Information Processing & Management, In Press, 2016

Finding Cultural Heritage Images through a Dual-Perspective Navigation Framework

Yi-ling Lin

National Sun Yat-Sen University

Peter Brusilovsky and Daqing He

Univerisity of Pittsburgh

Author Note

Yi-ling Lin, Department of Information Management, National Sun Yat-Sen University; Peter Brusilovsky, School of Information Sciences, University of Pittsburgh; Daqing He, School of Information Sciences, University of Pittsburgh.

The work presented in this paper was performed when the first author was a PhD student at the School of Information Sciences, University of Pittsburgh. This research was partially supported by the National Science Foundation under Grant 0840597. The authors would like to acknowledge our editor, Brianne J. Evans and Jon Walker, for their proofreading and formatting assistance.

Correspondence concerning this article should be addressed to Yi-ling Lin, Department of Information Management, National Sun Yat-Sen University, Taiwan. No. 70, Lienhai Rd., Kaohsiung 80424, Taiwan, R.O.C. Contact: yllin@mis.nsysu.edu.tw

Information Processing & Management, In Press, 2016

Abstract

With the increasing volume of digital images, improving techniques for *findability* of images is garnering heightened attention in recent times. The cultural heritage sector, with its vast resource of images, has realized the value of social tags and proposed the integration of folksonomies and controlled vocabularies for increasing the odds of users in finding items of interest. The main goal of this paper is to develop the Dual-Perspective Navigation Framework (DPNF) to integrate controlled vocabularies and social tags to represent more comprehensive aboutness of an item, which can maximize the information scent to facilitate resource finding without changing their nature or forcing users to choose one means over the other.

DPNF utilizes the mechanism of faceted browsing and tag-based navigation to offer a seamless interaction between experts' subject headings and public tags during the image searching experience. In a controlled user study, participants effectively completed more exploratory tasks with the DPNF interface than with the tag-only interface. DPNF is more efficient than both single descriptor interfaces (subject heading-only and tag-only interfaces). Participants spent significantly less time, fewer interface interactions, and less back tracking to complete an exploratory task without an extra workload. In addition, participants were more satisfied with the DPNF interface than with the others. The findings of this study can assist interface designers struggling with what information is most helpful to users and facilitate searching tasks. It also maximizes the end users' chances of finding target images by engaging image information from two sources, the professionals describing items in a collection and the crowd assigning social tags from their own classification and access needs.

Keywords: social tags, subject headings, indexing, navigation, interface, and user study

Information Processing & Management, In Press, 2016

Finding Cultural Heritage Images through a Dual-Perspective Navigation Framework

1. Introduction

Image search has been an important problem in the area of information access. Over the years, museums, news archives, and other key stakeholders established and perfected two ways of helping users to find relevant images – keyword-based search and metadata-based search. Yet, the rapid rise of online sharing of digital images challenges both approaches. While keyword-based search is still popular as demonstrated by Google image search[1], extracting relevant keywords to describe an image has become increasingly more problematic since many images are now published online without any textual descriptions.

Classical image metadata (known as subject headings in the museum context) supports both search and browsing (e.g., faceted browsing) but requires significant manual generation effort that is a challenge for large-scale image collections. It is still difficult for automatically generated metadata to match the quality of that created by professional indexers. At the same time, professionally generated metadata suffers from the classic *indexer-user mismatch problem*: non-professional end-users usually perceive items in a different manner than professional indexers. As a result, it has become increasingly difficult for the majority of end-users to find even properly processed and indexed images. The growing volume of content combined with the pressures of time, money, and competition means that the need to improve techniques for *findability* (Morville, 2005) of images is now becoming a critical issue.

In this context, social tagging has emerged as an alternative crowd-powered mechanism to generate textual descriptors that bring out the *aboutness* of the images so that effective and efficient browsing and keyword-based access to images can be achieved. *Aboutness* indicates the

[1] https://www.google.com/imghp?hl=en&tab=wi

Information Processing & Management, In Press, 2016

subject or topic that an item refers to from a user's perspective (Fairthorne, 1969). Different users with different perspectives may generate different descriptors for *aboutness*. In this study, we use aboutness to capture the main concept(s) expressed in an image. It can be the most significant characteristics of the image, such as the theme, the main character(s), the obvious attributes, and so on.

