# Relationships among Category Semantics, Perceptions of Term Utility, and Term Length and Order in a Social Content Creation System

Corinne Jörgensen Besiki Stvilia Shuheng Wu
College of Communication and Information, Florida State University, PO Box 3062100, Tallahassee, FL 32306-2100
cjorgensen@fsu.edu bstvilia@fsu.edu sw09f@my.fsu.edu

## ABSTRACT
While there are increased efforts to extend existing controlled vocabularies through harvesting socially created image metadata from content creation communities (e.g., Flickr), questions remain about the quality and reuse value of this metadata. Data from a controlled experiment was used to examine relationships among categories of image tags, tag assignment order, and users' perception of usefulness of preassigned image index terms. Preliminary findings indicate that, on average, "Group" category terms were assigned first, and were also rated highest in usefulness. Other broad tag categories that were assigned earlier and rated more useful were Human Attributes and People, but others were more variable. However, the study found no correlation between tag length and assignment order, or term length and its perceived usefulness. The study's findings can inform the design of controlled vocabularies, indexing processes, and retrieval systems for images.

## Categories and Subject Descriptors
H.3.1 [**Information Storage and Retrieval**]: Content Analysis and Indexing – *indexing methods, thesaurus.*

## General Terms
Measurement, Documentation, Experimentation, Human Factors, Languages.

## Keywords
Image indexing, Knowledge Organization, Folksonomies, Social tagging, Metadata, Experimental design.

## 1. INTRODUCTION
There has been a large body of research suggesting that socially created metadata (e.g., user-generated tags) could be complementary to traditional knowledge organization systems [6, 7, 8, 10, 11, 12, 13]. This has been accompanied by an increase in efforts to harvest metadata for image or photo collections to extend existing controlled vocabularies by deploying these collections in social content creation communities (e.g., Flickr). However, different types of users create social metadata in different contexts and for different purposes [2, 10]. The terms that users select to search for images may differ from those they use to describe images [1]. The usefulness and quality of metadata is recognized as contextual [3, 9]. Furthermore, the construction of high-quality knowledge organization systems requires expensive knowledge engineering work. Adding new metadata may not necessarily lead to value increase or cost reduction for the activity [9]. Questions remain about the quality and reuse value of social metadata for image indexing. In particular, it is essential to determine what terms are useful to the user and how to identify useful terms for a particular image inexpensively – that is, automatically.

This study examines relationships among categories of image tags, tag assignment order, and users' perception of usefulness of tags. The findings of this study can inform the design of controlled vocabularies, indexing processes, and retrieval systems for images. In particular, the findings of the study can advance the understanding of image tagging practices, tag facet/category distributions, relative usefulness and importance of these categories to the user, and inexpensive mechanisms for identifying important terms.

## 2. METHODOLOGY
This study examines the following research questions:

- What is the relationship between the categories of tags users assign to images and the tag assignment order? (Free tagging – Description Task)
- What is the relationship between the categories of tags and users' perceptions of the usefulness of these terms as indicated by their ratings of a set of pre-assigned terms? (Evaluation Task)
- Is there a relationship between tag length and tag assignment order?
- Is there a relationship between tag length and user's perception of its usefulness?

The study employed a controlled experimental design, asking a sample of 35 participants to perform image description and pre-assigned term evaluation tasks using modified Steve tagger software (http://sourceforge.net/projects/steve-museum/). Participants were undergraduate and graduate students and staff members recruited from the College of Communication and Information at Florida State University. Pre-experiment survey questions indicated few participants were familiar with the concept of controlled vocabularies; more were familiar with tagging, but the majority had not done tagging before.

