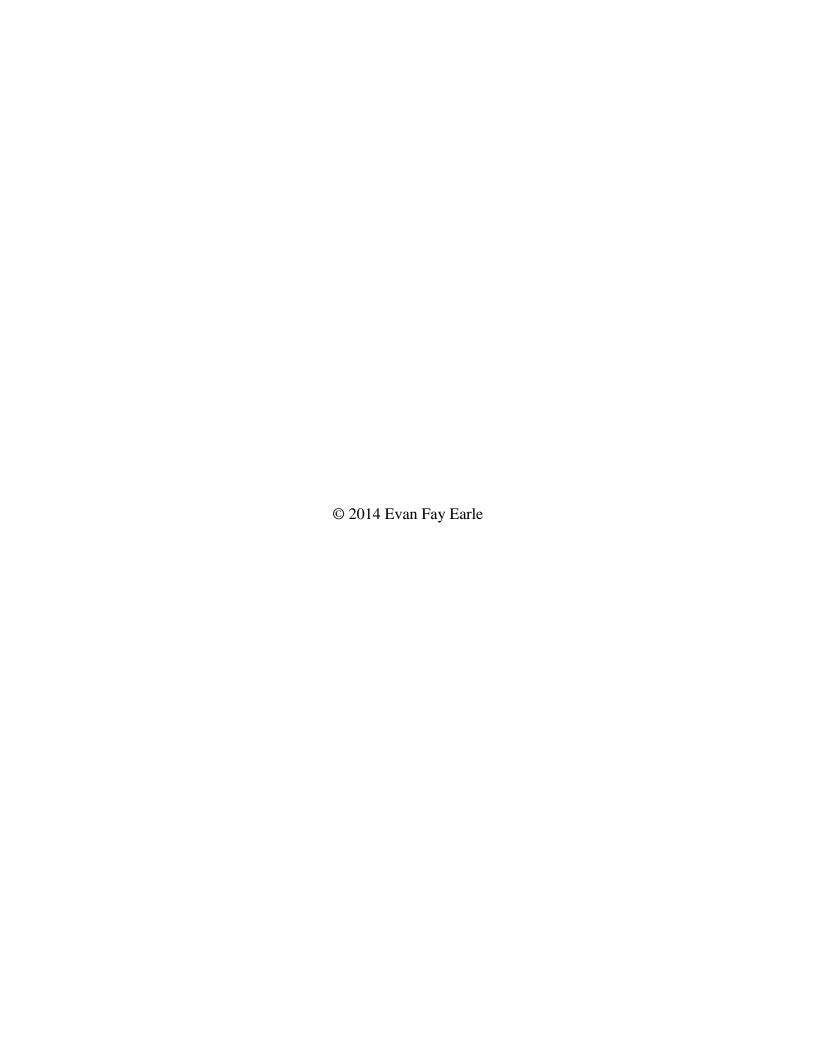# CROWDSOURCING METADATA FOR LIBRARY AND MUSEUM COLLECTIONS USING A TAXONOMY OF FLICKR USER BEHAVIOR

A Thesis

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Master of Science

by

Evan Fay Earle

January 2014

ABSTRACT

Library and museum staff members are faced with having to create descriptions for large numbers of items found within collections. Immense collections and a shortage of staff time prevent the description of collections using metadata at the item level. Large collections of photographs may contain great scholarly and research value, but this information may only be found if items are described in detail. Without detailed descriptions, the items are much harder to find using standard web search techniques, which have become the norm for searching library and museum collection catalogs. To assist with metadata creation, institutions can attempt to reach out to the public and crowdsource descriptions. An example of crowdsourced description generation is the website, Flickr, where the entire user community can comment and add metadata information in the forms of tags to other users' images.

This paper discusses some of the problems with metadata creation and provides insight on ways in which crowdsourcing can benefit institutions. Through an analysis of tags and comments found on Flickr, behaviors are categorized to show a taxonomy of users. This information is used in conjunction with survey data in an effort to show if certain types of users have characteristics that are most beneficial to enhancing metadata in existing library and museum collections.

BIOGRAPHICAL SKETCH

Evan Fay Earle is a 2002 graduate of Cornell University with a Bachelors of Science. His studies spanned a broad array of subjects with concentrations in education, business, communication, and information science. As an entrepreneur and avid collector of rare and unique books, he gained a greater understanding of the importance of metadata in managing large collections of books and disseminating information about them. His interest in the subject expanded through his experience managing data at a startup dot-com during the internet bubble of the late 1990s. While the technical and theoretical aspects of metadata management were intriguing, it was combining these with his passion for education and learning involving libraries and museums which lead to a fruitful connection.

While working with the Cornell University Laboratory of Ornithology, Evan completed a database of the over 1,000 pieces of artwork owned by the lab. The database, now available online with images of the artwork, contained over fifty different categories of metadata describing the artwork, its subjects, and its illustrators. Work with the Lab of Ornithology led to an appointment with Cornell's Division of Rare and Manuscript Collections at Kroch Library as the Archival Technical Services Coordinator where, for nearly a decade, he has managed and participated in a wide assortment of digital projects, involving massive amounts of metadata from collections held by this world class library. His work has helped provide high quality, rich descriptions of some of the University's most treasured collections for scholars all over the world. Research for Kroch Library has been in concert with his graduate studies. Several collections of digital materials have implemented crowdsourcing strategies that were developed over the course of his graduate work.

Evan also consults on digital projects as the archivist for the International Motor Racing Research Center in Watkins Glen, New York, where he is applying information gleaned from his graduate studies.

AKNOWLEDGEMENTS

Completion of this degree was made through Cornell University's Employee Degree Program, which, while not the easiest way to earn a degree, provided tuition assistance and allowed for studies while maintaining my work obligations as well as my paycheck. I am grateful to my colleagues in the Cornell University Library system for providing me the opportunity and flexibility to work towards the degree. Similarly, I am extremely thankful for the patience and guidance I received from the faculty within my graduate program for providing me with such a rich learning experience. Lastly this would not have been possible without the generous support of my family.

TABLE OF CONTENTS

INTRODUCTION

Libraries, museums, and other academic repositories that hold historic material are immense resources for research. Important collections around the world contain information to help us learn more about our past. On April 3, 2013, Professor Mary Flanagan, Sherman Fairchild Distinguished Professorship in Digital Humanities at Dartmouth College, stated in an interview with Trevor Owens: "There's no shortage of archival material across the world, as you know. In universities, archives, libraries and museum collections, millions of photographs, audio recordings and films lie waiting to be digitized. The British Library has warned that by 2020 vast quantities of legacy content will be undigitized and is in danger of being forgotten. But digitization is only part of the problem. Once digitized, someone has to tag the images properly. This takes significant staff time to input. There are many collections that are very well documented and just need to be brought into the digital age. There are, however, millions of artifacts in collections which have little or no informative descriptions aside from what may be written on the archival box or photo itself."

Library and museum staff members have tried to find ways to make more of their collections available to the public by digitizing them and placing them online. As technology has advanced and the costs associated with digitization have decreased, many institutions around the world are placing more of their holdings online for patrons to access. While funding staff to digitize materials will always be an issue, the time consuming and thus costly problem that curators and archivists face is creating metadata for digital images. The technical aspect of producing a high quality digital reproduction through scanning is no longer the main barrier to digital archiving. Similarly, having access to many terabytes of digital storage space to retain the archival digital images and serve them online has also become less cost prohibitive. In May 2013, Flickr began providing a terabyte of storage free for all its users, enough space to store 500,000 good quality images. Image scanners have become faster, storage media are more reliable and have higher capacity, and broadband allows high speed transfer of images over networks.

The creation of metadata for images remains the primary obstacle in the process of creating good online image databases.

Library and museum staff members are tasked with creating descriptions for large numbers of the relativity easily digitized items within their collections. Immense collections and a shortage of staff time prevents the description of collections using metadata at the item level. Large collections of photographs may contain important scholarly and research value, but this information may only be found if items are described in detail. Without detailed descriptions, the items are much harder to find using standard web search techniques, which have become the norm for searching library and museum collection catalogs. To assist with metadata creation, institutions can attempt to reach out to the public and crowdsource descriptions.

An example of where crowdsourced description generation takes place is Flickr, a website where the entire user community can comment and add metadata information in the form of tags to other users' images. Flickr has become one of the most widely used and largest image hosting databases on the web. In August 2011, the six billionth image was uploaded to their servers (Kremerskothen, 2011). The size and popularity of the website provides libraries and museums with an audience of potential workers. Despite the free-form nature of collaborative tagging, libraries and other academic institutions have begun looking into how collaborative tagging can help them in making more metadata from their collections available to typical users instead of just experts within the community. Highly respected organizations, like the Library of Congress, have shown that collaborative tagging can be successful in meeting these goals (Library of Congress. Prints and Photographs Division, 2008).

With the introduction of the Flickr Commons, several traditional institutions, like the Library of Congress, began exploring the impact of allowing the public to annotate digital image collections. The Commons pilot project was initially a collaborative effort between Flickr and the Library of Congress. Matt Raymond, the Director of Communications at the Library of Congress, described the project as an

effort to "help address at least two major challenges: how to ensure better and better access to our collections, and how to ensure that we have the best possible information about those collections for the benefit of researchers and posterity" (Raymond, 2008). The Flickr Commons contains large digital collections uploaded by member institutions, primarily libraries and museums. The goal the Flickr Commons is to share with the public the substantial photo archives that are stored at cultural heritage institutions around the world. Participants in the Flickr Commons may only upload content that has no known copyright restrictions and must actively curate and add to their collections. Currently the Library of Congress is working on the tagging of photographic collections of photos by Boyd Norton, Erik Calonius, Marc St. Gil, Frank Lodge and Charles Steinhacker.

In December 2013 the British Library uploaded over a million images from their collections into the Flickr Commons with the intent to find "new, inventive ways to navigate, find and display these 'unseen illustrations'" (O'Steen, 2013) Many of the images uploaded had minimal metadata and in some cases they admitted they may "know nothing about a given image." The British Library is launching a crowdsourcing application to help them learn what is contained within these images. This shows a trend of institutions deciding to take their minimally described digital collections to the public for help, rather than taking the traditional approach of waiting until image descriptions are detailed and complete before sharing collections online.

Information contributed to images on Flickr by users can readily be seen as adding useful description. Flickr users can be observed working together in sleuthing out details of an image that library and museum staff members may not have time to research. An example of this can be seen in the following image.

Flickr image title: **Even More Boons and Blessings**

Flickr ID #7340756974 from the National Library of Ireland

In this image, the comment section is filled with users working to identify the time, date and location information for this photo, which wasn't initially provided by the contributing institution. Comments include users with various skills approaching the image from different angles. One user links to a Google Streetview image of the location as it appears today. Another user suggests that the model of baby carriage means the photo comes after a certain date. A further user narrows things down by noting the clocktower and the length of the shadows to put the image in the summer. Finally, one user notices the newspaper headlines in the lower right mention an airplane disaster, researches airplane disasters, and adds the information that is discovered. Because of user contributions the photo is dated to the day and many additional tags are added describing the image. A tag is even added for the phrase "date

established," identifying this photo as successfully being dated. This is just one example of how users behave on Flickr and how their content may be useful for library and museum collections.

This thesis discusses some of the problems facing libraries and museums, particularly regarding metadata creation, and it provides insight on ways in which a better understanding of the crowdsourcing behaviors observed on Flickr can be beneficial to academic institutions. The main research question being addressed is how are certain behaviors more desirable for institutional repositories to use in an effort to harness the wisdom of Flickr users?

Through interviews, survey data, and experiments conducted with library and museum staff members, information was gathered regarding difficulties and preferences when working with collections. The surveys and interviews also provided an understanding of how those involved with academic institution repositories view aspects of crowdsourced metadata and digital image collections. Following this research, tagging and comment contributions were mined from collections within the Flickr Commons. The interviews and survey data provided guidance for analyzing the Flickr data. Using the analysis of tags and comments found on Flickr, insight into user behavior was categorized to show a taxonomy of users. This information is used in collaboration with other data in an effort to show if certain types of users have characteristic patterns. These patterns are then discussed to show how some user behaviors are more beneficial to enhancing metadata in existing library and museum collections than others. Certain patterns of behaviors grouped into a taxonomy can allow libraries and museums to better understand the motivations of taggers, tagger styles, etc. and seek out users who become the most useful in contributing metadata to their institutions' digital collections.

Machine created digital metadata versus hand-coded metadata and the backlog problem

We live in an era of born digital publications and images and a time when collaboration between individuals online has never been easier. Vast amounts of data are created with ease every day by professionals, scholars, and the average computer user. With born digital content, it is easy to index, search, or otherwise provide access to the information contained within the digital object. Software can automatically index large documents with a single mouse click. A document can be published across a wide range of web database and social media platforms where search engines crawl content with ease. Outside of the initial human effort to create the original object, human interaction needed to spread the information and make it available in understandable, searchable parts is minimal. While this notion pertains most clearly to textual information, even with digital photographs, metadata creation and dissemination can take place with limited human interaction. Smart cameras tag time and date information, as well as device specific information, such as make and model number and its settings. GPS enabled cameras provide specific location-based data, that when combined with software, can in many cases identify the image and properly add metadata automatically. This work flow obviously has not always been available and cannot work with such ease for older printed objects. A handwritten letter, even when scanned, is just an image with meaningless squiggles without human intervention to transcribe it. A printed photograph when digitized cannot have its full description shared unless it is identified so that it can become searchable. For printed material that has been scanned, the adding of metadata to items is "laborious," but without "rigorous standards of descriptions" collections, digital or not, are of little use to researchers (Prochaska, 2003). The smaller amount of metadata associated with an image, the less useful the image is to users and the less likely it is to be found during searches.

As quickly as we are creating new digital content to spread across the world online, we are losing historical items that if not preserved, such as by digitization, are lost to time forever. Great collections of

historic objects exist tucked away in the various attics of institutions in the world. Our lives are based on our history. Whether we want to acknowledge it or not, our pasts, both individually and collectively, are what make us who we are. History is what helps us understand current issues better by providing context. History is what helps us avoid repeating mistakes. It helps us learn and creates intrigue and excitement. The great Roman philosopher, Marcus Tullius Cicero, understood the importance of history over 2000 years ago when he stated, "Not to know what has been transacted in former times is to be always a child. If no use is made of the labors of past ages, the world must remain always in the infancy of knowledge." These notions are part of what makes it important to help share the knowledge that is contained in historic objects.

Libraries and museums are often thought of as a premier location for the dissemination of information. These repositories hold many historic items. Because of the large quantity of material, it has been impossible for all these items to be properly cataloged and thus discovered by the public. A library may hold vast amounts of information within its collections, but this information is useless if it is unknown to those who can use it. While a repository by definition (Merriam Webster, 2004) is a place that stores material, this is only the most rudimentary function of a library. For success, a library must provide for discovery and also for access. Throughout the world, there are millions of items with stories waiting to be told and mysteries waiting to be solved. Occasionally a treasure is uncovered in a library, and it makes the national news. A famous example of this are the Madrid Codices of Leonardo Da Vinci, found in the National Library of Madrid after being lost for 300 years (Reti, 1968) and more recently, September of 2008, the discovery of previously unknown work of Wolfgang Amadeus Mozart in a French library (McElroy, 2008).

It is often asked how such important works can go unknown within a library. There are several reasons. One is demonstrated by the case of the Da Vinci work, where the item may have been known at one point decades or centuries earlier and its cataloged record lost and the items' location becomes

unknown. In other cases, an item may be identified incorrectly by a staff member who may not be an expert on the item that they are cataloging. A most likely cause for unearthing undiscovered items is that they went into a repository and were never cataloged.

