

Facets of User-Assigned Tags and Their Effectiveness in Image Retrieval

by

Nicky Ransom

25th September 2009

DS30620 Undergraduate Dissertation

BSc Econ Information and Library Studies

Aberystwyth University

Acknowledgements

I would like to thank my supervisor, Dr. Pauline Rafferty, for her valuable support and encouraging feedback during the process of writing this dissertation, and also Marianne Taylor for her support during this time. I would also like to thank my son, Adam Ransom, who wrote and amended the computer program to download images and data from Flickr, without whose assistance I could not have collected the necessary data for the study. Finally, I would like to thank my family who have offered their support and advice, as well as all the friends who have listened patiently from start to finish.

Executive summary

The growth of user-tagging in online applications has the potential to overcome the problems of indexing the vast numbers of images that are now available online. It is unclear, however, whether user-assigned tags are useful for subsequent retrieval purposes. This study considers the value of user-assigned image tags by comparing the facets that are represented in image tags with those that are present in image queries to see if there is a similarity in the way that users describe and search for images. A sample dataset was created by downloading a selection of images and associated tags from Flickr, the online photo-sharing web site. The tags were categorised using the image facets from Shatford's matrix, which has been widely used in previous research into image indexing and retrieval. The facets present in the image tags were then compared with the results of previous research into image queries. The results reveal that there are broad similarities between the facets present in image tags and queries, with people and objects being the most common facet, followed by location. However, the results also show that there are differences in the level of specificity between tags and queries, with image tags containing more generic terms and image queries consisting of more specific terms. The study concludes that users do describe and search for images using similar image facets, but that measures to close the gap between specific queries and generic tags would improve the value of user tags in indexing image collections.

Table of Contents

[List of Figures](#)

[List of Tables](#)

[1 Introduction](#)

[1.1 Image indexing](#)

[1.2 User tagging](#)

[1.3 Research question](#)

[2 Literature review](#)

[2.1 Image indexing](#)

[2.2 User query analysis](#)

[2.3 User tagging](#)

[3 Methodology](#)

[3.1 Research environment](#)

[3.2 Research design](#)

[3.3 Data collection](#)

[3.4 Data analysis](#)

[4 Results](#)

[4.1 Statistical analysis of image tags](#)

[4.2 Categorisation of image tags](#)

[4.3 Comparison with categorisation of image queries](#)

[5 Discussion](#)

[6 Conclusion](#)

[Bibliography](#)

[Appendix A: Sample data](#)

[Appendix B: Example of a Flickr page](#)

[Appendix C: Coding sheet](#)

[Appendix D: Examples of coded tags](#)

[Appendix E: Examples of 'Other' tags](#)

List of Figures

Figure 1. Number of images with tags

Figure 2. Distribution of number of tags per image

Figure 3. Distribution of broad tag categories

Figure 4. Categorisation of tags using Shatford's matrix

Figure 5. Comparison of low and high tag usage

Figure 6. Categorisation of "Other" tags

Figure 7. Comparison of tags and queries by broad category

Figure 8. Comparison of tags and queries by individual facet

Figure 9. Comparison of previous query studies

Figure 10. Comparison of tags and queries without Van Hooland results

Figure 11. Comparison of tag and query facets with combined levels of specificity

Figure 12. Example of Flickr page showing image and tags

List of Tables

Table 1. Shatford facet matrix

Table 2. Previous studies into user queries using Shatford matrix

Table 3. Statistical analysis of tagged images

Table 4. Ranking of tag facets

Table 5. Sample data from Flickr download

Table 6. Coding sheet with coding examples from other research

Table 7. Example of how tags in the current study were coded

Table 8. Examples of tags that could not be categorised using Shatford's matrix

1.

Introduction

Advances in digital image technologies over the last few decades have resulted in an explosion in the availability of digital images. Whereas traditional picture libraries had intermediaries to help users access images (Enser, 2008), the advent of online image databases has resulted in users carrying out their own searches, and so accurate and comprehensive indexing has become critical for resource discovery (Matusiak, 2006; Trant, 2008a). In order to keep pace with the growing number of images online, researchers have been investigating ways to improve image indexing, but unlike the full-text searching techniques that can be utilized with written documents, accessing the intellectual content of images is much harder to achieve due to the “significant philosophical and practical problems” of creating indexing terms for visual content (Rafferty & Hilderley, 2005, p. 52).

1. Image indexing

Indexing is the process of assigning terms to represent a ‘document’ (textual or visual) in a retrieval system. When a user queries the system, the query terms are matched against the indexing terms and relevant results are presented to the user. The challenge of image indexing is how to represent visual information in such a way that the user can effectively query the system to get relevant results.

There are two indexing approaches that can be used to represent visual information: concept-based indexing and content-based indexing. Concept-based indexing uses textual descriptions to represent images, and traditionally is the method that has been used in image retrieval systems. However, it relies on human cataloguers to translate the semantic content of an image into indexing terms, which has been acknowledged as a difficult task and open to subjective interpretation (Enser, 2008), as well as being time-consuming and therefore costly. In addition, the indexing terms, which are usually selected from hierarchical classification schema and controlled vocabularies, may not reflect the natural language terms selected by users when they search for image (Mathes, 2004).

The problems of concept-based indexing have been multiplied with the development of the World Wide Web and the huge growth in the volume of images that have become available online. Search engines provide one solution to the problem by automatically indexing any images they encounter as they crawl the web, using data contained in an image’s filename or caption, or by using the text that surrounds an

image on a web page. Although this method of indexing has vastly improved access to the wealth of images online, it has significant limitations, particularly as the text surrounding an image is often not relevant in identifying the image itself.

The second approach, content-based indexing, was developed as an alternative to concept-based indexing by removing the need to represent an image using words, and instead using computers to index the actual content of the image itself through automatic processing. However, in spite of over a decade of research, it is still only possible to identify low level image attributes such as colour and texture using automatic methods and not the higher level attributes such as subject and meaning which are more important factors in image retrieval (Matusiak, 2006). It is therefore generally felt to be limited in its application for general image retrieval at this time (Chu, 2003; Enser, 2008).

The recent development of Web 2.0 technologies, with their emphasis on user contribution, active participation and the harnessing of collective intelligence (O'Reilly, 2005), has been seen as offering a potential solution to the problems of scalability of traditional techniques. In particular, interest has grown in the use of 'tagging' as a potential means of indexing online content.

2. User tagging

Tagging, also known as social classification, collaborative tagging, or user-tagging, is a Web 2.0 feature that allows users to add their own keywords or 'tags' to online documents and images so that they can organise resources for themselves, share them with others, and find resources that others have tagged. One of the earliest web sites to offer this feature was Delicious (<http://delicious.com>), an online social bookmarking application launched in 2003, followed a year later by Flickr (<http://www.flickr.com>), an online photo sharing web site: tagging has since been adopted by many other internet applications.

The vocabulary that develops as a result of user tagging, often referred to as a 'folksonomy' (Vander Wall, 2007), has been the subject of many studies over the last few years as researchers investigated its potential as an indexing language. Initially, there was speculation that user tagging would replace traditional indexing practices (Shirky, 2006) as it was flexible, abundant and cheaper to produce than more traditional methods (Hammond, Hannay, Lund & Scott, 2005). It was also suggested that its democratic nature that would "better reflect the peoples' conceptual model" (Munk & Mork, 2007, p.

17). However, research has revealed that tag vocabularies are full of inconsistencies and inaccuracies (Golder & Huberman, 2006), and the general consensus now is that user-tagging will compliment rather than replace formal classification systems (Hunter, Khan & Gerber, 2008; Macgregor & McCulloch, 2006; Matusiak, 2006; Voss, 2007).

Research into user tagging has focused on three main areas (Trant, 2008a): user behaviour and motivations to tag, the social systems in which tagging takes place, and the vocabulary of tags themselves. Much of this research has concentrated on the tagging of textual resources, such as bookmarks in Delicious or blogs in Technorati (<http://technorati.com>), with fewer studies related to the tagging of images. In spite of early speculation that tags would consist of “terms that real users might be expected to use in future when searching” (Hammond et al, 2005, Tag Soup, para. 7), little research has been undertaken into how user tags compare to the terms used in search queries, particularly in the context of digital images. It is this aspect of image tagging that this study will consider.

3. Research question

The aim of this study is to find out the value of tags for image retrieval by investigating whether the way that users describe images in tags is similar to the way they search for images. Specifically, this study will consider whether the image facets described in tags are similar to the facets that have been previously identified in image queries.

It has been decided to focus on image facets rather than looking at individual terms in this study in order that general patterns of tag use can be identified. Studies have shown that image tags contain a very large number of unique terms (Jansen, Spink & Pfaff, 2000; Ding & Jacob, 2009), making comparisons with image queries at the level of individual terms difficult to carry out. Image facets have also been the subject of numerous studies into image queries as the identification of the aspects of images that form user queries is important in ensuring that the indexing of images can meet user needs. It is of little use, for example, to index the colours present in an image if this is an image facet that does not appear in user queries. Similarly, if most user queries are concerned with finding images of a certain location, this would be an important facet to index comprehensively.

With current interest in the potential of user tagging for online indexing, particularly within the museum

community, this research will be of interest to those investigating ways of harnessing the collective indexing potential of user tagging as a means of coping with the challenge of indexing the abundance of image available online.

This study will therefore ask the following questions.

- Which image facets are described in user tags?
- How do these compare to those found in image queries?
- What implications does this have for the future use of tagging for online indexing?

The remainder of the paper is divided into five chapters. The next chapter will describe previous research into image retrieval, user queries and user tagging. Chapter 3 covers the methodology used to conduct a small case study of image tag use in Flickr, the online photo sharing web site, the results of which are presented in Chapter 4. Chapter 5 contains a discussion of the results, followed by the final chapter which draws conclusions from the study.

2. Literature review

There has been much research carried out into image retrieval over the past few decades. This section will review previous research relating to image indexing, user query analysis and user tagging.

1. Image indexing

Much of the research into image indexing has focused on providing subject access to images: the process of determining meaning in images and translating that meaning into indexing terms, traditionally carried out by experts. In particular, Enser notes the difficulty of translating visual content into verbal descriptions, especially as some messages contained in images “cannot be named” (2008, p. 534).

In discussing these issues, many writers have referred to the work of Edwin Panofsky (1955) whose discussions of meaning in art images have been influential in subsequent research on this topic (for example, Choi & Rasmussen, 2003; Peters & Stock, 2007; Rafferty & Hilderley, 2005; Shatford, 1986). Panofsky identifies three levels of meaning in a work of art: primary or natural subject matter

(pre-iconographical description of objects, events or emotions); secondary or conventional subject matter (iconographical descriptions of themes or concepts); and intrinsic meaning or content (iconological interpretation of the previous two levels, such as symbolic meanings) (1955, p. 28-31).

Sara Shatford (1986) built on Panofsky's work by re-interpreting his first two levels as describing what an image was 'of' and 'about' respectively, noting that the first level could include both generic and specific elements which need to be indexed separately. She combined these three levels – Generic Of, Specific Of, and About - with four facets of image description – Who, What, Where, When – to construct a matrix of indexing possibilities. While Shatford acknowledges that “subjectivity [...] enters into almost every aspect of picture indexing” (1986, p. 57), she felt that Panofsky's highest level was too subjective to index with any degree of consistency and it was therefore disregarded in her model. This matrix has since “figured prominently in the literature” (Enser, 2008, p. 533), suggesting its continuing value in image indexing.

Issues of subjectivity in image indexing have also featured in the work of subsequent researchers. Whereas traditional image indexing assumes that there is only one interpretation of an image, theories of cognitive psychology suggest that meaning does not reside in the image itself but is constructed by the user in the process of viewing an image (Greisdorf & O'Connor, 2002; Rafferty & Hilderley, 2007). An image can therefore mean different things to different people, resulting in the difficulty of capturing all the “impressions invoked by an image” (Greisdorf & O'Connor, 2002, p. 7).

