Critical Literature Review

Michelle Caples

University of Wisconsin-Milwaukee

INFOST 571G-202 Information Access and Retrieval

Abstract

Technology has influenced the way users access information retrieval. Now more than ever with multimedia information retrieval, users are retrieving information through visual representation. Curating collections based on research about user knowledge and how the user knowledge impacts the way users search for information, leads one to the phenomenon of digital images information retrieval. Considering digital images as thematical finding aids or “emotion” finding aids is fraught with determining how to process these finding aids. Does a digital image evoke the same emotion from users? Are linguistically descriptive finding aids tools of the past? Should archivists automatically create digital tagging finding aids? Words still matter, so there should always be a commitment to creating finding aids with imaginative and descriptive information. At the same time, digital images that evoke emotions impact the user experience for an overall balanced user retrieval experience. An image’s value is based on how the user comprehends the image, through their sense of current knowledge (or lack of) about the image before the information retrieved can be of use to the user. Developing metadata for multimedia information retrieval that is intuitive, and fluid can enhance digital finding aids. Finding aids based on algorithms is one way to achieve this goal, because users are participating in providing clues as to how people perceive images in finding aids. The use of cloud tagging is a tool in which the vocabulary for a digital image expands through the words provided by users to describe an image. Users can help curate a digital image which helps benefit the overall user community at large.

Keywords: multimedia information retrieval, digital image, finding aids, user knowledge, cloud tagging.

Critical Literature Review

# Introduction

This review will examine three articles that look at the ways digital images impact information retrieval for users and the importance of digital finding aids to information users. The goal of this examination is to provide a solid foundation of what digital images are and the importance of descriptive language as a necessity for digital images. A critical analysis of the articles and the authors’ findings is crucial to determine if authors insights and assessments are correct or miss the mark. The motivation for this critical review is to stress the importance for archivists and librarians to bring energy to providing creative ways of indexing digital metadata, which includes the use of algorithms and metadata tagging. The objective of this critical review is to stress digital images do not exist without supportive language, otherwise, the image fails to evoke meaningful emotions for the user.

## Significance

The discussion about digital images is significant because as more archivists and librarians are creating or working with digital records, collections, or content (social media) it is crucial to understand how the user searches for this type of content. It is also essential to understand how much digital images impact users daily and that the language they use to retrieve digital images is helpful with the development of metadata for finding aids. Finally, due to virtual assistants like Suri or Alexa, users get their questions answered with *minimal* input from the user, so it is important to remember users may expect information retrieval to work in the same manner.

The article The Effects of Background Information and Social Interaction on Image Tagging (Bar‐Ilan et al., 2010) focuses on an experiment that looked at the results of social interaction and background information on image tagging. The authors concluded that if an individual lacked knowledge about an image, with the use of social interaction, the social community would provide the individual with a context of what the image was about:

When viewing the image only, in some cases the users were not able to correctly identify what they saw in some of the pictures, but they overcame the initial difficulties after interaction. We conclude from this experiment that social interaction may lead to convergence in tagging and that the “wisdom of the crowds” helps overcome the difficulties due to the lack of information. (Bar‐Ilan et al., 1993, p.940)

Social interaction in the example above is not people communicating via Facebook or Twitter, rather in this instance, social interaction is the use of “tagging” to communicate information about a tagged image. Tagging is providing language and information to a user who was not able to identify the image based on their own knowledge. Tagging also helps build metadata behind the scenes which helps improve finding aids.

How did the authors gather this information? They conducted an experiment with 130 students form the Department of Information of Science of Bar-llan with one focus in mind:

In this experiment, our aim was to understand the effect of

different tagging modes on the resulting tag sets of images.  
The participating users were assigned to one of three groups  
depending on the type of background information available to  
them: group A had no additional information, just the image  
itself; group B saw the image and the title; while group C  
saw the image, the title, and a link to the page where the  
image appeared. The users of all three groups were asked  
to tag 12 images in the domain of Jewish cultural heritage (p.942)

The authors also wanted the experiment to answer the question of if user produced tags would provide reliable retrieval information in the future. The authors also were curious about group tagging negating individual tagging during the experiment. For example, if a person looked at an image of Cartier watch, they may just assign watch as a tag. However, a person familiar with luxury watches may assign the tags, luxury, luxury watch, tank, expensive, or the future information retrieval can be found in the following summation: The concept of the “wisdom of the crowds” as explained by Surowiecki (2004) is based on the assumption that the aggregated information leads to decisions that are superior to decisions made by any single member of the group could have made on his or her own. (p. 942)

