A Comparative Study of Flickr Tags and Index Terms in a General Image Collection

Abebe Rorissa
Department of Information Studies, State University of New York at Albany, Albany, NY 12222.
E-mail: arorissa@albany.edu

Web 2.0 and social/collaborative tagging have altered the traditional roles of indexer and user. Traditional indexing tools and systems assume the top-down approach to indexing in which a trained professional is responsible for assigning index terms to information sources with a potential user in mind. However, in today's Web, end users create, organize, index, and search for images and other information sources through social tagging and other collaborative activities. One of the impediments to user-centered indexing had been the cost of soliciting user-generated index terms or tags. Social tagging of images such as those on Flickr, an online photo management and sharing application, presents an opportunity that can be seized by designers of indexing tools and systems to bridge the semantic gap between indexer terms and user vocabularies. Empirical research on the differences and similarities between user-generated tags and index terms based on controlled vocabularies has the potential to inform future design of image indexing tools and systems. Toward this end, a random sample of Flickr images and the tags assigned to them were content analyzed and compared with another sample of index terms from a general image collection using established frameworks for image attributes and contents. The results show that there is a fundamental difference between the types of tags and types of index terms used. In light of this, implications for research into and design of user-centered image indexing tools and systems are discussed.

Introduction

The bulk of digital information produced every year is visual in nature (Lyman & Varian, 2003). It is estimated that there are over 10 billion images on the World Wide Web alone, most of which are created and published by individuals. This estimate excludes those in stock or publicly inaccessible personal photo collections. It is no surprise that only a small fraction of these images is indexed for effective and efficient retrieval. Social/collaborative tagging and photo sharing services such as Flickr (http://www.flickr.com) pose new challenges and present tremendous opportunities for designers of indexing and retrieval systems while demanding dynamic sets of solutions (Enser, 2008). End users are also creators and indexers of multimedia information sources through activities such as tagging. Therefore, partly because users are becoming accustomed to this level of involvement, the problem of how to incorporate user-generated tags into the process of indexing and retrieval needs urgent attention.

Evidence from previous research shows that, though images are visual information sources with little or no associated text, image users mainly use text to formulate their queries (O'Connor, O'Connor, & Abbas, 1999; Roberts, 2001; Tjondronegoro, Spink, & Jansen, 2009), which is a result of the lack of systems that offer other options. The vast majority of Web search engines that allow image searches provide search interfaces that accept only text queries. Keywords, subject headings, and text annotations are integral parts of both concept-based and content-based image indexing and retrieval systems (Enser, 2000, 2008; Jörgensen, 1995, 1996, 1998, 2003).

Some of the annotations by content-based image retrieval systems can be achieved through automated means, reducing expenses. However, even with the apparent limitations of concept-based image retrieval, such as the cost (Layne, 1994) and the variable precision of manual annotations, content-based image retrieval systems have failed thus far to bridge the semantic gap between indexer and end user (Datta, Li, & Wang, 2005; Enser, 2008; Lew, Sebe, Djeraba, & Jain, 2006; Neumann & Gegenfurtner, 2006; Tjondronegoro et al., 2009). Because words are the only means for describing the semantic content of images, at least for the foreseeable future, image retrieval systems will depend on text annotations, especially for image collections on social networking and photo-sharing services such as Flickr where users add textual tags and descriptions to images.

The nature and underlying structure of image tags and index terms assigned to images in a general image

collection is the focus of this study. Wolfram, Olson, and Bloom (2009) emphasized the importance of this sort of work when they stated that "it is worthwhile to explore potential parallel characteristics between indexing and tagging as these characteristics could influence the form of tagging offered in a given situation and how tagging influences retrieval" (p. 1996) and called for a better understanding of the nature of tags. Others (Bar-Ilan, Shoham, Idan, Miller, & Shachak, 2008) have also called for studies on the strengths and shortcomings of tags and conventional indexing methods. To clearly define the research problem, a review of the phenomenon of social tagging and literature on indexing in general and image indexing in particular is necessary.

Background and Previous Research

Social Tagging and Flickr

Social tagging, also called collaborative tagging, is the process of assigning tags (keywords) freely to organize individual and shared content for future retrieval (Golder & Huberman, 2006). In social tagging, no one person plays the librarian's role, and services such as Flickr allow users to not only organize their own resources, but also to share tags as well as resources among themselves, a characteristic of social networking (Golder & Huberman, 2006). Nevertheless, users tag mostly their own images, doing so for their own benefits and future use.

Flickr (http://www.flickr.com) is a popular photo sharing service that allows individuals to upload and tag photos and videos. These individuals may invite others (such as friends or family) to view and tag the images as well. It should be noted that Flickr tags are not mandatory, and that the Flickr system allows the modification of previously assigned tags. Users often create *folksonomies* (folk classifications or folk taxonomies), which are not true classifications in the sense that terms do not have clearly defined, hierarchical relationships (Jacob, 2004). The problem with social tagging (as it relates to traditional indexing and classification schemes) and user-generated descriptions is, in general, the imperfect semantic relations of words to their referents (Golder & Huberman, 2006). On the other hand, social tagging provides indexing-system designers with the opportunity of a rich end-user vocabulary, and the bulk of the workload is shared among users in a distributed and networked environment. Its multicultural and multilingual set of users and tags is an ideal test bed for cross-language image retrieval, and information professionals can take advantage of the inherent motivation of tagging for oneself to involve users in the indexing process (Matusiak, 2006).

Despite recent critical assessments of the optimistic views about and impact of social tagging and Flickr (Cox, 2008), social tagging is more useful and crucial for some collections than others. For instance, images in the cultural heritage domain could be tagged and annotated by users knowledgeable about local history and language (Matusiak, 2006). A few museums have seized the opportunity presented by social tagging (Bearman & Trant, 2005; Trant, 2006) to improve indexing and access to their image and art collections. Established cultural heritage institutions such as the Metropolitan Museum of Art are already testing tagging by untrained professionals (Bearman & Trant, 2005; Trant, 2006).

Some (e.g., Macgregor & McCulloch, 2006) argue that social tagging and professional indexing will have their own separate but complementary roles to play in information organization, access, and retrieval in the future. For instance, tagging and tags could be beneficial to users when they need to browse for information resources, a process more serendipitous than either a directed search or a known item search. An information organization system that relies on controlled vocabularies for indexing purposes might yield different results (Mathes, 2004).

Indexing and the Change in the Dynamics of Indexing

Indexing is a process of creating document surrogates and its main goal is to facilitate retrieval of information relevant to users' queries. There are two broad categories of the process of indexing: document-oriented and user-oriented. Thesauri built for user-oriented indexing contain concepts that are collected (a priori) from potential users or their expected queries (Soergel, 1995). Whereas the document-oriented approach is widely used, the utility of the user-oriented approach is limited partly because it is not practical to involve users in a priori indexing (Fidel, 1994). However, today's information users are not passive consumers of information. They are actively creating and organizing their own resources for personal use and for use by others.

Advocates of user-centered indexing suggest involving users in the indexing and organization process by encouraging them to supply index terms reflective of their future information need (Fox, 2006). When it comes to user involvement in indexing, one of the major obstacles is motivation on the part of the user. Creating an environment similar to services that offer social tagging could be a possible solution. This is, at least, a much better alternative to user-centered approaches that rely on a description of potential users to select suitable index terms (Fidel, 1994). Social tagging works because individuals are invested in tagging their own resources (e.g., photos) for their own retrieval purposes.

