



Recommendations

EnTag D5.1: Recommendations briefing paper

Document details

Author:	Koraljka Golub, Catherine Jones, Marianne Lykke Nielsen, Brian Matthews, Jim Moon, Douglas Tudhope
Date:	23 February 2009
Version:	Final
Document Name:	EnTag-D5.1-recommendations.doc
Notes:	Revised after JISC feedback

Summary

This document lists recommendations for JISC derived from the EnTag project. For other deliverables and related documents, please visit <http://www.ukoln.ac.uk/projects/enhanced-tagging>.

Acknowledgements

This research was funded by the Joint Information Systems Committee (JISC) of the Higher and Further Education.

Table of contents

1	Summary	3
1.1	Intute conclusions	3
1.1.1	Intute study recommendations	4
1.2	STFC conclusions	4
1.2.1	STFC recommendations	5
1.3	Comparison of Intute – STFC findings	5
1.4	Improvements to both systems	6
1.4.1	Further analyze, experiment and pilot test tools derived from both Intute and STFC demonstrators	6
1.4.2	Ongoing detailed qualitative analysis of Intute study data	6
2	General recommendations	6
2.1	Retrieval implications for JISC IE	6
2.2	Investigate feasibility of evaluating operational versions of both systems	6
2.3	Investigate different types of vocabularies and different domains	7
2.4	Augment existing vocabularies from tagging data	7
3	Technical recommendations	7
3.1	Automated suggestions as options in a tagging tool	7
3.2	Auto-completion	7
3.3	Integrate automatic classification in suggestions	7
3.4	Retrospective system ‘improvement’ of user tags	7
3.5	Improve tag cloud functionality	8
3.6	Improve controlled vocabulary presentation	8
4	User study recommendations	8
4.1	Further study of user behaviour	8
4.2	Further study of user motivation for tagging within JISC IE	8
4.3	Educational implications for JISC IE	9
5	References	9

Recommendations

This report forms part of the deliverables for the EnTag (Enhanced Tagging for Discovery) project, funded by JISC as part of the Repositories and Preservation Capital Programme. The aim of EnTag was to investigate the combination and comparison of controlled and folksonomy approaches to semantic interoperability by developing and evaluating two demonstrators that combined social tagging facilities with the resources of a controlled vocabulary. The project explored two communities of use: at Intute, focusing mainly on tagging by readers (users annotating resources with tags); and at STFC¹, focusing mainly on tagging by authors (when they deposit in the repository).

For details of the functionality of the two demonstrators and the user studies, see the report on each study and also the EnTag final report. All conclusions should be seen as somewhat provisional, bearing in mind the relatively limited resources available for the work and user studies. Further studies of use in operational contexts are recommended.

1 Summary

This document first briefly summarises and compares overall conclusions from the two parts of the project. It then goes on to make recommendations within the JISC IE context of repositories and digital collections.

1.1 Intute conclusions

Both simple tagging and enhanced tagging provided additional entry points beyond original indexing. There was some evidence that vocabulary-based suggestions, in particular, provided additional access points beyond the literal text.

User experience and task completion showed that both Simple and Enhanced Demonstrators were usable with little prior training. However comments showed that the interface, particularly in the Enhanced system was experienced as complex. By design the interface was cluttered, as we wished to test a variety of tagging features. An operational system should have a simpler, less cluttered user interface, focusing on the key functionality and with user interaction streamlined. Appreciation was shown by users of the need for consistency.

The global tag cloud (all tags) was little used and had mixed response in the questionnaires. Part of the problem was excessive scrolling due to the size of the cloud. This could be reduced by filtering the cloud on some criteria. However, it is not clear that displaying all global tags is useful for primarily retrieval purposes with large collections. Much more useful would be a single document (all user) tag cloud. Community-of-use tag clouds (or tag cloud filters), based on friends, communities or work groups would be another possibility.

While users appreciated the 'direct' suggestions and made some use of the disambiguation interface element, the analysis shows that they did not browse the Dewey hierarchy very much. Further work is needed to explore when browsing functionality is desirable in this context.

The suggestion facility was used. Generally the comments reported that the enhanced suggestions were sometimes useful but sometimes very wide of the mark and not helpful. There was evidence of support for automated (optional) suggestions, provided that suggestions are high quality and oriented to user. There are various possibilities to improve the suggestions (see Intute study report).