The diversity of input sources and the engagement of end users in the process of image description give modern social tagging systems tremendous power to assist users in finding images. Yet, tag-based access has its own problems, such as the lack of structure, semantic ambiguity, and wrong assignment, which may decrease the accuracy of the aboutness represented by social tags.

It's easy to see that components of both the metadata-based and tag-based approaches to image access have their own merits, which is why our paper develops a hybrid approach, Dual-Perspective Navigation Framework (DPNF), that includes both experts' and general users' descriptors to represent more comprehensive *aboutness* of an item. We argue that this hybrid approach can increase the completeness of the *aboutness* from diverse points of view and enhance the item's *findability*.

The cultural heritage sector, with its vast resource of images, has realized the value of social tags and proposed the integration of folksonomies and controlled vocabularies for increasing the odds of users finding items of interest (Hayman & Lothian, 2007; Rolla, 2009; Steele, 2009). However, the majority of the research on integrating experts' annotations and social tags has focused on how to utilize controlled vocabularies to structure folksonomies, which are taxonomies created by multiple users (Peters, 2009). A smaller thread, known as the multiple interface approach (McGrenere, Baecker, & Booth, 2002), explored the idea of using

Information Processing & Management, In Press, 2016

both professional index terms and social tags independently, but in parallel with one another, by offering multiple types of navigational support for users' various information needs.

By contrast, the approach presented in this paper focuses on a true integration of these two *aboutness* descriptors to facilitate resource finding. That is, tags and metadata are integrated in our approach in such way that they reinforce their strengths without changing their nature or forcing users to choose one means over the other. Our DPNF approach is an interface-level integration of tag-based and metadata-based information access mechanisms.

DPNF utilizes the mechanism of faceted browsing and tag-based navigation in its design of seamless interaction between experts' subject headings and public tags to maximize information scent (Pirolli, 2007) and facilitate the image search. Users are able to start their search via a traditional keyword query, a specific subject heading, or a tag, then progressively narrow down the search results using both subject headings and tags. The presence of hierarchical facets of subject headings and a flat cloud of tags allows users to search with more flexibility and to specify their interests more precisely using the structure of different types of information descriptors. To assess whether DPNF does, in fact, support efficient and effective user-oriented image finding, we performed a controlled user study, which is reported in this paper.

The remainder of the paper is organized as follows: Section 2 reviews the related work, including the applied background theories and relevant research in the area of image finding; Section 3 introduces our approach -- the Dual-Perspective Navigation Framework (DPNF); Section 4 demonstrates our research process including our data collection, system design, and experimental design, the research tasks involved, and the procedure of the user study; Section 5

Information Processing & Management, In Press, 2016

discusses research results; and Section 6 offers a discussion with a further analysis. Section 7 concludes the important findings of this research.

2. Related Work

2.1 Background theories

Aboutness of a document is a term coined by R. A. Fairthorne (1969), which indicates the subject or topic described in the document, and can be expressed by assigned or extracted index descriptors. Aboutness is known to be hard to capture accurately (Hjørland 1992), and Maron and Studies (1977) provided an operational definition in relation to search behavior: S-about (subjective aboutness), O-about (objective aboutness), and R-about (generalized objective aboutness for a specific community). In the image finding domain, several studies (Armitage & Enser, 1997; Hollink, Schreiber, Wielinga, & Worring, 2004; Jörgensen, 1998) have explored the importance of image descriptions from different perspectives. Our DPNF model is situated within the broader context of the aboutness paradigm. It suggests that both experts' and general users' descriptions can provide the dual-perspectives of aboutness, which can then be integrated into an interface to support image finding. Although information retrieval based on the aboutness of a document may use the title, description, keywords, headings, links, and various other information sources, the images we focus on lack this type of textual content; therefore, we focus on investigating the aboutness of links to guide a searcher to the target item in an efficient and effective manner.

Information scent (Pirolli, 2007) plays an important role in a user's ability to find desired items. A strong information scent can quickly lead a user to the target information, whereas a weak information scent causes a user to spend more time evaluating the options. Nielson (2003) presented the concept of information scent as cues and proposed guidelines on increasing

Information Processing & Management, In Press, 2016

information scent and fostering faster interaction. Spool, Perfetti, and Britten (2004) investigated the scent-following behavior of Web users, and argued that information foraging (Pirolli, 2007) is useful as a base for providing guidelines for Web design. They found that users searched for a scent trail and followed it toward their content. When they lost the scent, they backtracked until they picked it up again. Our dual-perspective navigation framework has been motivated and informed by the past research on the information scent. In particular, the idea to integrate seamlessly two types of information descriptors (subject headings and social tags) stemmed from the need to increase information scent at each step of a user's search for target items. We believe that providing both experts' and general users' interpretations of information items increases the information scent and therefore provides better support to users.