When calls for user-centered indexing were made in the past (Fidel, 1994; Hastings, 1999; Jörgensen, 1998; Layne, 1994), the assumption, in most cases, was that information professionals would choose index terms and maintain indexing tools and systems, while users would have a minimal role or no role. These assumed dynamics have changed with the proliferation of social networking services and the Web 2.0 phenomenon (Peters, 2009). Users are now creators of content and, through the creation of social tags and folksonomies, they have also assumed the role of indexers, albeit in an uncontrolled environment. The tags and folksonomies used on services such as Flickr are not the keywords and hierarchies assigned by professionals using controlled vocabularies (e.g., thesauri, subject headings) such as those used to index library collections. This reveals new challenges and

opportunities—for designers of information indexing and retrieval systems and users alike.

Although traditional indexing is done by professionals (e.g., librarians) for the benefit of potential users, social tagging is done by users for their own benefits. Indexers who rely on controlled vocabularies are reluctant to incorporate users' vernacular language. Impediments include indexers' training and experience (Bearman & Trant, 2005) and their skepticism of social tagging (Trant, 2006). Because users' language does not match that of the indexers, the utility and precision of indexing systems cannot be maximized (Jörgensen, 1999).

The use of user-generated tags or descriptions of images and other documents for indexing purposes has been criticized because of the imprecision and inaccuracy in users' language. It has been acknowledged that this weakness is somewhat unavoidable due to the free (not controlled) and open nature of social tagging (Guy & Tonkin, 2006; Hammond, Hannay, Lund, & Scott, 2005). On the other hand, this weakness is also the strength of social tagging; it captures the width of the users' vocabulary. A serious drawback to the use of user-generated tags for indexing, a result of its free nature, is increased recall at the expense of precision in retrieval. However, though repeatedly used tags increase indexing exhaustivity and recall, those that are rarely used increase indexing specificity and precision in retrieval. The unstructured nature of tags is seen by some as something that makes them useful. For instance, "tags are probably easier to enter than picking categories from an ontology; tags allow for greater flexibility and variation; and tags may naturally evolve to reflect emergent properties of the data" (Rattenbury, Good, & Naaman, 2007, p. 103). This is an important property of user-generated tags, which should be taken into consideration if they are to be used for indexing purposes, as is the *power law* of social tags (most used tags are likely to be used by others and the majority of tags are likely to be used by few; Mathes, 2004).

Tagging on Flickr is even being suggested as a third alternative type of image indexing, in addition to content-based and concept-based indexing (Schmidt & Stock, 2009). Evidence of the change in the dynamics of indexing is apparent in how social tagging addresses some of the concerns raised by Fidel (1994) regarding indexing policy. For instance, issues concerning sources of index terms and user language are readily resolved because tags are supplied by users. As for the specificity and accuracy of translating concepts into index terms, an assumption could be made that no translation is necessary because users are likely to use their tags as index terms for search and retrieval purposes. Further evidence of the change in the dynamics shows in the changing role of the indexer and requester. Social tagging has the potential for improving what Cooper (1969) terms "indexer-requester consistency" because the user who, in the indexer role, supplies the tags (to be used as index terms) is most likely to be the requester as well as others who share interests and, perhaps, vocabulary for those interests.

As social tagging evolves, the quality of tags may improve and features could be added to indexing systems in order to correct misspellings and avoid inclusion of inappropriate tags (Guy & Tonkin, 2006). Through user-generated tags, there is a real opportunity for a bottom-up approach to indexing, rather than the predominantly top-down approach previously taken by digital libraries. What is more, user-generated tags are more robust and descriptive than index terms assigned using vocabulary control tools (O'Connor & O'Connor, 1999). For the first time, enough user-generated tags are also available for training automatic indexing algorithms (von Ahn & Dabbish, 2004). The assumption may be made that, even though user-generated tags are sometimes imprecise and ambiguous, the same tags assigned to images are likely to be used at the time of searching. The claim that players of an image labeling game, *ESP*, (von Ahn & Dabbish, 2004) supplied about 85% useful and descriptive terms that correctly described images should be encouraging news and is contrary to what some critics were pointing to as weaknesses of social tagging as an indexing alternative.

For the purpose of the current work, image indexing is defined as the representation of images through textual descriptions of their contents, what they are *of* or *about* (Shatford, 1984), rather than their *isness* (Schmidt & Stock, 2009). An image may be of and about two different things. Although the *ofness* of an image is mostly concrete and objective, its *aboutness* is more abstract and subjective (Layne, 1994). A picture of a building represents an object, the building. Many viewers of a picture attach meanings to what the picture represents beyond the concrete object, and any indexing tool and system must be able to cater for the full range of attributes, perceived and concrete or otherwise. The language of end users should be incorporated as well.

Unlike the problems inherent in indexing text documents, those in image indexing are compounded by their visual nature. One of the limitations of concept-based image indexing and retrieval is the cost of manual assignment of index terms (Layne, 1994). Given that a picture is worth a thousand words, the question of how many of those would it take to adequately index an image is still open to debate. Still, textual descriptions of image attributes are popular methods of indexing (Enser, 2000; Jörgensen, 1995, 1996, 1998, 2003), and most image users formulate their queries using words, though this is neither always the necessary nor the only way (O'Connor et al., 1999). Furthermore, images, works of art included, are information sources first. In the words of Shatford (1984, p. 15), a work of art's "aesthetic value is a fortuitous by-product." We are not suggesting that the aesthetic value of images is any less important. Images, as sources of information, require indexing tools and systems that incorporate the language of the end users. Social tagging may provide solutions to the problems of creating such tools and systems.

The Problem

Previous research has explored the types and categories of attributes and their values (the actual terms) used to describe images (Jörgensen, 1995, 1998). Those used to search for

images have also been explored (Armitage & Enser, 1997; Choi & Rasmussen, 2003; Enser & McGregor, 1992; Hastings, 1995; Jörgensen & Jörgensen, 2005). Previous researchers have also questioned the adequacy of traditional indexing tools and systems (Jörgensen, 1998, 1999). However, due to the relative novelty of social tagging, research projects similar to the current study are few and far between (Spiteri, 2007), although this is beginning to change. While a few researchers have tackled the topic of social tagging of images in depth (e.g., Stvilia & Jörgensen, 2009), some only focused on general issues and considered only a few images and tags for analysis (e.g., Matusiak, 2006). So far, only a handful of researchers (Jörgensen, 1998, 1999; Trant, 2006) have embarked on research similar to the current study, in which the focus of the investigation is the difference between user-generated tags or descriptions of images and the index terms assigned to images in general image collections using traditional indexing tools. Even though these studies have all made significant contributions, few were based on large-scale image collections and a large number of users.

Despite the phenomenon's popularity, the nature of social tagging and tags is not a well-researched area and comparative studies with respect to image tags such as those on Flickr and indexing in general image collections are nonexistent. According to Macgregor and McCulloch (2006), "Several authors have documented their thoughts on collaborative tagging but few have done so via the scholarly literature" (p. 295).

What is more, the general problem of representation of and access to images still remains only partially solved due to the inadequacy of current image indexing tools and systems. Visual content and meaning extracted from images are rarely considered in traditional indexing systems (Jörgensen, 1998, 1999, 2003). Traditional indexing systems follow the top down approach and were designed neither with the users' language in mind nor to cover the full range of image attributes (Jörgensen, 1998; Roberts, 2001). According to Jörgensen (1998, p. 161), "the vast majority of image indexing systems are text-based systems that have been crafted from a number of pragmatic or theoretical considerations and assumptions about user needs and search strategies." Not much has changed in the last 10 years.