Greisdorf and O'Connor go on to acknowledge that the higher levels of meaning contained in images, such as symbolic values, abstract concepts and emotions, are particularly difficult to index successfully, and that these attributes are currently “untapped by traditional indexing techniques” (2002, p. 9), yet are often present in users' descriptions of images. Further difficulties in indexing occur because some information, particularly specific image aspects such as people's names or geographic locations, cannot be determined from the image content itself but requires information extrinsic to the image (Enser, Sandom, Hare & Lewis, 2007; Hastings, 1995; Jaimes & Chang, 2000). It has also been noted that the meaning of an image can change over time (Rafferty, 2009).

Subject analysis is only one part of the indexing process and other image access points have been noted as important for indexing (Hollink, Schreiber, Wielinga & Worrying, 2004; Jaimes & Chang, 2000; Layne, 1994; Turner, 1997). These include biographical attributes, such as those relating to the creation and subsequent history of an image, as well as physical attributes, such as size and technique. With advances in technology and the development of content-based image processing, image retrieval has also become concerned with accessing lower level content-based attributes, such as colour, shape and texture (Enser et al, 2007).

While the inclusion of content-based features has potentially widened the range of image attributes that could be captured during the indexing process, Enser et al emphasise that:

In general, users' interest in images lies in their inferred semantic content, and a retrieval facility which returns candidate images lacking in semantic integrity has little or no operational value (2007, p. 468).

The assigning of indexing terms to describe images is only part of the information retrieval process. The success of an image retrieval system depends on how well the assigned index terms match the terms provided by users in their searches. Knowledge about how a user formulates queries is therefore an important aspect of research into image retrieval, and the next section will consider this aspect of the information retrieval process.

2. User query analysis

In order to ensure that indexing practices support user needs and reflect “user warrant” in image retrieval (Rafferty & Hilderley, 2005, p. 46), research has been conducted into how users search for images and in particular the image attributes that feature in user queries.

The first major study into user queries was undertaken by Enser and McGregor in 1992 (reported in Chen & Rasmussen, 1999). The study investigated queries received by the Hulton Deutsch photograph library, a collection of images used primarily by the press. After analysing the queries, the researchers developed a scheme that classified queries in a matrix of four criteria based on the properties of uniqueness and refinement, with the majority of queries falling into the unique, non-refined category.

Although Pu (2003) noted that this framework could be useful for identifying trends in user queries, other studies have found difficulties in distinguishing between unique and non-unique attributes (Armitage & Enser, 1997; Chen, 2001).

Jorgensen (1996) analysed user queries in the domain of art history as part of her investigation into how humans perceive images. She identified 12 classes of image attribute, grouped into three main types: perceptual (resulting from visual stimulus, such as colour); interpretive (requiring interpretation of perceptual information using general knowledge); and reactive (resulting from personal reactions to an image). The results indicated that interpretive attributes were most common, with searches for objects occurring most frequently. This framework has been used in later research, but both Chen (2001) and Jansen et al (2000) found that it did not easily accommodate the queries in their studies, particularly in a web environment.

Unlike Jorgensen's narrow subject domain, Armitage and Enser (1997) analysed queries from seven different picture libraries. After finding difficulties in applying Enser and McGregor's property of uniqueness and in order to accommodate the high incidence of refiners that were present in their results, they used a facet matrix based on Shatford's adaptation of Panofsky's theories. Their results showed that a high incidence of queries related to specific people or things, specific locations and general people or things, but their report noted significant differences between the results of some of the participating libraries.

The Shatford matrix was also used by Choi and Rasmussen (2003) in their study of queries posed by students of American History. Their findings indicated a greater use of generic terms with generic people or things, events and locations occurring most frequently. Other studies using Shatford's matrix have reported differing results: Westman and Oittinen (2006) report a high incidence of both specific and generic people or things in their study of queries from journalists, while the results of Van Hooland's study (2006) of queries posed to a national archives web site show a high incidence of queries related to specific locations.

The studies described so far have shown that different user groups and subject domains have differing

information needs. It is also clear that many different query analysis frameworks have been developed, making it difficult to identify general patterns in user queries. One attempt to amalgamate these frameworks was made in research undertaken by Conduit and Rafferty (2007) as part of their work in developing an indexing framework for the Children's Society. They analysed queries from a broad range of previous research to identify the most commonly occurring image attributes, using Shatford's matrix as a common framework into which the queries from the other studies were mapped. This framework was chosen as it was "generally accepted as a useful foundation for modelling image retrieval systems" (Conduit & Rafferty, 2007, p. 901). Their results indicated that the most commonly used facets from this matrix were generic people or things, and specific people or things.

One area of general agreement amongst the research is that queries relating to abstract image facets, such as colour, texture, or emotion, did not occur very often. For example, Hollink et al (2004) found that only 12% of queries were for perceptual image attributes, such as colour or texture, Van Hooland (2006) found that none of Shatford's abstract facets were present in the queries posed to the National Archives, and a "low incidence" of abstract queries was noted in Armitage and Enser's study (1997, p. 294). Pu's 2003 research into web queries also found a low incidence of perceptual and reactive queries (7.2% and 9.25% respectively).

A newer area of research in user queries relates to user interaction with web-based resources, an area in which "little intelligence has been gathered" (Enser, 2008, p. 535). There have been some studies of search engine query logs, but these have mostly focused on how image queries differ to textual queries, (Goodrum & Spink, 2001; Pu, 2003), or on the linguistic aspects of queries (Jansen et al, 2000). However, search engine logs can only provide limited information about user behaviour and to date there have been no qualitative studies into web queries (Enser, 2008).

Other factors affecting the characteristics of image queries have been noted in the research, such as the familiarity with information retrieval systems or the use of intermediaries in the search process (Chen, 2001; Hollink et al, 2004). It has also been noted that there appear to be differences in the way that users search for images compared to the way that they describe them (Hollink et al, 2004; Jorgensen, 1996). The next section will review the literature relating to how users describe images in user-assigned

tags.

3. User tagging

Research into the way that users describe images is not a new phenomenon (see Jorgensen, 1996; Turner, 1995), but the appearance in 2004 of image tagging applications on the web has prompted a flurry of discussion and research into user tagging and its potential for indexing the mass of digital images now available. Early discussions focused on descriptive accounts of tagging systems (Hammond et al, 2005; Mathes, 2004), comparisons with more traditional forms of indexing (Matusiak, 2006; Peterson, 2006), and discussions of the strengths and limitations of user tagging (Golder & Huberman, 2006; Mathes, 2004; Matusiak, 2006; Munk & Mork, 2007; Peters & Stock, 2007).

The literature suggests that some general patterns of tag usage can be identified. All of the studies confirm Mathes' (2004) early predictions that tagging systems would conform to the Zipfian power law, where a small number of tags are used by a large number of users, leaving a 'long tail' of infrequently used tags. There was concern that the number of infrequently used tags would overrun tagging systems but Guy and Tonkin's study (2006) of Flickr and Delicious tags found that single use tags only accounted for 10-15% of the total number of tags, confirming the formation of a general consensus on tagging terms that had been identified by other studies (Angus, Thelwell & Stuart, 2008; Golder & Huberman, 2006; Kipp & Campbell, 2007). However, Munk and Mork question whether this agreement in terminology shows "a consensus of reflective meaning" or is the result of "a consensus to invest the fewest possible cognitive resources" (2007, p. 31).

Other research has looked more closely at the vocabulary of tags and their linguistic properties. It has been found that most tags describe what is depicted in the image, rather than what the image is about (Angus et al, 2008; Balasubramanian, Diekema & Goodrum, 2004; Bischoff, Firan, Nejdil & Paiu, 2008; Peters & Stock, 2007), and that the vast majority of tags are nouns (Grefenstette, 2008; Guy & Tonkin, 2006; Heckner, Muhlbacher & Wolff, 2008; Peters & Stock, 2007), with geographic places being the most commonly reported category in studies on Flickr (Beaudoin, 2007; Bischoff et al, 2008, Ding & Jacob, 2009; Overell, Sigurbjornsson & van Zwol, 2009).

Much of the research into tag vocabulary has concerned itself with vocabulary control and the problems of dealing with “uncontrolled and chaotic” tagging practices (Guy & Tonkin, 2006, para. 3). Some writers have suggested that educating users would improve the quality of tags, either through formatting guidelines, tag recommendations or other improvements to the user interface (Guy & Tonkin, 2006; Sigurbjornsson & van Zwol, 2008; Voss, 2007). Other studies have investigated how to add structure to tags through the use of tag clusters, ontologies or faceted data (Peters & Stock, 2007; Schmitz, 2006; Specia & Motta, 2007; Weller & Peters, 2008). However, concerns have been expressed about the loss of nuanced meanings that could result from the restriction of tagging vocabulary (Shirky, 2006).

Motivations for tagging fall into two basic categories (Golder & Huberman, 2006; Hammond et al, 2005; Marlow, Naaman, Boyd & Davis, 2006): those that are extrinsic to the tagger, such as social or altruistic motivations, and those that are relative to the tagger, such as organisational or selfish motivations. Whereas early studies hypothesised that users generally tag for their own benefit (Golder & Huberman, 2006; Hammond et al, 2005), more recent research suggest that social motivations are now more common (Ames & Naaman, 2007; Angus et al, 2008; Cox, Clough & Marlow, 2008).

Various authors note the enthusiasm of users towards image tagging (Angus et al, 2008; Beaudoin, 2007; Marshall, 2009), and that user tags are often “appropriate, thorough, and in many cases, authoritatively chosen” (Winget, 2006, p. 15). Research into the tagging of museum objects has been particularly enthusiastic about the value of user tags, with one study reporting that over 80% of user tags were new and valuable additions to the museum’s existing metadata (Chun, Hiwiller, Trant & Wyman, 2006), and another study stating that user tags “were of a quality that had not been experienced in other tagging projects” (Clayton, Morris, Venkatesha & Whitton, 2008, p. 12). However, some studies have noted that tags are more useful for clustering data than providing indexing terms as they tend to indicate user interests or goals rather than provide accurate content description (Ding, 2009; Razikin, Goh, Lee & Chua, 2008; Zeng & Li, 2008).

To date, few studies have compared user tags with user queries (Trant, 2008b), and those that have (Bischoff et al, 2008; Trant, 2008b), have done so at the level of individual terms rather than a more general consideration of the image facets that are described in user tags. While comparisons at the level

of individual terms may be appropriate in the narrow domain of a museum's collections, comparison of the image facets appearing in queries and tags would provide a broader picture of whether the terms assigned by users are useful for subsequent searching in the wider domain of the internet.

3. Methodology

- 1.
- 2.
- 3.

1. Research environment

Flickr, an online photo-sharing web site, has been chosen as a source of data for this study. Launched in 2004, Flickr allows users to upload images for storage online and annotate them with titles, descriptions or tags. Contributors have the choice of making the images publicly visible, only visible to family and friends, or to keep them private. Users can also choose to contribute their images to a Flickr ‘group’, thereby increasing the exposure of their images to a wider audience. In addition, a user can choose to permit others to add tags or comments to their images, although this is not a widely used function of the site (Marlow et al, 2006).

Flickr has been used in previous research investigating various aspects of social networking and folksonomy (for example, Angus et al, 2008; Guy & Tonkin, 2006; Overell et al, 2009; Schmitz, 2006). Its benefits to researchers are that data from the publicly available part of the site is readily accessible through the site’s application programming interface (API), and the use of system data allows researchers to study how end users behave in a real world setting (Nov, Naaman & Ye, 2008), rather than in an artificially-constructed laboratory situation, such as used in research by Cunningham & Masoodian (2006), Hastings (1995), or Hollink et al (2004). In addition, images in Flickr are not limited to particular subject domains unlike some earlier research focusing on art history (Chen, 2001; Hastings, 1995), history (Choi & Rasmussen, 2003), or cultural heritage (Trant, 2006; Van Hooland, 2006).

2. Research design

As Cox notes, the “sheer scale and flow of Flickr is daunting for analysis” (2008, p. 494), with well over 3 billion images in its dataset (Marshall, 2009). For the purposes of this study, a smaller sample of 250 images was collected.

Outcomes of previous research were used to develop three criteria to select these images. Firstly, as

Marlow et al note (2006), tagging practices have changed over time as users become familiar with the concept of tagging, and so the images chosen must have been recently uploaded to ensure that the research reflects current tagging behaviour. Secondly, the sample set should not be biased towards the tagging habits of any one individual, so only one image from each user was selected. Finally, only tags written in English were included in the study to enable the researcher to classify them correctly.