The analysis comparing the original Intute indexing with the tags added from both Simple and Enhanced showed little overlap. This suggests that tagging (both Simple and Enhanced) can provide additional access points to conventional indexing. The analysis comparing user tags to the document content shows slightly fewer tags per each element assigned in the Enhanced Tagger are found in the documents, compared to the Simple Tagger. When purely DDC-based

¹ CCLRC, the original EnTag partner, merged with PPARC to form STFC in April 2007.

tags are isolated, the overlap is much lower. This suggests that when used in combination with a free text search engine, vocabulary-based suggestion tags may be likely to provide more additional access points for retrieval than simple tagging – where they are used, they may offer significant ‘value for money’. Preliminary examples from initial qualitative analysis suggest that the vocabulary-based suggestions may prompt taggers to escape the literal text. They may potentially both encourage the description of resources by more facets (increased exhaustivity) and the vocabulary-based suggestions may also afford the capability of describing resources at a higher level of generalisation (the activity of classification). In other words, the suggestions may encourage both indexing and classification.

One element of feedback on EnTag at DC 2008 (concerned the possibility that non-subject (non-topical) tags might prove useful for general retrieval purposes, in some contexts. For example, tags might express the genre or utility of a document for a user’s purposes – see also Golder and Huberman (2006), Kipp (2006), and Kipp and Campbell (2006). While much of this tagging activity is for personal bookmarking purposes, some of it may have retrieval potential for third party retrieval.

1.1.1 Intute study recommendations

- Conduct further evaluation of a revised version, including a retrieval test
- Conduct further, longitudinal user studies
- Conduct further studies of tags to identify what topical facets are represented in the tags
- Investigate potential of ‘non-topical’ user tags
- Develop a streamlined enhanced system with improved interface and auto-completion
- Investigate vocabulary browsing interfaces and behaviour in tagging context
- Refine matching of user tags with vocabulary
- Improve quality of automatic suggestions
- Widen scope of suggestions, investigate incorporating automatic classification
- Develop provision for other vocabularies, develop import facility and underlying vocabulary web services
- Develop) a SKOS based vocabulary service (building on STAR project work and W3C standardisation work)

Please see the Intute study document (EnTag-D4.1-Intute-study) for more details.

1.2 STFC conclusions

There was 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 terms for information retrieval purposes, and could see the advantages of consistency of retrieval if the terms used are from an authoritative source. Most subjects claimed that they would be willing to use a tool similar to the one provided, albeit with some reservations about how this could be realised in practice, and proposed suggestions on the interface, and additional automation.

The suitability of the proposed controlled vocabulary was also an issue with most subjects. While they recognised that this was a well-known vocabulary used in computer-science publishing, they questioned its usefulness to accurately index their work for retrieval. Several wanted choices of vocabularies. Another possibility here might be to help develop vocabularies.

The Tag cloud was not a success. Most did not use it, and those that considered it either found it confusing to use or presenting too many options. It is interesting to speculate why this may be. Whilst an experienced group of IT professionals, they also claimed not to be major users of tags in most applications, and perhaps a larger sample could determine whether different groups with different types of experience would use the tag cloud more. However, there were

also clear problems with the realisation of the tag cloud concept in the STFC tool, and a revised interface may have more potential to be used more effectively.

Most depositors had a strong preference for the way they interact with the system, and how they used the variety of tools on offer, and have a clear preference not to use it in other ways. We could identify three main groups:

- a. Free text taggers: enter many free-text terms, and don't care about or use the controlled vocabulary.
- b. Thesaurus Browsers: Carefully and systematically browse the hierarchy of the controlled vocabulary, and only enter free-text term when the controlled vocabulary did not have a term they were comfortable with.
- c. Thesaurus Searchers: prefer to interact with the controlled vocabulary via a search tool, then move to browsing and only enter free-text term when the controlled vocabulary did not have a term they were comfortable with.

1.2.1 STFC recommendations

- Conduct larger studies of users across a wider range of disciplines, especially those not involved in Information Technology.
- Investigate whether increased familiarity with social networking tools changes the style of the interaction.
- Support alternative modes of interaction within the same system, to support different styles, while trying to avoid a too cluttered interface.
- Investigate the provision of more powerful tools for automatic assistance.
- Provide alternative vocabularies for different disciplines, preferably open vocabularies, using the SKOS format to encourage interchange and reuse.
- Allow cross-searching of free-text and controlled vocabularies for suggestions.
- Further investigate the use of Tag Clouds, considering issues of *Searching*, (especially integrated searching with other controlled vocabulary), *Personalisation* (to individuals and groups), *Structure* (so that the tag cloud can be systematically navigated) and *Presentation* (so that the Tag Cloud can be used more in a more visually appealing manner).
- Investigate the use of tag clouds to support the development of specialist or community vocabularies (*structured folksonomy*).
- Provide a more graphical/visual representation of the controlled vocabulary to assist navigation.

