participants from both studies claimed that they would be willing to use similar tools in real life. This can be applied in both repository contexts and collections, such as Intute. Given the patchy distribution of coverage in any single university repository today, some form of known item search or author-based search may be the most likely current option. Subject-based access would be highly desirable for various types of aggregated repositories in the future.

Recommendations

We recommend that social tagging be allowed in the JISC IE context, enhanced with suggestions from a controlled vocabulary. More findings are needed so it is important to further analyze, experiment and pilot test tools derivative from both Intute and STFC demonstrators. It was shown that further developments and improvements are needed in the following major aspects: automated suggestions, controlled suggestions, tag input features such as auto-complete and spelling checking, controlled vocabulary presentation, other controlled vocabularies, and user interface.

Please refer to the recommendations document (EnTag-D5.1-recommendations) for further details.

References

Connotea. (2008). Retrieved March 19, 2008, from http://www.connotea.org/.

Del.icio.us. (2008). Retrieved March 19, 2008, from http://del.icio.us/.

Entity Describer. (2007). Retrieved March 19, 2008, from http://www.connotea.org/wiki/EntityDescriber.

Flickr. (2008). Retrieved March 19, 2008, from http://www.flickr.com/.

Guy, Marieke; Tonkin, Emma. (2006). Folksonomies: Tidying up tags? *D-Lib Magazine, January 2006*. Retrieved March 19, 2008, http://www.dlib.org/dlib/january06/guy/01guy.html.

Hayman, Sarah. (2007). Folksonomies and tagging: New developments in social bookmarking. *Ark Group Conference: Developing and Improving Classification Schemes 27-29 June, Rydges World Square, Sydney.* Retrieved March 19, 2008, from http://www.educationau.edu.au/jahia/webdav/site/myjahiasite/shared/papers/arkhayman.pdf.

LibraryThing. (2008). Retrieved March 19, 2008, from http://www.librarything.com/.

Noruzi, Alireza. (2007). Folksonomies: Why do we need controlled vocabulary? Editorial. *Webology, Vol. 4, No. 2*. Retrieved March 19, 2008, from http://www.webology.ir/2007/v4n2/editorial12.html.

Raymond, Matt. (2008). My friend Flickr: A match made in photo heaven. *Library of Congress Blog*. Retrieved March 19, 2008, from http://www.loc.gov/blog/?p=233.

Smith, Tiffany. (2007). Cataloging and you: Measuring the efficacy of a folksonomy for subject analysis . In Lussky, Joan, Eds. *Proceedings 18th Workshop of the American Society for Information Science and Technology Special Interest Group in Classification Research*, Milwaukee, Wisconsin. Retrieved March 19, 2008, from http://dlist.sir.arizona.edu/2061/.

Spiteri, Louise F. (2007). Structure and form of folksonomy tags: The road to the public library catalogue. *Webology, Volume 4, Number 2*, June, 2007. http://www.webology.ir/2007/v4n2/a41.html

STFC ePublication Archive. (2008). Retrieved March 19, 2008, from http://epubs.cclrc.ac.uk/.

Tonkin, Emma; Corrado, Edward M.; Moulasion, Heather Lea; Kipp, Margaret E. I.; Resmini, Andrea; Pfeiffer, Heather D.; Zhang, Quiping. (2008). Collaborative and social tagging networks. *Ariadne 54*. Retrieved March 19, 2008, from http://www.ariadne.ac.uk/issue54/tonkin-et-al/.

Weller, Katrin. (2007). Folksonomies and ontologies: two new players in indexing and knowledge organization. *Online Information 2007*, 108-115.