Member Activities and Quality of Tags in a Collection of Historical Photographs in Flickr

Besiki Stvilia and Corinne Jörgensen

College of Communication and Information, Florida State University
Tallahassee, FL 32306-2100, USA
{bstvilia, cjorgensen}@fsu.edu

Abstract

There is growing interest in, and an increasing number of attempts by, traditional information providers to engage social content creation and sharing communities in creating and enhancing the metadata of their photo collections to make the collections more accessible and visible. To enable and guide effective metadata creation, however, it is essential to understand the structure and patterns of the activities of the community around the photographs, resources used, and scale and quality of the socially created metadata relative to the metadata and knowledge already encoded in existing knowledge organization systems. This article presents an analysis of Flickr member discussions around the photographs of the Library of Congress photostream in Flickr. The article also reports on an analysis of the intrinsic and relational quality of the photostream tags relative to two knowledge organization systems: the Thesaurus for Graphic Materials and the Library of Congress Subject Headings. Thirty seven percent of the original tag set and 15.3% of the preprocessed set (after the removal of tags with fewer than three characters and URLs) were invalid or misspelled terms. Nouns, named entity terms, and complex terms constituted approximately 77% of the preprocessed set. More than a half of the photostream tags were not found in the TGM and LCSH, and more than a quarter of those terms were regular nouns and noun phrases. This suggests that these terms could be complimentary to more traditional methods of indexing using controlled vocabularies.

Introduction

Knowledge organization and representation systems (e.g., lists of terms, taxonomies, thesauri, ontologies) traditionally have been essential parts of the information organization and retrieval infrastructure in libraries and museums, and they have now become increasingly important on the Web to support entity and concept identification, semantic annotation, information retrieval, and question answering (e.g., Perez, 2009). Not surprisingly, there has been considerable research on controlled vocabulary and ontology construction, including research identifying quality index terms and on automatic concept and

relationship identification (e.g., Chen, Yim, Fye, & Schatz, 1995; Lancaster, 2000; Soergel, 1974). Nevertheless, the construction of high-quality knowledge organization systems (KoSs) involves expensive knowledge engineering work. Furthermore, quality is being recognized as contextual and dynamic (Jörgensen, 1995b; Strong, Lee, & Wang, 1997; Stvilia, Gasser, Twidale, & Smith, 2007). With changes in domain culture, activity systems, knowledge and technology, and user expectations, the quality of these systems can quickly become outdated and require regular intensive maintenance and upkeep. There is a need to identify sources of—and to define methods and mechanisms for—inexpensive, dynamic acquisition and integration of new knowledge into traditional KoSs.

There is a growing body of research on enhancing ontologies and thesauri by acquiring knowledge and concepts from socially created representative textual documents and reference sources such as Wikipedia (e.g., Agirre, Ansa, Hovy, & Martinez, 2000; Medelyan & Milne, 2008). This has been accompanied by an increase in efforts from traditional information providers to acquire metadata for existing image and photo collections by deploying these collections in existing social tagging communities (e.g., Springer et al., 2008) or by bundling image collections with social tagging systems and mechanisms (e.g., gaming interfaces) to jump-start new communities and motivate users to contribute metadata (e.g., Trant, 2008; von Ahn & Dabbish, 2004).

The Flickr Commons project (http://www.flickr.com/commons), which began in 2008, is a collaboration between Flickr and U.S. cultural heritage institutions such as the Library of Congress and the Smithsonian Institute. The goal of the project is to increase access to the institutions' publicly held photo collections and gain help from the Flickr community in acquiring additional knowledge and metadata for the photographs. The usefulness and quality of metadata, however, is contextual (Greenberg, 2001b; Stvilia & Gasser, 2008; Taylor, 1986). The importance and quality of member-generated tags in the Flickr Commons can be conditioned by the importance of the current goals and objectives of the information-related activities of both Flickr members and the target user communities of the institution deploying the collection in the Commons. In addition, as with the quality of any other product, the quality of tags can be influenced by calibrating the process of their generation, or a through quality-based selection and weeding of previously generated tags (Stvilia & Gasser, 2008). Therefore, to enable and motivate effective metadata creation and acquisition in the Flickr Commons, it is essential to understand the objectives, structure, and patterns of member activities around the photographs and resources used to support those activities. In addition, it is important to assess the nature and scale of metadata and knowledge created by the community relative to the metadata and knowledge encoded in the existing cataloging records and KoSs.

This paper reports the results of an exploratory study of Flickr member activities around photographs from the Library of Congress photostream on Flickr. The study also explored the intrinsic quality as well as the scale and types of photostream tags and compared those with two controlled vocabularies used by the Library of Congress: the Thesaurus for Graphic Materials (TGM) and the Library of Congress Subject Headings (LCSH).

Literature Review

Addressing the "vocabulary problem"—the problem of people using different words when describing or searching for the same concepts and entities, or alternatively, using the same words for different concepts and entities—has been one of the oldest problems in knowledge organization and continues today to be a very active area of research and practice (e.g., Buckland, 1999; Furnas, Landauer, Gomez, & Dumais, 1987; Klavans et al., 2009; Svenonius, 1986; Tan, Kan, & Lee, 2006). Controlled vocabularies (i.e., lists of terms, thesauri) are used to address the vocabulary problem between the user and the system by translating user terms into the indexing language used by the system. To accomplish that effectively, however, a controlled vocabulary needs to be aligned well with the information needs and language of the user (Soergel, 1974).

Considerable prior research has been undertaken on index and query term evaluation (Blair, 1996; Brooks, 1993; Greenberg, 2001a; Hersh, Pentecost, & Hickam, 1996). More recently, Wacholder and Liu (2006, 2008) developed and experimentally tested a method for evaluating the quality of index terms. There is a consensus that quality is contextual (Strong et al., 1997; Stvilia et al., 2007). A known entity or object identification task may require the use of high-entropy or high-specificity metadata (e.g., identifiers), whereas for a general relevance-based (i.e., "aboutness"-based) or other attribute-based (e.g., format-based) selection task, the use of a low-specificity term may suffice (Stvilia, Gasser, Twidale, Shreeves, & Cole, 2004).

Researchers have emphasized the importance of indexing documents based on end-user information needs and search queries (e.g., Soergel, 1974). It has been shown that different groups of end users could have different information need structures for and perspectives on the same document (Hjørland, 2008; Lancaster, 2000). A significant amount of research has been done on identifying and categorizing nonexpert index terms for images (Jörgensen, 1995a, 1995b, 1996, 1998; Jörgensen & Jörgensen, 2002; Lin, Beaudoin, Bui, & Desai, 2006; Mathes, 2004; Schmitz, 2006). Jörgensen (1995b) used different types of descriptive tasks with a wide range of participants and derived 10 broad classes related to descriptive image content from approximately 14,000 terms. Another study found that users group or categorize images by broad concepts, whereas they describe images by using more specific concepts and terms (Rorissa & Iyer, 2008). In addition, it has been suggested that the granularity of terms used to describe images may not be at the same level as the granularity of the ones used in a search. Chung and Yoon (2008) sampled and classified Flickr tags and user image queries obtained from Excite 2001 query logs into the Shatford (1986) categories (abstract, generic, and specific). A comparison of the tag and query categories revealed that whereas most of the tags were of the generic type (63%), most of the queries were specific (51%).

It has been suggested that collaboration between intellectual content creators and information organization experts may lead to a higher level of quality in metadata creation. Creators have intimate knowledge of their creations, whereas indexers and catalogers can assist them with knowledge of metadata schemas and classification systems (Greenberg, 2002). With the establishment and increase in popularity of social content creation and tagging systems such as Wikipedia and Flickr, researchers gain access to large sets of nonexpert-created metadata (e.g., tags, classification strings). Access to these data sets may help researchers better understand user needs for metadata, maintain and enhance the quality of existing controlled vocabularies, or generate new ones. For instance, after matching a large sample of

Flickr tags to a nonspecialist controlled vocabulary, Jörgensen, Stvilia, and Jörgensen (2008) found that Flickr folksonomies could be helpful in completing the term list for the vocabulary. A thorough review of folksonomy-related research can be found in Trant (2009). Recently, there have been several studies which matched member generated tags from Flickr and Delicious to expert created controlled vocabularies. (e.g., Yi & Chan, 2010, Yoon, 2009,) to identify term overlaps and differences. Also, in a recent study, the traditional Precision and Recall based evaluation framework from the field of information retrieval was used to evaluate effectiveness of non-expert generated tags in image indexing and retrieval (Bar-Iian, Zhitomirsky-Geffet, Miller, & Shoham, 2010).