There is no consensus as to which attributes of images should be indexed (Chen & Rasmussen, 1999; Jörgensen, 1995; Layne, 1994). The main reason is that little is known about how users describe and tag images. The proliferation of social tagging services specifically for images presents an ideal opportunity for researchers and practitioners alike to attempt to solve the image representation problem by involving users in the design of indexing tools, thereby bridging the semantic gap between index terms assigned by professional indexers and terms used by laypersons (Trant, 2006).

Irrespective of whether some in the image research community see social tagging on Flickr as a challenge (Matusiak, 2006), even to the extent proportional to the number of Flickr users, or an opportunity (Matusiak, 2006; Trant, 2006) to improve image indexing, the nature of social tagging is

a research area that needs immediate attention due to its dynamic nature, its popularity, and the tremendous opportunity it affords those striving to design user-centered image indexing and retrieval systems. To inform system design for the indexing and retrieval of images from collections across the board (the Web, social networking services, digital libraries, etc.), a continued investigation as to what types of features are described through tagging is crucial.

The constant evaluation and improvement of current image indexing tools and systems can only be achieved through empirical research on the nature of image attributes, including user-generated tags, users' language, and user's indexing systems. Social tagging could be better understood with the help of comparative studies that also include either controlled vocabularies themselves or information and knowledge organization systems based on controlled vocabularies (Macgregor & McCulloch, 2006). In light of this, the current study seeks answers to the research question: What is the extent of the similarities and differences between Flickr tags and controlled index terms used in general image collections? Therefore, the main objective of the current study is to identify the structure of tags used for describing images on Flickr and empirically test the difference between that and the structure of index terms in general image collections according to categories of attributes of images in frameworks established by previous research.

Method

Materials

Approximately 186,000 unique Flickr users/owners uploaded 2.2 million images and assigned 4.8 million tags to them during the first seven days of June 2007. This pool served as a population of Flickr tags for the current study, and a Flickr API (application programming interface) was used to download records which were later imported into a database for preprocessing and data analysis. At the same time, there were approximately 300,000 images in the University of St. Andrews Library Photographic Archive (http://special.st-andrews.ac.uk/saspecial/index.php). Please see Reid (1999) for a detailed discussion of the University of St. Andrews Library Photographic Archive.

Materials for this work included a random sample of 4,441 Flickr tags, mainly English-language tags, assigned to 1,000 photos from Flickr. These were compared to a random sample of 3,709 index terms assigned to 996 photos from the University of St. Andrews Library Photographic Archive. Not all Flickr tags were legible or recognizable as words or known abbreviations (e.g., del, dêm, Đình, viau, w04). As a result, only those tags that the author and a research assistant could identify were coded and the final set consisted of 4,159 tags assigned to 975 photos.

Data Analysis

A content analysis of the Flickr tags and index terms was undertaken to code them according to two of the various

frameworks of image attributes often used in image research (Enser & McGregor, 1992; Jörgensen, 1998). Although automatic analysis of tags is common, content analysis of tags done manually is rare. Although comparison of individual tags and index terms may not be appropriate because of the difference in how they were assigned to images and because the collections of images were different, established frameworks previously utilized such as those by Jörgensen (1998) and Enser & McGregor (1992) serve as a basis for comparison. The two frameworks were chosen because each has been widely applied by image researchers to study descriptions of images (Jörgensen, 1998) and image queries supplied by people (Enser & McGregor, 1992) in a number of contexts and environments.

Enser and McGregor (1992) classified image queries and attributes into four categories: unique, unique with refiners, nonunique, and nonunique with refiners. They built these through a bottom-up approach where they analyzed about 2,700 queries submitted by a variety of users to the Hulton Deutsch picture archive (which was later incorporated into the collections of Getty images). Their framework was originally devised to investigate users' image search queries, and it is appropriate specifically for the study of tags because tagging an image or any other resource for future discovery and use is similar to thinking of keywords to be used in a search (Chi & Mytkowicz, 2008; Furnas, Landauer, Gomez, & Dumais, 1987).

Jörgensen's (1998) 12 categories of image attributes were a result of an analysis of data obtained through a number of tasks by users (a description/viewing task, a search task, and a memory task). Whereas the description/viewing task yielded unconstrained image descriptions from human subjects, the search task required them to view an image and come up with a textual query to retrieve the image from an ideal image retrieval system. The memory task required the subjects to recall the images in the viewing task and write their descriptions. Jörgensen's (1998) final analysis of data from the three tasks resulted in 12 categories of image attributes, namely objects, people, color, visual elements, location, description, people-related attributes, art historical information, abstract concepts, content/story, external relations, and viewer response. Jörgensen (1998) further categorized the 12 categories into perceptual (objects, people, color, visual elements, location, and description), interpretive (people-related attributes, art historical information, abstract concepts, content/story, and external relations), and reactive (viewer response) attributes.

The use of classification frameworks of image attributes to study tags assigned to visual resources has been recommended by others (Angus, Thelwall, & Stuart, 2008). For instance, Jansen (2008) analyzed image search queries on the Web by mapping them onto image attribute classification frameworks by, among others, Enser & McGregor (1992) and Jörgensen (1995, 1998). McCay-Peet and Toms (2009) also utilized Jörgensen's 12 categories to study users' image needs and use. Laine-Hernandez and Westman (2006) put together their classification of image attributes by

modifying and adding to Jörgensen's (1998) 12 categories of attributes.

Content analysis. In most cases, content analysis involves three steps: creation and testing of a coding scheme, including the definition of recording units and categories; assessment of the accuracy of coding (through intercoder agreement analyses); and revision of coding rules, and coding the entire text (Weber, 1990). However, because coding schemes and categories used by previous research in constructing the respective frameworks were adopted in this work, only the second and third steps were relevant. Flickr tags and index terms in the University of St. Andrews Library Photographic Archive were the recording units (basic unit of text to be coded or categorized).

Using the adopted coding schemes, all Flickr tags and index terms were coded by a research assistant. Coding of the tags and index terms was conducted by evaluating the corresponding images in Flickr and the University of St. Andrews Library Photographic Archive. The author then coded, independently, a random sample (20%) of Flickr tags and index terms. Intercoder agreement between the research assistant and the author was assessed using percent agreement and Cohen's (1960) kappa. For the sample of Flickr tags, computed values of the percent agreement and Cohen's (1960) kappa were 0.93 and 0.91 for the framework by Jörgensen (1998) and 0.99 and 0.98 for the framework by Enser and McGregor (1992), respectively. Percent agreement and Cohen's (1960) kappa values for index terms were 0.96 and 0.94 for the framework by Jörgensen (1998) and 1.00 and 0.99 for the framework by Enser and McGregor (1992), respectively. Some consider kappa values between 0.61 and 0.80 substantial, and values above 0.81 almost perfect (Landis & Koch, 1977). We concluded that the coding was highly consistent (reliable), given the fact that values over 0.70 for both measures are considered satisfactory in most cases (Neuendorf, 2002).

In addition, the chi-square and *t*-tests were used to determine whether there were similarities and/or differences in terms of the frequency and average number of Flickr tags and index terms according to the categories in the two frameworks and to determine the difference in their underlying structure.