The description of the experiment is very clear about three groups participating in the experiment, but you have to read closely to realize the experiment was also dependent on individual tagging versus group participation tagging, and the latter included a maximum of three people in a group, after a round of individual tagging. The two phases of tagging are individual and group. Once in the group setting, individuals could see their own tags and the group tags. The goal of fostering group conversation led to editing or deletion of individual tags and the creation of new individual tags. Each stage of the experiment lasted two weeks. An interesting side note- the longer the list unique tags, the less consensus among the group about the tags. The article mentions the following:

We define the consensus score of each tag as percentage

of users in the group who assigned the tag for the

specific image and specific interaction mode. For each image,

we compute the average score of all the tags that were

assigned at least with multiplicity two. Higher average

consensus score for a tag set reflects higher agreement

between participants regarding the overall relevance of

the tag set. In terms of image retrieval, this criterion

might improve the precision (p.943)

The complexity of user behavior and information need is often a complication of information retrieval; however, this article shows a way to solve this challenge. The fruition of the experiment is that proves its proposed hypothesis. User information from individuals and groups generated a tag cloud (illustrated in the article) with the most popular terms. If an information retrieval is a system that facilitates the matching of an information request with the information stored. This thought becomes a tangible process in the collection of metadata that creates tagging that is individual and interactive at the same time and contributes to elevating the information retrieval process. Social tagging aspects reflected in this article include content, context, attribute, subjective, and organizational tags, which are all helpful, but keep in mind users may not be experts in the subjects they are providing tags for.

In the article Analysis of Image Search Queries on the Web: Query Modification Patterns and Semantic Attributes (Choi, 2013) focuses on the following:

This study investigated query modification patterns and

semantic attributes in queries executed during user searches for images on the web. Its purpose was to identify whether query modification patterns were related to users’ contextual factors and content sources as well as whether the patterns characterize the use of semantic attributes expressed in users’ search queries in an interactive web-searching process (p.1423)

The study in the article is based on 978 image search queries by twenty-nine college students in three different sessions. The focus is primarily on semantics and how users use semantics. The study mentions average users typically spend more time searching images than other types of searches. This research concludes that image searches are more complex than other searches. Briefly, users tended to start image searches based on their own knowledge and contextual information. If the approach did not yield results, the user refines their search and uses terms that that depict the image as much as possible. The participants in the query were a small group (college media students) because the author wanted to reflect a limited scope. On of the main terms focused on in this study was the process of query remodification:

Continuation: A query on a topic searched for earlier by a

user; one or more keywords appeared in earlier queries, but

not in the last query.

• Generalization: A query on the same topic as the previous

query, but seeking more general information than the previous

query; one or more keywords have been deleted from

the query.

• Specialization: A query on the same topic as the previous

query, but seeking more specific information than the previous

query; one or more keywords have been added to the query.

• Reformulation: A query on the same topic that can be viewed

as neither a generalization nor a specialization, but a refor-

mulation of the prior query, including the following cases:

• Adding a preposition or an article (such as “the”) classified

as Reformulation as opposed to Specialization even if it

adds a word.

• Changing a word from acronym to full name or from full

name to acronym.

• Changing the order of words.

• Changing spelling and/or verb tense.

• Changing words from singular to plural or plural to singular.

• Changing words in a different language.

• Request: A request for another set of results on the same

query from the search service. Duplicate queries appear in the

data when a person requests another set of results for the

query. Removal of a quotation mark was classified as Request

rather than Reformulation. (p.1427)

The process above is the opposite of aspects of visual information retrieval. Plutchik’s wheel of emotions provides adjectives that are good search descriptors. However, this article focuses on the challenges of multimedia information retrieval, which include query types, feature extraction and representation, indexing and retrieval, and evaluation. All these things in a search engine must be at peak performance for the user to gain any benefits and this also depends on the user’s knowledge.

An interesting aspect of this study is that it is based on image search engines and general search engines. The order of continuation, generalization, specialization, and reformulation could change based on if the user were using the world wide web, or image search engines. Users would modify their queries in image search engines while constructing new search entries on general search engines. While this information is useful, the validity of the search is based on the user’s knowledge prior to starting an image search. This article also shows how natural language can get in the way of the information retrieval process. Contextual ambiguity -terms retrieved that do not relate, semantic ambiguity, words spelled the same but do not mean the same thing, and lack of hierarchal relationships, where terms are too broad, really come into play with this article and image searches.