Libraries are constantly battling with backlogs of material and do not have the resources, both staff and funding, to properly catalog. Having a large back log of material is not necessarily a deterrent to making new acquisitions, so sometimes these backlogs continue to grow instead of shrink. An influential paper, authored by Greene and Meissner in 2005, provided survey results illustrating the backlog problem. They showed that 78% of archival institutions in one survey were bringing in more material than they were processing. This same survey indicated that only 44% of the surveyed repositories allowed access for researchers to their unprocessed collections. In 1998 a survey conducted by the Association of Research Libraries asked participants about the percentage of their collections they felt were uncataloged. The survey revealed that one third of the respondents believed that more than 25% of their manuscript collections were not properly cataloged; 13% of respondents stated that greater than half of their collection holdings were unprocessed (Panitch, 2001). When this question was asked again in a survey conducted in 2003-2004, the percentage of respondents with more than half of their holdings unprocessed had risen to 34% (Green, 2005) When thinking about the amount of material that is cataloged in libraries around the world, even if the amount of uncataloged material was only a quarter of that which has been done properly, there is an enormous amount of material. While more historic material isn't being made, it is continually discovered and purchased by the repositories. This type of situation leads to even larger backlogs of material as staff members attempt to provide ample cataloging information to the collections. The amount of material that is contained in the holdings of large libraries is often overwhelming for the staff and therefore much of the material may remain hidden in so called "dark archives" (Jones, 2004).

Content disseminators and user generated information

When placing an image online, in a large image database, it becomes most useful only by having accurate metadata so that it can be found when a user of the system searches the collection. If an image has no, minimal, or wrong, metadata, digitizing it and placing it online within a collection is fairly useless. Unless a patron is planning on browsing vast amounts of images, which would take a long time, they are not likely to be able to easily find the specific type of image that they need.

With the emergence of Flickr, YouTube, and Wikipedia as giant content and information repositories and disseminating systems, content owners have the ability to reach huge audiences. The websites present an opportunity for feedback from users, and when embraced by brick and mortar libraries and museums, they provide an opportunity for increased education and efficiency. Allowing users to help annotate and contribute to the metadata of collections can greatly assist libraries and museums by enriching their collections, creating more exposure for items and decreasing staff time needed to describe the millions upon millions of images that exist. One of the most public early adapters of experimenting with a way to increase exposure of their large image collection holdings on Flickr was the Library of Congress. Their initial use of Flickr was studied carefully to see how the public would behave. The Library of Congress project to use Flickr started January of 2008 (Flickr Library of Congress, 2013). Archivists and library staff members have followed the project with great enthusiasm and waited with anticipation for the release of their report about the project. The final report released, December 11, 2008, showed that in the nearly year long period the initial pilot was run, the images uploaded, 4,615 of them, had 67,176 tags added (Library of Congress Prints and Photographs Division, 2008). This is an astounding amount of additional information that was made available to the library with minimal additional work on their part. When the images were initially uploaded, they were only tagged with "Library of Congress." Opening up an image collection to allow the public to tag the images

with additional content can greatly increase the richness of a collection along all these areas without a great deal of staff work. These collaborative tagging environments can be referred to as folksonomies, as described by Thomas Vander Wal (2007). More tags are generally seen as better than a few tags. However, in order for images to be found in an online system, it is necessary to have tags be descriptive at a level beyond an initial fairly useless basic description.

It is not necessarily as simple as immediately incorporating these systems into existing libraries; greater speed of metadata creation is not the only difficultly regarding image description. As with anything on the Internet, there are skeptics and problems with inaccuracies. There are different competencies of contributors as well, and better understanding behavior can perhaps enable courting of the particular users who provide the most useful content.

Importance of needing to find efficient, accurate metadata creation

There is a dialectical tension between the traditional stoic view of libraries and museums and the user generated content world of sites like Flickr. It is because of this view that libraries and museums are cautious to incorporate user generated content into their own collections. For the most accurate description of any given image, an expert or experts on the subject are needed to describe what the image contains. For example, if a photo of a statue was loaded into a digital repository, a non-expert may catalog descriptive information based solely on what is seen, words like statue standing or maybe something about the material the statue is made of would be all that was entered. Alternately an expert or someone familiar with the statue could look at the same image and perhaps recognize it, and by doing so, add information that goes beyond what is simply observed in the photo. Interpreted contextual information such as date, location, why the item is important, or what other events or activity surround the image could be added. The difference in the richness of the image descriptors, or tags, used in these

two examples is clear. One will provide searchers of an image collection a much better chance of finding an image related to what they are searching. It also will provide browsers of an image collection knowledge and information they may not have known otherwise.

On the other hand, perhaps the word, statue, is the most relevant term in describing the image where as additional tags, such as name and place, are not as relevant. Unless there was a special reason to hire an expert for a specific image collection, it is more than likely that the individual adding initial image metadata will not be an expert on the subject of the image. This is particularly the case, since the task of image metadata entry is often assigned in libraries to student workers to allow curatorial staff to work on other matters. If a staff member is asked to spend some time researching an image so that more than just "statue" is added as tag, it may be hard for them to know where to start research when beginning. Should their time be spent trying to research the date the image was taken, the name of the person in the image, or place the image originated? If they focus on just one of these aspects, one will get some richness in the metadata, but not sacrifice spending an inordinate amount of staff time adding the metadata. The metadata for images can include descriptive information, such as what can easily be drawn from the picture. It also can include description that may not be readily identifiable from the image, such as the location and date the image was taken. Details about format, such as slide, or negative, or even camera settings, may also be included.

If items are described with an overabundance of keywords or description, the description becomes less useful. Searches will return too many results and non-relevant results if every item within an image is tagged. Curators of digital collections would prefer to have smaller amounts of relevant metadata generation rather than large amounts of non-relevant metadata.

Expectations of libraries from staff and researchers

In visiting a library or museum, a researcher using a card catalog or reading a description about a collection expects what is read to be accurate and pertinent. As mentioned earlier, the lack of information about collections is common. Finding descriptions of collections that are blatantly inaccurate is far less of a problem. Researchers citing information from a library normally don't believe they need to question its truthfulness. Workers in libraries make an effort to assure accuracy by using reputable bibliographic sources. Because of time constraints, often the amount of factual information that can be gathered to form a quality record by a staffer is limited. It is critical though that in branding a website or information with the identity of an academic institution, that the information is accurate. If it is not, it can reflect badly on the organization, but even more likely, it will perpetuate the inaccuracy as researchers site the information as coming from a reliable source. Any library or museum wanting to implement public tagging of their images has to be aware of the accuracy issues and naturally wants to ensure that inaccurate information added to a collection is minimized and not interpreted as being endorsed by the institution.

LITERATURE REVIEW

To better be able to understand how to analyze tagging behavior and create a taxonomy, it is possible to look at some of the literature related to tagging. There is also an importance in looking for information that could be useful for libraries and museums in understanding what they should look for in a tagging environment. Tagging happens in many places across the web. While there are differences in how one may use tags, if they are tagging a blog post versus an image, there are quite a few similarities as well. It has been observed that the tagging behaviors, which take place in an image or photo context, appear to differ from other social media tagging environments (Marlow, Naaman, Boyd, & Davis, 2006). The most basic summation of a collaborative tagging environment in earlier studies was defined by agreement that user generated tagging systems succeed compared to other metadata systems because of their ease of use and limited amounts of time an effort needed to add tags. These tagging systems often produce too many false positives in searching, and thus if not moderated, are not a good tool for discovery and precision searching (Riddle, 2005). Further study has elaborated and clarified these views.

Libraries and museums have become familiar with the use of tagging systems and have published observations and recommendations based on their own findings independent of and in collaboration with tagging researchers. The Library of Congress Flickr report, *For the Common Good: The Library of Congress Flickr Pilot Project*, provides an excellent early analysis of the use of tags in an academic photo collection environment. While some recommendations were made in the report, it was mostly descriptive as to what had happened. More recently reports, including one prepared by the Online Computer Library Center OCLC, one of the standard bearers for library research, titled Social *Metadata for Libraries, Archives, and Museums*, offer fairly extensive survey results and recommendations. The report showed how tagging was the most used on library-related websites that allowed some form of social crowdsourcing, surpassing the use of comments alone. Recommendations

from OCLC included further library specific research into the use of social metadata sites (Smith-Yoshimura & Holley, 2012).

Quantity of Tags

A consideration that influences tagging quantity of the user includes tenure and the effects that leaders have on other newer users. For instance, studies have found that Flickr users who have a longer tenure using the site, average fewer tag postings per year than newer Flickr members. However, these longer term members have a greater enjoyment of tagging images. For new members, their enjoyment initially increases with the more they tag (Nov, Naaman, & Ye, 2010). The implication of this affects how one would want to recruit and promote projects for tagging. Rather than solely relying on an established community of taggers for a project, new projects should go out of their way in terms of advertising and marketing to bring in new users to Flickr so as to take advantage of their enthusiasm. One would not want to alienate the more established long time members as well since, though they eventually slow down in the amount of tagging they do, they do average a higher enjoyment in tagging and that enthusiastic participation can be beneficial. For libraries and museums to continue to receive input from a tagging community, the community must stay engaged and vibrant. To motivate users along these lines, keeping an institutional photostream updated with new collections would be one key. If a photostream became stagnant, without new content added regularly, some users may lose interest.

Nov et al. also showed that these longer term members can have a contagious effect by increasing commitment with newer users. This is accomplished by users taking advantage of the social aspects of Flickr, joining groups and sending messages. These are the very same social measures that, as indicated earlier, increase tagging. The longer term members, who feel more commitment to projects,

increase the feelings of commitment among younger users by interacting with them and bringing them into groups.

The notions of tenure on a site and experts or leaders guiding unfamiliar users, who then themselves can turn into experts, has been frequently observed on Wikipedia and studied on other websites. The reader to leader framework explained by Preece and Shneiderman is certainly applicable to Flickr users. The framework illustrates a narrowing in numbers as the user body self-selects through reader, contributor, collaborator, and leader stages. In Flickr these stages can respectively be aligned with a browser of the images, a user who tags on occasionally, a user who joins and works within a Flickr group community, and a user who helps dictate tasks, offers suggestions to novices, or otherwise leads by example. The patterns that align to these stages fall into a broader taxonomy of Flickr users. The reader to leader framework allows insight into the different types of Flickr users and what motivates or encourages them to tag images.

The quantity of tags and level of completeness on an item also affects the tagging behavior of subsequent taggers. The more tags that are present, the more likely tags will be added. There is also evidence that at a certain point, the number of tags results in more of the subsequent taggers correcting information rather than adding new information (Moltedo, Astudillo, & Mendoza, 2012). Information from a 2008 study concluded that more than 50% of the images on Flickr have only three or fewer tags (Sigurbjörnsson & Van Zwol, 2008). These ideas provide reason for excitement and concern with a crowdsourced metadata project. It shows the importance of having an active user community to generate enough content to meet the richness in description that is desired. An institution may have a fear of devoting resources to establishing a crowdsourced project only to have no one use it. This notion aligns with Hardin's tragedy of the commons (1968). This is the idea that in a community, if individuals feel their own needs are greater than those of the community, the whole system fails. An example of tragedy of the digital commons could be a crowdsourcing project where many users view and use content for

their work, but no one contributes metadata, creating a stagnant site. Interestingly some research has shown that this effect can be greater in an academic environment, where contributions may be seen as more for the good of the private institution and not for the common good of the public (Huberman, Romero, & Wu, 2009).

Motivations of Taggers

One of the most important areas to examine, which helps categorize users, is having an understanding of tagging motivations. Motivations can be useful when establishing which types of taggers are most beneficial in a collaborative tagging environment. If a certain type of motivation results in the output of a user being most useful, that is the type of user that should be observed and courted for their metadata contributions.

Previous studies looking at motivation by the likes of Ames (2007), Naaman (2010), Nov et al. (2010), and Trant and Wyman (2006) have all indicated that driving factors for one to start and continue to tag images online have strong roots in personal motivations. Some personal motivations are easily identified, such as a user having a desire for better organization of their own images so that they may find them again easily or so they can browse them in a manner they wish. These organizational or searchabilty motivations can be quite separate from those of communication, in which a user is tagging to add context to an image to help tell a story to others rather than to help another user or themselves find the image. A library or museum looking for user generated metadata might want to harness the work of a user who does mainly tagging for searchability, or in another circumstance, perhaps a user that annotates images to help tell a story is more useful. In collaboration they can be very useful. For example, having a user who tags images with appropriate tags to add descriptive metadata that makes an image more discoverable is useful. If that same user also adds to the comment section more narrative

information, then they are providing further explanatory text about the image. The terminology used in the comment sections, if used as tags, quite possibly would not be as useful if placed there alone.

Motivational factors may be observed by following a user behavior, but personal motivations can be quite different if a user is tagging their own photos versus tagging someone else's. Naaman's continued work on tagging motivations, including experiments in collaboration with Oded Nov, have added further insight into motivational factors on Flickr in terms of the private or public good and its relationship with social presence (Naaman, 2010). They were able to show that user's levels of motivations, both for self and for public, increase with the number of unique tags a user adds to images in their collections. They also showed that the greater the number of contacts or groups a user was involved with on Flickr, the greater the number of tags they applied to images. What this indicates is that there is a cyclic social aspect to tagging that is critical for a successful tagging environment. In examining different Flickr user types, a user who is shown to have greater connections or a wider array of tagging interests, may be more productive in quality metadata generation. If users are not getting many views on their images, they may not be getting those views because they don't have many tags. However, because they don't get views, they aren't motivated to tag more.

Types of Tags

Studies have given examples of the differences between the metadata a museum or library may use to describe an item and the terms that the public users would use if they were describing an item. The terms that the public would use to tag or describe an item are similar to the terms that they would use if they were trying to search out similar images. Terms a museum might use to describe a photograph, like "black and white," "gelatin print," or "art nouveau," don't relate well at all to the types of descriptive information that users are more likely to use for searching (Trant, 2006).

17

Research on grouping tags into types based on categories of word meanings have been able to show that different types of tags are valued more than others to taggers (Stvilia & Jörgensen, 2007; Sigurbjörnsson & Van Zwol, 2008; Beaudoin, 2007). A tagger who is productive may rely more heavily on the categories of tags that they feel is important. Their ideas of what is important may not align with those of a library or museum trying to enrich their metadata. Categories of tags such as adjective, person, place, event, emotion, humor have vastly different usefulness to different types of users. Studies have shown that tags relating to person or location are thought by users to be most important, and the high prevalence on Flickr and similar sites of location-based tags confirms this (Sigurbjörnsson & Van Zwol, 2008).

This can be seen as a positive or negative depending on the type of image metadata a library or museum may need. For instance, person and place descriptive information may be considered tags that library staff members can easily add when uploading an image, and they may be seeking to add other, deeper, description. In that case, a tagger whose behavior is seen as consistently adding tags beyond simple person and place information would be preferred by an institution to get involved adding metadata. If, on the other hand, very little is known about an image or institution, staff just don't have time to add any metadata, they can be fairly well assured that taggers will attempt to add person or place information. Similarly institutions that struggle with the idea of needing a more controlled vocabulary may wish to find taggers who tag specific objects or persons. Taggers that behave this way are likely to use terms that have consensus of other users as opposed to users who tag with more abstract ideas and emotions (Eleta & Golbeck, 2012).