Results

Descriptions of the Samples of Tags and Index Terms

For the current work, two separate samples were selected from the popular photo sharing service Flickr and the University of St. Andrews Library Photographic Archive during the first week of June 2007. The two samples had comparable numbers of images as well as comparable numbers of tags and index terms. One thousand nine hundred seventy-one images (975 and 996, respectively) and 7,868 Flickr tags and index terms (4,159 and 3,709, respectively) were considered for final analysis and are the basis for subsequent results

FIG. 1.   Clouds of (a) Flickr tags and (b) index terms.

TABLE 1.   Descriptive statistics for the samples of Flickr tags and index terms.

Sample	Frequency (%)	Minimum	Maximum	M	SD
Flickr tags	4159 (52.86%)	1	41	4.27[a]	4.04[a]
Index terms	3709 (47.14%)	1	10	3.73[b]	1.48[b]

[a]$n = 975$. [b]$n = 996$.

and discussions. Examples of the Flickr tags and index terms are in Figure 1 in the form of clouds; descriptive statistics for both samples are presented in Table 1.

Even though the number of tags and index terms in the sample were comparable, the number of Flickr tags assigned to an image ranged between 1 and 40; the number of index terms per archival image ranged between 1 and 10. In addition, the number of Flickr tags per image was more variable (coefficient of variation = 94.6%) than the number of index terms per image (coefficient of variation = 39.68%). This might be expected, because Flickr users have the freedom to assign as many tags as they wish to their images; professional indexers are limited by standard indexing procedures when it comes to selection of index terms. Another potential contributing factor to this difference is the fact that tagging on Flickr is social and collaborative with multiple taggers and professional indexing often involves a single individual (Schmidt & Stock, 2009).

Comparison of Tags and Index Terms Using Jörgensen's Twelve Categories

The top five attributes described by Flickr tags were location (30.63%), content/story (including events) (17.14%), people (14.98%), description (13.51%), and object (10.89%), whereas the most important attributes described by the index terms were location (43.35%), content/story/event (25.48%), object (22.51), people (4.83%), and art historical information (2.37%). Both Flickr tags and index terms rarely described the following attributes: color, visual elements, and people-related attributes. Table 2 presents the distribution of Flickr tags and index terms that describe Jörgensen's (1998) 12 image attribute categories. This contrasts with what other studies that analyzed multimedia search terms on the Web showed. Some of these studies found that the most frequent searches were for people, locations, or things/objects (Tjondronegoro et al., 2009).

Flickr tags and index terms from the St. Andrews Library Photographic Archive were generally significantly different with respect to Jörgensen's (1998) 12 individual categories of image attributes (χ^2 (9, $N = 7868$) = 1314, $p < 0.05$). On the other hand, at the aggregate level, that is at the two broad categories of perceptual and interpretive attributes, they were not significantly different with respect to the overall frequencies (χ^2 (1, $N = 7,868$) = 3.187, $p > 0.05$). However, the results for individual groups of attributes are mixed. Even though the average number of Flickr tags ($M = 1.14$) and index terms ($M = 1.09$) that describe interpretive attributes were about the same, $t(7866) = 0.863$, $p > 0.05$, there were significantly more Flickr tags ($M = 3.13$) that described perceptual attributes than index terms ($M = 2.63$) that did the same, $t(7866) = 4.43$, $p < 0.001$.

Although some previous researchers (e.g., Laine-Hernandez & Westman, 2006) found that more interpretive than perceptual descriptions were used for the images that they studied, more perceptual terms than interpretive ones were used to describe sample images in both Flickr and the St. Andrews Library Photographic Archive (see Table 3). One of the key factors contributing to this discrepancy, in addition to the sample of images, may be the type of users, at least as regards the Flickr sample. Laine-Hernandez and Westman (2006) collected their data from professional users (journalists) whereas most Flickr users are individuals tagging either photos for informal social purposes.

Comparison of Tags and Index Terms Using Enser and McGregor's Four Categories

Of the 4,159 Flickr tags, there were almost exactly the same percentage of unique (51.6%) and nonunique (48.4%) tags. In comparison, of the 3,709 index terms from the St. Andrews Library Photographic Archive, the percentage of nonunique index terms (71.66%) was much higher than the percentage of unique index terms (28.34%). In terms of the use of refiners, Flickr users added many fewer refiners (8.59%) than did indexers at the St. Andrews Library

TABLE 2. Distribution of Flickr tags and index terms according to Jörgensen's (1998) categories of image attributes.

	Flickr tags ($n = 975$)			Index terms ($n = 996$)		
Category	Frequency (%)	Cumulative %	M (SD)	Frequency (%)	Cumulative %	M (SD)
Location (P)	1274 (30.63%)	30.63	1.31 (1.57)	1608 (43.35%)	43.35	1.61 (0.96)
Content/Story (I)	713 (17.14%)	47.77	0.73 (1.01)	945 (25.48%)	68.83	0.95 (0.70)
People (P)	623 (14.98%)	62.75	0.64 (1.12)	179 (4.83%)	73.66	0.18 (0.48)
Description (P)	562 (13.51%)	76.26	0.58 (1.22)	0 (0.00%)	73.66	0.00 (0.00)
Objects (P)	453 (10.89%)	87.15	0.46 (1.05)	835 (22.51%)	96.17	0.84 (0.98)
Abstract concepts (I)	218 (5.24%)	92.39	0.22 (0.64)	25 (0.67%)	96.84	0.03 (0.16)
Art historical information (I)	66 (1.59%)	93.98	0.07 (0.29)	88 (2.37%)	99.21	0.09 (0.32)
People-related attributes (I)	111 (2.67%)	96.65	0.11 (0.37)	29 (0.78%)	100	0.03 (0.19)
Visual elements (P)	75 (1.80%)	98.45	0.08 (0.35)	0 (0.00%)	100	0.00 (0.00)
Color (P)	64 (1.54%)	100	0.07 (0.35)	0 (0.00%)	100	0.00 (0.00)
Total	4159 (100%)		4.27 (4.04)	3709 (100%)		3.73 (1.48)

Note. P = Perceptual; I = interpretive.

TABLE 3. Distribution of Flickr tags and index terms according to two broad categories of Jörgensen's (1998) categories of image attributes.

	Flickr tags ($n = 975$)		Index terms ($n = 996$)	
Category	Frequency (%)	M (SD)	Frequency (%)	M (SD)
Perceptual	3051(73.36%)	3.13(3.30)	2622(70.69%)	2.63(1.24)
Interpretive	1108(26.64%)	1.14(1.43)	1087(29.31%)	1.09(0.81)
Total	4159(100%)	4.27(4.04)	3709(100%)	3.73(1.48)

Photographic Archive (76.87%), even though indexers at the Archive used refiners with more nonunique index terms than unique ones. Flickr users assigned more than 100 times as many unique tags than unique index terms assigned to images in the St. Andrews Library Photographic Archive. On the other hand, there were only 4.3 times as many unique index terms with refiners as Flickr tags with refiners. Table 4 presents the frequencies, percentages, means, and standard deviations for numbers of Flickr tags and St. Andrews Library Photographic Archive index terms within Enser and McGregor's four categories.