In addition, the researcher would need access to the images themselves in order that the correct meaning of a tag could be deduced. As discussed by Angus et al, it might be difficult to accurately determine the meaning of each tag “due to the complex nature of image interpretation” (2008, p.93), and so it was decided that the tags would not be looked at in isolation but their categorization would be informed by the image itself. In this way, any ambiguity in terms used, such as homonyms or abbreviations, could be accurately interpreted.

In order to test the quality of data that could be retrieved, a pilot study was carried out to download details about the 250 most recent images from Flickr, selecting only one image per user. A program using Flickr’s API was written by an experienced programmer to download images and associated metadata based on the above criteria. This pilot study revealed two issues that required attention. Firstly, only 23% of images had tags, so it was decided that a total of 2000 images would be downloaded in order to ensure that enough images with tags would be available for the study. Secondly, it was discovered that 24 hours after an image was initially uploaded, additional tags were added to around 7% of the images, but that few users added more tags once an image had been online for more than 24 hours. It was decided therefore that the program would need to select images that had been online for at least 24 hours, and this was achieved by asking the program to search for images from the 1000th page of Flickr’s recently uploaded stream.

3. Data collection

The program to collect the data from Flickr was therefore amended by the programmer to reflect the following requirements:

- A total of 2000 images were to be downloaded.

- The images were to be selected from the 1000th page of Flickr’s most recently uploaded stream.
- Only one image per user would be selected.
- For each image, the program would download the image and the associated metadata, including image ID, image title, tags, user ID, the date of upload, and the URL of the image.

The program was executed on 26 June 2009 and the results were saved to an Excel spreadsheet. The dataset was checked to ensure its compliance with the selection criteria. A sample of the data collected is shown in Appendix A, with an example of a page in Flickr showing an image and its associated tags in Appendix B.

4. Data analysis

Initial analysis of the dataset was undertaken to determine the percentage of users who tag images and the average number of tags assigned per image. This allowed some conclusions to be drawn about general tagging practices which informed discussions about the usefulness of tags for image retrieval. Next, all the images without tags were removed from the dataset, and any images with non-English tags were also removed. A random selection of 250 images from the remaining dataset was then selected for the tag categorization process.

Each tag was categorised using the facets from the Shatford matrix (see Table 1). This framework was chosen because of its frequent use in image retrieval research, making it “well established for the analysis of image content”, and because of its “dual applicability [...] to the characterisation of both images and the queries which address those images” (Armitage & Enser, 1997, p. 294). A coding sheet was devised using data from previous

	Specific	Generic	Abstract
Who?	Individually named person, group, or thing (S1)	Kind of person, group or thing (G1)	Mythical or fictitious being (A1)
What?	Individually named event or action (S2)	Kind of event, action or condition (G2)	Emotion or abstraction (A2)
Where?	Individually named	Kind of place:	Place symbolised (A3)

	geographical location (S3)	geographical or architectural (G3)	
When?	Linear time: date or period (S4)	Cyclical time: season or time of day (G4)	Emotion or abstraction symbolised by time (A4)

Table 1. Shatford facet matrix (source: Armitage & Enser, 1997, p. 290)

research (Armitage & Enser, 1997; Conduit & Rafferty, 2007; Shatford, 1986) with additional coding decisions noted as the categorisation process proceeded. Appendix C gives details of the coding sheet used, and sample tags from the current study are shown in Appendix D.

Each tag was categorised with reference to its accompanying image to ensure that its meaning was captured as accurately as possible, and reference was also made to other areas of Flickr and to general reference materials to help deduce the meaning of some tags, such as geographic locations or tags denoting membership of a Flickr group. In general, each tag was assigned to only one category within the matrix as each tag represented only one image attribute. However, judgements were made by the researcher regarding the need to assign multiple word tags to more than one category if more than one facet was contained within the tag, such as the tag ‘20090511alaska’ which was separated into a date and a location tag. In addition, some single word tags were combined if that would more accurately reflect the probable meaning, such as the tags ‘for’ and ‘sale’. Any tags that could not be assigned to one of the facets were recorded in a separate category for later analysis.

The results of the categorisation process were then compared to those found in previous studies into image queries to determine if the facets that are used to describe images match the facets used in image queries. A range of previous research into user queries using the Shatford matrix was consulted and the results were combined and normalised to produce an overall picture of user queries for comparison with the outcome of the tag analysis. Details of the previous studies consulted are given in Table 2. The findings of the data analysis and comparison with user queries are presented in the next chapter.

Name of study	User group of study
Westman & Oittinen (2006)	Journalists and archivists in newspaper editorial office
Van Hooland (2006)	Users of a national archive web site

Conduit & Rafferty (2007)	Combination of 7 previous studies covering art history, history, publishing, local history, and history of medicine
---------------------------	---

Table 2. Previous studies into user queries using Shatford matrix

4. Results

1. Statistical analysis of image tags

The initial download of 2000 images included 1246 images (62%) without any tags, and 754 images (38%) with at least one tag (see Figure 1). The total number of tags assigned to these 754 images was 3941, giving a mean average of 5.2 tags per image (see Table 3). However, as has been found by other studies, (Ames & Naaman, 2007; Angus et al, 2008; Sigurbjornsson & van Zwol, 2008; Winget, 2006), there was a wide difference in the number of tags per image, ranging from 1 to 49 tags, but the majority of users (66%) assign five or less tags to their images (see Figure 2).

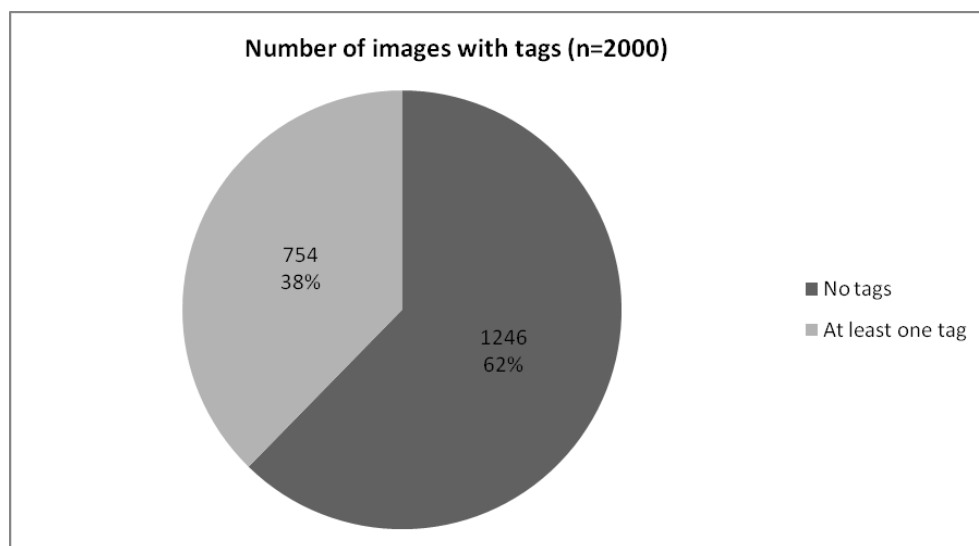


Figure 1. Number of images with tags

	All tagged images (n=754)	Images in sample set (n=250)
Total number of tags	3941	1121
Mean number of tags per image	5.2	4.5
Median number of tags per image	4	3
Modal number of tags per image	1	1

Maximum number of tags per image	49	28
----------------------------------	----	----

Table 3. Statistical analysis of tagged images

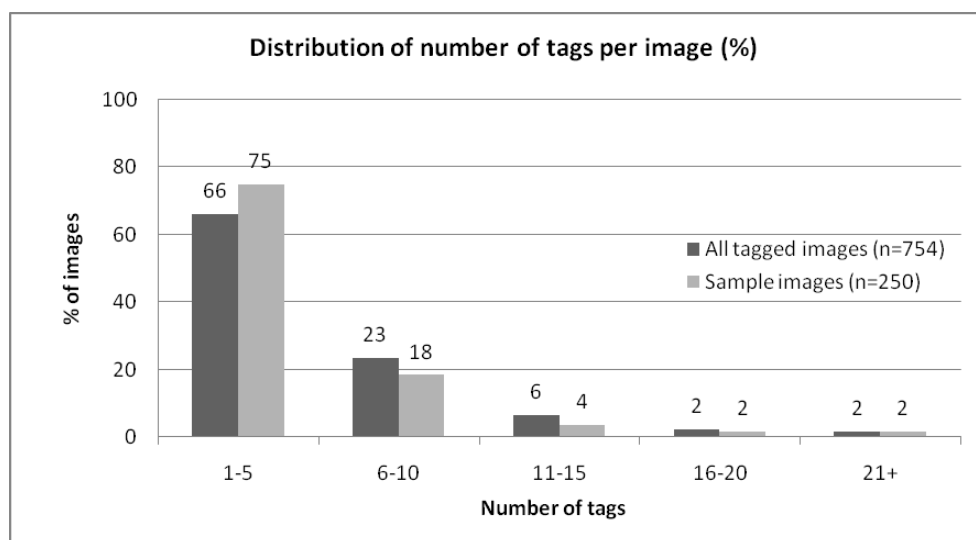


Figure 2. Distribution of number of tags per image

Of these 754 images, 199 had tags that included foreign words (26%), leaving a total of 553 images with tags in English. From these, a random selection of 250 images was chosen to create a sample set for the tag categorisation process. Analysis of this set showed a similar tag distribution to the larger dataset, although the largest number of tags for a single image was 28 which marginally lowered both the mean and the median averages (see Table 3).

2. Categorisation of image tags

The tags were then categorised using Shatford’s matrix. A total of 78% (839 tags) were able to be categorised, leaving 22% of tags which could not be assigned to any of the categories.

As can be seen from Figure 3, generic tags were used most frequently, accounting for more than half of all tags (59%), with a third (32%) of tags related to specific categories, and only 9% of tags assigned to abstract categories. Looking at the results in more detail, Figure 4 shows that most tags were categorised as G1 Generic Who (38%), with 19% categorised as S3 Specific Where, and 12% each as

S1 Specific Who and G2 Generic What. A list of the facets in order of ranking is given in Table 4.

