JISC

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	*Completion 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-D1.7-Completion-report.doc
URL	*n/a*	
Access	☑ Project and JISC internal	

Document History		
Version	**Date**	**Comments**
Final	14 November 2008	Submitted to JISC

JISC

EnTag Completion Report (2008)

Project Sign-off

1. Project Outputs

Please refer to the recommendations document (EnTag-D5.1-recommendations) for detailed outcomes and recommendations.

The following project outputs have been agreed on and submitted to JISC:
D1.1 Formal project plan
D1.2 Consortium agreement
D1.3 Project Website
D1.4 Progress report
D1.6 Final report
D1.7 Completion report
D4.1 Intute evaluation analysis report
D4.2 STFC evaluation analysis report
D5.1 Recommendations briefing paper

Deliverable **D1.5 Interim final report** has not been submitted as no need was recognized for it as the final report was due a month later.

Two remaining deliverables, **D.3.1 Pilot Intute demonstrator** and **D.3.2 Pilot STFC demonstrator** are software packages and we are waiting to hear from JISC how to best deliver them.

Other outputs listed in the workpackage document, such as

- specifications for demonstrators' interfaces,
- evaluation methodology for both studies, and,
- questionnaires,

are integrated in three of the above reports: **D1.6 Final report**, **D4.1 Intute evaluation analysis report**, **D4.2 STFC evaluation analysis report**, and **D5.1 Recommendations briefing paper**.

Formal publications are available at the project's Web site and at least several are planned.

2. Intellectual Property Rights

The STFC demonstrator code is public domain, although it has some connection with some ePubs code (which is not). The ACM classification is available freely for "personal or classroom use" and not for profit purposes.

EnTag
Final draft submitted to JISC
k.golub@ukoln.ac.uk
14 November 2008

3. Project Staff

Michael Day, Project Director, 0.0 FTE
Koraljka Golub, Project Manager/Research Officer, 0.56 FTE
Sally Lewis, Project Administrator, 0.0 FTE
Douglas Tudhope, Intute Glamorgan demonstrator leader, 0.1 FTE
Jim Moon, Software developer, 0.7 FTE
Debra Hiom, Intute data and users coordinator, 0.0 FTE
Chris Stephens, Intute software developer, 0.16 FTE
Brian Matthews, STFC demonstrator leader, 0.2 FTE
Catherine Jones, SFTC user study leader, 0.1 FTE

Non-funded supporting partners:
Diane Vizine-Goetz, Supporting officer for DDC
Andrew Houghton, Software supporter for DDC
Marianne Lykke Nielsen, Research officer

4. Dissemination Plan

Publications and speeches

Unless a URL is given, all publications are available at the project Website,
http://www.ukoln.ac.uk/projects/enhanced-tagging/dissemination/.

Golub, K; Tudhope, D; Lykke Nielsen, M; Moon, J (2008) EnTag: Enhanced Tagging for Discovery.
Dublin Core Special NKOS session, 24 September 2008. [ppt]

Matthews, B; Golub, K; Jones, C; Moon, J; Lykke Nielsen, M; Tudhope, D (2008) Enhancing social
tagging with a knowledge organization system. ALISS Summer Conference 2008. [ppt]

Golub, K; Jones, C; Lykke Nielsen, M; Matthews, B; Moon, J; Tudhope, D (2008) Enhancing social
tagging with a knowledge organization system. ALISS, Vol 3, No 4, July 2008, pp. 13-16. [manuscript
PDF]

Golub, K (2008) Information science and digital information management research. Presentation with
emphasis on EnTag at University of Bath's Bridging the Gap seminars, 8 May 2008, Bath
URL: http://www.bath.ac.uk/research/bridging/docs/seminars/KoraGolub.pdf

Golub, K; Jones, C; Lykke Nielsen, M; Matthews, B; Moon, J; Tudhope, D (2008) EnTag. Presentation
at JISC MDR SIG, 12 February 2008, Birkbeck
URL: http://wiki.cetis.ac.uk/12th_February_2008%2C_Birkbeck

Golub, K (2008) Talk given for Japanese visitors as part of a Repositories meeting at UKOLN, 23 Jan
2008 Presentation: [PDF]

Tudhope, D (2007) Problems of interoperability. Keynote at
The Challenge of the Electronic Environment to the Organization of Knowledge - Second International
Seminar on Subject Access to Information
Helsinki, Finland, 29-30 November 2007
URL: http://tds.terkko.helsinki.fi/dspace/handle/10250/132229

EnTag
Final draft submitted to JISC
k.golub@ukoln.ac.uk
14 November 2008

Publicity

1) CETIS workshop

-------- Original Message --------
Subject:RE: Invitation to connect on LinkedIn
   Date: Tue, 4 Mar 2008 06:33:10 -0800 (PST)
  From: Neil Fegen <n.fegen@hw.ac.uk>
    To: Koraljka Golub <k.golub@ukoln.ac.uk>

```
Neil Fegen wrote:
-------------------------------------------

Hi Kora

Your presentation was very informative. Of the 18 feedback responses, two-
thirds found your presentation useful, which was only slightly lower than
Scott's FeedForward, which was considered the best.

Hope this helps.

Regards

Neil
```

2) CETIS workshop

-------- Original Message --------
   Subject:RE: Greetings from Bath
    Date: Fri, 15 Feb 2008 15:28:41 -0000
     From:L.Whitelaw <L.Whitelaw@open.ac.uk>
      To: Koraljka Golub <k.golub@ukoln.ac.uk>
References:<47B5A80B.2090800@ukoln.ac.uk>

```
Hi Kora,

Your presentation was really interesting and it was lovely talking to
you. […]

Hope this helps

Lara
```

3) ISKO UK news

```
-------- Original Message --------
Subject:       [ISKOUK] KOnnect Digest #3 - What's happening at KOnnect
Date:   Tue, 22 Apr 2008 20:32:04 +0100
From:   Bob Bater <bbater@INFOPLEX-UK.COM> <mailto:bbater@INFOPLEX-UK.COM>
Reply-To:       UK Chapter of the International Society for Knowledge
Organization <ISKOUK@JISCMAIL.AC.UK> <mailto:ISKOUK@JISCMAIL.AC.UK>
Organisation:  InfoPlex Associates
To:     ISKOUK@JISCMAIL.AC.UK
```

EnTag
Final draft submitted to JISC
k.golub@ukoln.ac.uk
14 November 2008

Dear list member,

Please find below our third digest of activities at the KOnnect blog.