Researchers also have looked at member objectives and motivations to share content and metadata on Flickr. Ames and Naaman (2007) interviewed 13 'heavy' users of a Flickr application and identified four types of motivations for tagging: Self-Organization, Self-Communication, Social-Organization, and Social-Communication. A study of flickr collections by Stvilia & Jorgensen (2007) listed eight motivations members might have when organizing photographs into groups: (1) to enable easy finding, (2) for easy sharing, (3) for archiving, (4) vanity, (5) "bibliographical," documenting a particular subject or concept (e.g., a sunrise), (6) supporting group or community activities (e.g., playing a game), (7) supporting an individual activity (e.g., documenting a process of setting up a computer for later use), and (8) no particular motivation - the collection was a product of the sum of many individual one-time activities. In a followup study, based on survey data, Nov, Naaman and Ye (2010) grouped member motivations to contribute tags in Flickr into four types: self-development, reputation within community, enjoyment, and commitment to the community.

The preceding studies provide valuable insight into index and query term evaluation, vocabulary construction, and, member motivation to contribute image metadata. Additional research, however, is needed to investigate the context and process of metadata creation by members, and the quality of these nonexpert metadata *relative* to expert-constructed controlled vocabularies in general, and to controlled vocabularies used for image indexing in particular.

Research Questions

This study examined user discussions around the Library of Congress photostream of historical photographs on Flickr to identify the types of user activities around photographs, and the information and knowledge resources used in those activities. The research sought to use these conversations and metadata samples to identify and investigate two facets of metadata and knowledge creation and evaluation by users: 1) incidents of term assignment and the activity context of those incidents; and 2) the intrinsic and relational quality of member-generated tags for photographs in the Library of Congress photostream in Flickr.

In particular, the study aimed to address the following research questions:

1. What are the types of activities around photographs in Flickr? Can these activities lead to metadata creation or evaluation? What are some of the information and knowledge resources used in those activities?

2. What is the intrinsic quality of the folksonomy (i.e., the set of tags) of the Library of Congress photostream in Flickr?

3. What is the relational quality of the folksonomy? - Can the folksonomy be a source of new terms for traditional controlled vocabularies? If so, how different from or similar to the traditional controlled vocabularies are the terms in the folksonomy, and what is the nature of those differences or similarities?

Methods

The study selected the Library of Congress photostream of historical photos in the Flickr Commons for several reasons. The primary reason was the historical and cultural value of the Library of Congress collections as a whole and of the individual photos, which attract considerable interest and attention from Flickr members. In addition, the Library of Congress is one of the largest, if not the largest, traditional producers of metadata and metadata tools, including controlled vocabularies used for subject indexing such as the TGM and the LCSH. These tools are freely accessible from the Library Web site and can be downloaded and used by anyone. Finally, Flickr is one of the largest content-sharing systems and boasts a community with millions of members and billions of photographs (Flickr, 2009; Flickr Blog, 2009). Hence, the Library photostream is a perfect benchmark for evaluating the quality of socially created image metadata relative to the Library's existing controlled vocabularies such as TGM and LCSH.

Several kinds of metadata are generated by Flickr members. Each photo has a title and may have one or more tags. In addition, Flickr allows users to attach notes or annotations to photos and to discuss or make comments about them. In addition, the Flickr system defines two major types of collections, individually created Photosets and collectively developed Group collections, which have a defined (although sometimes quite broad) purpose. Each collection usually has a title and a description, which may specify the concept or relationship members used to organize the collection.

It has been suggested by the literature that activities of document tagging and indexing involve sensemaking (Golder & Huberman, 2006). Hence, the current study used metaphors, principles, and concepts from activity theory (Engestrom, 2000; Kuutti, 1991; Leontiev, 1978) and the theories of sensemaking (Dervin, 1992; Weick, 1995) to guide the content analysis of member conversations in Flickr. Activity theory helps conceptualize a macro view of the structure and relationships in the community's activity system, whereas sensemaking can provide metaphors and templates for analyzing micro interactions and action scripts of the activity system. Sensemaking assumes that a systematic pattern exists in an individual's information behavior: gap defining followed by gap bridging (Dervin, 1992). A gap, discontinuity, breakdown, or disruption in an activity occurs when the perceived state of the world is different from the expected one, when the current information about the world is insufficient to complete the activity, or both (Weick, Sutcliffe, & Obstfeld, 2005). Sensemaking is social because it relies on and is influenced explicitly or implicitly by social mechanisms and structures, including knowledge systems, norms, and conventions. In addition, sensemaking builds plausible interpretations in retrospect, after the action. Because an interpretation follows the action, activity, or event in time, the interpretation may not be accurate, but rather may be an approximation justified by the current knowledge, beliefs, and overall context of the interpreter (Garfinkel, 1967). These relationships, however, have not been articulated clearly in sensemaking theories. Activity theory can help fill this gap and provide a helpful model for reasoning about the interaction between the objectives and goals of the

interpreter and the components of the activity context—the culture, norms, conventions, and rules of the community.

The data in this study consisted of tags and comments associated with photos from the Library of Congress photostream in Flickr. A total of 28,303 unique tags and 43,152 comments associated with 7,192 photos were downloaded on September 13, 2009, from the Library of Congress Flickr photostream.

The harvested Flickr tags were preprocessed before matching with the TGM and LCSH. In particular, multiterm concatenated tag sets were recursively split into individual terms based on the terms and inflections from the WordList (http://wordlist.sourceforge.net/). In addition, the set was cleaned of all URLs and tags with fewer than 3 characters. This reduced the number of tags in a set to 20,946. Finally, both the Flickr tags and the controlled vocabulary terms from the TGM and LCSH were stemmed using the Porter Stemmer algorithm (Porter, 1980). To collect, preprocess, match, and analyze the data sets, the study used a Flickr Java application programming interface (http://sourceforge.net/projects/flickrj/), the Atlas.ti and Stata software, and Java codes developed by one of the researchers.

Previous studies of information work organization in a social content creation system such as Wikipedia (Stvilia, Twidale, Smith, & Gasser, 2008) showed that communities often use discussion pages not only to converse about issues related to a particular document, but also to coordinate their work or negotiate activities around the documents. The current study content analyzed 257 comments Flickr members had made about 20 photos in the Library of Congress photostream. This was a convenience sample. The researchers selected photos with a significant number of member activities associated with them. At the same time, the researchers excluded photos about which member conversations were dominated by discussions on controversial metadata (e.g., racially insensitive titles) or the content of the photo. Some of the longest conversations threads in the photostream were about controversial titles of some of the photos, or controversial actions of some of the politicians from the past depicted in the photos. However, those conversations were more about the members expressing their personal beliefs and political views, and they added little factual knowledge or metadata to the photos.

The conversations were coded for the types of activities in which members engaged around the photographs, information gap or discontinuity incidents they experienced during those activities, and the strategies and knowledge resources they used to bridge those gaps. The researchers performed the coding by using Atlas.ti software. Each researcher open coded the complete sample independently. After coding was completed, the resultant schemas were aggregated and differences were resolved. The resultant codes were iteratively clustered to develop categories or a typology (Bailey, 1994). The researchers then used the typology to recode the sample.

The study also investigated the intrinsic and relational quality of the folksonomy. Evaluation of the intrinsic quality of the tags was guided by a previously developed framework of information quality assessment (Stvilia et al., 2007). Intrinsic quality measures the internal characteristics of a tag itself in relation to some general reference standards in a given culture, such as the WordNet (a general-purpose comprehensive ontology of word senses; http://wordnet.princeton.edu/). In contrast, relational quality measures relationships among the tag and some aspects of its usage context, and the reference source of that usage context. The relational quality of the photostream's folksonomy was evaluated relative to the context of maintaining traditional controlled vocabularies. In particular, the study assessed the suitability of the folksonomy as a source of terms for the TGM and LCSH.