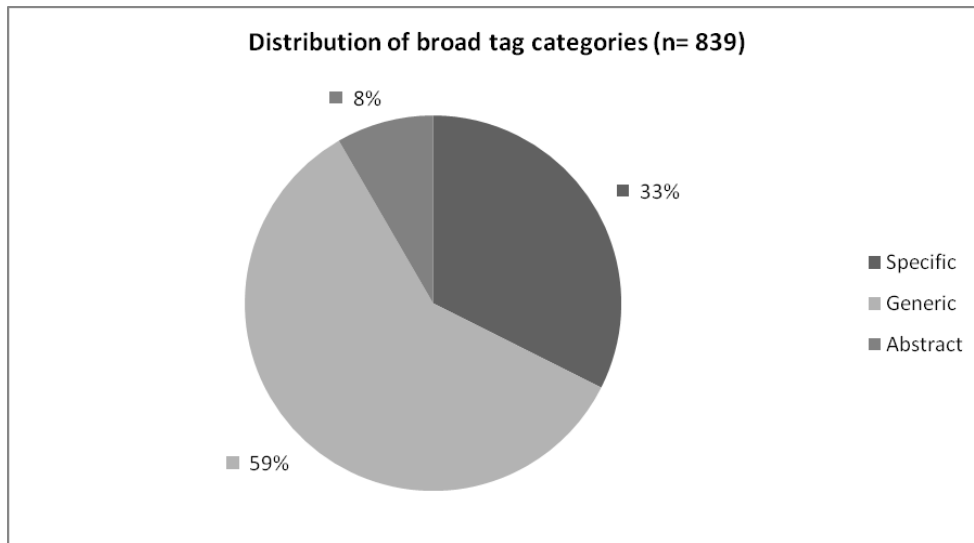


Figure 3. Distribution of broad tag categories

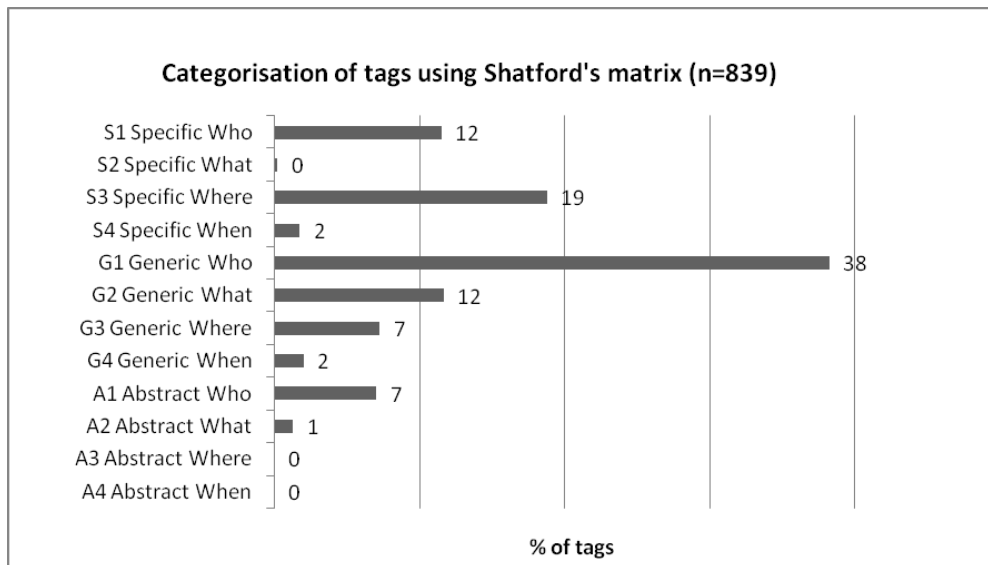


Figure 4. Categorisation of tags using Shatford's matrix

Rank	Facet	% of tags
1	G1 Generic Who	38.1
2	S3 Specific Where	18.8
3	S1 Specific Who	11.6
4	G2 Generic What	11.6
5	G3 Generic Where	7.3
6	A1 Abstract Who	7.3
7	G4 Generic When	2.0
8	S4 Specific When	1.8
9	A2 Abstract What	1.3

10	S2 Specific What	0.2
11	A3 Abstract Where	0
12	A4 Abstract When	0

Table 4. Ranking of tag facets

In order to ensure that the images with large numbers of tags did not dominate the results, those images with 15 or more tags (12 images) were removed from the statistical analysis to see if this affected the results. However, it was found that the percentages in each category remained broadly the same as before. A further analysis of the data was made to see if the behaviour of users who assigned a larger than average number of tags was different to those who assigned a smaller number. The results, shown in Figure 5, show that users who assign five or more tags tend to use marginally more generic tags than users who assign four or less tags, but in general, the pattern remain the same.

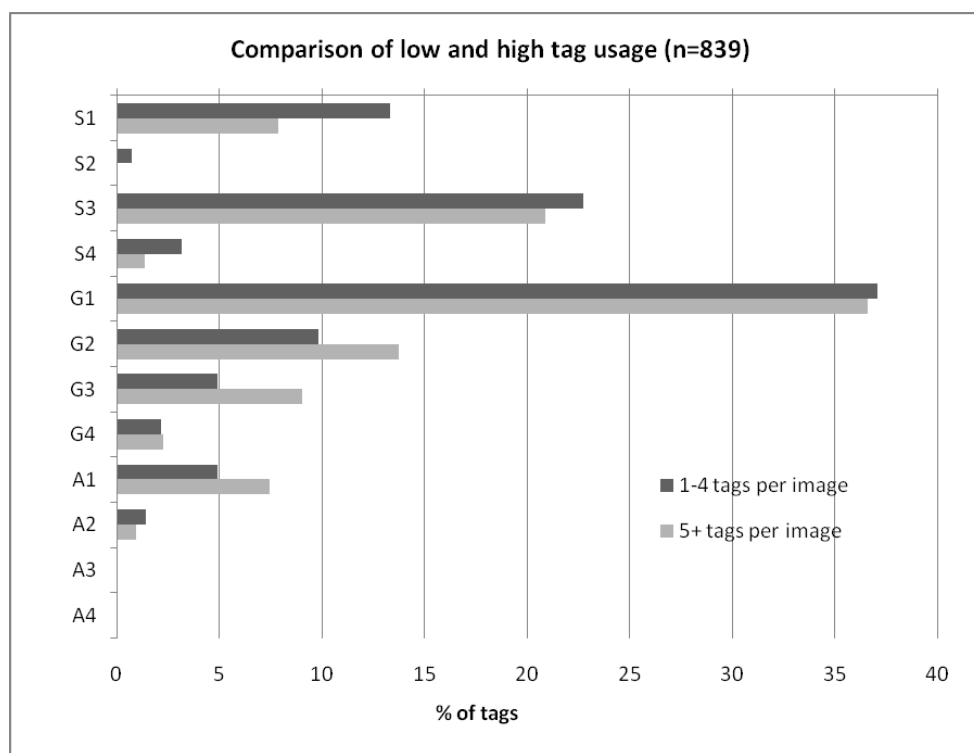


Figure 5. Comparison of low and high tag usage

The tags that could not be categorised using Shatford's matrix were analysed to identify any common themes and the results are presented in Figure 6. The majority of the tags (170 out of 234 tags) relate to the production of the image and its eventual use (the first five categories on the graph in Figure 6),

accounting for 15% of the total number of tags in the sample set. See Appendix E for examples of the tags that were coded in the ‘Other’ categories.

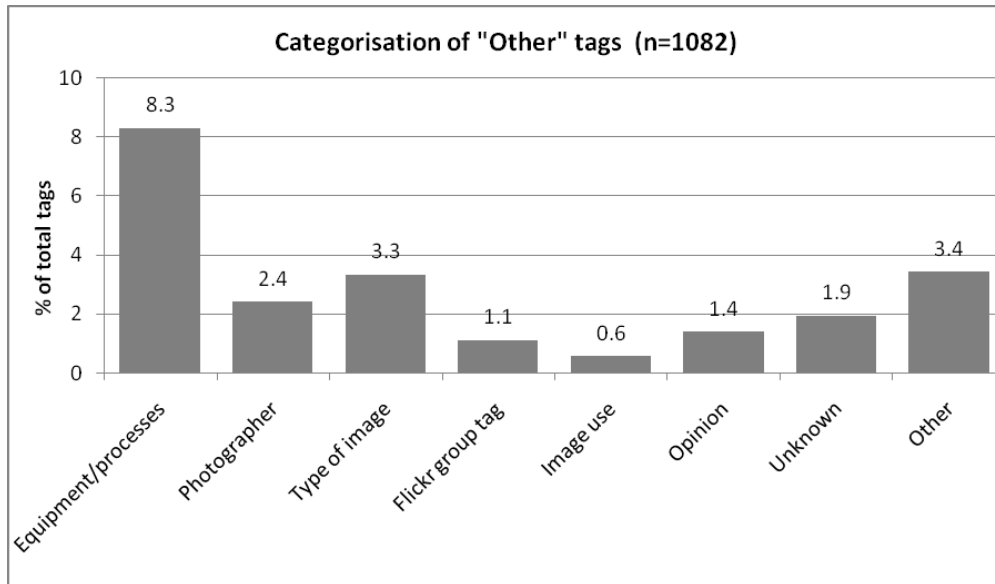


Figure 6. Categorisation of ‘Other’ tags

3. Comparison with categorisation of image queries

The results of the categorisation process were compared with the results of previous research into image queries. Figure 7 shows that while the majority of queries are for specific categories (62%), the majority of tags fall into generic categories (59%), with abstract categories accounting for very few queries or tags (3% for queries and 8% for tags).

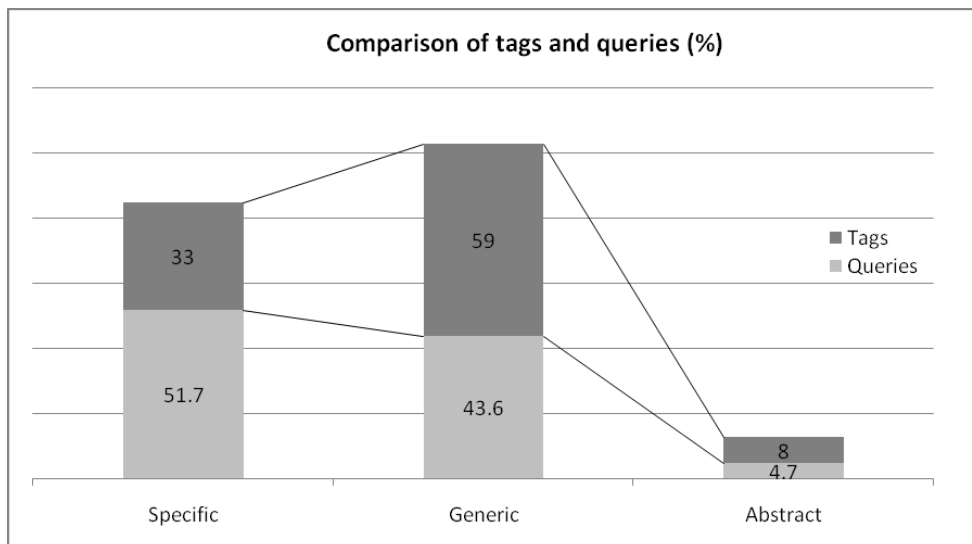


Figure 7. Comparison of tags and queries by broad category

A more detailed comparison of queries and tags using Shatford’s matrix is given in Figure 8. From this, it can be seen that there are significant differences in some of the categories. For example, there are more queries in the S1 Specific Who and S4 Specific Where categories, but significantly more tags in the G1 Generic Who category.

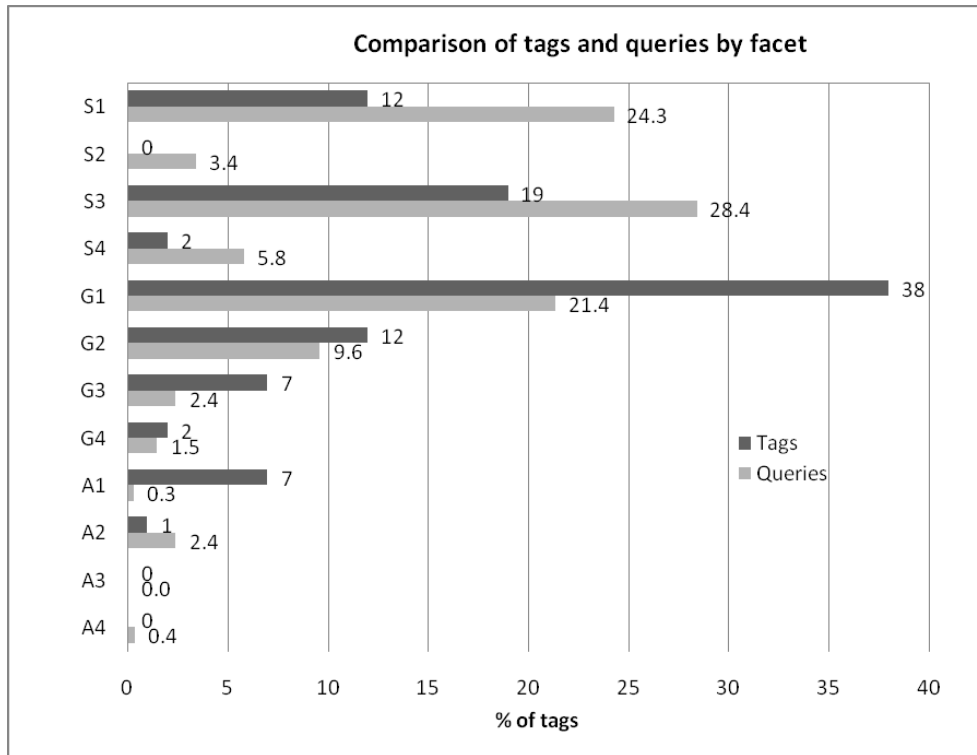


Figure 8. Comparison of tags and queries by individual facet

In order to see whether any of the individual query studies were overly skewing the figures for the user queries, the results of the individual studies were compared (see Figure 9). From this comparison, it appears that the results of the Van Hooland study are quite different to the other two studies, particularly in the case of the S3 Specific Where and G1 Generic Who categories. Therefore, the Van Hooland study was removed, and the remaining query studies were compared with the results of the tag analysis: the results are shown in Figure 10. While the difference between the incidence of tags and queries for S3 Specific Where category has been reduced, there are still significant differences between the results in the S1 Specific Who and G1 Generic Who categories.