Other quantitative research into the usefulness of comments and tags for libraries on Flickr suggest that tag and comment types that are useful to users would also be useful to institutions. By collecting usefulness judgments on commentary acquired from the Flickr Commons, there can be observed definitive categories of what is judged as useful and what is not (Momeni, Tao, Haslhofer, &

Houben, 2013). More useful data contained contributions of factual information about the image rather than factual, preferential or thoughtful information about the user making the comment. The same study by Momeni et al. also noted that comments that provided links to references were deemed more useful. As the authors of this study state, developing an automated system for judging and determining the most useful comments would assist academic institutions in knowing which type of metadata they should be aware of for inclusion in their own catalogs or for future digital collections. In the meantime, looking for users who tag and comment usefully as defined by Momeni et al. would in turn be useful for libraries and museums.

Another interesting note, combining tag types and motivations, surrounds the seemingly altruistic nature of tagging and often have underlying factors tied to personal satisfaction, particularly in an environment like Flickr. For example, in Flickr, an uploader is notified whenever their photos are favorited, commented on, or otherwise recognized, by a user. Because of the satisfaction that uploaders felt when this happened, the uploaders would "game" the system by applying tags to their images meant to attract more views or otherwise promote their images (Ames & Naaman, 2007). This type of gaming increases the number of tags in a system, but also can hurt accuracy. These types of gamed tags, similar to what might be referred to as keyword spamming, might not be seen in an institutional setting since users would not be tagging their own photos. If the behavior was seen in users who are used to tagging that way, an attempt could be made to dissuade attracting that type of user. There is a fine line between keyword spamming an image with tags and appropriate search engine optimization for discovery. Extremely popular tags of this nature have particularly limited usefulness in an academic setting. Indeed two of the most popular tags across all of Flickr are "instagramapp" and "iphoneography." These two tags in many cases are auto-generated when uploaded. It is clear that tags like this provide little content that would be valuable regarding the description of the object within the image.

Tag Accuracy

The idea of accuracy of a description in how it relates to an image may seem to be a bit ambiguous. It is possible to have two individuals rate the degree of accuracy of an image's description completely differently because of their separate past experiences. Research has shown that when comparing tags used by the general public against tags used by experts, generally there is a great deal of consistency (Bischoff, Firan, Nejdl, & Paiu, 2008; Library of Congress Prints and Photographs Division, 2008).

Guidance of user tagging behavior regarding accuracy has been shown to be possible by seeding images with certain types of tags during the initial upload process. If staff time permits, a certain number of seeder tags could be added to images that would then direct user behavior. Unfortunately for library and museum staff members, a study which illustrated that starter tags influence quantity and types of tags, showed that one way to increase output of suitable tags was to have seeder tags be unsuitable to the image (Maier & Thalmann, 2008). This approach clearly would not be appropriate in an academic setting. Academic staff are not likely to seek out taggers who have a behavior of tagging incorrectly in hopes that it would provoke a response from other taggers to correct the behavior. Research suggests this would work, but unsuitable information would then remain associated with the image. Despite concerns from library staff members regarding public tagging, accuracy studies have shown results contrary to the concerns.

Another example of how user generated metadata is thought of as predominately accurate, is that libraries are now integrating it into their own library catalogs. The Powerhouse Museum in Australia has used Flickr for some of their images (2013). They then pull that data that users place in tags into their own library catalog. They believe this shows the information is accurate. They do place a caveat on the field in their catalog where the tags are placed saying that it is generated from Flickr and not from

library staff members. The Library of Congress is also considering this. This takes the idea of user generated metadata a step further as you are now having it ingested directly to a library's catalog that has predominantly been a private sanctum, the realm of professional catalogers and trained librarians. This idea would be unsuccessful if inaccuracies in the tags were overwhelming.

Despite what studies have shown about the accuracy of user generated metadata, there are ways to make it even better. Previous Flickr research has shown some areas of concern. A 2006 study suggested that tags appearing in Flickr images had only a 50% chance of being relevant (Kennedy, Chang, & Kozintsev, 2006). Research into interface design and tagging systems that incorporate assistive tagging elements were surveyed by Wang, Ni, Hu, and Chua in 2012. This work often showed that by offering users tag recommendations or by providing tagging refinement, a richer and more accurate description of an image can be realized than without these systems. There are difficulties with machine based assistive tagging, as algorithms that make recommendations and supplementation may not be applicable to all types of collections. The specifics of the type of metadata that an institutional repository wants to have gathered for a certain collection has to be considered.

Consistency and Controlled Vocabulary Folksonomies

Certain types of tags and the users that use them can be more beneficial to a library and museum looking for metadata. Looking at work done in the area of tagging types can allow for seeking out the types of users that tag in categories that are most useful.

When cataloging items, libraries and museums often strive to use controlled vocabularies. These set and limited phrases are useful because they provide consistency. Consistency in metadata description allows for easier searching of items by ensuring all similar items can be found when searching that one single descriptive term. There are large curated authority files of approved terms from the Library of

Congress Subject Headings list (Library of Congress, 2013) to the Art and Architecture Thesaurus (Getty, 2013). These databases are used religiously and even fanatically by many in academic cataloging positions. Along these lines folksonomies are broken into two main categories, broad and narrow (Vander Wal, 2005). The narrow folksonomy consists of the generator of the content to be tagged using a small number of tags, often from a set vocabulary to describe an item. The broad folksonomy takes advantage of users adding many tags to the item from their own vocabularies, which would rarely be controlled. A narrow folksonomy takes too much staff time to expand to include rich description, while a broad system will be inconsistent image-to-image, and perhaps it will be too complex. Finding the balance in user behavior will yield the most useful results. The pros and cons of folksonomies in a library setting continue to be discussed, and while there are recurring themes in the rebuttals by librarians regarding lack of control and usability, the views are softening (Gerolimos, 2013).

In a Flickr environment, the users are not likely to limit themselves to a controlled vocabulary in either tagging or searching. Indeed studies have shown that students working in groups to tag images could often agree on what aspects of the image were worthy of tagging, but without discussion would not generate the same words (Achananuparp, McCain, & Allen, 2007). Another study cited by metadata librarians as reason for concern in using folksonomies showed that a full 40% of a sampling of tags from Flickr were not recognized as properly spelled or formed words when compared to the dictionaries of the Aspell open source spell checker (Guy & Tonkin, 2006). A problem here is that multi word phrases, slang, or concatenated phrases as seen in things like Twitter hashtags, are of course not going to get seen as correct in a spell checker, but very well may be excellent terms for tagging and thus discovering items. These studies provided fuel to arguments against public generation of metadata for intuitions with strict controls. They didn't fully take into consideration how people search, which is similarly to how they tag. Later studies have shown that some fears about the ambiguity and lack of controlled vocabulary are less. An examination of several popular sites using tags showed that the vast majority of

the user generated tags matched up the National Information Standards Organization guidelines (Spiteri, 2013). Spiteri's work did confirm that polysemy, when a word has two meanings, and synonymy, different words that mean the same thing, are existent, but perhaps not necessarily a problem. The concern should not be with strict controlled vocabularies, issues of polysemy and lack of automatically generated cross references for things like abbreviations are a better focus for attention. Polysemy is a more difficult problem to solve, but looking out for a tagger who consistently adds both an abbreviation and its full spelled out variant is useful for ensuring that images are found. Work into ambiguity of tags has shown that much ambiguity of this nature can usually be resolved or explained with the inclusion of a single additional tag (Weinberger, Slaney, & Van Zwol, 2008).

When thinking about controlled vocabularies, languages may also be considered. A certain set of images may lead an institution to seek out a tagger whose behavior exhibits the use of foreign language. A foreign language speaker may be better at properly adding location data or using terminology that contains specific alphabetic characters. An initial reaction may be that additional language can cause concerns for those who desire controlled vocabulary, the introduction of another language would seemingly introduce more terms needing control. Museum specific research has shown that the addition of a new language into a tagging system does not dramatically increase the amount of consensus users have with tags (Eleta & Golbeck, 2012).

Evidence points to having libraries and museums abandon at least concerns about the lack of controlled vocabularies. To understand what it can mean to have a rigid controlled vocabulary, take the example of the terms movie and cinema. Clay Shirky who has done extensive and enjoyable writing about crowdsourcing and tagging, has analyzed the ideas about having two words to describe different things and how this relates to controlled vocabularies and tagging. He used that example of the two words, movie and cinema. He states:

" 'Oh my god, that means you won't be introducing the movies people to the cinema people! ' To which the obvious answer is 'Good. The movie people don't want to hang out with the cinema people.' Those terms actually encode different things, and the assertion that restricting vocabularies improves signal assumes that that there's no signal in the difference itself, and no value in protecting the user from too many matches." (Shirky, 2005)

He also says that when tagging is restricted to categories defined by vocabularies and authority files, it minimizes the usefulness of the tagging for someone searching. Shirky believes that purely collaborative tagging systems will replace controlled vocabularies.

For librarians striving for at least a somewhat controlled vocabulary in a Flickr environment would seem ideal, but difficult, an institution can at least look for users who are consistent in their behavior to satisfy some needs of control. Users who do not deviate in how they describe an item, in terms of phrase, words and spelling, would be seen as more beneficial when tagging items across a collection. Abandoning completely on efforts to try and enforce a controlled vocabulary is something that would be very difficult at many institutions. Others, who have read Shirky's work, argue that controlled vocabularies are growing in parallel with the increase in collaborative tagging. They believe that with increased collaborative tagging, there becomes more of a need than ever to have some sort of hierarchy and order for systems to be useful in a library context. Without a basic set of rules governing tagging, they believe that searchers will not be able to discover material because there are no standards that will guide them (Macgregor & McCulloch, 2006). Macgregor does later go on to say that collaborative tagging systems can have a place in the library world, but seem to strongly believe that librarians should stop spreading the belief that controlled vocabularies are not user friendly. The average user though, who puts the first thing that comes into their mind into Google, is likely to search for images thinking about a controlled vocabulary.

METHODOLOGY

There were several components for the data gathering for this research. They include research through analysis of interview responses, a study of user tagging perceptions, and an analysis of data mined from the Flickr website.

Interviews

Qualitative, field work, style research was conducted through face to face interviews with two dozen university faculty and academic staff involved in the library and museum community. Insight was also gained through conversions with a smaller number of employees who worked for digital image websites and undergraduate students. These interviews provided direct input about views on digital collections, collection metadata tagging and what is most needed or desired for institutions to provide quality access for patrons using digital collections. Interviews occurred formally and informally. These interviews took place with staff at several schools and institutions in the northeastern United States. Because of the nature of the questions being asked, interviewees needed to be individuals familiar with digital collections issues surrounding metadata creation. Ideals regarding random sampling and generalizability in qualitative study (Mayring, 2007) were considered and best practices followed, though interviewees were actively sought out and not pulled randomly from a pool of potential candidates. The word of mouth, snowball effect, also assisted in attracting candidates. Participants were asked to recruit other individuals whom they believed could be useful in contributing knowledgeable conversation. Insight was also gathered from individuals not directly associated with the library or museum world, but who had familiarity as end users with digital collections websites.

Interview strategies followed guidelines by Becker as well as Glaser and Strauss to help ensure consistency and that the most pertinent information was captured during the process. Some of these strategies including pushing interviewees on responses where it seemed that perhaps they were reciting a corporate line or responding based on what they had heard before rather than what they truly believed. (Becker, 2008) Grounded theory memoing allowed for unique details to be noted amongst the more obvious similarities being compared across interviews (Glaser & Strauss, 2009). Particular effort was made to annotate similarities to previous heard information. This allowed for the formation of category ideas as the process unfolded, but formalized categories were not created until later. The constant comparative method allowed ideas that initially had not seemed relevant to come back to the forefront. Analytic induction analysis of transcripts from the interviews helped in categorizing and establishing recurring patterns beyond the surface level of the response. Determinations were made based beyond the face value of a response and included why it was thought an interviewee responded the way in which they did. This also allowed for delving into instances where perhaps individuals who happened to come to the same conclusion had come to that conclusion in different ways.

Transcripts from the audio recordings of the interviews were made and then coding was completed from the transcripts. Transcripts did use some subjective judgments to speed up the process, by eliminating writing up vocal pauses or conversations that were completely off topic. Data from the transcripts were coded in part using property space analysis as outlined in Becker in an effort to look for patterns across the different groups of interviewed individuals. By making tables with distinctions between general internet user, faculty, library employee, etc. and then adding columns for different characteristics, an attempt was made to seek patterns.

An initial attempt at qualitative comparative analysis truth tables proved difficult to implement. While comparison of responses between each other is helpful, responses to questions required more analysis or had more parts than a truth table could handle effectively. Attempting to compare answers to

a single question across the board between users was not optimal. This was mainly because answers given to different questions related to each other. In other words, the answer one respondent gave for a portion of question two could end up being similar to an answer another respondent gave to a portion of question four. This situation led itself to themes emerging from throughout the data across questions. Information from the interview process was used as grounding for the later analysis of the Flickr data.

Interview Questions

Questions asked during the interview process varied, but followed an interview guide similar to what can be seen in the appendix. Order was changed if the conversation pertained to a question further down the list. Earlier questions were then returned to later. The questions were not always spoken exactly as written in the interview guide, but changes to the language were minimal. The questions were occasionally followed up with explanation or examples for clarity if an interviewee needed this. There was a level of flexibility in the interviews to allow for the casual incorporation of new questions and pertinent thoughts in an effort to pursue points an interviewee had that seemed particularly important. Understandably, these differences could have an impact on generalizability of the data, but it was decided that this effect would be minimal, and that the greater data garnered from pursuit of specific points outweighed the need for extreme consistency in the interview process. Not all the questions asked yielded data that is specifically discussed within this paper, but each question did help provide insight regarding the individuals' views towards metadata and tagging that contributed to broader conclusions made from the portions that are analyzed in more detail here.

Tagging Perceptions and Preferences Study

A tagging study, completed in collaboration with Karin Patzke and Justin Cheng, was used to provide insight into some basic user preferences and motivations when in an image tagging environment. This study included data from surveys and experimental design to understand how users annotate images in different tagging environments and how they view annotations made in these different environments. This experiment looked to differentiate user behavior and perceptions based on a variable online image annotation system designed by the investigators. The tagging annotation system provided environments that allowed for either single word tagging, multiword tagging, or full commentary on images. Metadata provided by participants in the first part of the experiment were then presented to a second set of participants for addition evaluation.

In the first part of the experiment twenty-one undergraduate participants were given access to the online image annotation system. They were presented with the same set of thirty images, which had been selected to represent a wide range of topics and categories from creative commons licensed photos found on Flickr. The order in which these photos were presented was randomized for each participant. An annotation environment, single word, multiword, or comments, was randomly assigned to each image and the user asked to annotate the image following the environmental constraints.

Upon completion of the tagging portion of the experiment, participants were given a survey that defined the different tagging environments, and they were asked to provide rankings of the different annotation methods. Rankings were based on the participants' feelings towards how accurate, creative, interesting, and useful they thought each condition was when annotating the images. These ratings were done using a five point Likert scale. The survey questions can be seen in the appendix.