The TGM is a specialized controlled vocabulary used in image indexing by the Library of Congress. The TGM is developed in a bottom-up way by the Library of Congress Prints and Photographs Division to support its image indexing needs, following the principle of "library warranty." That is, the TGM development team does not strive to construct comprehensive conceptual hierarchies of terms. New terms and relationships are added to the TGM only if needed in the Library's operations. This property makes the TGM an excellent baseline for assessing the quality and relative value of user-generated terms and relations. Similar to the TGM, the LCSH is constructed in a bottom-up way to support the Library's subject cataloging. The scope and the scale of the LCSH, however, are different from those of the TGM. The LCSH provides authoritative (preferred) and alternative subject access headings for bibliographic records. In addition, the number of subject terms in the LSCH is more than one order of magnitude higher than the number in the TGM—hundreds of thousands of terms versus tens of thousands.

The intrinsic quality of the folksonomy was evaluated as the ratio of valid terms, whereas the relational quality was evaluated as the ratios of valid terms not present in the baseline controlled vocabularies (the TGM and LCSH). The complete set of tags from the Flickr photostream (20,946 tags) was matched to the terms from the TGM (13,317 terms, including both subject and genre terms; http://www.loc.gov/rr/print/tgm1/), the LCSH (696,358 preferred and alternative terms or key phrases; http://id.loc.gov/authorities), and also the WordNet ontology (version 2.1; 207,016 terms) to identify the scale and the nature of overlaps and differences between the tags and the terms from these KOSs. The study used a subsumption operator to match the Flickr tags to the controlled vocabularies. To be a match, a Flickr tag had to be a subset of the vocabulary term.

Finally, to obtain a more nuanced description of the intrinsic quality of the tags and their suitability in updating or enhancing the controlled vocabularies, the study content analyzed random samples of 300 terms from each set: the folksonomy, the TGM, and the LCSH. The samples were coded for grammatical types and the cognitive categories of nouns (Basic, Subordinate, and Superordinate) using Prototype Theory (Rosch, 1973, 1978; Rosch, Mervis, Gray, Johnson, and Boyes-Braem, 1976):. These term type distributions from the folksonomy were then compared with and contrasted to the term type distributions in the TGM and LCSH samples.

Prototype Theory suggests that objects and concepts are identified and categorized at different levels. Thus, different terms are used to refer to the same concept, and these terms provide different levels of abstraction at which a concept can be indexed. The Basic level maximizes information gain as it represents the highest number of shared properties of the members of a category (Rosch 1978). In contrast to the Superordinate level, the Basic level is the most inclusive level at which most of the culturally justified functions and uses for the members of a category are defined, and a single mental image for the category can be formed (e.g., furniture vs. chair). The Basic level is usually highly distinctive and of a medium specificity. These distinguish the Basic level from the Subordinate level at which concepts are highly specific, but are of low distinctiveness from neighboring concepts (e.g., top hat vs. beret) Basic level concepts are usually labeled with most commonly occurring, contextually neutral, shortest terms. Frequently, Basic level terms are count nouns; Superordinate terms are mass nouns, and Subordinate terms are compound or complex forms (Croft & Cruse, 2004, p. 82-86; Lakoff, 1988, p. 133-134; Rosch, 1978; Tanaka & Taylor, 1991). Finally, Basic-level terms are important to knowledge organization systems. The literature shows that pictures of objects are identified most accurately and categorized fastest at the Basic level, for instance, "hat," rather than "beret" or "clothing" (Lakoff, 1988, p. 133-134;

Rosch, 1978). Furthermore, according to the theory of controlled vocabulary construction (Lancaster, 2000; Soergel, 1974), preferred terms for concepts in a controlled vocabulary usually are selected from the most frequently used terms. Hence, Basic level category nouns can be a valuable source of preferred terms in controlled vocabulary construction and maintenance.

Findings

Activities around photographs

The analysis of the member discussions identified 7 types of member activities around photographs (see Table 1). One of the most frequently occurring activities was disambiguation and resolution of uncertainties about the content of the photo: identifying people, objects, events, locations, and dates; or understanding the overall content and context of the photograph.

Table 1. Activity types

Linking and Grouping resources related to the content of the photo: people, objects, organizations, locations, nations, and ethnic groups. The resources could be other photos, audio and video files, books, blogs, encyclopedia and newspaper articles, maps, and so on.

Musing or Reminiscing over the content of the photograph, including telling personal stories and uploading personal and family photos or images

Discussing issues related to or triggered by the content and context of the photograph

Evaluating the art and quality of the photo, the specific photographic technique and medium, or metadata

Disambiguating and Resolving the components and context of the photo content. These could be identification and resolution of uncertainties about events, people, places, objects, time periods, and photo techniques.

Suggesting and Negotiating new metadata, and corrections to existing metadata (e.g., suggesting tags, variant titles)

Asking and Answering questions about the photo or use of Flickr tools

The analysis showed that although activities in Flickr in general were initiated by members as separate instances and were driven by individual motives, these individual activities were often expanded in time and space and formed processes—a set of interconnected activities. For instance, the analysis revealed that disambiguation and knowledge creation were often the outcomes of members' musing or reminiscing about their personal experiences and telling personal or family stories. Reminiscence was triggered by the member looking at a photograph that reminded her or him of people, places, times, or experiences, as well as by socializing with other members and reading their comments. The following quotation was taken from a conversation thread in which members argued whether the women was dressed up and posed for the photograph, or whether she was wearing the clothes she would regularly wear at work:

> Ditto Handpainter! My Mom worked right alongside Rosie, they don't get enough credit for their hard work! You know why they were so invaluable? Their hands were smaller and could fit in places their male counterparts couldn't reach ;)

Another conversation led a member to draw on his professional experience and make a useful connection between two terms for the same concept: "aircraft inspector" and "quality controller":

> I was in the Air Force for 13 years and work in a plant similar to this one. As her armband says Vega Aircraft Inspector her job may not be a dirty one. Her job was probably to inspect newly made wire harnesses for the avionics systems or bomb bay systems. Today they are called Quality Control and they rarely get dirty unless they have to crawl into a tight space somewhere on the aircraft itself.

To determine components of the photo content, members sought information on and cross-examined information from both other related Flickr photos and outside reference sources. Consequently, the comments contained pointers to relevant resource-specific reference and knowledge sources, as well as to related photographs in Flickr:

> I matched the spire to this photo captioned 950 Broad on newarkstreets.com. The City of Newark landmarks page for visitors indicates that this is Grace Episcopal Church—here's a photo of Grace on Flickr.

Members also used URLs to make more explicit and "actionable" connections and associations among different entities and resources. Indeed, one of the most frequently occurring activities was members' linking or grouping related photos within Flickr. In addition, members would search for and link to documents and reference sources of different kinds outside the Flickr database: photos, video files, blogs, and newspaper and encyclopedia articles (see Figure 1).

The study extracted 8,391 URLs from the comments and grouped them by domain. As expected, the majority of the URLs pointed to related photos and groups within Flickr itself ($\approx 38\%$), although members linked to reference sources such as Wikipedia ($\approx 20\%$), digital libraries and archives such as the New York Times archives ($\approx 10\%$), and other reference sources and digital libraries (Google Books, Google Maps, Google Patents, Library of Congress, Smithsonian, New York Public Library, YouTube, etc.). These explicit links can be very valuable for constructing and mining webs of associations among related terms, concepts, and resources automatically.

In the absence of an explicit community objective and work specification for the activities, collaboration in the Flickr photostream was spontaneous and emergent. One member would define a gap in the knowledge or metadata about an entity in or the context of the photograph, whereas others might decide to respond to the challenge and search for information to bridge the gap:

> A few trifling questions are raised: How are Mr. & Mrs. Lawson connected to Marr's Pekingese Kennels? Did they own it? Was that company their breeding agent? Did they lease Noswal Neto for breeding purposes?

In some cases, members made more explicit requests for assistance from the community:

> Can anyone post that photo from the National Geographic with Woman with GREAT Eyes that was on the cover. I Think she was from Afghanistan.

The analysis also found instances of members evaluating the quality of metadata of the photographs. Indeed, in 20% of the sample (4 photos), members identified problems in the existing metadata and suggested corrections that were accepted by the provider, the Library of Congress. The types of metadata

quality problems identified by members included intrinsic quality problems, such as spelling errors, and relational quality problems, such as when members suggested corrections to the title of a photo when it conflicted with the norms and conventions of the present time:

> The abbreviation chosen is offensive to many and reverts to the dismissive terminology used during WWII and is out of place in this day and age.