A closer look at the frequencies and the mean numbers of tags and index terms that are either unique or nonunique and are with or without refiners reveals an even more apparent difference according to Enser and McGregor's (1992) framework of image attributes. These differences were confirmed by results of a chi-square analysis and subsequent t-tests. The chi-square analysis resulted in a significant chi-square value, $\chi^2 (3, N = 7,868) = 4261$, $p < 0.05$. An independent group's t-test (with unequal variances) was also conducted to see if the mean number of unique and nonunique Flickr tags and index terms were comparable. The t-tests resulted in significant differences ($p < 0.05$). These findings are in contrast with Enser and McGregor's (1992) analysis of 2,700 requests submitted to the Hulton Deutsch collection (a general-purpose picture archive that is part of Getty Images). Their analysis found that the queries varied in their focus of topics. They showed that the bulk of the requests were for either specific objects or events and often refined by a number of facets such as spatial and temporal facets.

In terms of the uniqueness and refiners used, the two samples (tags and index terms) differ vastly as well (see Table 5). As described above, about half of the Flickr tags were unique, whereas less than a third of the index terms were unique. However, the largest group of the unique Flickr tags (24.19% of the 4,159 tags) was location tags followed in size by unique tags describing people (8.78%), and content/story/event tags (7.41%). On the other hand, despite the fact that the bulk of the index terms (71.66% of the 3,709 index terms) were nonunique, most of the unique index terms (27.67% of all index terms) were locations as well. A significant number of the nonunique index terms (25.26% of all index terms) were content/story/event terms.

Although Flickr users used refiners such as *first* (birthday), *June 2007* (baptism), *Chicago* (skyline), and *cute* (couple); refiners used with index terms included *training* (ships), *Roman* (ruins), *rock formations–coastal*, and *processions– military*. For both samples, the refiners reflected the general nature of image attributes described by the tags and index terms. For instance, there were more refiners of locations (geographic names), events (birthdays, parties, parades), people (their appearance/feature/action), and objects (brands and models).

At the aggregate level, there were more unique Flickr tags describing perceptual attributes (42.38% of all tags) than unique tags describing interpretive attributes (9.22% of all tags). More nonunique tags (30.97% of all tags) seemed to have been used by Flickr users to describe perceptual attributes as well. The same is true for the index terms, though there are more nonunique index terms. The percentage of index terms that are nonunique and used to describe perceptual attributes (42.75% of all index terms) is greater than the percentage of nonunique index terms describing interpretive attributes (28.91% of all index terms). This, once again, demonstrates the fact that at the highest level of classification, both Flickr tags and index terms have similar structures, despite their differences at individual tag or

TABLE 4. Number of Flickr tags and index terms according to Enser and McGregor's (1992) categories of image queries.

Category	Flickr tags ($n = 975$)			Index terms ($n = 996$)		
	Frequency (%)	Cumulative %	M (SD)	Frequency (%)	Cumulative %	M (SD)
Unique	1875 (45.08%)	45.08	1.92 (1.83)	16 (0.43%)	0.43	0.02 (0.15)
Unique with refiners	271 (6.52%)	51.6	0.28 (0.68)	1035 (27.91%)	28.34	1.04 (0.50)
Nonunique	1927 (46.33%)	97.93	1.98 (2.91)	842 (22.70%)	51.04	0.85 (0.91)
Nonunique with refiners	86 (2.07%)	100	0.09 (0.34)	1816 (48.96%)	100.00	1.82 (0.97)
Total	4159 (100%)		4.27 (4.04)	3709 (100%)		3.72 (1.48)

index term level. In addition, both Flickr users and indexers were focusing on the description of perceptual attributes and visual characteristics of the images more than interpretive and reactive attributes. These findings are in line with previous research that showed naïve image users using more terms that describe perceptual attributes when viewing individual images (e.g., Jörgensen, 1995, 1998; Laine-Hernandez & Westman, 2006) and using more terms that describe interpretive attributes when labeling groups of images (e.g., Rorissa, 2008; Rorissa & Hastings, 2004; Rorissa & Iyer, 2008).

Discussion

Among the few researchers who address the topic of social tagging of images, some focused only on general issues, considering only a few images and tags for analysis (e.g., Matusiak, 2006). Only a handful of researchers (Jörgensen, 1998, 1999; Trant, 2006) have embarked on research similar to the current study, investigating the difference between social tags or user-generated descriptions of images and professionally assigned indexing terms.

As the findings above show, location, content/story (including events), people, and objects are among the top image attributes described by Flickr tags and index terms. Sigurbjörnsson and van Zwol (2008), in an analysis of 52 million photos that were assigned 188 million tags with the help of WordNet, also found that the most frequently assigned type of tag is location followed by artifacts or objects, people or groups, actions or events, and time. It is clear that on Flickr, "users do not only tag the visual contents of the photo, but to a large extent provide a broader context in which the photo was taken, such as, location, time, and actions" (Sigurbjörnsson & van Zwol, 2008, p. 330).

A number of other researchers have also arrived at similar conclusions. For instance, tags describing location attributes appeared with high frequency in a similar study by Yoon (2009) who offered the possible explanation that people who upload their pictures onto Flickr frequently upload travel or vacation photos. Yoon (2009) also found tags that describe event/action and time attributes appeared with higher frequencies, perhaps because Flickr users often upload birthday and wedding photos taken during a specific period or date. In an earlier study, Laine-Hernandez and Westman (2006), who modified Jörgensen's (1995, 1998) 12 categories of attributes,

concluded that the objects attribute was by far the most widely used by photojournalists. Photojournalism professionals have also been found to prefer searching image collections by events and objects represented in the images (Neal, 2008).

Findings from previous research with regards to image search queries and requests are mixed. Enser and McGregor (1992) noted that about 70% of the search requests were for unique events, persons, and time or unique events, persons and time with refiners whereas less than 10% were nonunique. An example of a unique query is "Bill Clinton" whereas "Bill Clinton 1980" would be an example of a unique with a refiner request. An example of a nonunique request is "middle-aged man"; the same request with a refiner would be "middle-aged man sitting by desk." A recent analysis of image search queries on the Web by Jansen (2008) with the help of Enser and McGregor's (1992) and Jörgensen's (1998) image attribute classification frameworks showed that an overwhelming majority (81.9%) was nonunique or nonunique with refiners.

In this work, the chi-square analysis and t-tests revealed that the makeup of the two types of terms (Flickr tags and index terms) were different with respect to the two frameworks used to analyze them. Results of previous research that compared image descriptions by naïve users (Jörgensen, 1998, 1999) and naïve indexers (Trant, 2006) to traditional indexing tools such as thesauri and index terms assigned by professional indexers reveal that, generally, there is a disparity. For instance, Chung and Yoon (2009), using a revised version of Shatford's (1986) classification of contents of images as a framework for analysis, found statistically significant differences between Flickr tags and search-query terms. Trant (2006) found more interparticipant consistency among untrained professionals (naïve indexers) than among their trained counterparts. Even though a high percentage (88%) of user-supplied terms did not match descriptions by professionals in The Metropolitan Museum of Art's database, manual evaluation of the nonmatching user-supplied terms revealed that more than 75% of them were appropriate descriptions of the works of art (Trant, 2006). With respect to the utility of traditional image indexing tools, results of a study that looked at the Art and Architecture Thesaurus (AAT) suggest that only about 16% of entries in the AAT could be used for the purpose of indexing a general image collection (Jörgensen, 2003).

TABLE 5. Number of Flickr tags and index terms according to categories of both frameworks.

