

# Tagging behaviour with support from controlled vocabulary

Marianne Lykke  
Aalborg University, Denmark

Anne Lyne Høj  
Royal School of Library and Information Science

Line Nørgaard Madsen  
Royal School of Library and Information Science

Koraljka Golub  
UKOLN

Doug Tudhope  
University of Glamorgan

## Abstract

The paper investigates how knowledge structures from a controlled vocabulary affect tagging. The study is a comparative analysis of tags assigned in two tagging systems - a simple tagging system (control system) that provides suggestions from two tag clouds (all users tags and my tags), and an enhanced tagging system (experimental system) that additionally offers suggestions from the Dewey Decimal Classification system (DDC). In the experimental study, 28 political students completed four tagging tasks, each comprising 15 documents. The focus was to examine how suggestions from the enhanced tagging system affect tags as regard tag specificity, exhaustivity, and novelty. Generally, there were no big differences between assigned tags from the two systems. The largest difference was a higher degree of tag specificity in the enhanced system indicating that suggestions from a controlled vocabulary might help taggers in being more specific in their tagging, allowing more precise information searching based on user tags. In addition, the results indicate that structured controlled suggestions might encourage taggers to add synonym variations enhancing the variety and number of access points. Furthermore, controlled vocabularies might be useful for automatic spell checking. Future study should explore in what direction the different kinds of suggestions lead the tagger and whether it is possible to identify scope or patterns between related tags from the two systems.

## Introduction

Social tagging is a way for users to provide user-oriented access to information resources on the Web. In information retrieval, tagging may be complementary to traditional indexing methods. However, many of the existing social tagging applications have not been designed with information retrieval in mind. Users use tags to organize their own private documents, and tags are rather personal than social (Guy & Tonkins, 2006). Folksonomies often lack basic control of spelling variants, synonyms, homonyms, and polysemes (Guy & Tonkins, 2006; Spiteri, 2007). Studies of the

del.icio.us folksonomy shows that tags are often broad and much like title terms and terms from tag clouds (Munk & Mørk, 2008ab). A retrieval study by Morrison (2008) confirms that folksonomies result in lower precision compared to search engines and subject directories due to general tags.

In the EnTag project (**Enhancing Tagging** for Discovery) we addressed the problems of lack of specificity and novelty of tags. The project investigated ways of enhancing social tagging via controlled vocabularies, with a view to improving the suitability for information discovery and retrieval (Golub et al., 2009). The project studied tagging in two different contexts: tagging by readers using documents from the Intute subject gateway and tagging by authors submitting papers to the STFC ePublication Archive.

This paper builds on data from the Intute user study and investigates how knowledge structures from a controlled vocabulary affect end-users' assignment of tags. The study is a comparison of tags assigned using a traditional tagging system that provides suggestions from two tag clouds (all users' tags and each user's tags) and an experimental, enhanced tagging system that additionally offers suggestions from the Dewey Decimal Classification system (DDC) (Dewey Service, 2011).

We specifically address the following research question:

- Is it possible to identify differences between tags regarding specificity, exhaustivity, and overlap with title terms and assigned controlled and uncontrolled keywords when using only social tagging versus when using social tagging in combination with suggestions from a controlled vocabulary?

## Related work

Weller (2007) compares ontologies and folksonomies, suggesting that they are not to be seen as rivals but complementary to each other, representing different points of view and alternative access points. The key challenge is to find the right combinations of approaches. Noruzi (2007) emphasizes that it is impossible to maintain consistency over time or across folksonomy users without a thesaurus. Hayman & Lothian (2007) present a proof of concept where users can tag resources by choosing from an established taxonomy or by entering their own terms. Users' own terms are used later to feed back into the taxonomy to improve its quality. Marchetti et al. (2007) describe SemKey in which they use WordNet and Wikipedia to introduce unique concepts. Trant & Bearman (2006) compare end-user tagging of art against existing professional terms and find that the former increases access points. Several social tagging projects have recognized the problem and tried to implement solutions (Golub et al., 2009). However, as known to the authors, no research has investigated the enhancement of social tagging with suggestions from a controlled vocabulary in a user evaluation with existing retrieval applications.

## Methodology

The first subsection describes the overall test design, and the second describes the methodology for tag analysis.

## Overall test design

28 politics students took part in the study. The tagging took place remotely at participants' homes using their own computers. A pre-study questionnaire showed that they were experienced Web users. Half of them had used tagging applications before but conducted little tagging. Almost a third had some acquaintance with DDC; there were none with any acquaintance with LCSH or other controlled vocabularies. The majority had never used Intute.