Entity resolution activities often led to corrections of the photo's metadata. Members were keen in noticing anomalies or ambiguities in the photo's metadata. By triangulating from multiple sources, they were able to determine the identities of entities depicted in photos and, as a by-product, to suggest changes and additions to the photo's metadata:

> The date suggested by the Bain ordering (Reginald was clearly in the news in the summer of 1913), and the fact that another picture related to US–Mexico relations is next in the series, make me feel that this is Reginald.

Some members were willing to coordinate their metadata activities with the community and sought feedback from the community and assistance on tags before adding them to the photo:

> I was trying to think of a tag that said something like "beginning of women's liberation movement" but more concise—ideas?

In addition, in a few instances, members expressed frustration with Flickr's limit on the number of tags, which precluded them from adding important metadata to the photo:

> Would there be room for more tags, I would have added "workforce."

> I tried to add a tag Kodachrome but there is a 75 tag limit.

Members also used comments to reveal relationships among different tags from a particular cultural perspective, which would not be explicated otherwise:

> The "tent" is called a lavvo. Just because it looks a bit like a yurt or a teepee, this is of course not evidence of any relation.

Not only did Flickr members use outside sources, such as Wikipedia, to cross-examine and resolve ambiguities about the content of a photo and reconstruct its context, but they also contributed knowledge to those sources. The photos and related member conversations of 10% of the photostream sample were used to begin new articles in Wikipedia:

> I started a Wikipedia entry for her, and all are welcome to add information there:
> http://en.wikipedia.org/wiki/Glenna_Smith_Tinnin

Figure 1. An example of tag and comment metadata in the Library of Congress photostream on Flickr

Quality of the folksonomy

The study examined the intrinsic and relational quality of the folksonomy tags as well as their grammatical and cognitive types. A comparison of the complete sets revealed that only 21% of the user-generated tags (4,477 tags) had a match in the TGM. The degree of overlap with WordNet was slightly higher, 33% of the tags (6,824 tags). The degree of overlap with the LCSH, however, was 45% (9,575 tags).

A manual analysis of the random sample provided a more nuanced snapshot of the quality and structure of the photostream's folksonomy. A full 15.3% of the Flickr tags were misspelled or invalid terms, including foreign words (see Figure 2). Noun terms (excluding personal nouns) and complex terms constituted approximately 49% of the set, of which 26.3% were of the Basic type (see Figure 3, Table3).

In addition, the distribution of the term types in the folksonomy differed from the term type distributions in the TGM and the LCSH. In particular, chi-square tests on the aggregated set showed that the distributions of the grammatical types were significantly ($p < 0.001$) dependent on the source (i.e., Flickr, TGM, LCSH). A multinomial regression of the types into the source confirmed significant differences for named entities and multipart or complex terms (model fit likelihood ratio: $\chi^2 = 300.35$; $p < 0.0001$; the Flickr sample was used as a baseline). In particular, the analysis showed that the chances of complex terms occurring were higher in the TGM than in the folksonomy ($p < 0.001$) but were still lower than in

the LCSH ($p < 0.001$). The folksonomy, however, contained a greater share of named entities than the TGM ($p < 0.002$) but still contained fewer than in the LCSH ($p < 0.001$; see Table 2).

Table 2. Relationships between the part of speech categories and the term source ([**]p<0.005).

Part of Speech	TGM	LCSH
Adjective	−	−
Complex term	+[**]	+[**]
Named entity	−[**]	+[**]
Not valid & Misspelled & Foreign	−	−
Noun	+	−
Verb	−	−

Note. A plus (+) sign indicates a positive relationship between the category and the source, whereas a minus (−) sign means that the term being of a particular type decreases the odds of the term being from a particular vocabulary. The Flickr sample was used as a baseline.
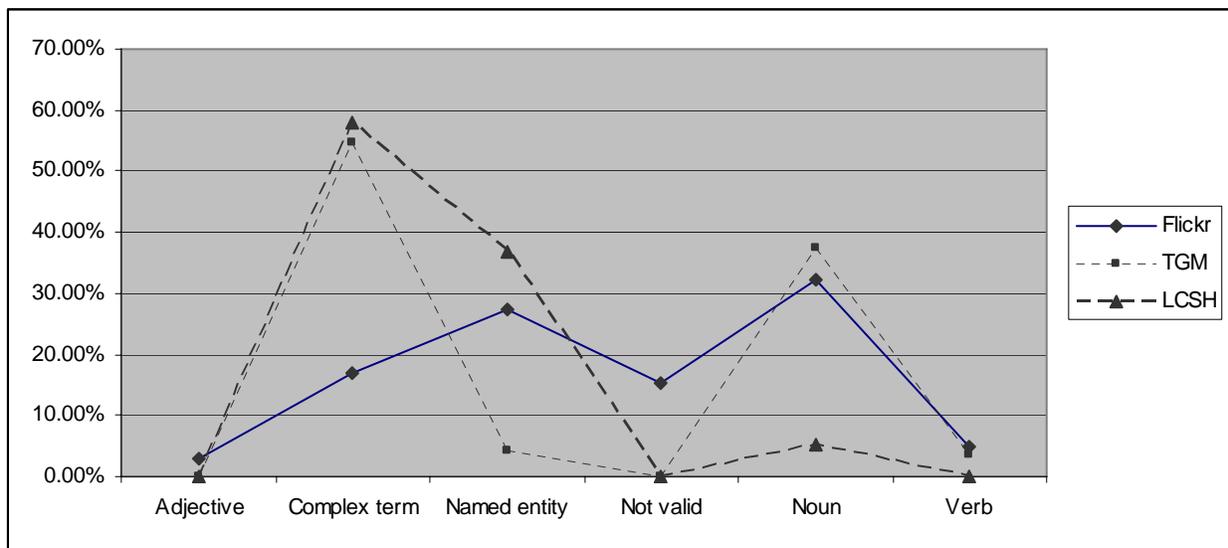

Figure 2. Part of speech categories of the terms from Flickr, the Thesaurus for Graphic Materials (TGM), and the Library of Congress Subject Headings (LCSH; based on random samples of 300 tags from each of the data sets).

To test for relationships between individual cognitive categories and source (i.e., Flickr, TMG, LCSH), the researchers used a model consisting of five categorical variables – one variable per cognitive category (i.e., Basic, Subordinate, Superodinate, NA) and one variable for source. A binary category variable was set to '1' if the term belonged to the corresponding cognitive category and '0' if it did not. Chi-square tests of the cognitive categories on source (i.e., Flickr, TGM, LCSH) showed significant dependencies for the Basic and Subordinate categories ($\chi^2 = 46.8$; $p < 0.001$; $\chi^2 = 97.7$; $p < 0.001$), but not for the Superordinate category..The results of a multinomial logistic regression of the cognitive categories into

source confirmed (model fit likelihood ratio: $\chi^2 = 187.18$; $p < 0.0001$) that the distributions of the Basic and Subordinate noun types in the folksonomy differed significantly from the distributions of these types in the TGM and the LCSH. In particular, the term being of the Basic type increased the odds of the term being from the Flickr set relative to the TGM and the LCSH ($p < 0.04$, $p < 0.03$). The term being of the Subordinate or Superordinate types, however, increased the odds of its source being from the TGM ($p < 0.01$, $p < 0.05$) or the LCSH ($p < 0.001$, $p < 0.007$) rather than the folksonomy. A comparison of the TGM and LCSH samples (with the TGM sample used as a baseline) showed that there were no statistically significant differences in the noun type distributions between these vocabularies, except for the Subordinate type. The odds of a Subordinate noun being from the LCSH were significantly higher compared with the odds of it being from the TGM ($p < 0.001$).