The study makes the astute observation:

It further suggests that a conceptual facet-based browsing option related to a query and search results would be useful, allowing users to interact with search results, rather than requiring the user to constantly compose a query. As available images on the web increase, the level of sophistication required for users to perform effective searches is likely to increase correspondingly. Thus, developing intuitive and intelligent search assistance for web image searching is essential. (p.1433)

Creating intuitive, responsive, and intelligent search assistance or finding aids is essential to ensuring users are not stuck doing futile searches, which are compelling reasons to create finding aids that meet the needs of high or low knowledge users.

Taking into consideration that most users want their information needs satisfied instantaneously, again think Alexa or Suri, digital images create a special informational retrieval need that must sate the user so that the user feels they can put forth the least amount of search effort and still yield a positive search result. The user will feel satisfied and will usually end their search on the spot. This article does acknowledge that the user typically starts off their search with cursory knowledge of what the image is about. A weak point of this article is it does not seem to research image information seekers with no concept of what the image stands for. The author does defend this action by stating more research in the future should include diverse types of digital information retrievers. The overall goal of the research was to acknowledge information retrieval studies can vary between general search engines and image search engines through using natural language. Keeping this in mind, finding aids will be in a constant state of evolution and word clouds and tagging will be on the front lines of refining and improving finding aids when it comes to digital images. Finding aids that can self service itself through technology is also a tremendous help to archivists and librarians.

Images are a prevalent part of a user’s daily life because users retrieve information from various sources such as phones, television, computers, smartwatches, or smart phones. The article Understanding Image Needs in Daily Life By Analyzing Questions in a Social Q&A Site (Reddit, Quora, or Yahoo, etc.) (Yoon, 2011) is a deep dive into a study of this form of information retrieval. Yoon states, “The purpose of this study is to enhance understandings of ordinary searchers’ image needs in their daily life, so that the findings provide practical suggestions for image indexing, image retrieval systems, and services for ordinary image searchers.” (p.2202) Yoon details a study that includes 474 answers from a social question and answer site. The study provided the following insights:

The study found that image needs reflected through the natural language questions contain several components: context of image needs (motive and intervening variables), image attributes (descriptive metadata, syntactic, and semantic attributes), and associated information (information on known/similar/comparative images and related stories). Characteristics of each component of image needs were analyzed, and accordingly image-indexing guidelines were suggested. Because image needs comprise diverse attributes, a single indexing approach might not support all complex needs for images. Therefore, this study proposes that different indexing approaches should be integrated for enhancing keyword search and browsing effectiveness. Such approaches include descriptive metadata assigned by a creator and/or automatic algorithms, user-assigned tags (or users’ reactions), indexing through associated text, and content-based image (p.2201)

This paper has touched on user tags and algorithms, but not descriptive metadata. What is descriptive metadata? Descriptive metadata is “descriptive attributes about an entity.” (Learn With Me, 2022, 00:06:13) Many finding aids live and die by descriptive attributes and this is where creating metadata that covers the basic attributes of an entity and thinks outside of the outside of the box by seeking natural language that one wouldn’t think about at first thought, but is a descriptive attribute nonetheless. Working with gamers to create metadata about a game would be a way to ensure that the metadata is using the current language that games use and create user tags and help generate an algorithm. Yoon’s study concludes by noting that most users tend to use a Q &A website when they cannot retrieve information through a general search engine. Yoon also stated the data may be slightly skewed because the image may have been an image not easily retrieved in a general search engine search. Nevertheless, Yoon does produce interesting findings image retrieval and associated or similar images and the need to consider how to elevate metadata, algorithms, to generate user tags that will the demands of this type of information retrieval. However, I am not sure if this study brings to light any new revelations about image retrieval that the other two articles mentioned in this paper cover. Yoon assumes users go to Q&A websites when they cannot find image information in a general search engine. Yet, we cannot assume that a user may go to a Q &A website first to get a community answer about an image, especially if the user is unfamiliar with the image, then go to the general search engine with the knowledge gained from the Q &A site. One must consider there are users who are more comfortable with getting a community consensus on a topic they have little knowledge about, even if the information from the Q &A site is not accurate. Yoon must remember users consume information daily that is misinformation fueled by inaccurate metadata, algorithms, and tagging. The user must be vigilant about where their sources come from as Q&A sites can provide information based on conspiracy theories, political or religious affiliations, racism, malicious intent, or out of date information. Tools used to create information retrieval are good or bad depending on who owns the tools.