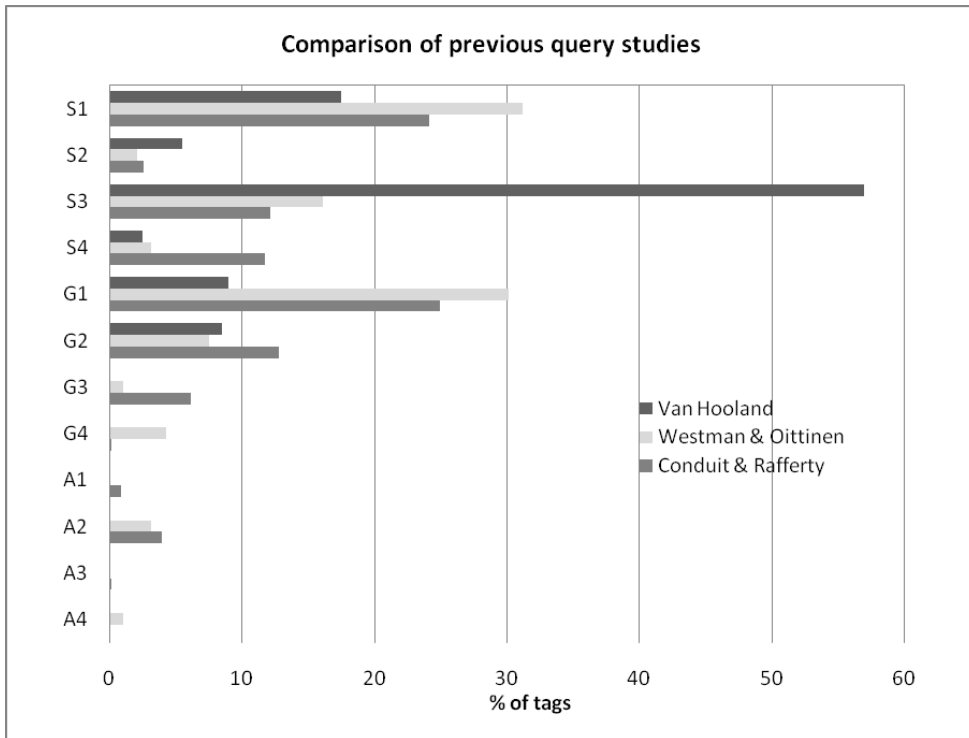


Figure 9. Comparison of previous query studies

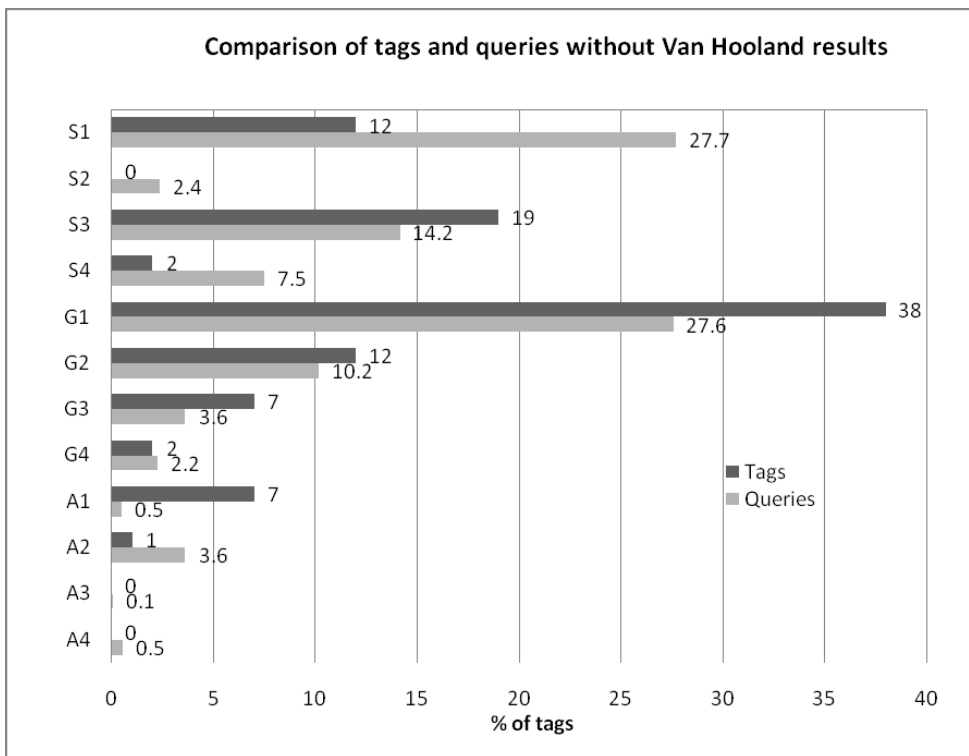


Figure 10. Comparison of tags and queries without Van Hooland results

A further analysis was made to amalgamate the specific, generic and abstract levels of the four facets of Who, What, Where and When to see if any patterns could be deduced. Figure 11 shows the results of this analysis. This shows that in the current study there is a strong similarity in the way that users tag images within these four facets when compared to previous research on image queries. The Who facet (people or objects) is the most commonly occurring in both tags and queries, accounting for 57% of tags and 56% of queries. The next most common facet is Where, but there is less correlation between tags and queries, with 26% of tags relating to this facet but only 18% of queries. The What facet is represented in similar quantities in both tags and queries (13% and 16% respectively), with the final facet, When, accounting for the smallest number of occurrences: 4% of tags and 10% of queries.

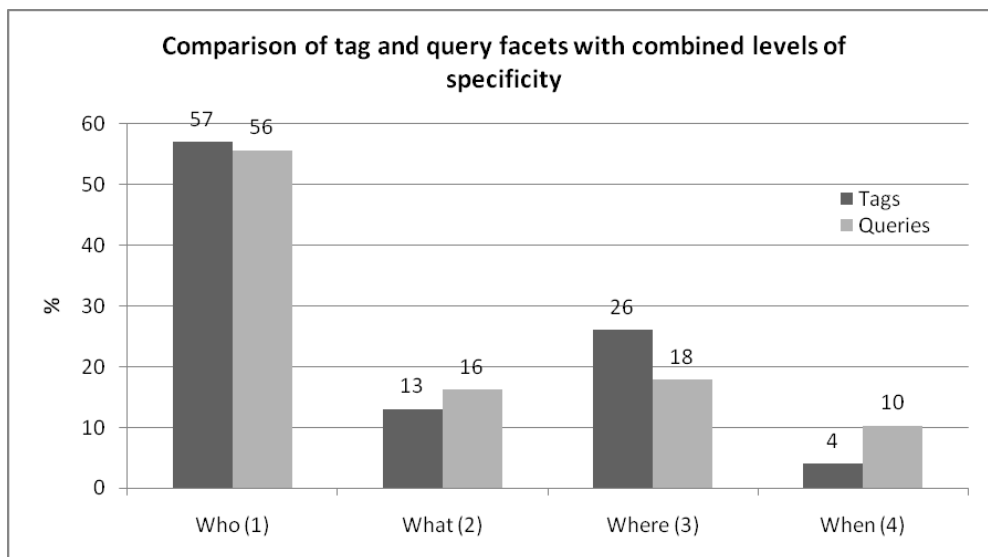


Figure 11. Comparison of tag and query facets with combined levels of specificity

From this analysis, it appears that, while users describe and search for images using similar facets, it is the level of specificity or abstraction that differs between image tags and image queries. A further discussion of the issues raised by the results will appear in the next chapter.

5.

Discussion

This research has considered the value of tags for image retrieval by investigating whether the way that users describe images in tags is similar to the way that users search for images, specifically looking at the image facets that are present in tags compared with those present in search queries.

The results confirm earlier research into image tagging (Hollink et al, 2004; Tamburrino, Schonmann, Vandewalle & Susstrunk, 2008), which showed that generic terms are more commonly used in image descriptions than specific terms, particularly by novice users such as the current user group (Beaudoin, 2008). The study also confirms that abstract terms are rarely used when describing images (Balasubramanian et al, 2004; Beaudoin, 2007; Hollink et al, 2004; Jorgensen, 1996).

The most commonly occurring image facets identified in the tags in this study were people and objects (Who facet), followed by locations (Where facet). This is in line with findings in previous studies into how users describe images, although there are disagreements as to which of these two facets has the highest incidence: Bischoff et al (2008), Overell et al (2009) and Jorgensen (1998) suggest that people and objects appear most often in image descriptions, but Beaudoin (2007), Marshall (2009), Schmitz (2006) and Sigurbjornsson & van Zwol (2008) found that location attributes were more prevalent.

When comparing image tags with query terms, the results of this research show that there are clear similarities in the facets that are present in tags and in query terms, with terms relating to people and objects, or locations being the most commonly used in both cases. This would suggest that the image attributes included by users in their image tags are the same as those that are of interest to users in their searches. On the other hand, the study noted differences in the specificity of terms used, with generic terms being more widely used in image tags and specific terms appearing more often in search queries, which could point to a limitation of image tags in that they may not contain the particular terms that appear in user queries.

While other research has also found similarities between the way that users describe and search for images, (Bischoff et al, 2008; Jorgensen, 1996), some studies have concluded that this is not the case

and that users search for and describe images in different ways (Goodrum, 2005; Trant, 2009). Enser (2008), in acknowledging the differences between the results from various studies, suggests that they may be explained by the different domains in which the studies were carried out. For example, Marshall's (2009) study concentrated on images of one particular location which may explain the high incidence of location tags.

The domain used in this study was Flickr, an online photo-sharing web site that was chosen as it offered the opportunity to study real end-users' tagging behaviours. There are several factors relating to this particular study that should be taken into account when considering the results in a wider context.

Firstly, the results of this study suggest a low level of interest in tagging on Flickr, with only 38% of images having tags. This is lower than the levels of tagging found in other studies on Flickr where over 60% of images were tagged (Marshall, 2009; Winget, 2006). One possible reason for the higher incidence of tagging in Winget's study is that the images were chosen from Flickr's 'interestingness' photostream: one of the criteria for inclusion in this photostream is that an image should have a large number of tags. Studies into tagging in the cultural heritage collections also report higher levels of tagging than the current study, with researchers noting the public's enthusiasm for tagging museum objects (Chun et al, 2006; Clayton et al, 2008; Trant, 2008b). The tagging behaviour of the Flickr users in this study may therefore not be representative of all taggers. The reasons for the low incidence of tagging in this study are unclear, although it has been suggested that tagging is not important for Flickr users as most browse the collection rather than carrying out a direct search (Cox et al, 2008).

Another factor that may have influenced the results is the limited sample size in this study, and a larger data set may have given different results. However, several other studies have noted that there appears to be a "significant consensus" (Marshall, 2009, p. 241) between users when they assign terms (Balasubramanian et al, 2004; Von Ahn & Dabbish, 2004) and so the results from this study may be more representative of a wider range of Flickr users than the sample size would suggest.

Some aspects of tagging behaviour may have also affected the results, particularly the inclusion of synonyms, such as 'cat' and 'cats' or 'woods' and 'forest', in a tag set. This was particularly

problematic in the S3 Specific Where category where several versions of a place name may appear in one tag set, for example, 'lauderdale, fort lauderdale, ft lauderdale, fl'. While the use of synonyms increases the likelihood of a match between search terms and image tags, it does affect the percentage of tags in each category in this study. A further investigation would be needed to see the full effect of this practice on the results. There are many other aspects of tagging behaviour that could also affect the value of tags for retrieval, such as homonyms, misspelling, and the use of foreign words, but as this research is concerned only with the facets represented by the tags, the 'correctness' of each tag has been ignored in this study.