Thus, in general, the photostream's folksonomy contained a relatively larger share of Basic nouns and smaller shares of Superordinate and Subordinate nouns and noun phrases than the TGM and the LCSH. In addition, nouns in the TGM were more general than those in the LCSH (see Table 4).
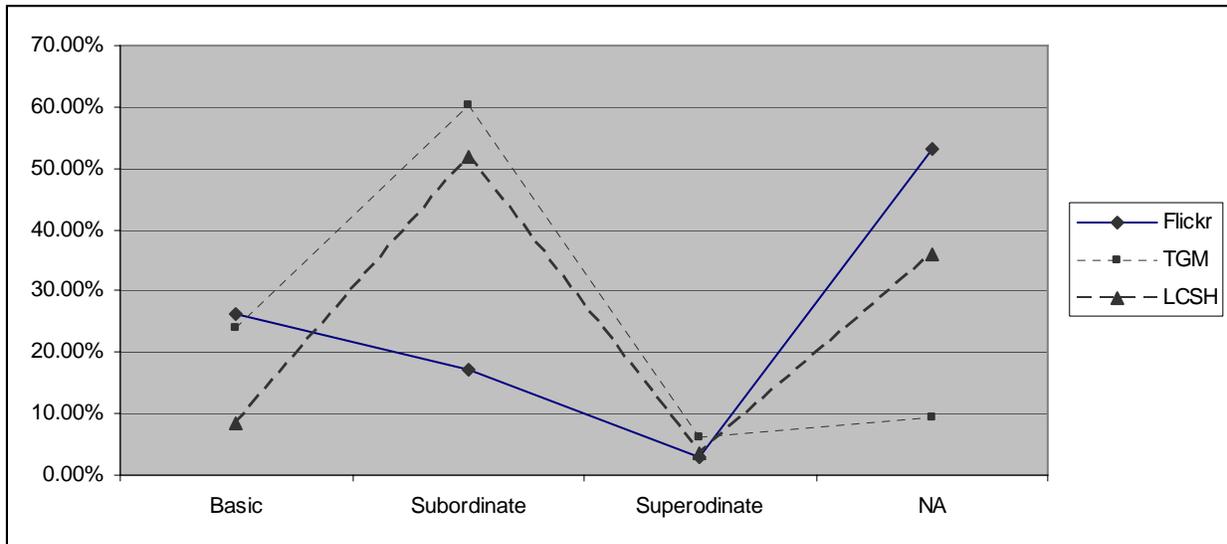

Figure 3. Cognitive categories (based on random samples of 300 terms from the Library of Congress photostream tags, the Thesaurus for Graphic Materials [TGM] and the Library of Congress Subject Headings [LCSH]—900 terms in total). Verbs, adverbs, adjectives, named entities, or invalid terms were categorized as not applicable (NA).

Table 3. Frequencies of the cognitive categories.

Cognitive categories	Flickr		TGM		LCSH	
	Frequency	%	Frequency	%	Frequency	%
Basic	79	26.3	72	24.0	25	8.3
Subordinate	52	17.3	181	60.3	156	52.0
Superordinate	9	3.0	19	6.3	11	3.7
NA	160	53.3	28	9.3	108	36.0

Note. Based on random samples of 300 terms from the Library of Congress photostream tags, the Thesaurus for Graphic Materials [TGM] and the Library of Congress Subject Headings [LCSH]—900 terms in total. Verbs, adverbs, adjectives, named entities, or invalid terms were categorized as not applicable (NA).

Table 4. Relationships between the cognitive categories and the term source relative to the folksonomy (*p<0.05).

Cognitive categories	TGM	LCSH
Basic	$-^{*}$	$-^{*}$
Subordinate	$+^{*}$	$+^{*}$
Superordinate	$+^{*}$	$+^{*}$

Note. A plus (+) sign indicates a positive relationship between the category and the source, whereas a minus (−) sign means that the noun being of a particular type decreases the odds of the term being from a particular vocabulary. The Flickr sample was used as a baseline.

In addition, the researchers manually analyzed random samples of 300 tags with no match in the controlled vocabularies. As expected, most of the no-matches (64%) to the TGM were proper nouns (named entities) and compound or complex terms (see Figure 4). The analysis also revealed regular nouns (8%) with no match in the TGM. Approximately 2% of those nouns (e.g., rod, pause, gauge) belonged to the Basic level category (see Table 5). The set of no-matches to the LCSH included 2% noun terms, none of which belonged to the Basic level category (see Figure 4, Table 5).

The distribution of grammatical types of the no-matches to the LCSH was very similar to that of the no-matches to the TGM (see Figure 4). The multiple regression (model fit likelihood ratio: $\chi^{2} = 597$; $p < 0.0001$) of the types of the no-matches into vocabularies showed significant differences only for misspellings and invalid terms. The no-matches to the LCSH had a greater shares of these types than the no-matches to the TGM ($p < 0.04$).

In addition, 5,241 tags (32%) with no match in the TGM (16,469 tags) were found in the LCSH. More than one-half of these matches were named entities. Eighteen percent of the matches were nouns, including nouns of the Basic category (based on the extrapolation from a random sample of 300 tags with no match in the TGM). In comparison, only 127 tags (1%) from the set of tags with no match in the LCSH (11,371 tags) were found in the TGM. Most of these tags were complex, multipart terms.
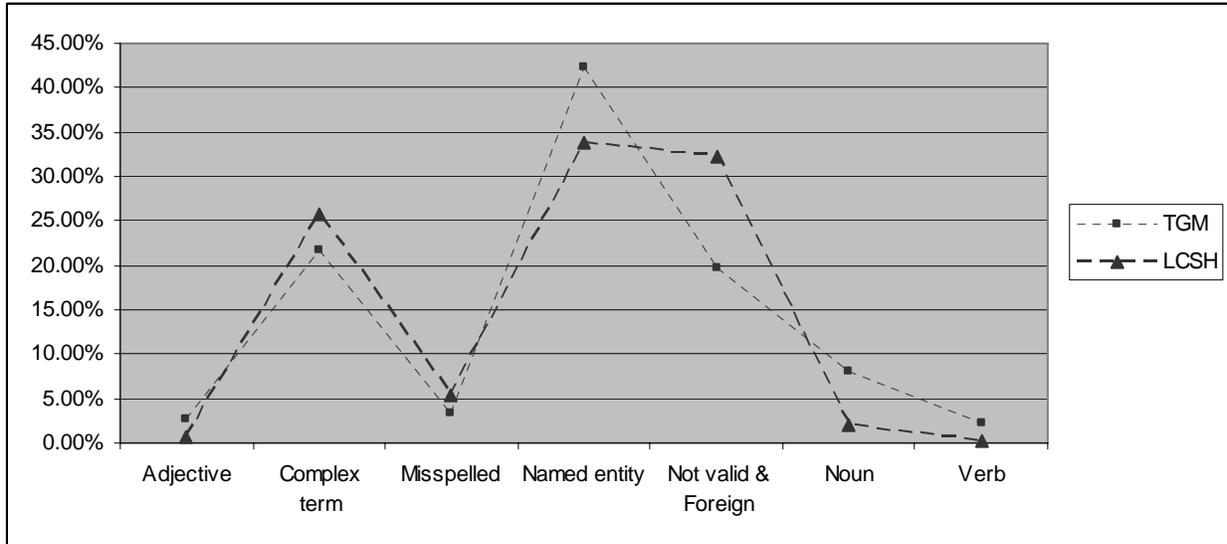
Figure 4. Grammatical categories of the tags with no match in the Thesaurus for Graphic Materials (TGM) and the Library of Congress Subject Headings (LCSH; based on two random samples of 300 tags with no match with the TGM and the LCSH).

Table 5. Categories of noun tags with no match in the TGM and LCSH.

Cognitive categories	TGM		LCSH	
	Frequency	%	Frequency	%
Basic	5	1.7	0	0.0
Subordinate	78	26.0	79	26.3
Superordinate	4	1.3	2	0.7
NA	213	71.0	219	73.0

Note. Based on a random sample of 300 tags with no match in the TGM and the LCSH. Only nouns and noun phrases were categorized. The rest were categorized as not applicable (NA). These included verbs, adjectives, named entities, misspellings, and foreign terms.

Finally, the analysis identified valid terms in the Flickr set that had no match not only in the TGM and LCSH, but also in the English Wikipedia. As expected, because of the historical nature of the Library of Congress collection, the terms represented entities and concepts from the past, such as Playograph, Grid Graph, and Aerocar.

Discussion

The first facet of the research sought to identify the types of activities Flickr members carried out around photographs, and information and knowledge resources used in those activities. The analysis revealed that, in addition to evaluating and critiquing artistic and technical aspects of photographs, members actively sought and collected information about the content of photographs. Often this was done through spontaneous collective sensemaking. As members tried to make sense of the past actions, events, and

activities depicted in the photos to reenact the context of a photograph and build and support plausible interpretations, in the process, they incrementally brought in and integrated different pieces of information. These pieces of information, then encoded in a form of metadata (comments, tags, notes), could be used to interpret other related photos.