The tags and comments from the first experiment were then loaded into the online system with the images. A second set of twenty-nine undergraduate students were shown the original set of thirty

images in a randomized fashion. For each image, the system randomly displayed the new participant the annotations created for one of the annotation methods, single word, multiword, or comments from the first experiment. The second participant, called an evaluator, was then asked to rank the annotations using Likert scales based on accuracy, searchability and interest for each image annotation pair. Upon completing all thirty image annotation pairs, the evaluator participants answered a survey similar to part one and were asked to rank the three annotation conditions in terms of accuracy, usefulness, interest, participation and preference.
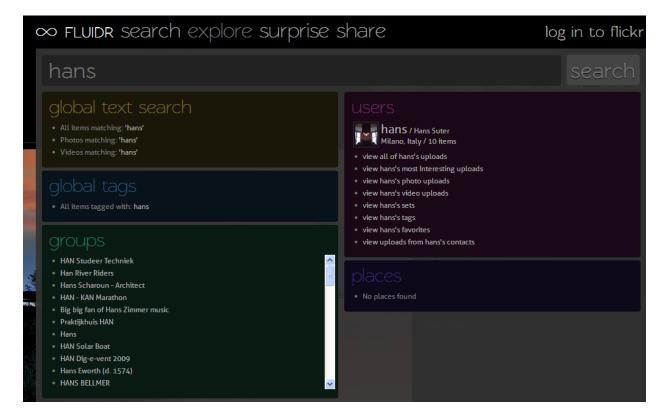
For both portions of the study, statistical analysis of the data using analysis of variance and linguistic inquiry and word count was completed to show the significant areas.
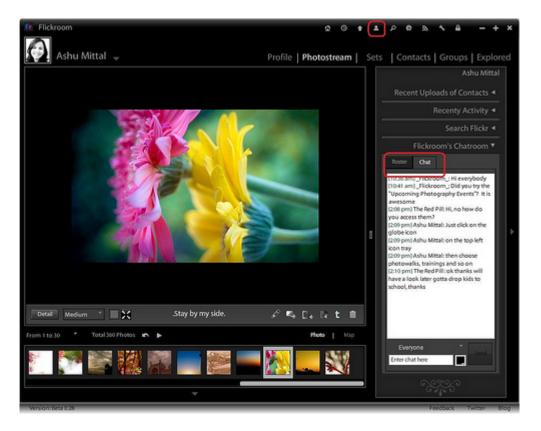
Flickr Tagging Data

With a baseline of knowledge gained from the interviews, regarding perceptions and desires surrounding crowdsourcing and tagging, data gathering from Flickr could begin. This data was captured to begin to look for patterns for the taxonomy of Flickr user behaviors. Information about user behavior and image tags on Flickr was harvested using the Flickr API (2013). Application programming interfaces on websites like Flickr allows users to access the vast quantities of data stored on a website by using a specific set of rules to make calls to the website. Developers can use APIs to design new and creative ways to interact with a website or to re-imagine or display the existing data in a manner differently than what is already available through the main website. Since the APIs provide access to the massive amounts of data stored on the site, they can also be used to scrape data for analysis and interpretation. The Flickr API system, www.flickr.com/services/api , is well developed, and Flickr provides a great deal of documentation and encouragement for developers to take advantage of making useful contributions. Each type of data stored in Flickr can be retrieved by using the appropriate code.

Examples of the type of data that can be called includes statistics like activity on a certain user's photos, such as additions of comments or tags, geographic location of images, or even type of camera used for a photograph. The range of uses of the Flickr API is almost as immense as the data. Some uses can be simple repackaging of search abilities into more user friendly layouts such as Fluidr:

http://www.flickr.com/services/apps/72157618934963200

Others can be complete desktop solutions for more efficient browsing, uploading, interacting and editing

such as Flickroom: http://www.flickr.com/services/apps/72157617247044853

Some uses of the API are purely comical such as Horse Lovers Camera, which allows users to place

horses in the foreground of any of their photos:

http://www.flickr.com/services/apps/72157626690796099

There are restrictions on use of the API, only non-commercial uses of the data are allowed

without a further vetting process. Flickr requires developers to apply for an API key and passphrase to

allow Flickr to track the usage of their data. Because of the nature of the data collection needed for this

work, only the simple developer application process was needed, and it was readily approved. The

tracking of developer usage allows Flickr to prevent excessive strain on the system by developers

making too many calls for certain types of data. For instance, the extensiveness of some the data means

that repeated calls could dramatically slow the system so limits are in place for how often certain types

of calls are allowed to be performed by a developer. Other types of data require that the developer gain

permission from the user whose photo data they are accessing. This process is built-in to the API and simply requires a developer to send notification to the owner of a photostream. When the owner approves, a token is given that allows the developer's key access. This system was used for some of the data gathering used here.

Data gathered from the API was restricted to users interacting with collections under the Flickr Commons area of the website. Data was pulled from a selection of institutions who had significant collections, 1,000+ photos and who have been active for more than a year, to ensure that their images had annotated content. Collections were randomly selected, in some cases using a twelve sided die. The shorter image tags made by users were collected along with the typically longer comments from the discussion section.

Some data was gathered doing manual test calls through a web browser, but the majority of the content was compiled into packages for analysis by creating scripts that generated the calls. Some institutions have staff who already compile the data needed for examination, and in those cases, data could be acquired directly. A script written in Python by Dianne Dietrich was used and modified as the basis for some data scraping; this script can be seen in the appendix. Calls through the Flickr API would target a particular institution in the Flickr Commons and then output data from the collection that included user behavior, such as tags, and comments, the photo that they were commenting on, and the date in which the activity took place. The script was written in a manner that allowed tags generated by the owner of the photo to be either included or omitted from the datasets, thus distinguishing the differences between owner and user generated tagging.

The data was gathered into large XML and text files and then added to a database where the information could be parsed into spreadsheets to allow searching and sorting analysis. Sample data examples are given in the appendix. These files each contained tens to hundreds of thousands of data points for review. Parsed data was sorted in several ways to observe user behavior. At no point was

identifiable user data that was not already public on the website captured. User identities were made anonymous, but given unique identifiers during the initial analysis. This step was not necessary from a confidential standpoint as this data is public, but assisted in not clouding the initial analysis based on a user name that was descriptive. User IDs were revealed later when observing the images that were tagged. A sorted list of a user's tagging and commenting behavior could be used to see patterns in activity. It also allowed counts of how often certain tags or phrases were used by a user. The data was sorted in several ways to identify significant contributions. For instance, counting tag or comment contributions allowed identification of the most prolific taggers or commenters. These individuals could then have all their contributions pulled out of the data and the content of the contributions analyzed. Counting also allowed popular phrases or tag types to be observed and traced. Specific words or phrases that raised interest could also be counted, traced, and matched to user accounts or images as needed for observation as well. Analysis of string length allowed the isolation of the longest comments for observation. When interesting or frequent patterns emerged, then the photos in which the user's comments or tags appeared were viewed for further analysis by associating the photo id with the user, tag, or comment data.

As mentioned earlier, the Flickr data was captured and analyzed after an examination of the interview data. The interview data had provided a framework of what attributes might be important in an academic crowdsourcing or digital image collection environment. These qualities gave the basis for some categories of users that were sought out, for instance ones who were prolific, ones who appeared to be using standard vocabularies, ones who were adding detailed academic content, ones who were adding gibberish. Other categories for the Flickr user taxonomy appeared from the analysis of the data and were not discussed during the interview portion of the research. These categories became named based on consistencies in the patterns of their contributions which were made apparent by exploration of the data collected from Flickr.

INTERVIEW ANALYSIS

Analysis of the interviews took place first in an effort to glean information that would be useful when looking for tagging behaviors of Flickr users. To ascertain the competencies and level of involvement of those being interviewed, initial questions focused around their experiences using or administering digital image collections. Questions were also asked to set baselines for whether the interviewee's familiarity came from experience with digital image collections from their work, personal lives, or a combination of both. These types of introductory questions were less important, but provided some framing during the interview process.

What was being pursued were thoughts regarding what could be looked for to increase the efficiency of metadata creation for academic institutions. Other aspects that were looked for included what could help show that user generated metadata would be useful or accurate and what had been observed as positive or negative behavior of users in a tagging environment.

Responses from the interviews provided various sets of behaviors that were seen as positive or negative by the interviewees. These behaviors formed a skeleton of characteristics that were used as a basis for the observations during the Flickr data analysis portion of the study.

Responses Regarding Tagging Motivations

An overwhelming response from interviewees is that tagging is useful, though library staff members had concerns, which will be discussed further on, there is no argument as to the core idea that tagging is beneficial for image collections. Comments consistently included phrases such as "very helpful," but also many individuals noted that tagging was useful in allowing them to find additional images, similar to ones they may find in a search, by following the tags from one image to another. It

also was the consensus that far more people use the tags than do the actual tagging. There were distinct differences that were apparent regarding why and with what a user will tag an image. These distinctions were most evident between staff at academic institutions whose only experience in tagging images was work related and those, like students, whose only experience is tagging personal photos on photo sharing sites like Facebook. While library and museum staff members are looking to tag images to provide richness in metadata and depth in description of the image, personal taggers were less likely to focus on the content of the image. They appear more likely to tag the images for reasons, like "funny captioning," "so my friends can find the images" or so "I can find my images later." This is significant for several reasons. First, it is a real example to that there is a difference between accuracy of tags and relevancy of tags. Second, the goal of the tagging varies significantly between institutional needs and the needs of the typical tagger.

Academic institutions are not going to be interested in taggers who behave in the manner they do when tagging personal photos. Conversations with interviewees illustrated why, and a few more examples of personal image tagging demonstrate this difference. Instead of focusing on who was in a picture or where an event took place or what was in a picture, a personal tagger may label the image with items like "[So-and-So] acting wicked stupid." or "Wasted dude" or "Partying gals." While this information may be accurate, it certainly is not very relevant to users searching a database looking for information about who was in the picture or where it took place. This type of tagging was described as taking place, because individuals were tagging images for themselves and weren't concerned that the tags may not be relevant to anyone but the tagger.

This personal, colloquial, tagging is not useful to an academic institution. Responses about tagging, from those who worked with collections in a library or museum environment, had much different thoughts on what needs to go into an image. Since libraries often hire student workers to do work adding metadata to images, it may mean that students were approaching their work by tagging in

this more colloquial manner. Perhaps this would explain why occasionally tags in academic institution digital collections do not seem very descriptive or relevant. Academic staff and image collection administrators being interviewed stated that the notion of being able to find items again was important to them, but that they valued it less so when compared to the ability of their patrons to use the tags to find images. Two individuals interviewed mentioned that they had thought that academic online image collections were initially thought of as storage places, rather than searchable repositories for patrons. If that is the feeling, it can be seen why users would be less inclined to tag images or to add rich metadata.

Responses Regarding Tagging Types

The most common response as to what was important for images to be tagged with was location or personal name. This was consistent across the interviewees, regardless of the background of the individual. One individual stated that, "People rely more on the place being right because it is easier to prove right." Meaning that in terms of tag types, at least this individual believed that if someone has identified a location in an image with a tag, they would think it was accurate, more so than a personal name or date. In terms of dates, staff in libraries and museums felt that dates were more important than other typical users. Reponses such as "we often get requests for images from a certain date range" were heard from that group. Interestingly, when pushing for more information about perceptions on dates, it appears that images tagged with a date range, such as 1940s or 1980s, are perceived as less accurate than images tagged with a specific date, such as 1945 or May 6, 1980. This is intriguing since it is much easier to specify an accurate date of something to within ten years than to within a year or specific day.

Place and personal name are thought of as most important was consistent, but as with date, library and museum staff members and administrators of image collection sites felt that other information was equally important. Tagging behaviors that just focus on place or person are not the only

thing that should be examined. An example that was given was of someone describing an image of a statue of a person. If we follow the reasoning that place and name are most important, then we would miss out on other information that someone may be searching for. Such as the person who would have found the image if it had been tagged "bronze statue" or the person looking for the artist of the statue and the name of the person portrayed in the statue. In other words, library staff members wanted as much information as possible added, the "perfect record" of the image.

The idea of technical metadata tags also came up in discussions with library and museum staff members and a user who was a photographer. This information is defined as things such as media or camera type, or terms like 35mm, cyanotype, size information, f stop, film speed, etc. This information was not mentioned once by general users of image collections. Those who did mention it said that they did not think it was of use to general patrons of digital image collections. This photography information in general can only be tagged by the creator of the media, (although most digital cameras do provide it along with the histogram of the image). If it appears, it does not have a great deal of usefulness to the general public. However library and museum staff members using a digital image collection as a catalog of record, strongly feel the need to include this information regardless of its usefulness to general patrons.

Responses Regarding Controlled Vocabularies

One of the concerns with user generated metadata illustrated several times was the notion of controlled vocabularies. Without some sort of control over terms, there was concern that items would be labeled incorrectly or inconsistently. However, interviewees responded that by not having controlled vocabularies, a system can have users tag objects in multiple ways so they are more likely to be found

by searchers looking for unusual features of the image. The example of "plane" versus "airplane" was given as well as abbreviations such as "NYC" or "New York City" or "The Big Apple."

Library and museum staff members familiar with proprietary image collection systems, ones that do not allow for publicly created metadata, stated that staff often feel locked into terms that may have been initially setup years ago in a relatively archaic system. This results in staff, either out of laziness or lack of time, to just keep reusing the terms that are already included in the system. This forces tags into a best fit, instead of the creation of new tags that would more accurately describe the material that staff are adding to a collection. If control of vocabulary is not possible, it was suggested, that at the very least, staff follow the journalistic process of what to include, who, what, when, where, how. These findings lead to searching for tagging behaviors in users that show variety and not a narrow set of tags that the user constantly reuses.

Responses Regarding Perceived Accuracy

Views on the perceptions of accuracy in online collections yielded several results. Some individuals stated that when they looked for items online, they did not give much thought to whether items were described accurately or not. Others said the opposite, that they were skeptical and usually tried to confirm things from multiple sources online when doing research. This relates to image research particularly for individuals who may use Google image search. Since Google image search uses words on the page where the image is contained to bring up results. The image displayed in the search could have nothing to do with keywords searched by a user. There have been several cases of journalists using online searches to find a picture they want and accidently choosing an image that did not relate to what they really intended.

Regarding accuracy, one interviewee stated, "People rely too much on accuracy of information online in general. Maybe this is changing, maybe now people are becoming more skeptical." Perhaps this notion that people are more skeptical now has lowered expectations and perceptions of accuracy. However, when talking with students and even staff who do image searching online, they seemed to believe what they found was accurate and tended to blindly accept the tags as fact. When pressed, users and individuals not associated with libraries and museums did respond to the notion that accuracy for images was better on user generated metadata sites, such as Flickr, compared to something like a Google image search. In addition several individuals noted that they believed that tags posted by "power taggers" or individuals who tag a lot of items could be trusted as more accurate than tags posted by someone who only tags a few things. One interviewee stated that they believed there were a lot of retired people who may know quite a bit about a certain topic and have lots of time to spend tagging items, and that these people would be trustworthy. A library staff member countered that the concept of basing accuracy on the prolific nature of an individual tagging is not necessarily a great idea unless the user is identified in some way as a knowledgeable or reputable expert. If a prolific tagger can be vetted as having a good track record for accuracy, that tagger would be more beneficial for adding metadata in an academic environment.

Library Views of User Generated Content

There were several observations that pertained specifically to the perceptions of library and museum staff members and public tagging systems. One individual quickly expressed that if a library or museum is simply interested in amassing a unique collection of material that they alone house, why would they want to place it online to begin with where anyone can take the image. This is a good question if an institution has the uniqueness and monetary value of an item as their priorities. However,

most institutions understand the culture heritage or educational value of an item and consider sharing it with the world as a priority.