The categorisation of the image tags and the use of Shatford's matrix also require some discussion. Time and resource limitations meant that the categorisation process was carried out by a single researcher and was therefore prone to a high degree of subjectivity. The researcher tried to compensate for this by reviewing the whole tag set more than once to ensure consistency in the categories assigned, but if the process had been carried out by more than one researcher a higher degree of objectivity could have been achieved in the results. However, other research has noted a relatively high level of agreement between reviewers when categorising queries (Bischoff et al, 2008; Chen, 2001; Hollink et al, 2004), indicating that the results may not change significantly if more than one reviewer had been involved.

As has been noted in other studies, difficulties were encountered in applying Shatford's categories to the tag set in this study (Armitage & Enser, 1997; Beaudoin, 2008; Bischoff et al, 2008; Hollink et al, 2004). One area of difficulty was in determining the difference between specific and generic facets, particularly relating to location (Where). The researcher's decision to assign any named location to S1 Specific Where meant that tags naming countries, for example, were labelled as Specific. However, as noted by Armitage and Enser (1997) and Marshall (2009), the use of such broad geographical tags has limited value in general retrieval as they could potentially return too many results. Reclassifying these tags as Generic would, however, have further widened the gap in specificity between image tags and query terms in this study. Problems were also encountered in categorising tags in the Abstract facets, as there was limited information in the literature to guide the researcher in determining the correct code. However, the low incidence of abstract tags meant that these difficulties did not greatly affect the final

results.

There were also issues in the applicability of Shatford's facets to the full range of tags present in the dataset: 22% of tags fell outside Shatford's 12 categories. The majority (15%) of these tags were related to the photographer, the camera or the intended use of the image. Angus et al (2008) suggest that tags such as these may only be useful for the individual uploading the photograph and are therefore not valuable in a general retrieval situation, so this limitation of Shatford's matrix may not be significant when considering the usefulness of tags for retrieval purposes.

Another observation about the use of Shatford's matrix concerns the broad nature of some of the facets. It was apparent that the majority (57%) of tags fell into the Who facets, but the structure of the matrix does not allow more detailed information to be discovered about the percentage of people, animals or objects that are included in this figure. As most studies support the view that people and objects form a large percentage of the tags and queries related to images (Bischoff et al, 2008; Jorgensen, 1998; Overell et al, 2009; Sigurbjornsson & van Zwol, 2008), a categorisation framework that allows for a more detailed look at these areas would give more useful information about the needs and behaviours of users.

Finally, an important point to note is that none of the query studies included in this research were carried out in an online environment as little such research exists (Enser, 2008; Jansen, 2008). It is therefore not clear whether the query studies used here reflect the way that users actually search online. However, Cunningham and Masoodian's 2006 study into general online image searching identified that a large proportion (70%) of image searches were for specific requests, so there may be some similarities between on- and off-line search strategies. This is certainly an area that would benefit from further research to determine the nature of user behaviour and image needs when searching for images online.

This study has concentrated on the use of image tags. Another avenue of research would be to investigate the effectiveness of other user-supplied descriptions, such as image titles or captions, to see how they would compare with image query terms. If, as has been suggested, they contain more specific descriptions than image tags, (Enser et al, 2007; Van Hooland, 2006), they could provide more specific

indexing terms which would in turn lead to greater retrieval success.

While the results of this research suggest that there are some limitations in using tags for searching because of their generic nature, it has been suggested that measuring the success of tags in terms of how well they perform in direct searches is to ignore their usefulness in browsing (Mathes, 2004). Browsing has been discussed in the literature as an important aspect of image retrieval due to the difficulties of successfully capturing the semantic content of an image in indexing terms (Enser, 2008; Layne, 1994). If the majority of image tags are generic, as this research has shown, their value may come from their ability to group together a range of relevant images from which a user can choose, rather than their ability to match with individual terms in a search query. Therefore, an area for future development could be in developing information systems that include more effective browsing mechanisms to make use of the information present in image tags.

This research has looked at one aspect of image tags, which has been to consider if there are similarities in the image facets that appear in user tags and user search queries. The results of this investigation have shown that similar image facets are present, but that there are differences in the level of specificity used. This limitation, together with the other problems caused by the uncontrolled nature of tags, supports the view that user tagging alone is not the solution to the problem of providing intellectual access to the millions of images available online. However, this does not mean that user tags are without value, and this study has shown that the tags provided by users do contain the same image attributes that appear in search queries. The challenge now is to find ways of incorporating tagging with other indexing methods to make better use of the potential that user tagging can offer in indexing images online.

6.

Conclusion

The aim of this research was to find out the value of tags for image retrieval by investigating whether the terms used in image tags are similar to the terms used in search queries. Rather than comparing individual terms, Shatford's matrix of image facets was used to compare the image attributes represented by the tags from images on Flickr with the attributes that had previously been identified in research into image queries.

The results of the research indicate that there are broad similarities between the image attributes represented by image tags and those appearing in image queries, with people and objects being the most common facet, followed by location. This would suggest that the aspects of images that users include in their image tags are the same as those that are of interest to users in their searches. However, the results also show that there are differences in the level of specificity of the terms used, with image tags containing a higher proportion of generic terms and image queries containing a larger number of specific terms. This is a potential problem for retrieval systems trying to create a match between image tags and query terms.

This research points to a need to develop retrieval systems that will bridge this gap between specific queries and generic tags in order to harness the collective power of user tagging and provide for more effective image retrieval. The study has suggested several areas for further research that may help to develop such systems, including the need for a greater understanding of search behaviour in an online environment, an investigation into the use of user titles or captions as a source of more specific indexing terms, and the development of methods to facilitate the use of browsing in image retrieval. It also supports the general consensus that user tagging needs to be incorporated with other indexing methods in order to overcome any limitations that are inherent in its uncontrolled nature.

The growth of online image tagging applications over the last few years and the enthusiasm for tagging that has been noted in many studies suggest that, in some environments, end-users enjoy tagging images and can provide tags of a high quality. This study has confirmed that, at a fundamental level, users tag images in a similar way to how they search for images, supporting the notion that user tagging could help

to provide a solution to the problems of scalability of traditional indexing techniques and so provide intellectual access to the millions of images that exist online.

Bibliography

- Ames, M., & Naaman, M. (2007, April). *Why we tag: motivations for annotation in mobile and online media*. Paper presented at CHI 2007, San Jose, California. Retrieved April 1, 2009, from <http://portal.acm.org/citation.cfm?id=1240772>
- Angus, E., Thelwell, M., & Stuart, D. (2008). General patterns of tag usage among university groups in Flickr [Electronic version]. *Online Information Review*, 32(1), 89-101.
- Armitage, L., & Enser, P. (1997). Analysis of user need in image archives [Electronic version]. *Journal of Information Science*, 23(4), 287-299.
- Balasubramanian, N., Diekema, A., & Goodrum, A. (2004). *Analysis of user image descriptions and automatic image indexing vocabularies: an exploratory study*. Retrieved April 3, 2009, from http://web.syr.edu/~diekemar/ruimte/Papers/cormedia2004_balasub.pdf
- Beaudoin, J. (2007). Flickr image tagging: patterns made visible [Electronic version]. *Bulletin of the American Society for Information Science and Technology*, (Oct/Nov), 26-29.
- Beaudoin, J. (2008). *A preliminary investigation of image indexing: the influence of domain knowledge, indexer experience and image characteristics*. Paper presented at the 19th Workshop of the American Society for Information Science and Technology Special Interest Group in Classification Research, Columbus, Ohio. Retrieved August 29, 2009, from <http://dlist.sir.arizona.edu/2462/>
- Bischoff, K., Firan, C., Nejdil, W., & Paiu, R. (2008). Can all tags be used for search?[Electronic version]. In *Proceedings of the 17th ACM Conference on Information and Knowledge Management* (pp. 193-202). New York, NY: ACM.
- Chen, H. (2001). An analysis of image queries in the field of art history [Electronic version]. *Journal of the American Society for Information Science and Technology*, 52(3), 260-273.
- Chen, H., & Rasmussen, E. (1999). Intellectual access to images - image database systems [Electronic version]. *Library Trends*, 48(2), 291-302. Retrieved April 28, 2009, from http://findarticles.com/p/articles/mi_m1387/is_2_48/ai_59473802/?tag=content;col1
- Choi, Y., & Rasmussen, E. (2003). Searching for images: the analysis of users' queries for image retrieval in American History [Electronic version]. *Journal of the American Society for*

Information Science and Technology, 54(6), 498-511.

Chu, H. (2003). *Information representation and retrieval in the digital age*. Medford: Information Today.

Chun, S., Hiwiller, D., Trant, J., & Wyman, B. (2006, March). *Steve.museum: an ongoing experiment in social tagging, folksonomy, and museums*. Paper presented at Museums and the Web, Albuquerque. Retrieved April 2, 2009, from <http://www.archimuse.com/mw2006/papers/wyman/wyman.html>

Clayton, S., Morris, S., Venkatesha, A., & Whitton, H. (2008). *User tagging of online cultural heritage items*. Retrieved April 2, 2009, from National Library of Australia: <http://www.nla.gov.au/nla/staffpaper/2008/documents/user-tagging.pdf>

Conduit, N., & Rafferty, P. (2007). Constructing an image indexing template for The Children's Society: users' queries and archivists' practice [Electronic version]. *Journal of Documentation*, 63(6), 898-919.

Cox, A. (2008). Flickr: a case study of Web 2.0 [Electronic version]. *Aslib Proceedings: New Information Perspectives*, 60(5), 493-516.

Cox, A., Clough, P., & Marlow, J. (2008, March). Flickr: a first look at user behaviour in the context of photography as serious leisure. *Information Research*, 13(1). Retrieved March 26, 2009, from <http://InformationR.net/ir/13-1/paper336.html>

Cunningham, S., & Masoodian, M. (2006). Looking for a picture: an analysis of everyday image information searching [Electronic version]. In *Proceedings of the 6th ACM/IEEE-CS Joint Conference on Digital Libraries* (pp. 198-199). New York, NY: ACM.

Cunningham, S., Bainbridge, D., & Masoodian, M. (2004). How people describe their image information needs: A grounded theory analysis of visual arts queries [Electronic version]. In *Proceedings of the 4th ACM/IEEE-CS Joint Conference on Digital Libraries* (pp. 47-48). New York, NY: ACM.

Ding, Y., & Jacob, E. (2009). Profiling social networks. *D-Lib Magazine* 15(3/4). Retrieved May 3, 2009, from <http://www.dlib.org/dlib/march09/ding/03ding.html>

Enser, P. (2008). The evolution of visual information retrieval [Electronic version]. *Journal of*

Information Science, 34(4), 531–546.