Although in one case in the sample, a member asked for feedback from the community on a tag before assigning a tag to the photograph, the Flickr work organization model does not encourage collaboration and work coordination. Similar to Wikipedia's work organization model (Stvilia et al., 2008), Flickr members self-select which photograph to contribute metadata to. In contrast to Wikipedia, however, the community does not own the content (photographs and attached metadata). The content in Flickr is owned by individual members and institutions that post photographs for sharing. In Wikipedia, on the other hand, there is no individual ownership of or authorship for articles, even for articles begun and maintained by a single member. Hence, the motivation of the community to contribute content could be lower in Flickr than in Wikipedia. In addition, in the Flickr Commons, members are restricted to providing metadata, and there is no direct community mechanism for metadata quality control. Members cannot delete each other's tags. Only the owner of the photograph can remove irrelevant, redundant, or inaccurate tags. This restriction on deleting or modifying tags also implies the absence of an interdependency mechanism and an explicit need for coordinating work. There is a large body of literature on work organization, including work adaptability, coordination, and articulation (e.g., Crowston & Scozzi, 2004; Gasser 1986; Scacchi, 2005; Schmidt & Bannon, 1993). It would be interesting to see whether modifying the Flickr model and allowing members to delete each other's tags could lead to more cooperation and coordination in member activities. In addition, the analysis showed that more than 37% of the tags (including the tags with fewer than 3 characters) were invalid. Future research could test how modifying the tagging model by allowing members to delete or edit each other's tags could affect tag quality.

This study identified types of activities that led directly to knowledge creation. Some members purposefully linked and grouped related photographs, and used third-party sources to disambiguate and resolve uncertainties about the content of a photo and the context of its creation. The analysis also revealed incidents when an "externalization" of useful knowledge was triggered by the members' reflecting and reminiscing about their past experiences. The members told personal stories and uploaded related family photographs, thus indirectly adding knowledge about the entities and events depicted in or related to the photo and its context, These findings point to the importance of creating and supporting the activity "triggers," and developing mechanisms for harvesting knowledge from those activities. The findings also point to the importance of libraries and museums in encouraging and facilitating a larger variety of user activities and experiences around their collections, not merely search and retrieval or browsing. There is promising research on matching member interests to work tasks in Wikipedia (Cosley, Frankowski, Terveen, & Ried, 2007), and triggering reminiscence by using archives of personal documents and impersonal prompts (e.g., Cosley et al., 2009; Gemmell, Bell, & Lueder, 2006). An interesting extension of that research might be studying the usefulness of social reminiscence theories and technologies to facilitate metadata and knowledge creation for collections of historical photographs.

The study revealed that Flickr members used a variety of reference resources, including digital libraries hosted by traditional libraries and institutions of cultural heritage. However, members did not use library KoSs. The analysis of member conversations suggested that library KoSs could be helpful in guiding

metadata creation and evaluation activities. At least in one instance, the member explicitly asked for help from the community to suggest an effective term for a concept. Future work might involve studying whether more awareness and easier access to knowledge encoded in library KoSs could increase their use in entity and context disambiguation activities.

The second facet was focused on assessing the intrinsic and relational quality of the Flickr tags. The analysis showed that 37% of the original set and 15.3% of the preprocessed set (after the removal of tags with fewer than three characters and URLs) were invalid or misspelled terms. The set of invalid terms included foreign words as well. In the preprocessing step, the study used an external dictionary to split and disambiguate concatenated multipart tags. A fruitful extension of this study could be exploring the utility of multilanguage dictionaries, other reference sources such as Wikipedia, and tag co-occurrence statistics to disambiguate some of the foreign terms and explicate relationships between the foreign terms and related terms in English. If successful, this could help enhance the intrinsic quality as well as the usefulness of the folksonomy.

Research question three aimed at examining the relational quality of the photostream's folksonomy relative to controlled vocabulary maintenance. In particular, the study investigated whether the folksonomy could be a source of new terms for the TGM and LCSH. The results of the statistical analysis suggest that Flickr sets could be an important source of noun terms and named entity information for both the TGM and the LCSH. Fewer than one-fourth of the Flickr tags had a match in the TGM. The LCSH had a significantly larger share of named entity terms than the folksonomy or the TGM. Indeed, the TGM guide suggests use of LCSH terms for personal names and geographic locations. Still, the TGM and LCSH together contained less than a half of the folksonomy tags. Although significant shares of the no-matches to the vocabularies were invalid and foreign terms, more than one-half of the no-matches to the TGM were valid nouns, named entities, and complex terms. The no-matches to the LCSH had a similar distribution, but with a larger share of misspelled and invalid terms. The difference in the shares of invalid terms can be attributed to the differences in the sizes of the vocabularies. The LCSH was an order of magnitude larger than the TGM. Therefore, the total share of valid term no-matches in the LCSH sample had to be lower than the share in the TGM.

The analysis revealed named entities and terms for objects and concepts from the past that were not present in either the TGM and LCSH or in WordNet and Wikipedia. Indeed, some of the named entities and the knowledge encoded in related member discussions were used as seeds to begin new Wikipedia articles. A straightforward extension of this research could be a study of the metadata of current photographs in Flickr to determine how the change in temporary coverage of the collection could affect the relational quality of the collection's folksonomy—that is, whether the folksonomy could be a useful source of contemporary or emerging concepts and terms as well.

There has been an increase in efforts by libraries and archives to acquire metadata for their image and photo collections by deploying these collections in existing social tagging communities. However, metadata in social tagging systems are generated by different types of users in different contexts and for different purposes (Cunningham & Masoodian, 2006; Stvilia & Jörgensen, 2009). Prior studies of end-user image-searching behavior, thesaurus use, and query expansion have shown that search vocabularies of users and their perceptions of thesaurus and metadata usefulness may vary with task type and user characteristics, such as domain knowledge and familiarity with the system (e.g., Choi, 2008; Cunningham

& Masoodian, 2006; Efthimiadis, 2000; Greenberg, 2001b). Hence, in addition to assessing the number and types of suitable terms a Flickr Commons collection can add to a traditional controlled vocabulary, it is important to assess the added value of these terms to the users of that controlled vocabulary.

Conclusion

There is growing interest in, and efforts by, traditional information providers to engage social content creation and sharing communities in creating and enhancing the metadata of their photo collections to make the collections more accessible and visible. To enable and guide effective metadata creation by these communities, however, it is essential to understand the structure and patterns of member activities around the photographs and the resources used in the activities. Furthermore, it is important to assess the quality and scale of the socially created metadata against the metadata and knowledge encoded in the existing metadata records and KoSs. This study advances our understanding of Flickr member activities around historical photographs, and the quality and types of member assigned tags relative to traditional vocabularies. The findings of the study can benefit not only the Library of Congress and Flickr, but also other information institutions interested in enabling and supporting social tagging communities around their collections, or using socially created metadata to extend their knowledge organization systems.

To the best of our knowledge, this is the first study that compared the distributions of grammatical types and cognitive categories of Flickr tags to term type distributions in both the TGM and the LCSH. The study's findings suggest that the Library of Congress's photostream-generated folksonomy on Flickr can help in vocabulary translation and increase the robustness of traditional KoSs to changes in user expertise, task, and culture. Less than a half of the folksonomy terms were found in the TGM and LCSH, and more than a quarter of those terms were regular nouns and noun phrases. Flickr members assigned not only a higher share of Basic level category terms, but they often indexed photographic content with all three level terms (e.g., animal, dog, bulldog) as well as with terms in different languages and vocabularies (e.g., scientific). Extending the TGM and LCSH with user-generated terms and category relationships can make these vocabularies more accessible to different user communities, not just to information professionals. Indirectly, these could also improve the recall and robustness of image retrieval systems which use these KoSs in indexing.

At the same time, previous studies of metadata collections have shown that not all metadata are likely to be equally important for supporting a specific activity or set of activities (Greenberg, 2001b; Stvilia et al., 2004). In addition, the value of changes in metadata or knowledge tools can be measured based on changes in the value or cost of the activity's outcome. Adding new metadata or enhancing existing metadata may not necessarily translate into a value increase or cost reduction of the activity (Stvilia & Gasser, 2008). Future work may include assessing the added value of the tags and relationships from the Flickr photostream relative to the TGM and the LCSH using controlled experiments.