1.3 Comparison of Intute – STFC findings

The two subject groups of depositors (STFC) and searchers/readers (Intute) clearly have different roles (though can overlap). Also regular depositors tended to be more mature than the searchers, due to more experience within a discipline being required before authoring papers. Nevertheless, number of similarities between the Intute and STFC users can be identified.

- In both cases the Global tag cloud proved problematic to use effectively.
- In both cases the user interface proved important along with the visual presentation and interaction sequence

- In both cases the quality and appropriateness of the controlled vocabulary proved to be important.
- In both cases there was evidence of support for automated suggestions if they are appropriate and relevant.

Users also appreciated the benefits of consistency and vocabulary control and were potentially willing to engage with the tagging system if clear benefits to the individual were seen.

1.4 Improvements to both systems

Both systems would require further work before they could be employed in any operational setting. The various recommendations to improve the performance of both demonstrators should be pursued (see above and the Intute and STFC reports for details).

1.4.1 Further analyze, experiment and pilot test tools derived from both Intute and STFC demonstrators

Both demonstrators and studies indicated that choosing terms from a controlled vocabulary was potentially beneficial. Further testing of new versions of the demonstrators based on the recommendations in this document, including a retrieval test, are needed.

1.4.2 Ongoing detailed qualitative analysis of Intute study data

Work is ongoing on more detailed qualitative analysis of the Intute study log data and this will be published in due course.

2 General recommendations

2.1 Retrieval implications for JISC IE

Both free tagging and vocabulary-based tagging can potentially serve to add access points compared to current indexing.

There is some evidence from 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. 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 and IRS (Intute Repository Search). 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. However, subject-based access would be highly desirable for various types of aggregated repositories, for example IRS and other future possibilities.

2.2 Investigate feasibility of evaluating operational versions of both systems

Initial evaluation of both systems suggests the potential for further study in operational settings. Revised/improved versions of both systems could be investigated in the Intute and STFC settings respectively. Extended (longitudinal) evaluation in live settings would give information of such systems in real life. It would also be possible to focus on particular aspects of each system in the short term.

2.3 Investigate different types of vocabularies and different domains

Even though Dewey Decimal Classification and Library of Congress Subject Headings proved useful, other controlled vocabularies should be experimented with for the same and other purposes. ACM was not seen as good enough for the purposes of the STFC group. Either the coverage was not seen as adequate, or it was seen as too theoretical, or not up to date.

Both studies employed classification systems. Each type of controlled vocabulary is suited for a different purpose and usage group. A different type of vocabulary, such as a thesaurus, may have different effects as a basis of tagging suggestions. Furthermore, the purpose of the particular vocabulary offered should be made clear to the tagger (e.g. classification, information retrieval).

Such developments could be linked to any future JISC terminology (vocabulary) registry (see TRSS project report).

2.4 Augment existing vocabularies from tagging data

There is significant potential to augment the entry vocabularies of the controlled vocabularies (Knowledge Organization Systems) used in a study where successful mapping of user tags has occurred. Vocabulary terminology could potentially be modernised and informal user terms introduced.

3 Technical recommendations

3.1 Automated suggestions as options in a tagging tool

Both studies, particularly Intute but also comments from STFC users, provided support for the potential of automated suggestions in tagging tools. If interaction can be streamlined and if the suggestions are seen as high quality then such utilities may be seen as useful additions to tagging interfaces. Suggestions can serve to encourage consistency and also to introduce new angles on topics to tag. Suggestions should be user-oriented as regards terminology, level of specificity, perspective and currency.

3.2 Auto-completion

An auto-complete feature would allow an easier entrée to a suggestion facility. It would also offer an option to shorten the sequence of interaction steps in choosing suggestions. This rich system support could include keyword proposal, dictionary lookup, interactive term disambiguation and visualizations.

3.3 Integrate automatic classification in suggestions

Suggestions can be based on automatic classification of document content, in addition to user topics. In this case, the match is between a given document and the underlying vocabulary. The Intute demonstrator implemented a very crude form of automatic classification, in that the title of a document selected for tagging was fed through to the DDC matching system. The top ranked match yielded suggestions which automatically appeared in the Suggestion tag cloud. In some cases, this worked remarkably well. In many cases, it did not. However, a more sophisticated automatic classification system would in our view be a very useful source of suggestions. There have been various automatic classification and indexing projects recently and there are now tools available that can be applied.