In the description task, participants were given 10 sampled photographs selected from a set of 7,192 photographs from the Library of Congress Flickr photostream; they described each of these photographs by assigning tags. Next, in the evaluation task,


*iConference '12, Feb 07-10,2012, Toronto, ON, Canada ACM 978-1-4503-0782-6/12/02.*

participants rated a set of pre-assigned terms for each of the 10 photographs by their usefulness for describing the content of the photographs on a five-point Likert scale (i.e., 'strongly disagree', 'disagree', 'neutral', 'agree', 'strongly agree'). The pre-assigned terms were obtained from the following sources: (a) the Thesaurus for Graphic Materials, (b) the Library of Congress Subject Headings, (c) tags assigned to the photographs by Flickr members, (d) a folksonomy generated from the Library of Congress photostream on Flickr, (e) the folksonomy from the complete Flickr database, and (f) the English Wikipedia. The pre-assigned terms were obtained as described previously in [12]. To examine the relationship between tag categories and tag assignment order, for each combination of tag, image, and participant, tags were assigned numbers indicating the order in which the participant assigned the tag to the image. Next, two of the researchers coded each tag according to Jörgensen's coding scheme of broad categories composed of several specific types of attributes [4, 5], with the addition of a new category, that of "Group," referring to organizations, institutions, or social groups with a particular purpose or goal, as opposed to a transient or spontaneous gathering of people. While terms with this meaning had previously been assigned to another category, their frequency in this particular set of data, especially where additional text information about the image appeared, suggested the utility of this addition.

## 3. FINDINGS

The Wilks-Shapiro test showed that neither tag assignment order nor term ratings were normally distributed. Hence, the researchers used a nonparametric Kruskal-Wallis test to examine relationships among tag categories, tag assignment order, and term rating. The Kruskal-Wallis test showed that both specific and general tag categories were statistically significantly different on the order of tag assignment ($\chi2 = 132$, df = 28, $p < 0.001$; $\chi2 = 58.6$, df = 11, $p < 0.001$). Table 1 shows that, on average, Group was the first assigned broad category of tags, followed by Color, Art Historical, Story, People, Human Attributes, and Abstract.

**Table 1. The mean and median of order of tag assignment for general codes (average tags per image = 5.7).**

| Categories | N | Median | Mean |
|---|---|---|---|
| Group | 40 | 1 | 1.85 |
| Color | 50 | 2 | 2.52 |
| Art Historical | 151 | 2 | 3.46 |
| Story | 438 | 2 | 3.67 |
| People | 185 | 2 | 3.74 |
| Human Attributes | 216 | 2 | 3.75 |
| Abstract | 68 | 2 | 3.90 |
| Description | 21 | 3 | 3.05 |
| Location | 32 | 3 | 3.72 |
| Object | 787 | 3 | 4.88 |
| Total | 1988 | 3 | 4.08 |

Similarly, two of the researchers also coded the pre-assigned terms according to Jörgensen's coding scheme [4, 5], with the same addition of Group. The Kruskal-Wallis test showed that both specific and general codes were statistically significantly different on usefulness ratings ($\chi2 = 549.8$, df = 24, $p < 0.001$; $\chi2 = 251.2$, df = 11, $p < 0.001$). Table 2 shows the average rating of Group terms was the highest, followed by Human Attributes, Abstract, People, and Description categories.

The study did not find significant associations between tag length and assignment order (Spearman's rho 0.06, $p<0.01$) and between pre-assigned term length and perceived usefulness (Spearman's rho 0.07, $p<0.001$).

**Table 2. Categories of terms by ratings (1:least useful to 5: most useful; based on 12,180 user ratings).**

| Category | Median | Mean |
|---|---|---|
| Group | 4 | 3.933 |
| Human Attributes | 4 | 3.769 |
| Abstract | 4 | 3.729 |
| People | 4 | 3.681 |
| Description | 4 | 3.678 |
| Story | 3.7 | 3.469 |
| Objects | 3.6 | 3.441 |
| Art Historical | 3.5 | 3.344 |
| Color | 3 | 3.015 |
| Viewer Response | 3 | 2.886 |
| Location | 3 | 2.880 |
| External Relation | 3 | 2.695 |

## 4. DISCUSSION

The preliminary findings of the study suggest that while in some instances users might assign broad categories of terms that they perceive most useful first, order is not necessarily an indicator of perceived utility. While term order could guide library and museum communities in identifying useful and important index terms for their image collections at a relatively low cost, these results indicate that at the broader level of category order, the results are more variable. Future steps in the study include examining the relationship between perceived usefulness and assignment order at the term, or tag level.