12 Categories	Unique		Unique with refiners		Nonunique		Nonunique with refiners	
	Tags	Index terms	Tags	Index terms	Tags	Index terms	Tags	Index terms
Location (P)								
Frequency	966	1	40	1025	260	239	8	343
% of total (tags/index terms)	23.23	0.03	0.96	27.64	6.25	6.44	0.19	9.25
% within category	75.82	0.06	3.14	63.74	20.41	14.86	0.63	21.33
% within unique/nonunique	51.52	6.25	14.76	99.03	13.49	28.38	9.30	18.89
Content/Story (I)								
Frequency	227	8	81	0	377	109	28	828
% of total (tags/index terms)	5.46	0.22	1.95	0.00	9.06	2.94	0.67	22.32
% within category	31.84	0.85	11.36	0.00	52.88	11.53	3.93	87.62
% within unique/nonunique	12.11	50.00	29.89	0.00	19.56	12.95	32.56	45.59
People (P)								
Frequency	360	5	5	1	251	36	7	137
% of total (tags/index terms)	8.66	0.13	0.12	0.03	6.04	0.97	0.17	3.69
% within category	57.78	2.79	0.80	0.56	40.29	20.11	1.12	76.54
% within unique/nonunique	19.20	31.25	1.85	0.10	13.03	4.28	8.14	7.54
Description (P)								
Frequency	189	0	92	0	271	0	10	0
% of total(tags/index terms)	4.54	0.00	2.21	0.00	6.52	0.00	0.24	0.00
% within category	33.63	0.00	16.37	0.00	48.22	0.00	1.78	0.00
% within unique/nonunique	10.08	0.00	33.95	0.00	14.06	0.00	11.63	0.00
Objects (P)								
Frequency	57	2	5	2	383	414	8	417
% of total(tags/index terms)	1.37	0.05	0.12	0.05	9.21	11.16	0.19	11.24
% within category	12.58	0.24	1.10	0.24	84.55	49.58	1.77	49.94
% within unique/nonunique	3.04	12.50	1.85	0.19	19.88	49.17	9.30	22.96
Abstract concepts (I)								
Frequency	16	0	12	1	175	24	15	0
% of total (tags/index terms)	0.38	0.00	0.29	0.03	4.21	0.65	0.36	0.00
% within category	7.34	0.00	5.50	4.00	80.28	96.00	6.88	0.00
% within unique/nonunique	0.85	0.00	4.43	0.10	9.08	2.85	17.44	0.00
Art historical information (I)								
Frequency	16	0	5	0	44	13	1	75
% of total (tags/index terms)	0.38	0.00	0.12	0.00	1.06	0.35	0.02	2.02
% within category	24.24	0.00	7.58	0.00	66.67	14.77	1.52	85.23
% within unique/nonunique	0.85	0.00	1.85	0.00	2.28	1.54	1.16	4.13
People-related attributes (I)								
Frequency	20	0	6	6	77	7	8	16
% of total (tags/index terms)	0.48	0.00	0.14	0.16	1.85	0.19	0.19	0.43
% within category	18.02	0.00	5.41	20.69	69.37	24.14	7.21	55.17
% within unique/nonunique	1.07	0.00	2.21	0.58	4.00	0.83	9.30	0.88
Visual elements (P)								
Frequency	23	0	25	0	27	0	0	0
% of total(tags/index terms)	0.55	0.00	0.60	0.00	0.65	0.00	0.00	0.00
% within category	30.67	0.00	33.33	0.00	36.00	0.00	0.00	0.00
% within unique/nonunique	1.23	0.00	9.23	0.00	1.40	0.00	0.00	0.00
Color (P)								
Frequency	1	0	0	0	62	0	1	0
% of total(tags/index terms)	0.02	0.00	0.00	0.00	1.49	0.00	0.02	0.00
% within category	1.56	0.00	0.00	0.00	96.88	0.00	1.56	0.00
% within unique/non-unique	0.05	0.00	0.00	0.00	3.22	0.00	1.16	0.00
Total								
Frequency	1875	16	271	1035	1927	842	86	1816
% of total (tags/index terms)	45.08	0.43	6.52	27.91	46.33	22.70	2.07	48.96
% within unique/nonunique	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00

Note. P = Perceptual; I = interpretive.

The figure looks worse given that the AAT was designed as an indexing tool for collections of text, images, objects, and other visual resources housed in archives, special collections, museums, and libraries (Soergel, 1995).

Conclusion

This work provides an insight into the differing natures of Flickr tags and traditional index terms assigned to images in a general-purpose collection. The findings suggest that user-generated tags and professionally assigned index terms have different underlying structures. Other researchers (e.g., Chung & Yoon, 2009; Jörgensen, 1998, 1999, 2003; Trant, 2006) have also found differences between users' tags and descriptions of images by professional indexers. The prevailing recommendation is that social tagging and traditional/professional indexing should be used together to complement each other. The importance of tagging as complementary activity to indexing is not a new idea. Furnas (1985) suggested the use of adaptive indices as a useful method to assist users in their efforts to access keyword-based systems.

To put the findings of this work into perspective, it should be noted that in social/collaborative tagging, tags are assigned freely without any restrictions as to type and number, whereas professional indexers adhere to guidelines that define types and minimize the number of terms assigned. Another major distinction between tags and index terms assigned with the help of controlled vocabularies is that people who tag, especially those tagging their own resources, do so within some context (social, personal, etc.), and the tags reflect that context and the perspective of the person doing the tagging. As a result, tags are richer in their semantic content than index terms, which are at times void of context. However, professionals who index information sources using controlled vocabularies may evaluate contents of the information source more thoroughly than a Flickr tagger, adding value to index terms in terms of their precision for retrieval purposes. This comes at a high cost compared to social tagging because professional indexing is resource intensive, especially when the size of the collection of information sources is large.

Social tagging may not be the "silver bullet" that will solve the problems of image representation and indexing. However, although social tagging and traditional indexing tools each have their pros and cons, they are complementary and effective when utilized together (Guy & Tonkin, 2006). The need for new means of image indexing and the need for constant evaluation and improvement of current image indexing tools and systems have already been emphasized (Jörgensen, 1998; Layne, 1994). Social tagging could serve as a platform on which to build future indexing systems. Knowledge of the underlying structure of user-generated tags is necessary to inform the future design of professional indexing tools such as taxonomies, thesauri, and ontologies for personal and institutional image collections. To these ends, the current study serves as a formal and empirical evaluation of current image indexing tools and systems.

Implications for Image Indexing and Retrieval

Indexing and annotating visual resources is often a time-consuming and labor-intensive task. Social and collaborative tagging could help to solve these problems. With this in mind, some researchers have not only analyzed the nature of user descriptions of images, but have also utilized results of their analyses to show how tags can be useful in the design of traditional indexing tools such as thesauri (e.g., Rorissa, 2008; Rorissa & Iyer, 2008; Yoon, 2009). The merits of indexing through collaborative tagging coupled with the popularity of folksonomies have already been recognized as well (Olson & Wolfram, 2008).

The significance of research on the nature of social tagging and tags lies in the assumption that users will be likely to use similar, if not the same, terms during searching. Understanding the nature of social tagging and tags will provide clues to the perceptions of image users. Research on social tagging should have the potential to bridge the semantic gap between indexer-assigned terms and users' search language.

According to Enser (2008), "social tagging has brought a new dimension to the representation of the semantic content of visual materials" (p. 534). When users assign personal tags to visual materials on popular photo and video sharing services such as Flickr and YouTube, it "challenges the supremacy of professionally sourced, authoritative subject representation, whilst introducing opportunities for beneficial enhancement of both exhaustivity and specificity in subject indexing" (Enser, 2008, p. 534). It is crucial to study the extent to which the indexing language used in digital libraries matches users' language. Findings from such studies could inform the design of information organization and retrieval systems that match users' vocabulary (Eerola & Vakkari, 2008).