Staff of academic repositories have concerns with allowing users to have control over content that is associated with their institution. The idea that it may cheapen the image of a reputable institute occurred on several occasions. Users of these collections, faculty and particularly students, on the other hand, did not have such strong feelings regarding this notion of prestige. Mostly they were happy to use an image where ever it came from regardless of the status of the individual posting the image. There were concerns that allowing users to argue with each other over correctness of tags on a website associated with an academic institute would reflect poorly on the institution. Staff also made it clear that unless the metadata they placed online were perceived as accurate, there was no sense in them incorporating any sort of system like this. When library and museum staff members were asked if they thought that branding their images with logos or other information to associate the images with a reputable institution would increase the perception of accuracy, they all agreed. However, this made a bit of a "Catch 22" since there was this fear of associating contentious images with their institution if there were going to be negative arguments or derogatory user comments online. Again, there was also a great fear that users would place inaccurate information on images. The idea of trolling, or people purposefully adding wildly inappropriate or vulgar tags, was discussed as a possible problem. Users acting with this type of behavior should easily be found when examining their tagging patterns. One individual, who had experience working with a digital image collection company, said that they do, but rarely, have what they call the "penis problem." This is when a childish person goes through and randomly tags images with totally irrelevant and inappropriate words. Even though it is a rare occurrence, it is a great concern for library and museum staff members.

Interestingly, individuals who had experience in dealing with image collections, where the content may be particularly controversial, suggested specifically that moderation of collections is

important. This view steers slightly away from the hands-off approach that most library staff members viewed as happening once images were uploaded. With staff time being one of the problems for institutions, having to hire a full time Flickr moderator and then having them carefully review all comments and tags, is a tall order.

TAGGING PERCEPTIONS STUDY RESULTS

Results of this study gave guidance for looking into behaviors during the later Flickr data analysis. A more detailed quantitative analysis of this data formed part of the paper "ManyTags – User Perception, Motivation and Behavior in Tagging Systems" (Earle, Patzke, & Cheng, 2011). The first part of the study showed that single word tags and multiword tagging was perceived as very similar in terms of comfort, ease, participation, creativity and overall preference. Tagging was preferred to commenting in all these aspects except for creativity. Feedback from the open ended questions in the survey revealed that the majority of participants liked the tagging conditions more than the commentary, because it provided them a way to describe the images more quickly. There was a very small percentage of participants that had the opposite sentiment, stating that the restriction of only using single words didn't allow them to expand on the description or be as creative as writing lengthier comments would allow. These users commented that being forced into using shorter tags would encourage them to participate in social forums outside of tagging.

When reviewing user motivation comments, 86% of participants said that their motivations for annotating images were dependent on what specific condition they were presented. There was a feeling that tagging was generally more for the community, in an effort to make images more easily found, but that commenting on images was for personal gain to express an opinion, and less useful to the community. One participant concisely explained these considerations:

"[In the] single word case I'm less likely to think about what the image means and more likely to deal with exclusively the visual aspects of the image. In the commenting case I'm more likely to discuss a very specific aspect of the image or its meaning. In the multiword case I'll most likely be able to convey what I think about the image without having to put much effort into trimming down the phrase or coming up with something to speak about (as in the commenting case)."

Another participant felt that comments are:

"used to describe an image to someone, [while] single word tags [are] used to match search queries, [and] multi word tags [represent] the middle ground."

Other items that came out of the analysis of the first experiment of the study included that comments were perceived as significantly more subjective than the other two types of tagging, and that single word tags tended to be more emotional.

The study two results confirmed much of the study one results that single word tags and multiword tags were viewed similarly and that commenting was seen as something different. The evaluators responded that in both tagging conditions the shorter tags allowed for more accuracy, because they felt it focused them to describe what they see rather than write commentary that is more subjective. While evaluators perceived the two tagging conditions as more searchable, the comments were perceived as more interesting. This distinction led evaluators to not show as strong an overall preference between the three different conditions, where in the first experiment, the two tagging conditions were preferred over commenting. This suggests that metadata creators would prefer to focus on the completion of a task, tagging with short or one-word phrases, to easily disseminate information.

However, users on the other end of the task, the evaluators, have no preference for where information is gathered in relation to their search queries as long as the information returned is relevant.

Results from both experiments showed that users viewed tagging as helping others by increasing the searchability of the images, where commenting was seen as a way to share personal information or opinions that might not be as useful for searching. Combining results across both experiments, participants had overall preference for tagging. Multiword tagging was preferred slightly over single word tagging and comments.

This study shows that in an academic crowdsourcing setting, that interface design can impact the type of content that users contribute and their perceptions of the task they are completing. Instructions and interface design can be crafted in a manner to attract what an institution may feel is most beneficial. For example, if managers of a digital image collection wanted to crowdsource metadata that was more objective and community-focused, they could design in a manner that encourages single word tagging from participants. A project that requires more creative, subjective input from participants should focus on providing a comments-driven interface.

FLICKER USER ANALYSIS AND TAXONOMY

Utilizing information gained during the interview process, groups of ideal behaviors and trouble areas regarding metadata creation and dissemination in an academic repository setting were made. Within the interview discussions and with coding later, these groupings were fitted to characteristics that could be applied to the behaviors of users who generate metadata. The behaviors were then searched for within the data scraped from Flickr. Through examination of the data there are some distinct categories of user behaviors. These categories, though based on the characteristics that came from the interviews, did not have their formal taxonomy appear until after the Flickr data was analyzed. It was only after users with similar recurring patterns were observed that the typing of these users was made clear. The user type groups were given names that help describe their observed behaviors. The individual group types that stood out the most and related closely to characteristics derived from the interviews can be described as hobbyists, experts, specialists, geocoders, proofers, uniformists, promoters, remixers, narrators and oddities. Each type of individual has a certain set of patterns in which they tend to engage with a significant portion of their tagging or commentary activity. These actions each can have their own separate benefits to a library or museum wanting metadata added to their digital image collections. These categories are applied to users commenting and tagging other individuals' photos. A taxonomy of behaviors of users describing their own images may look similar, but certainly would have some variations in categories.

Hobbyists

This is the broadest category of user, but it contains three subcategories. Most users, who tag consistently, could be seen falling in this grouping. These are the users that are probably creating the

45

bulk of the broad description content in the Flickr Commons. These individuals show familiarity with some specific topic areas, a loyalty to a specific photostream, or just a consistent tagging output with no discernible pattern. A hobbyist might generalize around certain subject areas or photostream, but tend not to be exclusive. Individuals in this category typically just tag, they are quiet in that they rarely add commentary. As such, they aren't claiming to be professionals on a particular subject area. The subjects tagged could just be something with which the person has a curiosity or some knowledge. An example of this would be the tagger who spends most of their effort tagging images related to a broad subject, such as political Americana or Human Ecology. An another example of this would be an individual that follows a certain Flickr photostream and tags images as they show up, regardless of the images content. An individual in this category might have a personal connection to that particular institution. Their productivity on other photostreams can be non-existent or greatly reduced. They show a loyalty. Other users that fall into this category appear to just enjoy the act of tagging images and show no distinct patterns in what types of images or what types of collections they tag.

A fun example of a hobbyist that, at first examination, might not make sense, was a user who worked specifically on train images. Odd numeric tags kept reappearing in this user's activity, they were always three digit codes separated by a hyphen. At first glance, this information seemed to be meaningless and gibberish. To someone who is a fan of trains, these three digit codes are obvious and helpful. This tagger was describing locomotives using a well-known in the industry and 100+ year old standard, called the Whyte notation, which counts wheel arrangement. The first number indicates the number of non-driven leading wheels, followed by the number of driving wheels, followed by non-driven trailing wheels. This user's diligence of viewing images, counting the wheels, and adding the tag, fits clearly into the hobbyist behavior. Below is an example of an image tagged 2-10-2.

Flickr image title: **O&W Engine #358**

Flickr ID #3738206497 from Cornell University Library

Experts

The term, experts, does not necessarily refer to a user's tenure or prolific nature on Flickr. It pertains to their content that sets them apart. An expert is someone who was noticed adding details to a tag that an average, hobbyist type user might not know. Expert users will add information in the form of tags or comments that can add a great deal of richness to the metadata. Again it is important to understand that a user exhibiting expert behavior does not necessarily have to be a professional in the field that they are commenting. They could be, but they also could just be adding expert level content. What sets this apart from a hobbyist is that the tagging and description that is added goes beyond what

can be determined just by looking at the physical description of items in the photo. Both the hobbyist and expert categorizations can be tied to levels of expertise as defined by Collins and Evans (2002). Briefly, two categories of expertise as explained by Collins and Evans include interactional and contributory. Interactional expertise is gained through research and conversations with experts within a certain area. This level of expertise allows for knowledgeable conversation about a topic, but does not necessarily mean the ability to complete tasks within the field. Contributory expertise includes interactional expertise, but carries it further to say that these individuals can complete tasks within the field. This difference can be described colloquially as "walk the walk" versus "talk the talk." In this user taxonomy, a hobbyist may exhibit interactional expertise, but not contributory, where an expert user may have both types.

An example of an expert tagger in the data that was explored was a user that clearly had a strong architectural background and spent time adding the names of architects to images of buildings. This additional information goes beyond the initial description of what the building is, and can be seen as quite useful for scholars. What was also interesting about this user was they were inconsistent in their usage of tags pertaining to these names. In some circumstances they would tag an individual's name using separate tags for first and last name, and on other photos, they would tag by creating a single tag for the entire name. As was discussed earlier, this type of lack of control is very frustrating for library staff members and may not be as useful for searching. It also is obvious how it can be fairly useless to tag an image with a single name tag with a very popular name like Louis.

Below are examples of this user tagging images with architectural names. This first image was tagged by user with the separate tags, "dankmar," "adler," "louis," "Sullivan" for the names of the building's architects Dankmar Adler and Louis Sullivan. The user tagged several images with these exact terms.
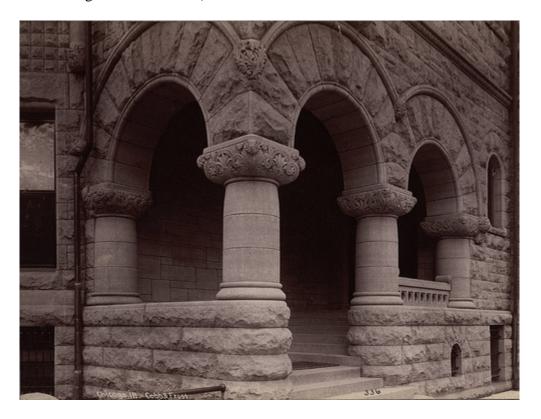
Flickr image title: **Interior, Auditorium Hotel**

Flickr ID #3678139789 from Cornell University Library

Here is another example of the same user showing expert behavior in the architect subject area. The user

tagged this image and others with the single tag "Henry Ives Cobb," the name of the architect.

Flickr image title: **Entrance, Ransom R. Cable House**

Flickr ID #3678118981 from Cornell University Library

Another user that demonstrated expert level content tagging relied on previous experience to add their content to the commentary section of images. Their consistent behavior across multiple images added information, but only validated it based on their memory, which is better than no validation, but some library and museum staff members would still be skeptical of information that isn't cited. Below are comments from fifteen images from this user from a single day. The content, if accurate, is quite a great amount of good added metadata to selection of photos on railroads. Note the recurring pattern of justifying validity using their previous remembered experience with "When I… ."

"When I saw this photo at Cornell in the late 1960s, it was captioned that this was the original Midland Shops in Oswego, near the site of the roundhouse."

"Scranton Br. bridge at Cadosia."

"When I viewed this photo at Cornell in the late 1960s, this was captioned as Oneida."

"This is the inspection car from which Master Mechanic John Minshull died near Trout Brook."

"When I saw these photos at Cornell in the late 1960s, this one was captioned that it is the crossing where Master Mechanic John Minshull died jumping from Inspection Engine 3."

"Trestle on left is D& H trestle served by O&W; right is O&W trestle."

"When I saw this photo in the late 1960s, this was captioned as a NY & Oswego Midland train, location unknown."

"This is part of the Fish's Eddy wreck repair."

"This is the D&H trestle in Oswego. Track underneath is the R&O to their freighthouse along the Oswego River."

"This is Port Jervis, Monticello & NY train on the Port Jervis Branch before O&W operations."

"This is Oswego, NY looking railroad southbound toward the roundhouse. Schuyler St. is on the right."

"This is Oswego, NY along Schuyler St. Note the slant board signal guarding the crossing of the O&W and NYC near the roundhouse."

This is Oswego NY, along Schuyler St. on right. View is railroad southbound, coal trestles are behind photographer."

"This is one of a series of photos of what appears to be a Dickson hog 2-8-0, probably in Mayfield Yard."

"This is the dock area in Oswego, NY."

Expert users who validate with greater explanations than memory could be more useful than the above example. These experts could fall into a subcategory that can appear to be know-it-all types. These are the users that may be professionals in the subject area for which they are adding information. They justify their responses by saying things like "I'm a professor of X at Y University" or "I received my doctorate in X field." These users feel the need to justify the validity of their responses by listing academic credentials. While this can be useful to library and museum staff members in trying to determine the accuracy of information they add, when it is presented in this manner, it can be a viewed as confrontational or egotistic and can make metadata needlessly complex and add information about the tagger unnecessary to the content of the image..

Other experts establish their credibility in ways that are received more positively. One way they may do this is by discussing employment at the place where the image was taken. The below exchange illustrates how a dialog of commentary can yield information that is useful. These questions occurred when a user noted that the picture of a building completed in 1796 was indicated as being designed by an architect born in 1937. Note the behavior shown in comment three.

Comment 1 "Why is Robert Adams the architect for a building completed in

1796?"

Comment 2: *"Good question! We're going to look into this, and we'll let you*

know what we turn up."

Comment 3: *"I work at the estate actually! There's a Robert Adams who was a*

Scottish Architect in the late 1700s who developed the style that the house was

built in, but the house was designed by William Hamilton who lived in the

mansion."

Comment 4: *"Thanks for the information! I'm passing it along to the cataloger*

who originally worked on these images."

Lastly, experts can also take the additional form of users who add content by linking to expert sources and references. These users have a mixed level of usefulness. They expose other users to more information about an image, but adding a link does not enhance the discoverability of the image. Linked sources also could have a varying degree of accuracy. Users who behave this way consistently link to sources that have been vetted as authoritative, can greatly enhance learning through linked content.

A typical behavior of this type can look like this:

"This building stands at the southwest corner of Erie and Wabash. Additional information and images can be found on the Art Institute Library website:
http://www.artic.edu/aic/libraries/research/specialcollections/cobb/res1.html"

Specialists

Specialists are some of the most interesting taggers on Flickr. One has to be careful not to confuse this type of Flickr user with hobbyist or expert. This tagger tags very specific items. The context of the photo surrounding the item they are tagging is not relevant to them. They are targeting a very specific observation. They may use very specific "expert" terms, but don't engage with other users as much. Unlike an expert, who may try to bring in as many details from a photo as possible when describing it, a specialist only cares about some detail of the photo, it could be a type of object, it could be a color, it could be a place. These users tend to not use the breadth of vocabulary in tagging that the hobbyist or expert use. They just use the small set of terms they value.

An example from one prolific tagger showed them focusing most of their energy around a narrow group of tags, which had seemingly limited use. This user used close to 100 different tags over many different images, but four tags made up nearly 50% of their tag usage. Nearly 100% of the images they tagged contain the word "woman," "outdoors," "portrait," and/or "standing." Woman, outdoors, and portrait have some degree of usefulness, but they are broad terms. What was odd was the continued use of the tag "standing," perhaps useful for someone who really did not want images of people sitting down.