2.2 Subject headings and tags

Subject headings are controlled vocabulary for capturing the *aboutness* of a subject. Although subject headings provide many benefits and opportunities for search or browsing, index terms created by applying controlled vocabularies have shown limited adequacy for online resource discovery (Macgregor & Mcculloch, 2006). Professional index terms limit users' expression of their needs to a set of professional terms that they might not understand, hindering them from generating efficient search queries (Furner, 2007).

Social tags may provide alternative search and navigational support. Several studies have proved that social tags can be used to improve search performance and support users to access information more efficiently (Bischoff, Firan, Nejdl, & Paiu, 2008; Kammerer, Nairn, Pirolli, & Chi, 2009). However, the freedom of annotation could make tag descriptions noisy, shallow, ambiguous, inconsistent, and sparse (Peters, 2006; Plangprasopchok, Lerman, & Getoor, 2010).

Information Processing & Management, In Press, 2016

Many researchers have realized the value of social tags and proposed the integration of folksonomies with controlled vocabularies. Rolla (2009) pointed out that tags could be used to enhance subject access but not replace controlled vocabularies like Library of Congress Subject Headings (LCSH). Lee and Schleyer (2012) also found that social tagging and controlled indexing couldn't be used as substitutes for each other since they represent different viewpoints/processes in the medical domain. Most proposed solutions for integrating social tags and professional index terms focused on how to utilize controlled vocabularies to structure folksonomies from different points of view (Koraljka, Lykke, & Tudhope, 2014; Syn & Spring, 2013; Wu, He, Qiu, Lin, & Liu, 2012), algorithms (Begelman, Keller, & Smadja, 2006; Hotho, J, Schmitz, & Stumme, 2006), feature combinations such as query expansion with WordNet (Laniado, Eynard, Colombetti, & Milano, 2007), and semantic relations between metadata and social tags (Al-Khalifa & Davis, 2007; Nowak, Liorente, & Motta, 2010). These studies did not integrate each of the two sources of information; they merely added terms from one into the other without considering the inherent differences between them. Another thread of studies applied multiple interface approach (McGrenere et al., 2002) to integrate subject headings and tags. Users can choose to start with either a subject heading interface or a tag interface when seeking relevant information, then switch to the other interface through a specific link; however, navigation between the two interfaces is not connected.

Our DPNF integrates subject headings and social tags with a more complete description of the items in the collection, it provides multiple accesses points for users, and it freely supports cross-navigation between subject headings and social tags.

Information Processing & Management, In Press, 2016

2.3 Faceted browsing and tag-based navigation

In interactive image search, information seeking tasks involve various levels of *exploration* in line with users' varying contexts (Liu, Little, & Ruger, 2011). White and Marchionini (2007) pointed out that in the aspect of finding an optimal path to an information resource, *exploratory search* is related to information foraging theory (Pirolli, 2007). While exploratory search is difficult to define, it is generally thought of as a search that requires both querying and browsing strategies in order to meet a user's information goals (Marchionini, 2006; White & Roth, 2009). Consequently, many exploratory search technologies focused on supporting navigation or browsing as part of the search process. Several methods integrating navigation structure with the keyword search have been proposed, including category systems (flat, hierarchical, and faceted), TOC views, and automated clustering techniques. Of particular interest in this paper are faceted browsing and tag-based navigation support.