A possible approach to incorporating social tagging into the indexing process is asking users to come up with tags that describe the information resource. These tags should preferably be unique to generate a richer description of the resource (Chi & Mytkowicz, 2008). Another possible application and use of collaborative tagging might be identifying relations between tags assigned by groups of people and aggregating the relations to create categories of tags and information resources to which they were assigned (Plangprasopchok & Lerman, 2009).

Involving users in the process of indexing through collaborative tagging could mean consensus on the tags and index terms built. Additionally, users' expertise could be crucial in generating relevant index terms. This would hold true especially when the users are expert users in a specific field of study (Stock, 2007). An opportunity provided by social tagging is that it could be used to both enrich and validate traditional indexing tools such as controlled vocabularies (Qin, 2008). The results of research into the nature of user-generated tagging and tags can assist designers

of tools such as thesauri in identifying and selecting concepts that need to be included and defining the semantic relations of the selected concepts (Yoon, 2009). However, freely assigned tags need "linguistic, syntactic, and semantic processing" (Chen, Liu, & Qin, 2008, p. 117), which could be intensive computationally or ineffective if attempted manually.

On the merits of incorporating social tagging into the process of indexing, Wetterstrom (2008) summed it up by stating that "user-assigned tags could provide additional access points, and the co-existence of tags and controlled vocabularies... could thus enhance the discovery of documents" (p. 297). Tagging and tags may also have an added value in that, compared to index terms assigned with the help of controlled vocabularies, tags are diverse, specific, current, and popular (Wetterstrom, 2008). The main appeal of tags and folksonomies is "their inclusiveness; they reflect the vocabulary of the users, regardless of viewpoint, background, bias, and so forth" (Spiteri, 2007, p. 14).

Like any uncontrolled vocabularies, tags have their problems. These include ambiguity (tags are assigned by different users in different ways), polysemy (e.g., port could mean a place where ships load and unload or an interface that links a device to another device in a computer), and synonymy (e.g., Mac, macintosh, and apple could all describe an Apple Macintosh computer; Golder & Huberman, 2006; Guy & Tomkin, 2006; Spiteri, 2007).

Limitations of Current Work

As one of the reviewers pointed out, a weakness of this work is the choice of a general image collection, the University of St. Andrews Library Photographic Archive, which has, in its collection, black and white as well as color photos taken mostly in a country (Scotland). A few aberrant photos are of places around the world. The selection of index terms from the University of St. Andrews Library Photographic Archive may have some bearing on the results obtained and results may differ if one picked a different image collection. However, I should point out that the University of St. Andrews Library Photographic Archive has recently been used for major image retrieval system evaluation campaigns such as ImageCLEF (http://ir.shef.ac.uk/imageclef/2005/). Despite this, results obtained in this study may apply to a more broadly general image collection and a collection of images built by individuals or groups of general users. However, future studies should compare this study to results obtained from specialized image collections (e.g., medical images) and users (e.g., doctors, journalists).

Acknowledgments

References

Angus, E., Thelwall, M., & Stuart, D. (2008). General patterns of tag usage among university groups in Flickr. Online Information Review, 32(1), 89–101.

Armitage, L.H., & Enser, P.G.B. (1997). Analysis of user need in image archives. Journal of Information Science, 23(4), 287–299.

Bar-Ilan, J., Shoham, S., Idan, A., Miller, Y., & Shachak, A. (2008). Structured versus unstructured tagging: A case study. Online Information Review, 32(5), 635–647.

Bearman, D., & Trant, J. (2005, September). Social terminology enhancement through vernacular engagement: Exploring collaborative annotation to encourage interaction with museum collections. D-Lib Magazine, 11(9). Retrieved January 30, 2007, from www.dlib.org/dlib/september05/bearman/09bearman.html

Chen, H., & Rasmussen, E.M. (1999). Intellectual access to images. Library Trends, 48(2), 291–302.

Chen, M., Liu, X., & Qin, J. (2008). Semantic relation extraction from socially-generated tags: A methodology for metadata generation. In Proceedings of the 2008 International Conference on Dublin Core and Metadata Applications (pp. 117–127). Singapore: Dublin Core Metadata Initiative.

Chi, E.H., & Mytkowicz, T. (2008). Understanding the efficiency of social tagging systems using information theory. In Proceedings of the Nineteenth ACM Conference on Hypertext and Hypermedia (pp. 81–88). New York: ACM Press.

Choi, Y., & Rasmussen, E.M. (2003). Searching for images: The analysis of users' queries for image retrieval in American history. Journal of the American Society for Information Science and Technology, 54, 498–511.

Chung, E., & Yoon, J. (2009). 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), Paper 408. Retrieved January 14, 2010, from http://InformationR.net/ir/14-3/paper408.html

Cohen, J. (1960). A coefficient of agreement for nominal scales. Educational and Psychological Measurement, 20(1), 37–46.

Cooper, W.S. (1969). Is interindexer consistency a hobgoblin? American Documentation, 20, 268–278.

Cox, A.M. (2008). Flickr: A case study of Web 2.0. Aslib Proceedings, 60(5), 493–516.

Datta, R., Li, J., & Wang, J.Z. (2005). Content-based image retrieval: Approaches and trends of the new age. In H. Zhang, J. Smith, & Q. Tian (Eds.), Proceedings of the Seventh ACM SIGMM International Workshop on Multimedia Information Retrieval (pp. 253–262). New York: ACM Press.

Eerola, J., & Vakkari, P. (2008). How a general and a specific thesaurus cover expressions in patients' questions and physicians' answers. Journal of Documentation, 64(1), 131–142.

Enser, P.G.B. (2008). The evolution of visual information retrieval. Journal of Information Science, 34(4), 531–546.

Enser, P.G.B. (2000). Visual image retrieval: Seeking the alliance of concept-based and content-based paradigms. Journal of Information Science, 26(4), 199–210.

Enser, P.G.B., & McGregor, C.G. (1992). Analysis of visual information retrieval queries. Report on Project G16412 to the British Library Research and Development Department. London: British Library.

Fidel, R. (1994). User-centered indexing. Journal of the American Society for Information Science, 45(8), 572–576.

Fox, R. (2006). Cataloging for the masses. OCLC Systems & Services, 22(3), 166–172.

Furnas, G.W. (1985). Experience with an adaptive indexing scheme. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (pp. 131–135). New York: ACM Press.

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

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

Guy, M., & Tonkin, E. (2006, January). Folksonomies: Tidying up tags?. D-Lib Magazine, 12(1). Retrieved January 30, 2007, from http://www.dlib.org/dlib/january06/guy/01guy.html

Hammond, T., Hannay, T., Lund, B., & Scott, J. (2005). Social bookmarking tools (I): A general review. D-Lib Magazine, 11(4). Retrieved January 21, 2010, from http://www.dlib.org/dlib/april05/hammond/04hammond.html

Hastings, S.K. (1995). Query categories in a study of intellectual access to digitized art images. In T. Kinney (Ed.), Proceedings of the 58th Annual Meeting of the American Society for Information Science (pp. 3–8). Medford, NJ: Information Today.

Hastings, S.K. (1999). Evaluation of image retrieval systems: Role of user feedback. Library Trends, 48(2), 438–452.