[…]

KOKO Briefs

Three brief news items of potential interest:

[…]

Enhancing Social Tagging using KO

ISKO UK member Koraljka Golub is involved with a UKOLN project to enhance social tagging.

Bob Bater

Vice Chair, ISKO UK

KOnnect Moderator

5. Exit Plan

UKOLN will continue to host the project Web site for three years after the project ends, and will assist JISC in archiving it subsequently.
All major project outputs will be made available at the project's Web site. The demonstrators are available on request.

6. Sustainability Plan

A revised version of the Glamorgan/UKOLN demonstrator will be available via the EnTag project website, with study users removed and, following the study, some simple heuristics for simplifying the DDC suggestions trialled as a very initial step in this direction. The demonstrator is part of the ongoing and future research programme at both institutions. Funding for further development, taking account of user study and longitudinal studies of use is being sought and potential partners are being contacted.

The STFC demonstrator is being considered within the development plan of the STFC institutional repository to provide support for tagging in a production setting.

7. Budget

Overall project expenditure fell short of the funding awarded by 3,973 on direct costs. This is planned to be spent on further dissemination activities as well as for trips involving meeting partners for work on publications.

EnTag
Final draft submitted to JISC
k.golub@ukoln.ac.uk
14 November 2008

Lessons Learned

8. Aims and Objectives

The aims and objective set at the start were achieved. The need originally envisaged was fulfilled.

9. Overall Approach

Given the resources available, the timescale was necessarily condensed for both development and evaluation and this curtailed longitudinal study possibilities. Even with the wide range of partners, it proved difficult to attract participants to the study. Communities with an operational use context would be ideal but are hard to locate.

10. Project Outcomes

- The project conducted significant developments and user studies.

- The project showed the importance of social tagging supported by controlled vocabulary suggestions, both at the time of tagging (indexing) and retrieval.

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

- The value and usefulness of the suggestions is very dependent on the quality of the suggestions. The suggestions must be user-oriented as regards level of specificity, perspective and currency.

- The Global tag cloud proved problematic to use effectively in retrieval oriented applications.

- The user interface proved important, along with the visual presentation and interaction sequence.

- The quality and appropriateness of the controlled vocabulary proved to be important.

- There was evidence of support for automated suggestions in tagging context if they are appropriate and relevant. 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.

- Users 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.

- 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.

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

- There was some evidence that automatic suggestions of vocabulary-based tagging has potential to offer additional access points beyond the literal text and thus can enhance access compared to free text search engines. This can be applied in both repository contexts and collections, such as Intute. Given the patchy distribution of coverage in any single university

EnTag
Final draft submitted to JISC
k.golub@ukoln.ac.uk
14 November 2008

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 in the future.

- There is significant potential to augment the entry vocabularies of controlled vocabularies where successful mapping of user tags has occurred. Vocabulary terminology could potentially be modernised and informal user terms introduced.

11. Stakeholders

We recommend that social tagging be allowed in the JISC context (e.g., repositories), supported by controlled vocabularies. The project showed the importance of social tagging supported by controlled vocabulary suggestions, both at the time of tagging (indexing) and retrieval. Controlled vocabulary suggestions help produce ideas what to tag, ensure consistency and retrieval, make it easier to find focus for the tagging, etc.

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.

There are obvious benefits in the repository context for users-as-authors if they feel confident that increased discovery of their publications will result. 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.

Another promising application area is investigating tagging activity with an explicitly pedagogical focus. Thus an extended EnTag could 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.

12. Project Partners

Working with both funded and non-funded project partners went rather smoothly; no conflicts occurred within the project. The case studies were independently organised, with sharing of ideas, approaches, evaluation criteria and the development of appropriate recommendations.

EnTag
Final draft submitted to JISC
k.golub@ukoln.ac.uk
14 November 2008

It was enriching for the project to have a wide ranging set of collaborators. Also, it was good to have the two different case studies even if comparison at a high level only.

We have been requested by HILT to supply user tags (anonymised) from the Intute study to help them augment their vocabularies and their resolving user input to vocabularies.

13. Project Management

As long as project planning has the flexibility of adjusting to findings along the way, it delivers in the end.

UKOLN had a crucial member of staff leaving and it took time to recruit a new one, who started 3 December 2007, while the project was to start 1 September. STFC was affected by unforeseen budget reductions in its core funding which affected its staffing and scheduling within the project. Also, as pointed out earlier, recruiting users proved a much bigger challenge and took much more time than initially predicted. In addition, two members of the team had a death in their families. Due to these and several other minor factors, the projects final deliverables were delayed from end of September to mid-November.

14. Programme Support

Programme support worked well for the EnTag project.

15. Future Work

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 user interface.

The following have also been recognized:

- 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); and,

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

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

EnTag
Final draft submitted to JISC
k.golub@ukoln.ac.uk
14 November 2008

Appendixes

Appendix A. Final Budget