The work of a specialist may not be as useful to library or museums or searchers of media, because the tags they add may only apply to a tiny portion of an image. These users deserve several examples, because they are quite fun. Tracking users that tag this way can give a sense that they are

playing some sort of hidden object game, encouraging other users to try and find why they tagged the image with that tag. It seems that some terms may lend themselves to specialists; they always seem to be objects. To explain, there may be a photograph of a wide bustling cityscape, and in one small portion of the photograph there might be an individual holding an umbrella. A tagger who specializes in umbrellas will see that tiny object and add the umbrella tag to an image and then move on to another image to do the same thing.

The first example is a tagger who showed a proclivity for two terms that attract specialists. The first example is a photo of many men in which the tag moustache has been added. Clearly a moustache is not the focus of the image; only two of the gentlemen appear to have small moustaches. The tag adds questionable enhancement to the metadata of the photo. Moustache is a popular specialist term, but often is used where the moustache is more obvious or unique.

Flickr image title: **Group [farewelling Eric Sheldon and band?] on wharf**

Flickr ID #4009413285 from the Powerhouse Museum

Another set of tags that appears to be popular with specialists are terms relating to laundry. The laundry specialists really begin to show the hidden item, a *Where's Waldo*, aspect of their behavior. Interestingly, the tagger who exhibited the above moustache tagging specialty also tagged laundry related items in a similar specialist style.

Below are examples showing some hidden image specialist tagging. The first example is an easy one to find; it shows a beautiful image of a cathedral. The cathedral clearly is the focus of the image, yet this image is tagged "laundry lines."

Flickr image title: **Durham Cathedral from Southeast**

Flickr ID #3611616846 from Cornell University Library

Here is a much trickier one tagged "laundry lines" by the same user as the above.

Flickr image title: **Old Nice**

Flickr ID #3485982005 from Cornell University Library

Give up? See below for the highlighted laundry.

Other examples of specialist behavior are the bicycle hunters. Below are two examples of this, first the easier one.

Flickr image title: **Westland Row? Absolutely not!**

Flickr ID #7628356832 From the National Library of Ireland

The bicycle example below is particularly interesting. In this case the tagger, once finding the very hidden bicycle, then used the style of bicycle to make a note about the age of the photograph. Apparently the gear style of this bike dates the image later than a certain year, and the specialist noted this. In this case, this odd specialization can be useful if no other information about dates is available. Libraries and museums could take advantage of date seeking specialists in this manner. This behavior can also be observed with individuals who use distant clock towers in pictures and then match times up with shadows on the ground to determine time of year a photo was taken.

Flickr image title: **Sibyl**

Flickr ID #9454984499 From National Library of Ireland

RIVER BLACKWATER.CAPPOQUIN.5845.W.L.

In some cases, the users who tag in this manner don't just tag or annotate the actual object, but related objects as well. Using the laundry and bicycle examples from above, here are two further examples illustrating text related to the object.

In this photo the sign on the left side of the image has been annotated with the word "laundry," the sign says "French Laundry."

Flickr image title: **Argyle Cut, The Rocks**

Flickr ID #3003576493 from the Powerhouse Museum

In the below photo, an annotation regarding bicycles has been added. It is quite hard to find, but when viewing the full resolution image, the signage on one of the trolleys is advertising a bicycle shop.

Flickr image title: **Born to Run**

Flickr ID #6978773963 from National Library of Ireland

Another entertaining tagger was an individual who was tagging items using the great phrase "Library Porn." Now this may bring up various connotations, but what it appears they were doing was tagging images that showed particularly wonderful old photos of libraries. This user also had another specific tag phrase called "Libraries Across the Commons." This phrase was used to show pictures involving libraries, but just within photos found within the Flickr Commons area. It is easy for users to search only library photos found within the Commons area of Flickr, but having a user develop a tag to specifically identify photos as such is interesting. These user-created tag phrases that aren't clearly defined, but hold a specific connotation to the user who created the tag, are not necessarily useful at first. A searcher looking to find spectacular photos of libraries might not think to try the phrase "library porn," but if a phrase like this gains popularity with other taggers, it can become quite useful if known.

Examples of a library as defined by this tagger, but not worthy of the Library Porn tag.

Flickr image title:

View of corridor, David Scott Mitchell's residence, c. 1907, by unknown photographer

Flickr ID #4415252053 from State Library of New South Wales

Flickr image title: **City, public library**

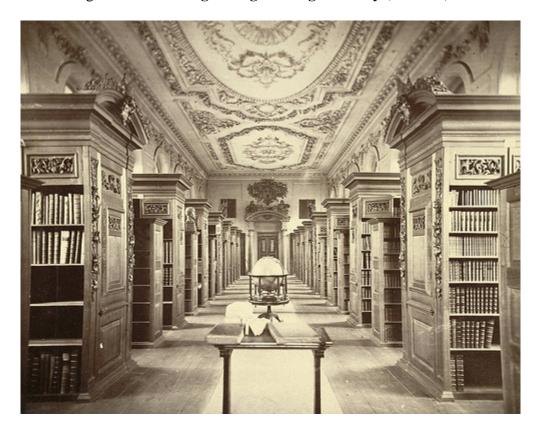Flickr ID #2899334394 from The Library of Virginia

Samples of Library Porn images as determined by the same tagger.

Flickr image title: **[Main Reading Room. View from above showing researcher desks. Library of Congress Thomas Jefferson Building, Washington, D.C.] (LOC)**

Flickr ID #8471102120 from The Library of Congress

Flickr image title: **Cambridge. King's College Library (Interior)**

Flickr ID #3610752603 from Cornell University Library

Also of note on the above photo is that its page shows the proofer behavior category. A user has commented that the image is actually Queen's College and then provides a link to a reference justifying their statement. This user shows proofer behavior in other comments as well.

Geocoders

This type of user is most interested in the location of images. Flickr allows users to add location information that places the photos on a map at their proper location. The users who fall into this category spend most of their time ensuring that this location description is accurate. The comments below were from a user who annotated several images with similar statements.

"mapped incorrectly, this was at the corner of Fifth Avenue and 52nd Street"

"this is not mapped correctly"

Other taggers who may have more familiarity with a locale are useful in adding correct native spellings of locations. Contributions from a tagger that exemplified this behavior can be seen below. Note that this user makes the extra effort to use proper punctuation and even alphabetic characters such as the Thorn, Þ, and Eth, Ð, which are not used in all languages.

Tags by a geocoder user

Austurhlíð	Reykholtsdalur	Svínafell
Eyjafjallajökull	Reykjavik	Þingvellir
Eyjafjörður	Rjómabú	Þórsmörk
Hvítá	Skeiðarárjökull	Tungufljót
Nikulásargjá	Stórólfshvoll	Vopnafjörður

Proofers

Proofers correct things. One of their distinctions is that proofers tend to stay quiet until they see things that need correction. The majority of their contributions are solely to point out and revise errors. Correcting information that is inaccurate is of course helpful for library and museums, but it can be hard for staff to blindly accept corrections from the public. Fortunately, another trait of proofing individuals is that they often cite their sources. They might provide a link to another source to help verify their response or just provide some form of validation beyond just this is wrong and here is the right answer. For instance, in the image below, a reasonable and accurate correction is made by a user, but for it to satisfy the professionalism of some metadata librarians, it would be nice if it included validation as part

of the correction. The image contained the description "From the left the names of the instruments are 'piri' (flute), 'haegum' (two-stringed fiddle), 'komungo' (six-stringed zither), and 'kayagum' (twelve-stringed zither)." Which was corrected by a user with:

"Two corrections: 1) The piri is an oboe, not a flute and 2) The trapezoidal string instrument on the right is a yanggeum (hammered dulcimer), not a kayageum."

Flickr image title: **A concert of Corean musicians**

Flickr ID #4095409537 from Cornell University Library

It is interesting that despite the correcting behavior on this image, a correction of the transliterated word Corean, has not been corrected to Korean to add more useful searching.

While the above is an example of corrective behavior, proofers provide a greater and more useful level of correction, with cited sources. A good example of a well cited source is this comment added by a user who showed proofer behavior; the below caption comes from the image **Cambridge. King's College Library (Interior)**, which is mentioned earlier in this paper.

"Mis-captioned - this is the library at Queen's College, Oxford (compare the admittedly tiny image at www.queens.ox.ac.uk/library/ - a Google search will bring up bigger images, but this is the College's website - they can't have got it wrong!) According to Anthony F. Kersting and John Ashdown, The buildings of Oxford (London: Batsford, 1980), p.147 the ceiling represents two phases of work, the primary scheme by James Hands completed in 1695 having been embellished in the rococo style by Thomas Roberts in 1756 (I have added a note to the photo to indicate one of the areas that I think this involved)."

These example comments represented one user's comments from a single day of commenting and illustrate this behavior clearly:

"I could be wrong, but I think your description may be inaccurate. I don't think this is a street with a stream running through it but the streamed [sic] *itself. It's possible it could be Ch'ŏnggyech'ŏn (Cheonggyecheon)."*

"I could be wrong, but I think it is unlikely that this is Seoul, at least not along the Han River, which is considerably wider than this (even back then)."

"I think Shariin (doesn't the hiragana say Sharirun?) is supposed to be Sariwon (Korean characters), now in North Korea, a bit south of Pyongyang. My Halmŏni was from there."

"This kind of thing was apparently a cultural no-no that shocked the Korean hosts: http://www.monster-island.net/2009/11/century-of-foreign-guests-behaving.html"

Uniformists

Uniformists are taggers that are all about consistency. They want tags to follow a standard. Some of their actions may seem like those of a proofer or specialist, but by following them, you can see the differences. They may spend time correcting items that already have a tag that they want to fix, or they may just be very deliberate in how they add new tags. Uniformists add tags to items that may not be obviously incorrect, just formatted in a way they think is wrong. The standards that a uniformist uses may come down from an expert user, in other words, if a user on Flickr is thought of as an authority and always tags the same way, a uniformist may pick up on this and base their corrections on this standard. Uniformists can be seen adding date tags in an attempt to get date formatting consistent, an example of this might be separating a tag that is listed as "October 1984" to two individual tags "October" and "1984." Or they may go the other way and combine the two separate terms into one, more useful, tag. Uniformists also can be seen expanding abbreviations or adding abbreviations of terms to assist in searchability. In an academic setting, particularly a library, where, as discussed earlier, there may be standards that seem archaic and non-user friendly in structure, a uniformist, if using proper standards, could be welcomed. However, a uniformist may be correcting content for searchability, and the terms they use may be in conflict with old established controlled vocabularies used in libraries.

An example of a user understanding controlled vocabularies can be seen in the pattern illustrated below. This user was deliberate in making sure that images they tagged would be found by all forms of an image. The chart below gives an example of just two of the images they tagged. Note the deliberate nature of ensuring that all the terms they used are listed in multiple ways so that a person may find them.

Uniformist Tagging Showing Multiple Vocabularies

Image 1 tags	Image 2 tags
Rhode Island	New York
RI	New York City
Narragansett	NYC
Narragansett Pier	City
Narragansett Pier, RI	Manhattan
Narragansett Pier Casino	59th and 5th
	59th Street
	Fifth avenue
	Planning
	Urban planning

Instead of adding a single term that is descriptive and moving on, a user like this is doubling, tripling or more their effort on a single image to ensure its discoverability.

Another uniformist was the individual who sought out images of Theodore Roosevelt and then consistently added the following set of tags to multiple images. Note here, as with the above, the user making a deliberate effort to include multiple variations of spellings to ensure the image is found.

Uniformist Tags Showing the Use of Name Variations

President Teddy Roosevelt
President Theodore Roosevelt
Teddy Roosevelt
Theodore Roosevelt
TR
Republican
Republican Party

This individual also falls into another category, the specialist, in that 96% of their tagging behavior involved the above tags alone.

Promoters

Promoters are easily explained and of minimal use to libraries and museums. Flickr allows users to add images to groups of photos and they do so by asking permission and posting a link in a photostream. While this may allow someone browsing to find more images that relate to the image they have found, these categories are often very subjective or odd, like "Cats eating weird things" http://www.flickr.com/groups/77507197@N00 . Promoters are users whose behavior is almost exclusively that of posting a link to a group they manage. There were many examples of a user's only activity on the Flickr Commons falling into this category. Examples from the understandable to the unclear to the odd include:

> *"Hi, I'm an admin for a group called ANYTHING GOES ENGLAND,VINTAGE*
>
> *PHOTOS,POSTCARDS,EPHEMERA etc..., and we'd love to have this added to the group!"*
>
> For the group: http://www.flickr.com/groups/534552@N23

> *"Hi, I'm an admin for a group called Clown will Eat me, and we'd love to have this added to the*
>
> *group!"*
>
> For the group:
>
> http://www.flickr.com/groups/clownwilleatme

> *"i love you just the way you are, please come h ♥ m e"*
>
> For the group: http://www.flickr.com/groups/1147976@N25 a group about home décor.

"Hi, I'm an admin for a group called Rhode Island Architecture, and we'd love to have this added to the group!"

For the group: http://www.flickr.com/groups/996802@N23 of note about this user is this:

was their avatar, not something that particularly associates well with an academic institution trying to forge a relationship with Flickr.

Remixers

Remixers are users who take an image from a collection and then use the image on something and then comment back on their accomplishment. These users are even less useful in terms of metadata content than promoters. They are, however, sometimes entertaining. A use that they might serve for a library or museum is advertising an image collection. They also can be used to show the breadth of users who find a Flickr Commons collection interesting, images aren't just for scholarship.

Below is a remixed image and the user's comment.

"I made a purse from this image;"

[http://www.flickr.com/photos/11418107@N02/5263738943/]

Flickr image title: **Flora Rose Purse**

Flickr ID #5263738943 from pennylrichardsca (now at ipernity)

The original image used by the remixer is below.

Flickr image title:

Oval portrait of Flora Rose in straw hat, dated by Beulah Blackmore at about 1916.

Flickr ID #3856344396 from Cornell University Library

Another example:

Flickr image title: **Air Ship**

Flickr ID #6094611275 from rpl47

Original image:

Flickr image title: **Freight Ship Pueblo**

Flickr ID #3738742491 from Cornell University Library

In another case, a user showed how they had used a collection of photos to take patterns found in the architecture of Medieval and Renaissance era buildings and turned the patterns into wallpaper in real life.

Narrators

Narrators fill comments with stories or commentary about the images. These are individuals who add comments to images that tell some sort of story. They may describe visiting the place or meeting the person in the photo. Sometimes they will post the same comment on multiple images, and the comments often are too subjective or personal to be of much use in explaining or describing an image. In some cases these stories contain content that may add descriptive information, but usually they are strictly personal anecdotes. Again, these comments often are not useful or helpful to others. Narrative information is important and sought out in the museum and library world, often in the form of oral histories. Oral histories are a great example of how hearing an individual's description of an experience can be enormously helpful in capturing a time or an event. In a setting such as Flickr, the narrative can stand out as out of place. Where they are most useful is in a place where library staff members are moderating so they really notice the comments. In this manner an individual that does have a particularly useful story to tell can be sought out for a formal oral history or description. This is opposed to haphazardly cramming information into a comment. With Flickr having a much bigger exposure, it is nice to have information there, but extensive oral history related comments are better suited for other websites. The website HistoryPin, http://www.historypin.com , has been designed to capture this sort of memory related content. HistoryPin also has great features allowing Flickr images to be aligned with Google Street View images for then and now style imagery.