The overall theme of each article is that digital images require more effort than traditional subject indexing. Digital imaging really involves the archivist, the librarian, and the user, who all may not know they are working together behind the scenes. Finding aids comprise of descriptive vocabularies that users can influence by the natural language they use when using a search engine to retrieve information about an image. Natural language uses words that are common and uncommon but serve to build a repository of information about images or multimedia information. Natural language is not lazy nor ambiguous and helps generate algorithms and user tags that benefit the information retrieval community. Metadata for images can become a never-ending story that constantly provides new chapters in the form of image description. However, the user must have critical thinking skills and the ability to interpret the information provided about an image. The user does not necessarily have to know anything about an image they are researching, but upon retrieval of the information, the user must have a way to interpret the information based on the local or world knowledge they possess. The world that contains Q &A websites, Suri and Alexa, is not honing a user’s critical thinking skills.

Yet, there are still ways archivists, librarians, and libraries can make the information retrieval process user friendly. Query suggestions, relevance ranking, and aesthetically pleasing and easy to navigate user interface are a solution to helping users get the information they are seeking, while also providing the user with a sense of confidence in their searching abilities. Designing interfaces based on user preference and behaviors is a terrific way to access data quickly and easily. Tailoring results based on a user’s profile and predicting what a user is looking for are all ways to encourage users and help them improve their searching and retrieval skills. Computers have the AI capability to do these tasks and even connect users to other users if they choose to do so.

This following image from Yoon’s article (p.2211) is a summarization of how retrieval should flow. A diagram of a diagram

Description automatically generated

We can see the searchers motive, and how intervening variables help the searcher achieve the goal of finding the image or information about the image they want. The ability to provide a known image, a similar image, and a comparative image is terrific because the user has options and may find the similar or comparative image to be the information the user was seeking. The user learns descriptive metadata and a related story. The user’s local or world knowledge expands through this information retrieval process.

In conclusion I would like to see future research about image date retrieval look at how ChatGPT factors into image information retrieval. Should libraries begin to pattern their search engines to mimic the chatbot and virtual assistant? As images become a larger part of users lives, I would like to see more articles about the competition between natural language and images. The articles cited in this paper agree that the information library sciences discipline is lagging in meeting the demands of image information retrieval. The articles are seeking a call to action to remain creative and vigilant about the need to foster positive user interactions, though doing so behind the scenes through technology.

Metadata must be basic words and encompassing phrases at the same time to describe imagery. Metadate cannot be a static description that will be relevant ten days or ten years from its creation. Technology is moving too fast for that type of thinking. Information retrieval must be a gateway to future users who rely on those in the present to create finding aids that will help them understand an image from the past. These articles had flawed research methods at times, but did provide decent things to consider and directions as to how to navigate future information retrieval systems.

As for the present time, I would like to have libraries reach out to niche segments of the community like gamers, pickle ball enthusiasts, collectors, to research how they respond to images of their interests and how they communicate with their community. This is the way to do good research in the future for metadata, social tagging, and creating engaging finding aids. Each article did support user community involvement, and this would be a wonderful way to access their needs and make them think about visual representation of their community. A What type of image can they provide to the future user that would give them a language and a history.

# Bibliography

Bar-Ilan, J. Z.-G. (2010). The Effects of Background Information and Social Interaction on Image Tagging. *JOURNAL OF THE AMERICAN SOCIETY FOR INFORMATION SCIENCE AND TECHNOLOGY, 61(5):*, 940–951.

Choi, Y. (2013). Analysis of Image Search Queries on the Web: Query Modification Patterns and Semantic Attributes. *JOURNAL OF THE AMERICAN SOCIETY FOR INFORMATION SCIENCE AND TECHNOLOGY, 64(7):*, 423–1441.

Learn with Me, March 14, 2022 **Types of Metadata | Descriptive | Structural | Administrative www.bing.com/videos/riverview/relatedvideo?&q=descriptive+metadata+examples&&mid=00854DAE85F16DB985C000854DAE85F16DB985C0&mmscn=mtsc&aps=37&FORM=VRDGAR**

Yoon, J. C. (2011). Understanding Image Needs in Daily Life By Analyzing Questions in a Social Q&A Site. *JOURNAL OF THE AMERICAN SOCIETY FOR INFORMATION SCIENCE AND TECHNOLOGY—62(11)*, 201–2213.

**www.bing.com/videos/riverview/relatedvideo?&q=descriptive+metadata+examples&&mid=00854DAE85F16DB985C000854DAE85F16DB985C0&mmscn=mtsc&aps=37&FORM=VRDGAR**