Information Processing & Management, In Press, 2016

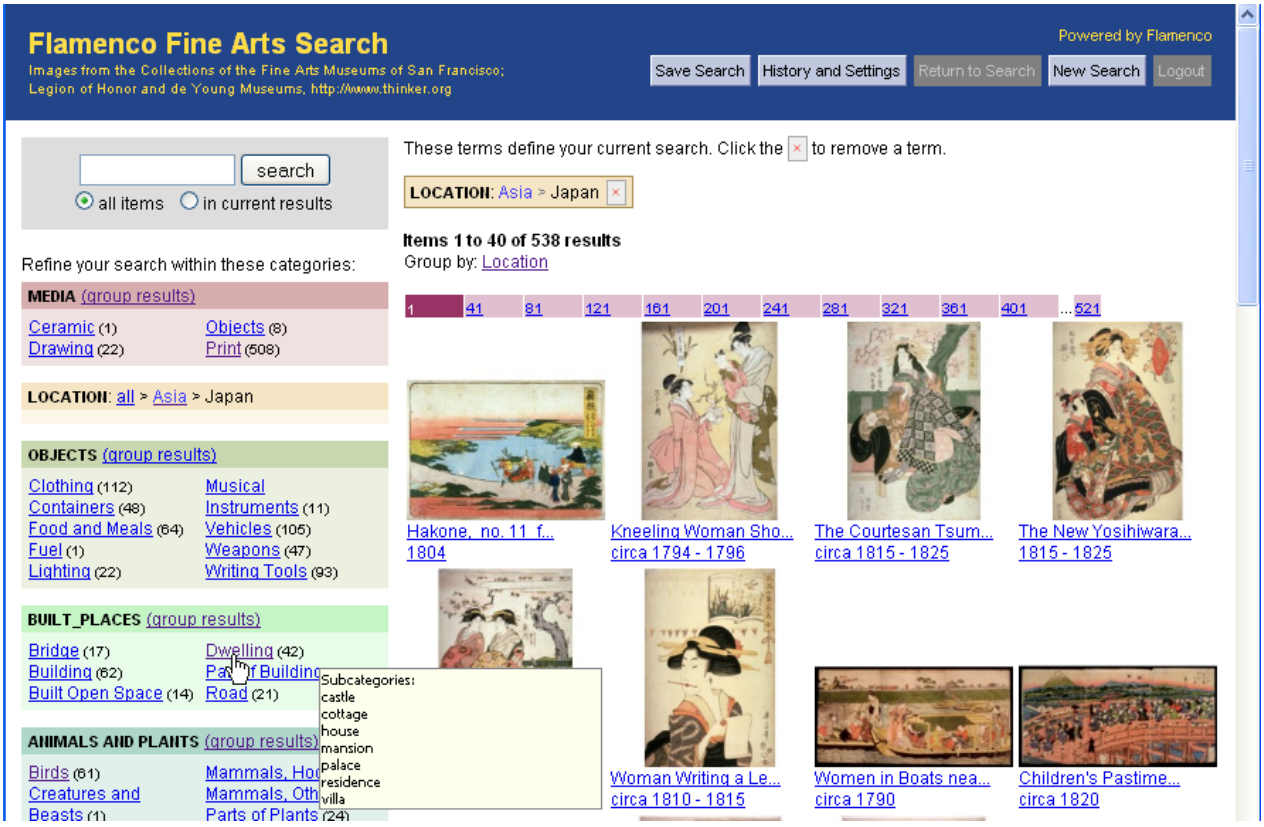

Figure 1. The Flamenco Interface

Faceted browsing is a popular exploratory search approach that provides an attractive alternative to "text box" search in situations when item metadata are available (Hearst, 2006; Karlson, Robertson, Robbins, Czerwinski, & Smith, 2006; Marchionini & Brunk, 2003; Yee, Swearingen, Li, & Hearst, 2003). In contrast to hierarchical browsing in which users navigate a single extensive hierarchy to narrow their choices, faceted browsing integrates browsing with the classification of objects along several dimensions called facets. With faceted browsing, users progressively narrow down the list of results, making choices in several taxonomies that classify different aspects of the objects of interest. The presence of these multiple facets allow the users to search more flexibly and to specify their interests more precisely than one dimension of classification. To further guide the users' choices and help them make sense of results, modern

Information Processing & Management, In Press, 2016

faceted browsing interfaces such as Flamenco [Figure 1] (Hearst, 2006; Yee et al., 2003) or Relation Browser (Capra & Marchionini, 2008; Marchionini & Brunk, 2003) display query previews that show the number of documents available for every facet category.

Faceted browsing interfaces have been shown to be helpful and preferred by users over the traditional search interface (Yee et al., 2003). However, its application in its standard form was limited to domains where objects of interest are classified along several dimensions of metadata (e.g., price, year, brand, and other object-specific aspects). Thus, classic faceted search cannot replace traditional search in domains where multiple classification facets are not established, or where the objects are not classified along multiple facets.