- Enser, P., Sandom, C., Hare, J., & Lewis, P. (2007). Facing the reality of semantic image retrieval [Electronic version]. *Journal of Documentation*, 63(4), 465-481.
- Fidel, R. (1997). The image retrieval task: implications for the design and evaluation of image databases [Electronic version]. *New Review of Hypermedia and Multimedia*, 3(1),181-199.
- Golder, S., & Huberman, B. (2006). Usage patterns of collaborative tagging systems [Electronic version]. *Journal of Information Science*, 32(2), 198-208.
- Goodrum, A. (2005). I can't tell you what I want, but I'll know it when I see it. *Reference & User Services Quarterly*, 45(1), 46-53.
- Goodrum, A., & Spink, A. (2001). Image searching on the Excite web search engine [Electronic version]. *Information Processing & Management*, 37(2), 292-311.
- Grefenstette, G. (2008, May). *Comparing the language used in Flickr, general web pages, Yahoo Images, and Wikipedia*. Paper presented at the International Conference on Language Resources and Evaluation. Retrieved April 2, 2009, from http://www.lrec-conf.org/proceedings/lrec2008/workshops/W3_Proceedings.pdf#page=10
- Greisdorf, H., & O'Connor, B. (2002). Modelling what users see when they look at images: a cognitive viewpoint [Electronic version]. *Journal of Documentation*, 58(1), 6-29.
- Guy, M., & Tonkin, E. (2006). Folksonomies: tidying up tags? *D-Lib Magazine*, 12 (1). Retrieved February 20, 2009, from <http://www.dlib.org/dlib/january06/guy/01guy.html>
- Hammond, T., Hannay, T., Lund, B., & Scott, J. (2005, April). Social bookmarking tools (I): a general review. *D-Lib Magazine*, 11(4). Retrieved April 18, 2009, from <http://www.dlib.org/dlib/april05/hammond/04hammond.html>
- Hastings, S. (1995). Index access points in a study of intellectual access to digitized art images. In D. Bearman (Ed.), *Multimedia Computing and Museums: Selected Papers from the Third International Conference on Hypermedia and Interactivity in Museums (ICHIM '95 / MCN '95): Vol. 1* (pp. 299-309). Toronto: Archives and Museums Informatics.
- Heckner, M., Mühlbacher, S., & Wolff, C. (2008). Tagging tagging. Analysing user keywords in scientific bibliography management systems. *Journal of Digital Information*, 9(27). Retrieved

April 21, 2009, from <http://journals.tdl.org/jodi/article/view/246/208>

- Hollink, L., Schreiber, A., Wielinga, B., & Worrying, M. (2004). Classification of user image descriptions [Electronic version]. *International Journal of Human-Computer Studies*, 61(5), 601-626.
- Hunter, J., Khan, I., & Gerber, A. (2008). HarvANA - harvesting community tags to enrich collection metadata [Electronic version]. In *Proceedings of the 8th ACM/IEEE-CS joint conference on Digital libraries* (pp. 147-156). New York, NY: ACM.
- Jaimes, R., & Chang, S. (2000, January). *A conceptual framework for indexing visual information at multiple levels*. Paper presented at IS&T/SPIE Internet Imaging, San Jose, CA. Retrieved April 20, 2009, from <https://eprints.kfupm.edu.sa/17006/>
- Jansen, B. (2008). Searching for digital images on the web [Electronic version]. *Journal of Documentation*, 64(1), 81-101.
- Jansen, B., Spink, A., & Pfaff, A. (2000, November). *Linguistic aspects of web queries*. Paper presented at the annual meeting of the American Society of Information Science, Chicago, IL. Retrieved March 31, 2009, from <http://jimjansen.tripod.com/academic/pubs/asis2000/asis2000.html>
- Jorgensen, C. (1996, October). *Indexing images: testing an image description template*. Paper presented at the ASIS Annual Conference. Retrieved April 2, 2009, from <http://www.asis.org/annual-96/ElectronicProceedings/jorgensen.html>
- Jorgensen, C. (1998). Attributes of images in describing tasks [Electronic version]. *Information Processing & Management*, 34(2/3), 161-174.
- Jorgensen, C., Jaimes, A., Benitez, A., & Chang, S. (2001). A conceptual framework and empirical research for classifying visual descriptors [Electronic version]. *Journal of the American Society for Information Science and Technology*, 52(11), 938-947.
- Jorgensen, C. (2007). *Image access, the semantic gap, and social tagging as a paradigm shift*. Retrieved April 1, 2009, from http://dlist.sir.arizona.edu/2064/01/Jorgensen_Update_SIG-CR_2007.doc
- Kato, M., Ohshima, H., Oyama, S., & Tanaka, K. (2008). Can social tagging improve web image search? [Electronic version]. In J. Bailey, D. Maier, K. Schewe, B. Thalheim, & X. S. Wang (Eds.), *Proceedings of the 9th International Conference on Web Information Systems*

- Engineering* (pp. 235-249). Berlin, Springer-Verlag.
- Kipp, M., & Campbell, D. (2007). Patterns and inconsistencies in collaborative tagging systems: an examination of tagging practices [Electronic version]. *Proceedings of the American Society for Information Science and Technology*, 43(1), 1-18.
- Layne, S. (1994). Some issues in the indexing of images [Electronic version]. *Journal of the American Society for Information Science*, 45(8), 583-588.
- Lee, C., Goh, D., Razikin, K., & Chua, A. (2009). Tagging, sharing and the influence of personal experience. *Journal of Digital Information*, 10(1). Retrieved February 25, 2009, from <http://journals.tdl.org/jodi/article/view/275>
- Lerman, K., & Jones, L. (2007, March). *Social browsing on Flickr*. Paper presented at the International Conference on Weblogs and Social Media, Boulder, CO. Retrieved April 20, 2009, from <http://www.icwsm.org/papers/3--Lerman-Jones.pdf>
- Macgregor, G., and McCulloch, E. (2006). Collaborative tagging as a knowledge organisation and resource discovery tool [Electronic version]. *Library Review*, 55(5), 291-300.
- Marlow, C., Naaman, M., Boyd, D., & Davis, M. (2006, May). *Position paper, tagging, taxonomy, Flickr, article, toread*. Paper presented in the Collaborative Web Tagging Workshop at the 15th International World Wide Web Conference, Edinburgh. Retrieved April 2, 2009, from <http://www.semanticmetadata.net/hosted/taggingws-www2006-files/29.pdf>
- Marshall, C. (2009). No bull, no spin: a comparison of tags with other forms of user metadata [Electronic version]. In *Proceedings of the 9th ACM/IEEE-CS Joint Conference on Digital Libraries* (pp. 241-250). New York, NY: ACM.
- Mathes, A. (2004). *Folksonomies - cooperative classification and communication through shared metadata*. Retrieved February 21, 2009, from adammathes.com: <http://www.adammathes.com/academic/computer-mediated-communication/folksonomies.html>
- Matusiak, K. (2006). Towards user-centered indexing in digital image collections [Electronic version]. *OCLC Systems & Services*, 22(4), 283-298.
- Merholz, P. (2004). *Metadata for the masses*. Retrieved May 26, 2009, from <http://www.adaptivepath.com/ideas/essays/archives/000361.php>

- Morrison, P. (2007, October/November). Why are they tagging, and why do we want them to? *Bulletin of the American Society for Information Science and Technology*, 34(1). Retrieved February 15, 2009, from <http://www.asis.org/Bulletin/Oct-07/morrison.html>
- Morrison, P. (2008). Tagging and searching: search retrieval effectiveness of folksonomies on the World Wide Web [Electronic version]. *Information Processing & Management*, 44(4), 1562-1579.
- Munk, T., & Mork, K. (2007). Folksonomy, the power law & the significance of the least effort. *Knowledge Organization*, 34(1), 16-33.
- Naaman, M., et al. (2004, October). Context data in geo-referenced digital photo collections [Electronic version]. In *Proceedings of the 12th Annual ACM International Conference on Multimedia* (pp. 196-203). New York, NY: ACM.
- Nov, O., Naaman, M., & Ye, C. (2008). What drives content tagging: the case of photos on Flickr [Electronic version]. In *Proceeding of the 26th Annual SIGCHI Conference on Human factors in Computing Systems* (pp. 1097-1100). New York, NY: ACM.
- O'Reilly, T. (2005). *What is Web 2.0: design patterns and business models for the next generation of software*. Retrieved May 26, 2009, from <http://www.oreillynet.com/lpt/a/6228?page=3#designpatterns>
- Overell, S., Sigurbjornsson, B., & van Zwol, R. (2009). Classifying tags using open content resources [Electronic version]. In *Proceedings of the Second ACM International Conference on Web Search and Data Mining* (pp.64-73). New York, NY: ACM.
- Panofsky, E. (1955). *Meaning in the visual arts: papers in and on art history* [Electronic version]. Garden City: Doubleday.
- Peters, I., & Stock, W. (2007). Folksonomy and information retrieval [Electronic version]. *Proceedings of the American Society for Information Science and Technology*, 44(1), 1-28.
- Peterson, E. (2006, November). Beneath the metadata: some philosophical problems with folksonomy. *D-Lib Magazine*, 12(11). Retrieved February 17, 2009, from <http://www.dlib.org/dlib/november06/peterson/11peterson.html>
- Pu, H. (2003). An analysis of web image queries for search [Electronic version]. *Proceedings of the American Society for Information Science and Technology*, 40(1), 340-348.

- Pu, H. (2008). An analysis of failed queries for web image retrieval [Electronic version]. *Journal of Information Science*, 34(3), 275-289.
- Rafferty, P., & Hilderley, R. (2005). *Indexing multimedia and creative works: the problems of meaning and interpretation*. Aldershot: Ashgate.
- Rafferty, P., & Hilderley, R. (2007). Flickr and democratic indexing: dialogic approaches to indexing [Electronic version]. *Aslib Proceedings: New Information Perspectives*, 59(4/5), 397-410.
- Rafferty, P. (2009, March). *Informative tagging of images: the importance of modality in interpretation*. Paper presented at Congress ISKO-Spain, Valencia. Retrieved June 3, 2009, from <http://dialnet.unirioja.es/>
- Rasmussen, E. (1997). Indexing Images. *Annual Review of Information Science and Technology*, 32, 169-196.
- Razikin, K., Goh, D., Lee, C., & Chua, A. (2008). Weighing the usefulness of social tags for content discovery [Electronic version]. In G. Buchanan, M. Masoodian, & S.J. Cunningham (Eds.), *Proceedings of the 11th International Conference on Asian Digital Libraries: Universal and Ubiquitous Access To Information* (pp. 51–60). Berlin, Springer-Verlag.
- Rorissa, A. (2008). User-generated descriptions of individual images versus labels of groups of images: a comparison using basic level theory [Electronic version]. *Information Processing & Management*, 44(5), 1741–1753.
- Schmidt, S., & Stock, W. (2009). Collective indexing of emotions in images. A study in emotional information retrieval [Electronic version]. *Journal of the American Society for Information Science and Technology* 50(5), 863-876.
- Schmitz, P. (2006, May). *Inducing ontology from Flickr tags*. Paper presented at the 15th International World Wide Web Conference, Edinburgh. Retrieved April 1, 2009, from <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.130.4081&rep=rep1&type=pdf>
- Shatford, S. (1986). Analyzing the subject of a picture: a theoretical approach. *Cataloging and Classification Quarterly*, 6(3), 39-61.
- Shirky, C. (2006). *Ontology is overrated: categories, links, and tags*. Retrieved February 19, 2009, from http://www.shirky.com/writings/ontology_overnated.html

- Sigurbjornsson, B., & van Zwol, R. (2008, April). *Flickr tag recommendation based on collective knowledge*. Paper presented at the 17th International World Wide Web Conference, Beijing. Retrieved April 1, 2009, from <http://www2008.org/papers/pdf/p327-sigurbjornssonA.pdf>
- Smith, G. (2008, August/September). Tagging: emerging trends. *Bulletin of the American Society for Information Science and Technology*, 34(6). Retrieved February 15, 2009, from http://www.asis.org/Bulletin/Aug-08/AugSep08_Smith.html
- Specia, L., & Motta, E. (2007). Integrating folksonomies with the semantic web [Electronic version]. In E. Franconi, M. Kifer, and W. May (Eds.): *Proceedings of the 4th European Semantic Web Conference. The Semantic Web: Research and Applications* (pp. 624-639). Berlin: Springer-Verlag.
- Spiteri, L. (2007, June). Structure and form of folksonomy tags: the road to the public library catalogue. *Webology* 4(2). Retrieved April 19, 2009, from <http://www.webology.ir/2007/v4n2/a41.html>
- Tamburrino, D., Schonmann, P., Vandewalle, P., & Susstrunk, S. (2008, January). *The flux: creating a large annotated image database*. Paper presented at IS&T/SPIE Electronic Imaging: Image Quality and System Performance V, San Jose, CA. Retrieved March 27, 2009, from <http://infoscience.epfl.ch/record/114219>
- Trant, J. (2006). *Understanding searches of an on-line collections catalogue*. Retrieved April 20, 2009, from <http://conference.archimuse.com/system/files/trantSearchTermAnalysis061220a.pdf>
- Trant, J. (2008a). Studying social tagging and folksonomy: a review and framework. *Journal of Digital Information*, 10(1). Retrieved January 29, 2009, from <http://journals.tdl.org/jodi/article/view/269/278>
- Trant, J. (2008b). Tagging, folksonomy and art museums: early experiments and ongoing research. *Journal of Digital Information*, 10(1). Retrieved January 29, 2009, from <http://journals.tdl.org/jodi/issue/view/65>
- Trant, J. (2009). *Tagging, folksonomy and art museums: results of steve.museum's research*. Retrieved August 30, 2009, from <http://conference.archimuse.com/files/trantSteveResearchReport2008.pdf>
- Turner, J. (1995, October). *Comparing user-assigned indexing terms with indexer-assigned*