This discussion below is an example of narrative commentary that does not provide any really useful information, but is worth noting. It is notable because the users engaging in the conversation both appear to have been security guards at the place where photo was taken. The photo they are commenting on is one of Westminster Abbey. While the content they discuss does not add much for other users,

library staff members who might be interested in learning inside information about working at the facility may find this an opportunity to reach out to these individuals to learn more.
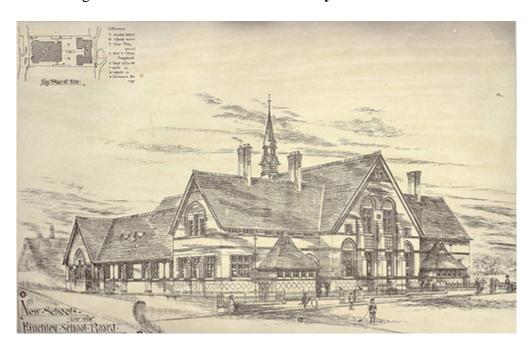
User 1: Starts by addressing a user who posted previously and then continues: *"Memories of 18 Protection Post whilst at Cannon Row, presumably introduced after the Stone of Scone went walk about on Christmas Day1950? Nice picture of the 'quieter' days."*

User 2: Address User 1 then continues: *"I don't remember that protection post butI was on duty the night when the stone was removed from the throne and half way down the aisle the wheels on the makeshift cart collapsed. It was sometime in the seventies and the alarm worked!"*

User 1: Addresses User 2 then continues: *"Yes, unforgettable really, early 1960's it was 18 protection post, pretty grim in the middle of winter on night duty! One night (3am) had to search the interior with others, talk about eerie, the hairs on the back of your neck stood up as you had no idea if there were intruders or not, thankfully none were present."*

Below is an example of extensive commentary. This user writes several sentences about the photo. In this case some of the information may be useful.

Flickr image title: **New Schools for the Finchley School Board**

Flickr ID #3611678698 from Cornell University Library

"Very interesting. Looking at the plan in the top left hand corner the road in the front of the school is Long Lane, East Finchley. I went to school at the Alder County Secondary School for Boys, Long Lane, East Finchley and the little map looks exactly right with a church (hall?) situated at the back of it. The build date is right and it had an infants section with separate entrances. Even the sites for the boys and (ex)girls toilets is right. The buildings at the bottom of the playground were different (the wood and metal work shops in my time). So was this an early draft of what was eventually built? The building was demolished in the 1970s I think but I have been unable to find a picture of the outside. I am certain a picture exists on the school's records on Friends Reunited"

Some examples of commentary that is very narrow in its usefulness can be seen below. These can be particularly cluttering if the user adds the identical commentary to many photos. When the commentary has a unique or rare word that gets picked up by the search engine, it can be frustrating to other users as the images get returned in the results, but it doesn't contain useful information.

"Do you have any information about the people in this pic? I'm from Stóri-Núpur and my family is very interested about this picture."

"I like these old photos. The frame itself is an artwork. I also keep this kind of photographs of unknown portraits from Riga, Latvia, that my husband's grandma left."

Oddities

Oddities need to be mentioned to describe some of the behaviors that clearly aren't particularly useful. One particular example is hard to explain, but it is a user type who creates quite a bit of clutter in a photostream. This type of user either is unfamiliar with how Flickr works or has some other odd reasoning for their behavior. In Flickr, users can build their own albums of favorite photos from other individuals' photos. This allows users to revisit photos easily on their own page. It is noted to the owner of the photo and to the public who has marked it as a favorite, but this indication does not clutter the searching for other users. What was observed were some users marking items as favorites, but then also tagging the photo with their user name. This seems like a completely unnecessary step, and it adds useless and potentially misleading metadata to an image. In at least one case the only term that one of these users tagged was their own username, they did not add any other type of tags. It is hard to say what the thought process was or if this was a misunderstanding of how to use Flickr. This user also had an appropriate Favorites page indicating that they did know how to use the system.

The problem of totally strange or vulgar posts by users, as mentioned earlier, was observed on some occasions. This behavior clearly is not wanted and is quite inexplicable, but is also apparently inevitable if the digital image collection is large enough and unmoderated. This type of behavior shows the need for library and museum staff members to be willing to accept that this may happen or to ensure they have enough staff time to moderate user contributions. Below are two examples of this behavior.

In this photo the image has been annotated by a user with the phrase "oh my penis"

Flickr image title: **[The road from Eide to Voss, Hardanger Fjord, Norway] (LOC)**

Flickr ID #3174178805 from Library of Congress

This following image has been annotated with the tag NSFW, the common abbreviation seen to indicate

Not Safe For Work. It also has annotation "help he's raping me!!!"

Flickr image title: **Collision between two engines, Bay of Quinte Railway, ON, 1892**

Flickr ID #2918567169 from Musée McCord Museum

ADDITIONAL DISCUSSION

Combining what was learned through the interviews about what traits are desirable with the details discovered through the Flickr user analysis, it is possible to form a simplified chart matching user types with characteristics. In this manner we can see what type of user is most likely to perform well within certain areas of metadata creation. The chart below shows each user type's performance level on goals that are of concern for academic institutions. A high level of performance will be seen as more desirable to an academic institution when putting an emphasis on a particular goal.

Searchability: User increases the ease of which images can be found.

Quantity: User is prolific and contributes to large numbers of individual items.

Accuracy: User's content is generally high quality in its relevancy and detail.

Depth: User's content provides details or specifications that an average user could not.

New Information: User provides content related to the item, not just description of what is seen.

Exposure: User's content increases the presence of the item, collection, and institution.

Flickr User Type Performance Levels

	Searchability	Quantity	Accuracy	Depth	New Information	Exposure
Hobbyists	Medium	Medium	Medium	Medium	Medium	Low
Experts	High	Low	High	High	High	Low
Specialists	High	Medium	Medium	Low	Low	Low
Geocoders	Medium	Medium	High	Low	Low	Low
Proofers	Low	Medium	High	Low	Low	Low
Uniformists	Medium	High	High	Low	Low	Low
Promoters	Low	Low	Low	Low	Low	High
Remixers	Low	Low	Low	Low	Low	High
Narrators	Medium	Low	Medium	Medium	High	Low
Oddities	?	?	?	?	?	?

Trying to create a perfect record with metadata describing all aspects of an image is time consuming. By allowing staff to focus on one particular portion of the metadata, we can increase efficiency. This research can help by allowing libraries and museums to develop strategies and interfaces to capture the most useful metadata and the most useful metadata taggers. There remains the difficulty that even if we do know what types of metadata are most useful. Do we have library staff members add that information or is that the type of information that is more likely to be added by users? Should they instead add what would be difficult for users to know about an image and then let the easier information that can be pulled right from the image itself be added by patrons? Either way, by doing a little work on the back end and getting the image online quickly for tagging, libraries are likely to be able to increase the richness of their collections while decreasing the amount of time they spend describing images.

There are clearly legitimate concerns with user generated content that can only be addressed by attempting trial runs and allowing institutions to come to their own conclusions. At this point there is no all-encompassing answer. Fortunately there are some good patterns and focus areas that can continue to be studied.

For an academic institution to successfully implement crowdsourced tagging programs, they should take note of this and facilitate the formation of an active community or group setting. These social aspects help feed motivations, encouraging more tagging, and thus more visibility of images for other users. This cycle can help ensure a self-sustaining environment.

Recommendations described within the taxonomy can be presented to institutions, but goals of individual libraries and museums may dictate what they believe is most useful for their collections. Approaching more library and museum staff members or system administrators with the defined categories of users and then doing further survey work on their opinions of usefulness could yield

helpful information. Implementation of systems and studying real world results is the best way to further study.

IMPLICATIONS FOR FUTURE RESEARCH

An immense amount has been written on tagging and crowdsourcing, but far less has been written applying these studies and findings to a library or museum specific setting. These areas clearly have specific characteristics that make them different from typical personal tagging. Further collaboration with cultural heritage repositories with research would be welcomed and beneficial to both the institutions and the researchers.

The most useful information for institutional users of image tagging would be to follow up on the motivation of specific tagger types and relate them to variations in interfaces. An institution could attract the type of user behavior they believe would be most beneficial for metadata creation by customizing an interface. Concurrent research to this thesis noted significant differences in user behavior, and perceptions of accuracy of tagging can be controlled by interface design (Earle, Patzke, and Cheng, 2011). Flickr could build in options for content owners so that the photostream owners could dictate different abilities of the interface for users when they tag or comment. There is a great deal of research on the accuracy of tagging, but understanding why some taggers focus on correcting tags rather than adding new information is also important. In previous taxonomies of human computational efforts (Quinn & Bederson, 2011), it would be important to understand if factors like altruism, enjoyment and/or reputation are motivators when doing tagging for a library and/or museum. Do these motivators differ from image tagging done for other reasons?

Another issue that reflects on the reputation of an institution is quality control. Again, library and museum specific studies regarding how users react to different variables regarding types of tag (date range, location, name, source of the image, etc.) or who does initial tagging could show more defined

differences in quantity and quality of tagging behavior. Studies modifying the variables or specificity in the task description given to the public when tagging could also show ways to yield better quantity and/or quality of tags. Is there a benefit to having institutional staff "seed" images with tags to increase tagging behavior and guide the accuracy and quality?

A separate direction in which the metadata problem is being addressed is through the gamification of digital collection interfaces. Research, as described above, can be applied both in directing the design of metadata games and in analyzing the contributions they receive from participants. There have been a number of projects that have accomplished metadata creation, transcription, and citizen science through online crowdsourcing games. One of the most popular examples is Zooniverse, http://www.zooniverse.org. These projects have shown that you can engage the public to complete research in a citizen science fashion.

Similar models can be applied to digital image collections. The ESP game developed by Luis Von Ahn provided an early example of how the public could be effective at identifying images and enjoy the task in a gaming environment (Van Ahn & Dabbish, 2004). Players of the game were presented images for which they had to describe with tags and would advance through the game by matching responses with each other. A similar Dartmouth based project, called Metadata Games, http://metadatagames.com, has made some inroads in developing tools specifically for library and museum collections. One problem with these tools is the game aspects haven't developed for or really immersive play; they are repetitive and this may impact long term enjoyment of participation.

A final approach, for tackling metadata creation by the public, can be in utilizing tasks that are already a part of everyday life. The reCaptcha project, http://www.google.com/recaptcha, is an excellent example of a web-based task that fulfills the need of human verification, while also digitizing the content of books. Further study and development of projects along these lines for images in digital repositories could help create massive amounts of metadata without the need to recruit or maintain a

captive audience. Incorporating design based on the different user behaviors into the development game or task type crowdsourcing systems can allow for future study of the effectiveness of varied designs.

CLOSING REMARKS

The vast amount of data that can be garnered from crowdsourcing sites allows for numerous studies. With interest in crowdsourced projects rising among academic institutions, results from these studies can be used in the design and implementation in new systems quite easily. Ideas and knowledge gained from the results can be tested and then further studied. Academic institutions, while naturally hesitant to reinvent the wheel when starting a crowdsourcing project, are often willing to partner with researchers, and at the very least, provide mountains of content needing description. Research and information gained during the production of this paper were quickly shared and applied to guide the development of crowdsourcing projects.

APPENDICES

Interview Guide

1. Initial questions regarded background in terms of whether participant was a library or museum staff member, faculty, student, other general internet user.

2. What is your experience with digital image collections? Probing to get a level of user experience.

   a. If person has experience with digital image collections in a work environment, probe them to distinguish how they used collections for work and for personal use.

3. After searching for an image, what are your experiences with tags in image collections?

   a. Have you tagged images?

      i. What is your thought process when tagging; does it differ depending on your goals or if the images are yours or someone else's?

      ii. Is there a certain type of tag you find yourself adding more than others?

      iii. Do you ever fix incorrect information in tags, if so, why and what types of inaccuracies were they?

   b. Do you think that certain types of tags are more or less accurate than others? For instance comparing dates versus names.

   c. What types of tags are most or least helpful?

4. What are your thoughts on the accuracy of information placed in user content generated image collection websites?

   a. When you look at users who add information to these websites, are there aspects that lead you to believe that their information is more accurate than others?

   b. What can academic institutions do to help convince you that their information is accurate?

5. What are your thoughts on user generated tags placed in academic digital image collections?

    a. Have you noticed users that do a particularly good or bad job in tagging?

        i. How do you characterize good or bad?

    b. What do like to see from tagging content generated by users?

    c. Are there particular types of tagging that you see as more useful, less useful?

Tagging Perception Study Survey Questions

Study 1:
Questions for the end of session survey:
1. How familiar are you with tagging images online? What platforms or websites are you most familiar tagging?
2. How comfortable were you in using the single-word tagging environment?
3. How comfortable were you in using the multi-word tagging environment?
4. How comfortable were you in using commenting environment?
5. Which environment did you like the least?
6. Which environment did you like the most?
7. How much did your motivations or intentions change when you were tagging in the single-word tagging environment?
8. How much did your motivations or intentions change when you were tagging in the multi-word tagging environment?
9. How much did your motivations or intentions change when you were tagging in the commenting environment?
10. When searching for images online how much attention do you give to tags already assigned to the images in the search results?
11. If you were performing an online search for an image is there a condition (single-word, multi-word, or commenting) you would prefer?
12. How willing would you be to participate in tagging the single-word tagging environment? If you believe this type of tagging influenced your willingness; explain why.
13. How willing would you be to participate in tagging the multi-word tagging environment? If you believe this type of tagging influenced your willingness; explain why.
14. How willing would you be to participate in tagging the commenting environment? If you believe this type of tagging influenced your willingness; explain why.
15. When considering the single-word tagging environment, do you feel it encourages or discourages your creativity when tagging? If you feel this encouraged or discouraged creativity; explain why.
16. When considering the multi-word tagging environment, do you feel it encourages or discourages your creativity when tagging? If you feel this encouraged or discouraged creativity; explain why.
17. When considering the commenting environment, do you feel it encourages or discourages your creativity when tagging? If you feel this encouraged or discouraged creativity; explain why.
18. What is your age?
19. What is your sex?
20. What is your major or occupation (if applicable)

Study 2
Participants in survey two were given the questions from Study 1 as well as the following questions presented with every image annotation pair:
1. These words accurately describe the image. (Descriptiveness)
2. These words would be useful for searching for this image. (Searchability)
3. These words offer a different or interesting perspective of the image. (Interestingness)