Jacob, E.K. (2004). Classification and categorization: A difference that makes a difference. Library Trends, 52(3), 515–540.

Jansen, B.J. (2008). Searching for digital images on the web. Journal of Documentation, 64(1), 81–101.

Jörgensen, C. (1995). Image attributes: An investigation. Unpublished doctoral dissertation, Syracuse University, NY.

Jörgensen, C. (1996). Indexing images: Testing an image description template. In S. Hardin (Ed.), Proceedings of the 59th Annual Meeting of the American Society for Information Science. (pp. 209–213). Medford, NJ: Information Today.

Jörgensen, C. (1998). Attributes of images in describing tasks. Information Processing & Management, 34(2/3), 161–174.

Jörgensen, C. (1999). Image indexing: An analysis of selected classification systems in relation to image attributes named by naïve users. Annual Review of OCLC Research. Retrieved January 30, 2007, from http://digitalarchive.oclc.org/da/ViewObject.jsp?objid=0000003507

Jörgensen, C. (2003). Image retrieval: Theory and research. Lanham, MD: Scarecrow Press.

Jörgensen, C., & Jörgensen, P. (2005). Image querying by image professionals. Journal of the American Society for Information Science and Technology, 56(12), 1346–1359.

Laine-Hernandez, M., & Westman, S. (2006). Image semantics in the description and categorization of journalistic photographs. In A. Grove & J. Stefl-Mabry (Eds.), Proceedings of the 69th Annual Meeting of the American Society for Information Science and Technology. Retrieved December 22, 2009, from http://eprints.rclis.org/8122/

Landis, J.R., & Koch, G.G. (1977). The measurement of observer agreement for categorical data. Biometrics, 33(1), 159–174.

Layne, S.S. (1994). Some issues in the indexing of images. Journal of the American Society for Information Science, 45(8), 583–588.

Lew, M.S., Sebe, N., Djeraba, C., & Jain, R. (2006). Content-based multimedia information retrieval: State of the art and challenges. ACM Transactions on Multimedia Computing, Communications, and Applications, 2, 1–19.

Lyman, P., & Varian, H.R. (2003). How much information 2003? Berkeley: University of California, Berkeley. Retrieved March 3, 2010, from http://www.sims.berkeley.edu/research/projects/how-much-info-2003/

MacGregor, G., & McCulloch, E. (2006). Collaborative tagging as a knowledge organisation and resource discovery tool. Library View, 55(5), 291–300.

Mathes, A. (2004). Folksonomies—cooperative classification and communication through shared metadata. Retrieved January 30, 2007, from http://blog.namics.com/2005/Folksonomies_Cooperative_Classification.pdf

Matusiak, K.K. (2006). Towards user-centered indexing in digital image collections. OCLC Systems & Services, 22(4), 283–298.

McCay-Peet, L., & Toms, E. (2009). Image use within the work task model: Images as information and illustration. Journal of the American Society for Information Science and Technology, 60(12), 2416–2429.

Neal, D. (2008). News photographers, librarians, tags, and controlled vocabularies: Balancing the forces. Journal of Library Metadata, 8(3), 199–219.

Neuendorf, K.A. (2002). The content analysis guidebook. Thousand Oaks, CA: Sage.

Neumann, D., & Gegenfurtner, K.R. (2006). Image retrieval and perceptual similarity. ACM Transactions on Applied Perception, 3, 31–47.

O'Connor, B.C. (1996). Explorations in indexing and abstracting: Pointing, virtue, and power. Englewood, CO: Libraries Unlimited.

O'Connor, B., & O'Connor, M. (1999). Categories, photographs & predicaments: Exploratory research on representing pictures for access. Bulletin of the American Society for Information Science, 25(6), 17–20.

O'Connor, B.C., O'Connor, M.K., & Abbas, J.M. (1999). User reactions as access mechanism: An exploration based on captions for images. Journal of the American Society for Information Science and Technology, 50(8), 681–697.

Olson, H.A., & Wolfram, D. (2008). Syntagmatic relationships and indexing consistency on a larger scale. Journal of Documentation, 64(4), 602–615.

Peters, I. (2009). Folksonomies: Indexing and retrieval in Web 2.0. Berlin, Germany: De Gruyter Saur.

Plangprasopchok, A., & Lerman, K. (2009). Constructing folksonomies from user-specified relations on Flickr (pp. 781–790). In Proceedings of the 18th International World Wide Web Conference. New York: ACM Press.

Qin, J. (2008, August). Controlled semantics vs. social semantics: An epistemological analysis. Paper presented at the 10th International ISKO Conference: Culture and Identity in Knowledge Organization, Montreal, Canada.

Rattenbury, T., Good, N., & Naaman, M. (2007). Towards automatic extraction of event and place semantics from Flickr tags. In Proceedings of the 30th Annual International ACM SIGIR Conference on Research and Development in information Retrieval (pp. 103–110). New York: ACM Press.

Reid, N.H. (1999). The photographic collections in St Andrews University Library. Scottish Archives, 5, 83–90.

Roberts, H.E. (2001). A picture is worth a thousand words: Art indexing in electronic databases. Journal of the American Society for Information Science and Technology, 52(11), 911–916.

Rorissa, A. (2008). User generated descriptions of individual images versus labels of groups of images: A comparison using basic level theory. Information Processing & Management, 44(5), 1741–1753.

Rorissa, A., & Hastings, S.K. (2004). Free sorting of images: Attributes used for categorization. In Proceedings of the 67th Annual Meeting of the American Society for Information Science and Technology (pp. 360–366). Medford, NJ: Information Today.

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.

Schmidt, S., & Stock, W.G. (2009). Collective indexing of emotions in images: A study in emotional information retrieval. Journal of the American Society for Information Science and Technology, 60(5), 863–876.

Shatford, S. (1984). Describing a picture: A thousand words are seldom cost effective. Cataloguing and Classification Quarterly, 4(4), 13–30.

Sigurbjörnsson, B., & van Zwol, R. (2008). Flickr tag recommendation based on collective knowledge. In Proceedings of the 17th International Conference on World Wide Web (pp. 327–336). New York: ACM Press.

Soergel, D. (1995). The art and architecture thesaurus (AAT): A critical appraisal. Visual Resources, 10(4), 369–400.

Spiteri, L. (2007). The structure and form of folksonomy tags: The road to the public library catalog. Information Technology and Libraries, 26(3), 13–25.

Stock, W.G. (2007). Folksonomies and science communication: A mashup of professional science databases and Web 2.0 services. Information Services & Use, 27, 97–103.

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

Tjondronegoro, D., Spink, A., & Jansen, B.J. (2009). A study and comparison of multimedia Web searching: 1997–2006. Journal of the American Society for Information Science and Technology, 60(9), 1756–1768.

Trant, J. (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.

von Ahn, L., & Dabbish, L. (2004). Labeling images with a computer game. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (pp. 319–326). New York: ACM Press.

Weber, R.P. (1990). Basic content analysis (2nd ed.). Newbury Park, CA: Sage.

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 & Information Management Journal, 50(4), 296–310.

Wolfram, D., Olson, H.A., & Bloom, R. (2009). Measuring consistency for multiple taggers using vector space modeling. Journal of the American Society for Information Science and Technology, 60(10), 1995–2003.

Yoon, J. (2009). Towards a user-oriented thesaurus for non-domain-specific image collections. Information Processing and Management, 45(4), 452–468.