The main data collection method was transaction logging. In order to understand the results further, three questionnaires were used: 1) a pre-study questionnaire, to collect background information; 2) post-task questionnaires, to gather experience about the tagging task, and, 3) a post-study questionnaire, to gather experience about the whole study.

After signing the participation consent form and completing the pre-study questionnaire, the participants were given a training document through which they were to acquaint themselves with the demonstrator and tagging. Then they received an instructions document describing the tagging tasks. The four tagging tasks were on topics from politics: two were pre-defined (controlled) and two were on a self-chosen topic (free). The controlled task for simple tagging was on the topic of 'European integration', and for enhanced tagging on 'peacekeeping'. A hypothetical group project scenario was outlined as a rationale and motivation for the tagging activity. The following box shows one of the controlled enhanced tagging tasks.

Imagine that, as part of one of your courses, you are asked to write a four-page essay on the **topic of European integration**, as a joint project in groups of four. The essay should critically discuss existing theories about the creation of the European Union and its institutions. Your lecturer has instructed you to look for resources in the EnTag system. Since you will be working together with three other students, you should tag the documents you retrieve with tags that would be useful to you but would also enable other students to find those documents in EnTag and understand from your tags what the documents are about.

In each task they were first instructed to search for documents and then tag 15 of them. In controlled tasks they were asked to tag the top 15 retrieved documents; in free tasks they could choose any 15 that they found relevant. The analysis showed that 53 documents were tagged at least once in the controlled tasks, and 41 in the simple tasks. This is because participants did not closely follow the top 15 instruction, and because some URLs were temporarily unavailable. They were instructed to spend between five and ten minutes on tagging each document, and to tag as many aspects and topics as they thought appropriate for the task. In the case of long documents, they were instructed to focus on the abstract, introduction, conclusion, headings and table of contents. When tasks were being completed in the enhanced interface, the instruction was to try to consider controlled vocabulary suggestions if appropriate. The order of tasks was rotated in order to reduce the learning influence.

## Coding and analysis of tags

The focus was on examining how suggestions from the enhanced tagging system affected characteristics of tags in the following aspects: specificity, exhaustivity, and novelty. In order to

make the comparison of tags from the simple and enhanced test systems, documents and related tags originating from the controlled tagging tasks that were tagged in both test systems were selected. Documents that were unavailable at the time of the analysis were removed, leaving twenty-one documents. From this subset we selected a random sample of twelve documents for the analysis.

Inspired by Rolling (1981) and Wennerstrøm (2008), we used a reference standard as the instrument to analyse the tags. The reference standard was taken to be existing Intute controlled and uncontrolled keywords. The controlled keywords comprise terms from DDC, International Bibliography of the Social Sciences thesaurus (IBSS, 2011), and the Humanities and Social Science Electronic Thesaurus (HASSET, 2011). Uncontrolled keywords mostly include names of countries. We used the standard to characterise the types of match between a tag and (un)controlled keywords and title terms as given in Table 1. We defined the degree of tag specificity as the percentage of narrower tag terms, and the degree of tag exhaustivity as the percentage of related tag terms. Tag novelty was analysed by calculating tag overlap with title terms and assigned controlled and uncontrolled keywords.

<b>Match type</b>	<b>Definition and measurement</b>
Title term (TT)	A tag is identical to title terms. If a tag consists of more terms from the title, the order of terms is not taken into account.
Preferred term (PT)	A tag is identical to one of the Intute keywords
Synonym term (ST)	A tag has the same meaning as one of the assigned Intute keywords
Narrower term (NT)	A tag is hierarchically subordinated to one of the assigned Intute keywords
Broader terms (BT)	A tag is hierarchically broader than one of the assigned Intute keywords
Related term (RT)	A tag is associatively related to one of the assigned Intute keywords
Spelling mistakes (SM)	A tag is a spelling mistake

Table 1. Study variables

The coding was carried out exclusively, i.e. each tag was coded as one single match type. The list of variables has been applied consecutively, in the order listed in Table 1. Because document title was immediately visible for the participants and presumably an inspiration for the tagger, we decided to prioritize title terms before other matches. We did not distinguish between upper-case and lower-case letters. The following tags were removed from the analysis, because they required further interpretation: tags resembling a sentence, e.g. ‘MoD articles on Iraq, government perspective – possible highly politically motivated’, and compound tags representing multiple concepts, e.g. ‘civil-military action in Afghanistan and Liberia’. However, DDC hyphenated keywords were included, where the coded part was what was taken to be the focus of the keyword.