- indexing terms for storage and retrieval of moving images*. Paper presented at the 58th ASIS Annual Meeting, Chicago. Retrieved June 7, 2009, from <http://www.mapageweb.umontreal.ca/turner/english/texts/asis95.htm>
- Turner, J. (1997, June). *Indexing pictures: some considerations*. Paper presented at the Annual Meeting of the Council on Botanical and Horticultural Libraries, Montreal. Retrieved June 7, 2009, from <http://www.mapageweb.umontreal.ca/turner/english/texts/cbhl97.htm>
- Van Hooland, S. (2006). *From spectator to annotator: possibilities offered by user-generated metadata for digital cultural heritage collections*. Retrieved March 27, 2009, from <http://eprints.rclis.org/7156/1/Usergeneratedmetadata.pdf>
- Vander Wall, T. (2007, February). *Folksonomy*. Retrieved March 27, 2009, from <http://www.vanderwal.net/folksonomy.html>
- von Ahn, L., & Dabbish, L. (2004, April). *Labeling images with a computer game*. Paper presented at CHI 2004, Vienna. Retrieved June 2, 2009, from <http://www.cs.cmu.edu/~biglou/ESP.pdf>
- Voss, J. (2007). *Tagging, folksonomy & co – renaissance of manual indexing?* Retrieved April 1, 2009, from Cornell University Library web site: http://arxiv.org/PS_cache/cs/pdf/0701/0701072v2.pdf
- Weller, K., & Peters, I. (2008, September). *Seeding, weeding, fertilizing - different tag gardening activities for folksonomy maintenance and enrichment*. Paper presented at I-Semantics, Graz. Retrieved April 1, 2009, from http://i-know.tugraz.at/blog/wp-content/uploads/2008/11/13_seeding-weeding-fertilizing.pdf
- Westman, S., & Oittinen, P. (2006). Image retrieval by end-users and intermediaries in a journalistic work context [Electronic version]. In *Proceedings of the 1st International Conference on Information Interaction in Context* (pp. 102-110). New York, NY: ACM.
- Winget, M. (2006, November). *User-defined classification on the online photo sharing site Flickr... or, how I learned to stop worrying and love the million typing monkeys*. Paper presented at the 17th Workshop of the American Society for Information Science and Technology Special Interest Group in Classification Research. Retrieved April 2, 2009, from <http://dlist.sir.arizona.edu/1854/01/winget.pdf>
- Zeng, D., & Li, H. (2008). How useful are tags? - an empirical analysis of collaborative tagging for web

pages [Electronic version]. In *Proceedings of the IEEE ISI 2008 Paisi, Paccf, and SOCO International Workshops on Intelligence and Security Informatics* (pp 320-330). Berlin, Springer-Verlag.

Appendix A: Sample data

CleanTags	Date Uploaded	Photo ID	Title	UserID	WebURL
	20/06/2009 17:45	3643785153	DSC08185	29928304 @N03	http://www.flickr.com/photos/29928304@N03/3643785153/
venice blackandwhite italy attesa	20/06/2009 17:45	3644590646	Lonely stone	24488377 @N08	http://www.flickr.com/photos/24488377@N08/3644590646/
kenya masaimara cbswt	20/06/2009 17:45	3643784541	Masai Shehab	24921083 @N03	http://www.flickr.com/photos/24921083@N03/3643784541/
	20/06/2009 17:45	3644590370		33986282 @N06	http://www.flickr.com/photos/33986282@N06/3644590370/
	20/06/2009 17:45	3643784471	P6190035	8528355@ N02	http://www.flickr.com/photos/8528355@N02/3643784471/
raw	20/06/2009 17:45	3643784503	DSC_0356.NEF	34929804 @N08	http://www.flickr.com/photos/34929804@N08/3643784503/
	20/06/2009 17:45	3643784243	DSC_5898	39018732 @N08	http://www.flickr.com/photos/39018732@N08/3643784243/
wool for ravelry	20/06/2009 17:45	3644590178	Acrylic mohair	24704271 @N06	http://www.flickr.com/photos/24704271@N06/3644590178/
	20/06/2009 17:44	3644590034	mi waton	39208073 @N08	http://www.flickr.com/photos/39208073@N08/3644590034/
	20/06/2009 17:44	3643784037	TorresCoto10	96412596 @N00	http://www.flickr.com/photos/96412596@N00/3643784037/
	20/06/2009 17:44	3644589918	IMG_2050.JPG	30028113 @N02	http://www.flickr.com/photos/30028113@N02/3644589918/
	20/06/2009 17:44	3644589832	IMG_3377	35042472 @N08	http://www.flickr.com/photos/35042472@N08/3644589832/
	20/06/2009 17:44	3643783619	P4121178	13348685 @N07	http://www.flickr.com/photos/13348685@N07/3643783619/
geek pr	20/06/2009 17:44	3643783637	DIY is pain	24352398 @N08	http://www.flickr.com/photos/24352398@N08/3643783637/
	20/06/2009 17:44	3644589540	CANVAS CIRCLES1	38324399 @N06	http://www.flickr.com/photos/38324399@N06/3644589540/
sparshjun20 09	20/06/2009 17:44	3643783313	Sparsh, Jun 2009 009	26571293 @N08	http://www.flickr.com/photos/26571293@N08/3643783313/

city cloud mist tree silhouette fog capetown	20/06/2009 17:44	3644589140	City of Clouds	32419885 @N07	http://www.flickr.com/photos/32419885@N07/3644589140/
--	---------------------	------------	----------------	------------------	---

Table 5. Sample data from Flickr download

Appendix B: Example of a Flickr page

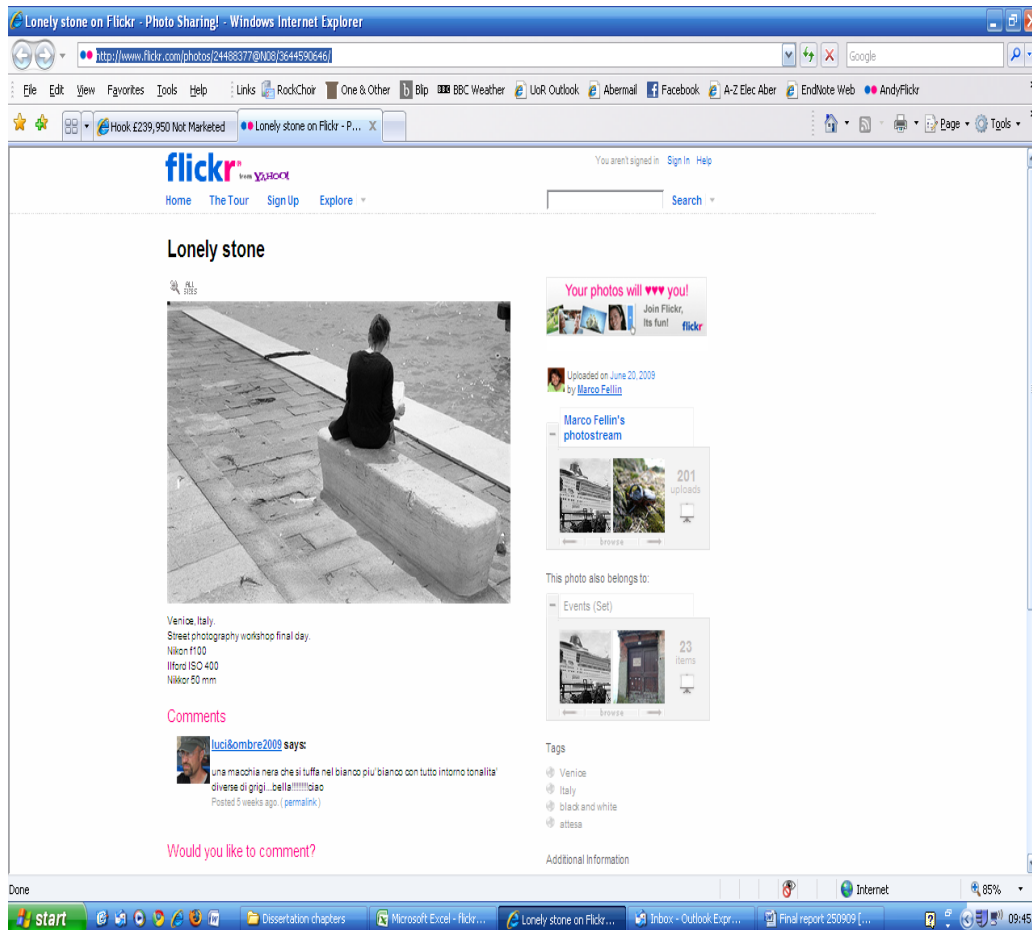


Figure 12. Example of Flickr page showing image and tags

Appendix C: Coding sheet

Shatford's categories as amended by Armitage & Enser	Shatford's examples	Armitage & Enser's examples	Conduit & Rafferty's examples		
S1	Specific Who	individually named person, group, thing	Guaranty Building	Napoleon	specific person or thing; brand name
S2	Specific What	individually named event, action	1980 Rose Bowl	Martyrdom of St. Lawrence	-
S3	Specific Where	Individually named geographical location	New York; Mars	Ranelagh Gardens; the moon	country; city; institution/building; OS grid reference;
S4	Specific When	linear time: date or period	June 1885; renaissance	1799; 50s; 20th C; Victorian	historical period; date image created; date object in image created
G1	Generic Who	kind of person or thing	skyscraper; actress; woman; oil painting	Suffragettes; writers	age range; gender; building material; colour; description of garment; self-portrait
G2	Generic What	kind of event, action, condition	football game; death; decay; sleep; buying	dolphin leap; anaesthesia; entertained; ceremony	weather; lighting conditions;

G3	Generic Where	kind of place: geographical, architectural	landscape, cityscape; interior; planet	game reserves; asylums; kitchen	indoors/ outdoors; urban/rural
G4	Generic When	cyclical time: season, time of day	spring; night	night; dawn; equinox	day/night; season; time of day;
A1	Abstract Who	mythical or fictitious being	academic art; modern architecture; muse	Janus' head; King Arthur	-
A2	Abstract What	emotion or abstraction	commerce; sorrow	iconography; allegory; phobia; vanity	concepts; emotion; mood;
A3	Abstract Where	place symbolised	paradise; Bethlehem; Mount Olympus; Jerusalem	-	-
A4	Abstract When	emotion, abstraction symbolised by time	rarely used - Father Time;	-	-

Table 6. Coding sheet with coding examples from other research

Appendix D: Examples of coded tags

	Specific	Generic	Abstract
Who?	S1: Alexandra Nine Inch Nails Harley Davidson	G1: woman tree toddler building	A1: art friends family Bulbasaur
What?	S2: Lemans2009	G2: sewing fundraiser 50 th birthday natural light	A2: business travel worry scary
Where?	S3: Grand Canyon Tel Aviv	G3: desert garden	A3: (none)

	Rum River USA	downtown street	
When?	S4: 2009 June 16-19 Miocene	G4: summer evening dusk	A4: (none)

Table 7. Example of how tags in the current study were coded

Appendix E: Examples of 'Other' tags

Category	Sample tags
Equipment /processes	nikond200 macro photoshop mobile shozu iso3200
Photographer	takenbymark joeynashoriginals photoboet pwillsportfolio jasonphillipsdjason
Type of image	photo bw scannedfromprints bestofcolourlowres screenshot
Flickr group tag	beinspired macrolicious tweaktoday
Image use	secondlife profile

	webdesign avatar
Opinion	cute cool fun pretty funky surreal
Unknown	ami8avis mhpc frm tanatan
Other	oneofmypics daytrip thingsthatareinspiringme ilovebirdies anotheraliceinwonderlandshotnextweekithink

Table 8. Examples of tags that could not be categorised using Shatford's matrix