3.4 Retrospective system 'improvement' of user tags

After initial user tagging, vocabulary based improvements could be applied (e.g. correct misspelling, specify the language, treat compounds properly and consistently, link between synonyms, create partial hierarchies, create facets).

3.5 Improve tag cloud functionality

Tag cloud search and browse functionalities could be improved via advanced clustering, exploring co-occurrence, other aggregations, filters, ranking, personalisation and visualization supporting navigation.

3.6 Improve controlled vocabulary presentation

In the STFC study, a significant minority wanted a more graphical/visual representation of the controlled vocabulary than supplied – probably to help them navigate. Some of the participants wanted more information about the semantic meaning of a term in the controlled vocabulary and even other people's tags.

In the Intute study the middle pane for browsing DDC hierarchy was not used very much, and reasons for need to be further explored. The best presentation of vocabulary context and hierarchies need be considered.

Further study whether it is useful to structure the suggestions so the structure of the vocabulary can yield a faceted, structured tagging/check list suggestions.

4 User study recommendations

4.1 Further study of user behaviour

Need for more extended study in operational settings (logging very useful)

Need for longitudinal study to observe behaviour over time

Need for focused study with appropriate performance measures of retrieval effectiveness (complex as indexing-search-evaluation-motivation interlinked)

Need for further investigation of user styles and types of tagging behaviour

In the STFC study three user groups were recognized, two of which make use of the controlled vocabulary: 1) those who enter many free-text terms, and don't care about or use the controlled vocabulary, 2) those who carefully and systematically browse the hierarchy of the controlled vocabulary, and only enter free-text term when the controlled vocabulary did not have a term they were comfortable with, and 3) those who prefer to interact with the controlled vocabulary via a search tool, then move to browsing and only enter free-text term when the controlled vocabulary did not have a term they were comfortable with.

4.2 Further study of user motivation for tagging within JISC IE

Both studies provided some evidence that users would consider using tagging tools if they were confident that personal benefits would follow. There also were indications that a suggestions facility could increase potential take up, provided that the suggestions were (mostly) relevant/useful.

A number of participants in both studies were conscious of the role of consistency in retrieval. In the Intute study, questionnaire results showed that a majority appreciated the suggestions and would consider using a similar system in real life. While 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 in the Enhanced Tagger (bearing in mind that users were encouraged to consider the suggestions if appropriate in their instructions).

For users as searchers/readers in the JISC IE, the rationale is less straight forward than with authors. In some situations, where a user is part of a natural community engaged in a joint endeavour (as in the scenario suggested in the study), tagging content will serve for mutual benefit. In some cases, users may be motivated to act as good (informed) citizens and tag based on their desire to share expertise or enthusiasm. The examples given by the numerous wiki and blogging applications suggest a willingness to orient and contribute (via tagging) to a

collaborative Web 2.0 framework. In educational settings, this could be explicitly part of the pedagogical process.

In many of the popular Web 2.0 applications, we may also argue that social tagging occurs as an extension of personal bookmarking activity. One possibility for applications such as Intute, in the JISC IE context, is to consider social tagging, as an extension of personalisation facilities. MyIntute currently offers a tagging interface for a user's personal tags only. It would be interesting to explore whether a personalisation tagging service augmented with automatic suggestions would draw users into tagging activity generally.

We recommend that further investigation of the possible rationale for tagging by different types of users be conducted in the JISC IE context. One aspect of this could be a consideration of the different types of tagging activity. For example, tags might express the genre or utility of a document for a user's purposes. To the extent that others share the same perspective, non-subject based tags might serve as useful access points for others, in addition to their potential in personalising access to a collection.

4.3 Educational implications for JISC IE

Another promising application area is investigating tagging activity with an explicitly pedagogical focus. Thus an extended EnTag could be used as part of a JISC IE/eLearning project to study the educational benefits of participatory tagging and annotation activity. This is related to the recommendations on study user tagging behaviour and user motivation for tagging. If feasible, tagging activity could be prescribed as one of the learning activities in a particular setting.

5 References

Golder, S A & Huberman, B A (2006). The structure of collaborative tagging systems. *Journal of Information Science*, 32 (2), 198-208.

Kipp, M E I (2006). Complementary or Discrete Contexts in Online Indexing: A Comparison of User, Creator and Intermediary Keywords. *Proceedings Canadian Association for Information Science*, York University, Toronto, Ontario, Canada

Kipp, M E I & Campbell, D G (2006). Patterns and inconsistencies in collaborative tagging systems. Paper presented at ASIS&T Annual Meeting, November 7, Austin, TX.