The coding was carried out by three members of the research team. First, each researcher conducted the analysis on her/his own, then the results were compared, and finally the differences were harmonized in a joint discussion. To illustrate the coding, an extract of the coding file is shown in Table 2. Each tag was characterized by whether it was taken from controlled DDC suggestions (‘tag from DDC’), or was a free-form tag (‘free tag’), and by which match type it belonged to.

Title: NATO'S assistance to the African Union for Darfur		
Controlled keywords: civil war; nato; international security		
Uncontrolled keywords: Sudan, Darfur		
User	Tag	Characteristics
3	North Atlantic Treaty Organization	tag from DDC <b>ST</b>
3	Humanitarian intervention	tag from DDC <b>RT</b>
3	NATO	free tag <b>PT</b>
3	Darfur human rights	free tag <b>NT</b>
3	Peacekeeping	free tag <b>RT</b>
3	International action – NATO	free tag <b>BT</b>
3	Sudan	free tag <b>PT</b>

Table 2. Extract of coding file

## Results and discussion

For the twelve documents, a total of 711 tags were used in the simple system (on average 59 tags per document) and 618 tags in the enhanced system (on average 52 tags per document). In both systems 71.6% of the tags were freely assigned tags, and the remaining 28.4% were tags selected from the DDC suggestions. Table 3 provides an overview of the average results for the tag match types. Results for the individual study variables are presented and discussed in subsections below.

Match types	Simple system	Enhanced system
TT	71 (10,0%)	61 (9,9%)
PT	21 (3,0%)	27 (4,4%)
ST	8 (1,1%)	17 (2,7%)
NT	126 (17,7%)	135 (21,8%)
BT	36 (5,1%)	28 (4,5%)
RT	438 (61,6%)	339 (54,9%)
SM	11 (1,5%)	11 (1,8%)
ALL	711 (100%)	618 (100%)

Table 3. Results for variables (number/percentage)

The number of title term tags is similar for the two test systems: 10.0% of tags assigned in the simple system, and 9.9% in the enhanced system. The number of title term tags is relatively high compared to the other variables. This may be explained by the fact that title was visible to the taggers. In addition, the fact that the coding procedure prioritized title term tags may have also contributed. The number of title term tags varies from one document to another, and depends on the significance of the title, e.g. for the title 'Operation Telic', no title tags were assigned compared to the title 'United Nations peacekeeping operations'. Overall, the analysis showed no large difference

regarding percentage of title terms: around 10% originated from the title in both systems. Depending how tags are exploited by the search engine, it might be relevant to avoid title tags, because they do not add to the overall document description.

In the simple system 3.0% of tags were preferred term tags compared to 4.4% in the enhanced system. This difference may be explained by the fact that in the enhanced system the taggers had been presented with the DDC suggestions. The number of preferred terms may have been actually higher, because several of the assigned tags were title term tags as well as preferred term tags, as such being coded as a title term in accordance to the coding principle. The number of preferred tag terms is slightly larger in the enhanced system. The results indicate that controlled suggestions may encourage taggers to control their vocabulary by assigning preferred terms. As a tagger commented in the post-questionnaire, 'Continuity in the exact labels is important since there are a bunch of different ways of labelling the same thing (i.e. Croatia vs. Republic of Croatia or European Parliament vs. Parliament of Europe, etc.)'.

Only a small number are synonym term tags: 1.1% in the simple system and 2.7% in the enhanced system. Most synonym term tags are acronyms or their full form versions, 13 out of 17 in the enhanced system and six out of eight in the simple system. While it seems that the taggers are able to apply acronym variations in both systems, the double number of acronym synonym term types in the enhanced system indicates that the DDC suggestions inspire taggers to assign synonym variations, providing more variation in access points. Depending on the aim of tagging this should be exploited when implementing a tagging system, e.g. by emphasizing synonyms for the tagger.

In the simple system 17.7% of the tags are narrower term tags compared to 21.8% in the enhanced system. The difference of 4.1% more narrower term tags in the enhanced system indicates that the specificity of tags is higher in the enhanced system compared to the simple system. In both systems the consistency of narrower terms is high compared to broader, synonym and related term tags.

In the simple system 5.1% of the assigned tags are broader term tags compared to 4.5% in the enhanced system. In both systems the assigned broader term tags are generally very broad: e.g. 'security', 'international', and 'law'. The number of broader tags was slightly larger in the simple system and the number of narrower tags larger in the enhanced system. It seems that the enhanced system, by showing the hierarchical structure, inspired the taggers to assign more specific tags. The finding suggests that suggestions from a hierarchical structured vocabulary might inspire taggers to assign specific terms.