One of the limitations of this study is that participants were self-selected from a single academic department. Replicating the experiment with a larger and more representative sample of participants and a larger sample of photographs would be desirable to strengthen the findings of the study.

## 5. ACKNOWLEDGMENTS

This research is partially supported by an OCLC/ALISE Research Grant, 2010.

## 6. REFERENCES

[1] Chung, E., and Yoon, J. 2008. A categorical comparison between user-supplied tags and web search queries for images. In *Proceedings of the American Society for Information Science and Technology*, 45, 1, 1-3. DOI= http://dx.doi.org/10.1002/meet.2008.1450450392.

[2] Cunningham, S., and Masoodian, M. 2006. Looking for a picture: An analysis of everyday image information searching. In *Proceedings of the 6th ACM/IEEE-CS Joint Conference on Digital* Libraries (Chapel Hill, North Carolina, June 11 - 15, 2006). JCDL '06. ACM, New York,

NY, 198-199. DOI= http://doi.acm.org/10.1145/1141753.1141797.

[3] Golbeck, J., Koepfler, J., and Emmerling, B. 2011. An experimental study of social tagging behavior and image content. *Journal of the American Society for Information Science and Technology*, 62, 9, 1750-1760. DOI=http://dx.doi.org/10.1002/ASI.21522

[4] Greenberg, J. 2001. Quantitative categorical analysis of metadata elements in image application metadata schemas. *Journal of the American Society for Information Science and Technology*, 52, 11, 917-924. DOI= http://dx.doi.org/10.1002/asi.1170.

[5] Jörgensen, C. 1995. *Image attributes: An investigation.* Unpublished Ph.D. thesis, Syracuse University.

[6] Jörgensen, C. 1998. Attributes of images in describing tasks. *Information Processing and Management*, 34, 2/3, (Mar. - May, 1998), 161-174. DOI= http://dx.doi.org/10.1016/S0306-4573(97)00077-0.

[7] Jörgensen, C., Stvilia, B., and Jörgensen, P. 2008. Is there a role for controlled vocabulary in taming tags? In *Proceedings of the 19th Workshop of the American Society for Information Science and Technology Special Interest Group in Classification Research*. Columbus, OH.

[8] Matusiak, K. K. 2006. Towards user-centered indexing in digital image collections. *OCLC Systems & Services: International Digital Library Perspectives*, 22, 4, 283-298. DOI= http://dx.doi.org/10.1108/10650750610706998.

[9] Rolla, P. J. 2009. User tags versus subject headings: Can user-supplied data improve subject access to library collections? *Library Resources and Technical Services*, 53, 3 (Mar. 2009), 174-184.

[10] Stvilia, S. and Gasser, L. 2008. Value-based metadata quality assessment. *Library and Information Science Research*, 30, 1 (Mar. 2008), 67-74. DOI= http://dx.doi.org/10.1016/j.lisr.2007.06.006.

[11] Stvilia, S. and Jörgensen, C. 2009. User-generated collection level metadata in an online photo-sharing system. *Library and Information Science Research*, 31, 1 (Jan. 2009), 54-65. DOI= http://dx.doi.org/10.1016/j.lisr.2008.06.006.

[12] Stvilia, S. and Jörgensen, C. 2010. Member activities and quality of tags in a collection of historical photographs in Flickr. *Journal of the American Society for Information Science and Technology*, 61, 12, 2477-2489. DOI= http://dx.doi.org/10.1002/asi.21432.

[13] Stvilia, S., Jörgensen, C., and Wu, S. 2011. Establishing the value of socially created metadata to image indexing. *Library and Information Science Research* (in press).

[14] Wetterstrom, 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 (May. 2008), 296-310.