The study's first facet has a limitation. To identify the member activity types and resources used, it relied on only a content analysis of the logs of member comments and conversations around the photographs. Adding an analysis of other kinds of empirical data (e.g., a survey of or interviews with the members, or both) might provide a more complete representation of members' information behavior in Flickr and

could increase the reliability of the study findings. A second set of limitations relate to the evaluation of quality: the cognitive level of terms is only well established for nouns (not including named entities); this eliminates a large number of non-noun terms from consideration (and proper nouns). The cognitive level can also change with the user community's level of expertise in the area. Only one method was used to evaluate terms: comparison to other existing vocabularies; it would be beneficial to gather user input on the term utility as well, and even better to investigate within the context of specific tasks. Nevertheless, the investigation was fruitful enough to warrant suggesting further research be done in the area.

Acknowledgements

References

1. Agirre, E., Ansa, O., Hovy, E., & Martinez, D. (2000, August). *Enriching very large ontologies using the WWW.* Paper presented at the ECAI Ontology Learning Workshop, Berlin, Germany. Retrieved April 26, 2010, from http://ol2000.aifb.uni-karlsruhe.de/final/EAgirre_14.pdf

2. Ames, M., Naaman, M. (2007). Why we tag: motivations for annotation in mobile and onlinemedia. In: B. Begole, S. Payne (Eds.), *Proceedings of the SIGCHI.* (pp. 971-980). New York, NY: ACM.

3. Bailey, K. (1994). *Methods of social research* (4th ed.). New York: The Free Press.

4. Bar-Iian, J., Zhitomirsky-Geffet, M., Miller, Y., & Shoham S. (2010). The effects of background information and social interaction on image tagging. *Journal of American Society for Information Science and Technology, 61*(5), 940-951.

5. Blair, D. C. (1996). STAIRS redux: Thoughts on the STAIRS evaluation. *Journal of the American Society for Information Science, 47*(1), 4–22.

6. Brooks, T. A. (1993). All the right descriptors: A test of the strategy of unlimited aliasing. *Journal of the American Society for Information Science, 44*(3), 137–147.

7. Buckland, M. (1999). Vocabulary as a central concept in library and information science. In T. Arpanac et al (Eds.), *Proceedings of the Third International Conference on Conceptions of Library and Information Science* (pp. 3–12). Zagreb, Croatia: Lokve. Retrieved June 28, 2009, from http://www.sims.berkeley.edu/~buckland/colisvoc.htm

8. Chen, H., Yim, T., Fye, D., & Schatz, B. (1995). Automatic thesaurus generation for an electronic community system. *Journal the American Society for Information Science, 46*(3), 175–193.

9. Choi, Y. (2008). Analyzing image searching on the Web: How do undergraduates search and use visual information? In *Final Report, 2008 OCLC/ALISE Library and Information Science Research Grant Project.* Retrieved August 17, 2009, from http://library.oclc.org/cdm4/item_viewer.php?CISOROOT=/p267701coll27&CISOPTR=276

10. Chung, E., & Yoon, J. (2008). Categorical and specificity differences between user-supplied tags and search query terms for images. An analysis of Flickr tags and Web image search queries. *Information Research, 14*(3). Retrieved March 17, 2010, from http://informationr.net/ir/14-3/paper408.html

11. Cosley, D., Akey, K., Alson, B., Baxter, J., Broomfield, M., Lee, S., & Sarabu, C. (2009). Using technologies to support reminiscence. In *Proceedings of the 2009 British Computer Society Conference on Human-Computer Interaction* (pp. 480–484). Swinton, UK: British Computer Society.

12. Cosley, D., Frankowski, D., Terveen, L., & Riedl, J. (2007). Suggestbot: Using intelligent task routing to help people find work in Wikipedia. In A. Puerta & T. Lau (Eds.), *Proceedings of the 12th International Conference on Intelligent User Interfaces: IUI '07* (pp. 28–31). New York: ACM Press.

13. Croft, W. & Cruse, D. (2004). *Cognitive Linguistics.* Cambridge, MA: Cambridge University Press.

14. Crowston, K., & Scozzi, B. (2004). *Coordination practices for bug fixing within FLOSS development teams.* Paper presented at the First International Workshop on Computer Supported Activity Coordination (CSAC 2004)*,* Porto, Portugal. Retrieved April 26, 2010, from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.84.2366&rep=rep1&type=pdf

15. Cunningham, S., & Masoodian, M. (2006). Looking for a picture: An analysis of everyday image information searching. In G. Marchionini & M. Nelson (Eds.), *Proceedings of the 6th ACM/IEEE-CS Joint Conference on Digital Libraries* (pp. 198–199). New York: ACM Press.

16. Dervin, B. (1992). From the mind's eye of the user: The sense-making qualitative-quantitative methodology. In J. D. Glazier and R. Powell (Ed.), *Qualitative research in information management.* (pp. 61–84). Englewood, CO: Libraries Unlimited.

17. Efthimiadis, E. N. (2000). Interactive query expansion: A user-based evaluation in a relevance feedback environment. *Journal of the American Society for Information Science, 51*(11), 989–1003.

18. Engestrom, Y. (2000). Activity theory as a framework for analyzing and redesigning work. *Ergonomics, 43*(7), 960–974.

19. Flickr. (2009). http://www.flickr.com

20. Flickr Blog. (2009). Retrieved June 28, 2009, from http://www.sims.berkeley.edu/~buckland/colisvoc.htm

21. Furnas, G., Landauer, T. K., Gomez, L. M., & Dumais, S. (1987). The vocabulary problem in human–system communication. *Communications of the ACM, 30*(11), 964–971.

22. Garfinkel, H. (1967). *Studies in ethnomethodology.* Englewood Cliffs, NJ: Prentice-Hall.

23. Gasser, L. (1986). The integration of computing and routine work. *ACM Transactions on Office Information Systems, 4*(3), 225–250.

24. Gemmell, J., Bell, G. & Lueder, R. (2006). MyLifeBits: A personal database for everything. *CACM, 49*(1):88–95.

25. Golder, S. & Huberman, B. (2006). Usage patterns of collaborative tagging systems. *Journal of Information Science, 32*(2): 198–208.

26. Greenberg, J. (2001a). Optimal query expansion (QE) processing methods with semantically encoded structures thesauri terminology. *Journal of the American Society for Information Science and Technology, 52*(6), 487–498.

27. Greenberg, J. (2001b). Quantitative categorical analysis of metadata elements in image applicable metadata schemas. *Journal of the American Society for Information Science and Technology, 52,* 917–914.

28. Greenberg, J. (2002). Semantic web construction: An inquiry of authors' views on collaborative metadata generation. In *Proceedings of the International Conference on Dublin Core and Metadata for e-Communities* (pp. 45–52). Florence, Italy: Firenze University Press. Retrieved April 26, 2010, from http://dcpapers.dublincore.org/ojs/pubs/article/view/693/689

29. Hersh, W., Pentecost, J., & Hickam, D. (1996). A task-oriented approach to information retrieval evaluation. *Journal of the American Society for Information Science, 47*(1), 50–56.

30. Hjørland, B. (2008). What is knowledge organization (KO)? *Knowledge Organization, 35*(2/3), 86–101.

31. Jörgensen, C. (1995a). Classifying images: Criteria for grouping as revealed in a sorting task. In R. Schwartz (Ed.), Proceedings of the 6th ASIS SIG/CR Classification Research Workshop: Vol. 6. Advances in classification research (pp. 45–64). ASIS Monograph Series. Medford, NJ: Information Today.

32. Jörgensen, C. (1995b). *Image attributes: An investigation*. Unpublished Ph.D. thesis, Syracuse University, Syracuse, NY.

33. Jörgensen, C. (1996, October 8). *Indexing images: Testing an image description template*. Paper delivered at the 59th Annual Meeting of the American Society for Information Science, Baltimore, MD.

34. Jörgensen, C. (1997). Challenges in image classification. In E. Efthimiadis (Ed.), *Advances in classification research: Vol. 8. ASIS monograph series: Proceedings of the 8th ASIS SIG/CR Classification Research Workshop* (pp. 92–98).