Related term tags are the most frequent tags, 61.6% in the simple system and 54.9% in the enhanced system. Overall, the related term tags are topical, although other types of tags are identified, for instance 'resource' type tags (e.g. 'EU e-journals' or 'European research papers'), 'ownership/source' tags (e.g. 'Blank' and 'Kermit Blank'), and 'opinion' tags (e.g. 'excellent website for research'). In particular, the 'resource type' tags appeared in the simple system. 'Ownership/source' and 'opinion' types of tags appeared both in the simple and enhance system. The number of related terms was larger in the simple system compared to the enhanced system. This difference should be further investigated.

In the simple system, 1.5% of the tags were spelling mistakes compared to 1.8% in the enhanced system. In percentage terms the difference between the systems is small. However, the type of spelling mistake differs. In the enhanced system around half of the mistakes appear in name tags, e.g. names of organizations ('Ministry of Defence'), countries ('Afghanistan') or cities ('Dafur'). This is surprising, because name tags are present in DDC, and could easily have been selected. Apparently, the taggers believed they knew how to spell names. It might be useful to use the controlled vocabulary for automatic spell checking.

It is a limitation that the present study only examined the subset of documents and tags that were tagged in both systems. A future study should examine all documents and related tags in order to verify the result with a larger data set.

## Concluding remarks

In general, the analysis showed no large difference in the assignment of tags in the two test systems. The largest difference was higher degree of specificity in the enhanced system. In addition, the results indicate that suggestions from a controlled vocabulary might encourage taggers to add synonym variations. Furthermore, controlled vocabularies might be useful for automatic spell checking.

The variety of related tags was large. Several taggers pointed out that the enhanced suggestions provided good ideas for the tagging, e.g. 'prompted tags that wouldn't ordinarily have thought of' or 'came up with things I would not have thought of if left to my own devices'. Future study should explore in what direction the suggestions led the tagger, whether it is possible to identify scope or patterns between related tags from the two systems.

## References

- Dewey Services: Overview* (2011). Retrieved from <http://www.oclc.org/dewey/overview/default.htm>
- Golub, K., et al. (2009) EnTag: Enhancing Social Tagging for Discovery. *Joint Conference on Digital Libraries (JCDL)*, Austin, TX, June 15-19.
- Guy, M., & Tonkin, E. 2006. Folksonomies: Tidying up tags? *D-Lib Magazine*, January 2006. Retrieved from <http://www.dlib.org/dlib/january06/guy/01guy.html>
- Hayman, S. & Lothian, N. (2007). Taxonomy directed folksonomies. *73RD IFLA General Conference and Council*, 19-23 August 2007, Durban, South Africa.
- Humanities and Social Science Electronic Thesaurus (HASSET)* (2011). Retrieved from <http://www.data-archive.ac.uk/search/hassetAbout.asp>
- IBSS thesaurus* (2010). Retrieved from <http://www.lse.ac.uk/collections/IBSS/about/thesaurus.htm>

Marchetti, A., et al. (2007). SemKey: A semantic collaborative tagging system. In *Tagging and Metadata for Social Information Organization, 16th International World Wide Web Conference*, May 8-12, Banff, Alberta, Canada.

Munk, T. B., & Mørk, K. (2007a). Folksonomies, tagging communities, and tagging strategies - an empirical study. *Knowledge Organization*, 34(3), 115-127.

Munk, T. B., & Mørk, K. (2007b). Folksonomy, the power law and the significance of the least effort. *Knowledge Organization*, 34(1), 16-33

Noruzi, A. (2007). Folksonomies: why do we need controlled vocabulary? Editorial. *Webology*, 4 (2).

Rolling, L. (1981). Indexing consistency, quality and efficiency. *Information Processing and Management*, 17(2), 69-76.

Smith, G. (2008). Tagging: emerging trends. *The Bulletin of the American Society for Information Science and Technology*, August/September, 34(6).

Spiteri, L. F. (2007). Structure and form of folksonomy tags: the road to the public library catalogue. *Webology*, 4 (2).

Trant, J., & Bearman, D. (2006). Exploring the potential for social tagging and folksonomy in art museums: proof of concept. *New Review of Hypermedia and Multimedia*, 12(1), 83 - 105.

Weller, K. 2007. Folksonomies and ontologies: two new players in indexing and knowledge organization. In *Online Information Conference Proceedings*, London 2007, 108-115.

Wetterström, M. (2008). The complementarity of tags and LCSH – a tagging experiment and investigation into added value in a New Zealand library context. *The New Zealand Library and Information Management Journal* 50(4), 296-310.

Corresponding author:

Marianne Lykke can be contacted at [mlykke@hum.aau.dk](mailto:mlykke@hum.aau.dk)