Flickr Data Scraping Script

The script was created by Dianne Dietrich, Physics & Astronomy Librarian at Cornell University, for data analysis of Cornell University Libraries Flickr Collections, July 19, 2010. The script has since been modified in 2011, 2012, and 2013 to optimize data collection.
The script requires Active Python and the Flickr api egg to be installed.

```python
#!/usr/bin/env python26

import time
import flickrapi, sys
import re
import lxml.etree as ElementTree
from datetime import datetime
import sys

# Insert API Key and Secret
api_key = 'REDACTED'
api_secret = 'REDACTED'

# Insert User ID and collection # of desired institution
# Example National Archives - Harry S Truman Presidential Library and Museum

user_id="35740357@N03"
collection_id="72157629840523993"

# Connection function
def create_session(api_key, api_secret):
    """Creates session"""

    session = flickrapi.FlickrAPI(api_key, api_secret, format='etree')
    (token, frob) = session.get_token_part_one(perms='read')

    # Needed to authorize the "app" to access the account
    if not token:
        # If the browser doesn't spawn a new window to auth...
        print session.auth_url('read', frob)

        raw_input("Hit return after authorizing this program with Flickr")
    session.get_token_part_two((token, frob))
    return session

# Create the session
flickr = create_session(api_key, api_secret)


psreply = flickr.collections_getTree(user_id=user_id, collection_id=collection_id)



# Set up the XML to write to
root = ElementTree.Element('flickrmetadata')


#from lxml import etree # nothing legacy here...
```

```python
# {Collection: [set id, set id, ... ]}
colls = {}

# {Collectionid: Collname} # To not have to make another API call
collnames = {}

# ...

pscolls = psreply.findall('collections/collection')



for pscoll in pscolls:
    colls[pscoll.get('id')] = []

    collnames[pscoll.get('id')] = pscoll.get('title')


    for set in pscoll.getchildren():
        colls[pscoll.get('id')].append(set.get('id'))

# Get ids per coll/set
for coll in colls.keys():
    for set in colls[coll]:
        psitems = flickr.photosets_getPhotos(photoset_id=set)
        psphotos = psitems.findall('photoset/photo')

        for psphoto in psphotos:
            photoid = psphoto.get('id')

            item = ElementTree.SubElement(root, 'item')
            fadmin = ElementTree.SubElement(item, 'fadmin')
            culadmin = ElementTree.SubElement(item, 'admin')

            # Collection ID
            flickrcollectionid = ElementTree.SubElement(fadmin,
'flickrcollectionid')
            flickrcollectionid.text = coll

            # Collection Name (for culadmin)
            culcollection = ElementTree.SubElement(culadmin, 'collection')
            culcollection.text = collnames[coll]

            # Set ID
            flickrsetid = ElementTree.SubElement(fadmin, 'flickrsetid')
            flickrsetid.text = set

            # Photo ID
            flickrphotoid = ElementTree.SubElement(fadmin, 'flickrphotoid')
            flickrphotoid.text = photoid

            # Get information for each photo
            usermetadata = ElementTree.SubElement(item, 'usermetadata')

            photoinfo = flickr.photos_getInfo(photo_id=photoid)


            # COMMENTS
```

```python
            # Comments requires a different API call
            try:
                photocomments = flickr.photos_comments_getList(photo_id=photoid)
            except:
                time.sleep(2)
                photocomments = flickr.photos_comments_getList(photo_id=photoid)

            allcomments = photocomments.findall('comments/comment')
            comments = ElementTree.SubElement(usermetadata, 'comments')

            for onecomment in allcomments:
                comment = ElementTree.SubElement(comments, 'comment')

                commentauthor = ElementTree.SubElement(comment, 'author')
                commentauthor.text = onecomment.get('author')

                commenttext = ElementTree.SubElement(comment, 'text')
                commenttext.text = onecomment.text

                # DD: New for date
                commentdate = ElementTree.SubElement(comment, 'date')
                xmldate = onecomment.get('datecreate')
                toconvert = int(xmldate)
                commentdatestamp =
str(datetime.fromtimestamp(toconvert).isoformat())
                commentdate.text = commentdatestamp


            # TAGS
            alltags = photoinfo.findall('photo/tags/tag')
            tags = ElementTree.SubElement(usermetadata, 'tags')

            # IDENTIFIER (in tags)
            identifier = ElementTree.SubElement(culadmin, 'identifier')
            identifiertags = []

            # Turn this on and off to pull tags added by the photo owner. If on it
ignores owner generated content

            for onetag in alltags:
                tagauthor = onetag.get('author')
                if tagauthor != user_id:
                    tag = ElementTree.SubElement(tags, 'tag')

                    tagauthor = ElementTree.SubElement(tag, 'author')
                    tagauthor.text = onetag.get('author')

                    tagtext = ElementTree.SubElement(tag, 'text')
                    # tagtext.text = onetag.text
                    tagtext.text = onetag.get('raw')


            #else:
            #   if (onetag.text.startswith('culidentifier')):
            #       # Create the list of just culidentifier tags
            #       identifiermachinetag = onetag.get('raw')
```

```python
            #         identifierpattern =
re.compile('culidentifier:(.*)=(.*)').match(identifiermachinetag)
            #
            #         identifierchild = ElementTree.SubElement(identifier,
identifierpattern.group(1))
            #         identifierchild.text = identifierpattern.group(2)

                # This could also be done with attributes




            # NOTES
            allnotes = photoinfo.findall('photo/notes/note')
            notes = ElementTree.SubElement(usermetadata, 'notes')

            for onenote in allnotes:
                note = ElementTree.SubElement(notes, 'note')

                noteauthor = ElementTree.SubElement(note, 'author')
                noteauthor.text = onenote.get('author')

                notetext = ElementTree.SubElement(note, 'text')
                notetext.text = onenote.text

                noteboundingbox = ElementTree.SubElement(note, 'boundingbox')
                notebb = (onenote.get('x'), onenote.get('y'), onenote.get('w'),
onenote.get('h'))
                noteboundingbox.text = str(notebb)


print ElementTree.tostring(root, pretty_print=True)
```

Sample Data From Script

Below is an example of sample data pulled from a single photograph in a collection. It shows the unique

identifier letting us know which photo the metadata comes from and then provides user specific

information showing what a user tagged or commented on an image.

```xml
<item>
- <fadmin>
  <flickrcollectionid>30510347-72157616674562400</flickrcollectionid>
  <flickrsetid>72157619422460331</flickrsetid>
  <flickrphotoid>3611960232</flickrphotoid>
    </fadmin>
- <usermetadata>
- <comments>
- <comment>
  <author>42826854@N00</author>
  <text>Hi, I'm an admin for a group called <a
    href="http://www.flickr.com/groups/892403@N23/">Heritage Key</a>, and we'd
    love to have this added to the group!</text>
    </comment>
    </comments>
- <tags>
- <tag>
  <author>35474119@N03</author>
  <text>Colosseo</text>
    </tag>
- <tag>
  <author>35474119@N03</author>
  <text>Roma</text>
    </tag>
- <tag>
  <author>35474119@N03</author>
  <text>Italia</text>
    </tag>
- <tag>
  <author>35474119@N03</author>
  <text>Archeology</text>
    </tag>
- <tag>
  <author>35474119@N03</author>
  <text>Archeologia</text>
    </tag>
    </tags>
  <notes />
    </usermetadata>
    </item>
```

Sample Parsed Data

Here is an example of the same data from the sample data from script as parsed into a sortable and searchable spreadsheet.

Collection ID	Set ID	Photo ID	User ID	Tags	Comments	Date
30510347-72157616674562400	7.21576E+16	3611960232	42826854@N00		Hi, I'm an admin for a group called Heritage Key, and we'd love to have this added to the group!	7/28/2009 9:15
30510347-72157616674562400	7.21576E+16	3611960232	35474119@N03	Colosseo		9/15/2009 4:56
30510347-72157616674562400	7.21576E+16	3611960232	35474119@N03	Roma		9/15/2009 4:56
30510347-72157616674562400	7.21576E+16	3611960232	35474119@N03	Italia		9/15/2009 4:56
30510347-72157616674562400	7.21576E+16	3611960232	35474119@N03	Archeology		9/15/2009 4:56
30510347-72157616674562400	7.21576E+16	3611960232	35474119@N03	Archeologia		9/15/2009 4:56

BIBLIOGRAPHY

Achananuparp, P., McCain, K. W., & Allen, R. B. (2007). Supporting student collaboration for image indexing. In *Asian Digital Libraries. Looking Back 10 Years and Forging New Frontiers* (pp. 24-34). Springer Berlin Heidelberg.

Ames, M., & Naaman, M. (2007, April). Why we tag: motivations for annotation in mobile and online media. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 971-980). ACM.

Beaudoin, J. (2007). Folksonomies: Flickr image tagging: Patterns made visible. *Bulletin of the American Society for Information Science and Technology*, *34*(1), 26-29.

Becker, H. S. (2008). *Tricks of the trade: How to think about your research while you're doing it*. University of Chicago Press.

Bischoff, K., Firan, C. S., Nejdl, W., & Paiu, R. (2008, October). Can all tags be used for search?. In *Proceedings of the 17th ACM conference on Information and knowledge management* (pp. 193-202). ACM.

Collins, H. M., & Evans, R. (2002). The Third Wave of Science Studies Studies of Expertise and Experience. *Social studies of science*, *32*(2), 235-296.

Earle, E, Patzke, K. & Cheng, J. (2011). ManyTags – User Perception, Motivation and Behavior in Tagging Systems.

Eleta, I., & Golbeck, J. (2012, February). A study of multilingual social tagging of art images: cultural bridges and diversity. In *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work* (pp. 695-704). ACM.

Flickr (2013). Flickr: The Commons. From: http://www.flickr.com/commons

Flickr (2013). Flickr: The Library of Congress. From: http://www.flickr.com/people/library_of_congress

Flickr (2013). Flickr Services API Documentation. From: http://www.flickr.com/services/api

Glaser, B. G., & Strauss, A. L. (2009). *The discovery of grounded theory: Strategies for qualitative research*. Transaction Books.

Gerolimos, M. (2013). Tagging for Libraries: A Review of the Effectiveness of Tagging Systems for Library Catalogs. *Journal of Library Metadata*, *13*(1), 36-58.

Getty Research Institute (2013). Art & Architecture Thesaurus. From: http://www.getty.edu/research/tools/vocabularies/aat

Greene, M. A., & Meissner, D. (2005). More product, less process: Revamping traditional archival processing. *American Archivist*, *68*(2), 208-263.

Guy, M., & Tonkin, E. (2006). Tidying up tags. *D-lib Magazine*, *12*(1), 1082-9873.

Hardin, G. (1968). The tragedy of the commons. *New York*

Huberman, B. A., Romero, D. M., & Wu, F. (2009). Crowdsourcing, attention and productivity. *Journal of Information Science*, *35*(6), 758-765.

Jones, B. M. (2004). Hidden collections, scholarly barriers: Creating access to unprocessed special collections materials in America's research libraries. *RBM: A Journal of Rare Books, Manuscripts and Cultural Heritage*, *5*(2), 88-105.

Kennedy, L. S., Chang, S. F., & Kozintsev, I. V. (2006, October). To search or to label?: predicting the performance of search-based automatic image classifiers. In *Proceedings of the 8th ACM international workshop on Multimedia information retrieval* (pp. 249-258). ACM.

Kremerskothen, K. (August 4, 2011). 6,000,000,000. *Flickr Blog.* From: http://blog.flickr.net/en/2011/08/04/6000000000

Library of Congress (2013). Library of Congress Subject Headings. From: http://id.loc.gov/authorities/subjects.html

Library of Congress, Prints and Photographs Division, Springer, M., Dulabahn, B., Michel, P., Natanson, B., Reser, D. W., ... & Woodward, D. (2008, October). For the common good: The Library of Congress Flickr pilot project. Library of Congress, Prints and Photographs Division.

Macgregor, G., & McCulloch, E. (2006). Collaborative tagging as a knowledge organisation and resource discovery tool. *Library review*, *55*(5), 291-300.

Marlow, C., Naaman, M., Boyd, D., & Davis, M. (2006, August). HT06, tagging paper, taxonomy, Flickr, academic article, to read. In *Proceedings of the seventeenth conference on Hypertext and hypermedia* (pp. 31-40). ACM.

Maier, R., & Thalmann, S. (2008). Institutionalised collaborative tagging as an instrument for managing the maturing learning and knowledge resources. *International Journal of Technology Enhanced Learning*, *1*(1), 70-84.

Mayring, P. (2007, September). On generalization in qualitatively oriented research. In *Forum Qualitative Sozialforschung/Forum: Qualitative Social Research* (Vol. 8, No. 3).
McElroy, Steven, "Mozart Score Is Found," *The New York Times*, September 19, 2008

Merriam-Webster Inc. (2004). *Merriam-Webster's collegiate dictionary*. Merriam-Webster.

Moltedo, C., Astudillo, H., & Mendoza, M. (2012, October). Tagging tagged images: on the impact of existing annotations on image tagging. In *Proceedings of the ACM multimedia 2012 workshop on Crowdsourcing for multimedia* (pp. 3-8). ACM.

Momeni, E., Tao, K., Haslhofer, B., & Houben, G. J. (2013, July). Identification of useful user comments in social media: A case study on Flickr commons. In *Proceedings of the 13th ACM/IEEE-CS joint conference on Digital libraries* (pp. 1-10). ACM.

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

O'Steen B. (December 12, 2013). A Million First Steps. *British Library Digital Scholarship Blog.* From: http://britishlibrary.typepad.co.uk/digital-scholarship/2013/12/a-million-first-steps.html

Panitch, J. M. (2001). *Special collections in ARL libraries: Results of the 1998 survey sponsored by the ARL Research Collections Committee*. Association of Research Libr.

Powerhouse Museum (2013) Powerhouse Museum Collection. From: http://www.powerhousemuseum.com/collection/database

Preece, J., & Shneiderman, B. (2009). The reader-to-leader framework: Motivating technology-mediated social participation. *AIS Transactions on Human-Computer Interaction*, *1*(1), 13-32.

Prochaska, A. (2003). Special collections in an international perspective. *Library trends*, *52*(1), 138-150.

Quinn, A. J., & Bederson, B. B. (2011, May). Human computation: a survey and taxonomy of a growing field. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 1403-1412). ACM.

Raymond, M. (January 16, 2008). My Friend Flickr. *Library of Congress Blog*. From: http://blogs.loc.gov/loc/2008/01/my-friend-flickr-a-match-made-in-photo-heaven

Reti, L. (1968). The Two Unpublished Manuscripts of Leonardo Da Vinci in the Biblioteca Nacional of Madrid-I. *The Burlington Magazine*, *110*(778), 10-2.

Riddle, P. (2005). Tags: What are they good for. *School of Information. University of Texas*.

Shirky, C. (2005). Ontology is overrated: Categories, links, and tags. *Clay Shirky's Writings About the Internet*.

Sigurbjörnsson, B., & Van Zwol, R. (2008, April). Flickr tag recommendation based on collective knowledge. In *Proceedings of the 17th international conference on World Wide Web* (pp. 327-336). ACM.

Spiteri, L. F. (2013). The structure and form of folksonomy tags: The road to the public library catalog. *Information technology and libraries*, *26*(3), 13-25.

Stvilia, B., & Jörgensen, C. (2007). End-user collection building behavior in Flickr. *Proceedings of the American Society for Information Science and Technology*, 44(1), 1-20.

Trant, J., & with the participants in the steve. museum project. (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.

Trant, J., & Wyman, B. (2006, May). Investigating social tagging and folksonomy in art museums with steve. museum. In *Proceedings of the WWW'06 Collaborative Web Tagging Workshop*.

Von Ahn, L., & Dabbish, L. (2004, April). Labeling images with a computer game. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 319-326). ACM.

Vander Wal, T. (February 21, 2005). Explaining and Showing Broad and Narrow Folksonomies. From: http://www.vanderwal.net/random/entrysel.php?blog=1635

Vander Wal, T. (February 2, 2007). Folksonomy Coinage and Definition. From: http://vanderwal.net/folksonomy.html

Smith-Yoshimura, K., & Holley, R. (2012). Social metadata for libraries, archives, and museums. Part 3: Recommendations and readings. Dublin, Ohio: OCLC Research.

Wang, M., Ni, B., Hua, X. S., & Chua, T. S. (2012). Assistive tagging: A survey of multimedia tagging with human-computer joint exploration. *ACM Computing Surveys (CSUR)*, *44*(4), 25.

Weinberger, K. Q., Slaney, M., & Van Zwol, R. (2008, October). Resolving tag ambiguity. In *Proceedings of the 16th ACM international conference on Multimedia* (pp. 111-120). ACM.