35. Jörgensen, C. (1998). Image attributes in describing tasks: An investigation. *Information Processing and Management*, *34*(2/3), 161–174.

36. Jörgensen, C., & P. Jörgensen (2002, January 22). Testing a vocabulary for image indexing and ground truthing. In G. B. Beretta & R. Schettini (Eds.), *Internet Imaging III: Vol. 4672. Proceedings of SPIE* (pp. 212–215). Bellingham, WA: SPIE.

37. Jörgensen, C., Stvilia, B., & Jörgensen, P. (2008). *Is there a role for controlled vocabulary in taming tags?* Paper presented at the 19th Workshop of the American Society for Information Science and Technology Special Interest Group in Classification Research, Columbus, OH.

38. Klavans, J. L., Sheffield, C., Abels, E., Lin, J., Passonneau, R., Sidhu, T., & Soergel, D. (2009). Computational linguistics for metadata building (CLiMB): Using text mining for the automatic

identification, categorization, and disambiguation of subject terms for image metadata. *Multimedia Tools and Applications 42*(1), 115–138.

39. Kuutti, K. (1991). Activity theory and its applications to information systems research and development. In H.-E. Nissen (Ed.), *Information systems research* (pp. 529–549). Amsterdam: Elsevier Science.

40. Lakoff, G. (1988). Cognitive semantics. In U. Eco, M. Santambrogio, & P. Violi (Eds.), *Meaning and mental representations* (pp. 119-154). Bloomington, IN: Indiana University Press.

41. Lancaster, F. (2000). *Indexing and abstracting in theory and practice.* Champaign: University of Illinois, Graduate School of Library and Information Science.

42. Leontiev, A. (1978). *Activity, consciousness, personality*. Englewood Cliffs, NJ: Prentice-Hall.

43. Lin, X., Beaudoin, J. E., Bui, Y., & Desai, K. (2006). Exploring characteristics of social classification. In J. Furner & J. T. Tennis (Eds.), *Advances in classification research: Vol. 17. Proceedings of the 17th ASIS&T Classification Research Workshop*. Retrieved August 17, 2009, from http://dlist.sir.arizona.edu/1790/01/lin.pdf

44. Mathes, A. (2004). *Folksonomies: Cooperative classification and communication through shared metadata*. Retrieved August 17, 2009, from http://www.adammathes.com/academic/computer-mediated-communication/folksonomies.html

45. Medelyan, O., & Milne, D. (2008). *Augmenting domain-specific thesauri with knowledge from Wikipedia.* Paper presented at the NZ Computer Science Research Student Conference (NZCSRSC 2008), Christchurch, New Zealand. Retrieved April 26, 2010, from http://www.cs.waikato.ac.nz/~olena/publications/nzcsrsc08_medelyan_milne.pdf.

46. Nov, O., Naaman, M., & Ye, C. (2010). Analysis of participation in an online photo-sharing community: A multidimensional perspective. *Journal of American Society for Information Science and Technology, 61*(3), 555-566.

47. Perez, J. (2009). Google rolls out semantic search capabilities. *PCWorld.* Retrieved August 17, 2009, from http://www.pcworld.com/businesscenter/article/161869/google_rolls_out_semantic_search_capabiliti es.html

48. Porter, M. (1980). An algorithm for suffix stripping. *Program, 14*, 130–137.

49. Rorissa, A., & Iyer, H. (2008). Theories of cognition and image categorization: What category labels reveal about basic level theory. *Journal of the American Society for Information Science and Technology, 59*(9), 1383–1392.

50. Rosch, E. (1973) Natural categories. *Cognitive Psychology 4*, 328-350.

51. Rosch, E. (1978). Principles of categorization. In E. Rosch & B. Lloyd (Eds.), *Cognition and categorization (pp. 27-48)*. Hillsdale, NJ: Erlbaum.

52. Rosch, E., Mervis, C., Gray, W., Johnson, D., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology, 8*(3), 382–439.

53. Scacchi, W. (2005). Socio-technical interaction networks in free/open source software development processes. In S. T. Acuña and N. Juristo (Ed.), *Software process modeling*. New York: Springer.

54. Schmidt, K., & Bannon, L. (1993). Taking CSCW seriously: Supporting articulation work. *Computer Supported Cooperative Work, 1*(1–2), 7–40.

55. Schmitz, P. (2006). Inducing ontology from Flickr tags. Paper presented at the World Wide Web 2006 (WWW2006): Collaborative Web Tagging Workshop. Retrieved April 26, 2010, from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.130.4081&rep=rep1&type=pdf

56. Shatford, S. (1986). Analyzing the subject of a picture: A theoretical approach. *Cataloging and Classification Quarterly, 6*(3), 39–62.

57. Soergel, D. (1974). *Indexing languages and thesauri: Construction and maintenance*. Los Angeles: Wiley.

58. Springer, M., Dulabahn, B., Michel, P., Natanson, B., Reser, D., Woodward, D., and Zinkham, H. (2008). *For the common good: The Library of Congress Flickr pilot project.* Retrieved August 19, 2009, from http://www.loc.gov/rr/print/flickr_report_final.pdf

59. Strong, D., Lee, Y., & Wang, R. (1997). Data quality in context*. Communications of the ACM, 40*(5), 103–110.

60. Stvilia, B., & Gasser, L. (2008). Value based metadata quality assessment. *Library and Information Science Research, 30*(1), 67–74.

61. Stvilia, B., Gasser, L., Twidale, M., Shreeves, S., & Cole, T. (2004). Metadata quality for federated collections. In S. Chengulur-Smith, L. Raschid, J. Long, & C. Seko (Eds.), *Proceedings of the International Conference on Information Quality—ICIQ 2004* (pp. 111–125). Cambridge, MA: MITIQ.

62. Stvilia, B., Gasser, L., Twidale M. B., & Smith, L. C. (2007). A framework for information quality assessment. *Journal of the American Society for Information Science and Technology, 58*(12), 1720–1733.

63. Stvilia, B., & Jörgensen, C. (2009). User-generated collection level metadata in an online photo-sharing system. *Library & Information Science Research*. *31*(1), 54–65.

64. Stvilia, B., Twidale, M., Smith, L. C., & Gasser, L. (2008). Information quality work organization in Wikipedia. *Journal of the American Society for Information Science, 59*(6), 983–1001.

65. Svenonius, E. (1986). Unanswered questions in the design of controlled vocabularies. *Journal of the American Society for Information Science, 37*(5), 331–340.

66. Tan, Y., Kan, M., & Lee, D. (2006). Search engine driven author disambiguation. In *Proceedings of ACM/IEEE Joint Conference on Digital Libraries* (pp. 314-315). New York, NY. ACM.

67. Tanaka, J. & Taylor, M. (1991). Object categories and expertise: is the basic level in the eye of the beholder? *Cognitive Psychology, 23*(3), 457–82.

68. Taylor, R. (1986). *Value-added processes in information systems*. Norwood, NJ: Ablex Publishing.

69. Trant, J. (2008). *Keynote address*. Paper presented at the Dublin Core Annual Meeting, Humboldt-University, Berlin, Germany. Slides retrieved August 17, 2009, from http://conference.archimuse.com/blog/jtrant/tagging_and_folksonomy_keynote_dc2008

70. Trant, J. (2009). Studying social tagging and folksonomies: A review and framework. *The Journal of Digital Information, 10*(1). Retrieved August 17, 2009, from http://dlist.sir.arizona.edu/2595/

71. von Ahn, L., & Dabbish, L. (2004). Labeling images with a computer game. In *Proceedings of the CHI 2004* (pp. 319–326). New York: ACM Press.

72. Wacholder, N., & Liu, L. (2006). User preference: A measure of query-term quality. *Journal of the American Society for Information Science and Technology, 57*(12), 1566–1580.

73. Wacholder, N., & Liu, L. (2008). Assessing term effectiveness in the interactive information access process. *Information Processing and Management, 44*(3), 1022–1031.

74. Weick, K. (1995). *Sensemaking in organizations*. Thousand Oaks, CA: Sage.

75. Weick, K., Sutcliffe, K., & Obstfeld, D. (2005). Organizing and the process of sensemaking. *Organizational Science, 16*(4), 409–421.

76. Yi, K., & Chan, L. M. (2009). Linking folksonomy to Library of Congress Subject Headings: An exploratory Study. *Journal of Documentation, 65*(6), 872-900