Social tags are different from classic metadata used for navigation. There are no parent-child structures, no hierarchies, no relationships between tags, and usually no categories or facets (Smith, 2008). Many social websites such as Flickr, Delicious, and others have the distinct advantage of adopting tags to support users' Web browsing and navigation. Figure 2 shows the results of a search on the tag "flower" from the Flickr website. In this view, the most recent pictures are shown and the related tags are shown beside the result set. A cluster function is provided as well. By selecting the function, users can view the results in different groups including nature, macro, color, and types.

A large volume of research has investigated diverse tag artifacts for information access such as tag clouds (Bateman, Gutwin, & Nacenta, 2008; Rivadeneira, Gruen, Muller, & Millen, 2007; Venetis, Koutrika, & Garcia-Molina, 2011), clustered/classified tag clouds (Hassan-Montero & Herrero-Solana, 2006; Knautz, Soubusta, & Stock, 2010; Zubiaga, García-Plaza, Fresno, & Martínez, 2009), and tag hierarchies (Candan, Di Caro, & Sapino, 2008; Helic & Strohmaier, 2011; Trattner, Körner, & Helic, 2011). Within this stream of work, our previous

Information Processing & Management, In Press, 2016

study (Trattner, Lin, Parra, & Brusilovsky, 2012) compared a traditional tag-cloud interface and a more advanced faceted tag-cloud interface in a controlled user study that included both lookup search and exploratory search tasks. From the users' perspective, both simple and faceted tag-based interfaces enhanced support for both types of search tasks and gave users higher confidence that they would find information that is more relevant. Yet, performance and log analysis demonstrated that only traditional tag-cloud interface offered significant improvement over the baseline search-only interface, in terms of time and actions. This result was critical for us in selecting the tag-based component for our DPNF proposal as explained below.

Figure 2. Flickr output for the tag flower (top) and cluster output for the tag flower (bottom)

3. Dual-perspective navigation framework

DPNF was developed as an attempt to integrate classic faceted browsing and tag-based navigation. Our goal was to develop a more efficient exploratory search approach that could be applied in situations when both metadata (subject headings) and tags are available. With many museums and archival sites now embracing social tagging, the presence of both kinds of image

Information Processing & Management, In Press, 2016

descriptions is becoming a regular case rather than an exception. When designing DPNF, we were driven by two opposite intentions. On the one hand, we wanted to design a true integration where the combination of tags and metadata perform better than the individual components alone. On the other hand, we wanted the traditional faceted browsing and tag-based navigation to remain recognizable and usable in the context of the new interface. We believe that this is important for efficiency (both approaches are known for their efficiency) and learnability (users can adapt to the new interface more readily since they may recognize faceted browsing and tag-based navigation, allowing them to draw on relevant past experience).

Thanks to the extensive research conducted by Hearst and others (reviewed above), the faceted browsing approach has emerged to become an efficient interface. Thus Flamenco-like faceted metadata browsing was a natural selection for DPNF. However, it is less clear which version of tag-based navigation approach should be integrated in order to achieve an efficient DPNF interface. Originally we intended to use the *faceted* tag interface that we previously developed (Y. Lin et al., 2010) since its faceted nature appeared to be most appropriate for combining with faceted browsing. However, a preliminary study (Trattner et al., 2012) indicated that a simple tag-cloud interface is both more efficient and less confusing for non-expert users. Given that the traditional tag-cloud is also the most popular and most familiar to users (Millen, Feinberg, & Kerr, 2006; Seifert, Kump, Kienreich, Granitzer, & Granitzer, 2008), we decided to build DPNF as a combination of the Flamenco-style faceted browsing and the traditional tag-cloud interface.

As shown in Figure 3, DPNF contains three important elements. Firstly, it provides a search text box for a basic keyword search that offers the look and feel of well-known search engines. Users can issue a query and get a thumbnail preview of the resulting images sorted by

Information Processing & Management, In Press, 2016

relevance. The backend search engine, Apache Lucene, indexes all the image-related content including subject headings, tags, and descriptions. Secondly, a faceted browsing interface along with subject headings is provided on the left side of the screen. The subject headings are classified into four facets: activities, objects, locations, and people. Each subject heading is associated with the number of images related to it. The subject headings in each facet are presented in descending order based on these numbers. Thirdly, a tag cloud is presented on the top of the screen. In this traditional tag cloud interface, tags are alphabetically ordered and displayed with different font sizes according to their frequency.

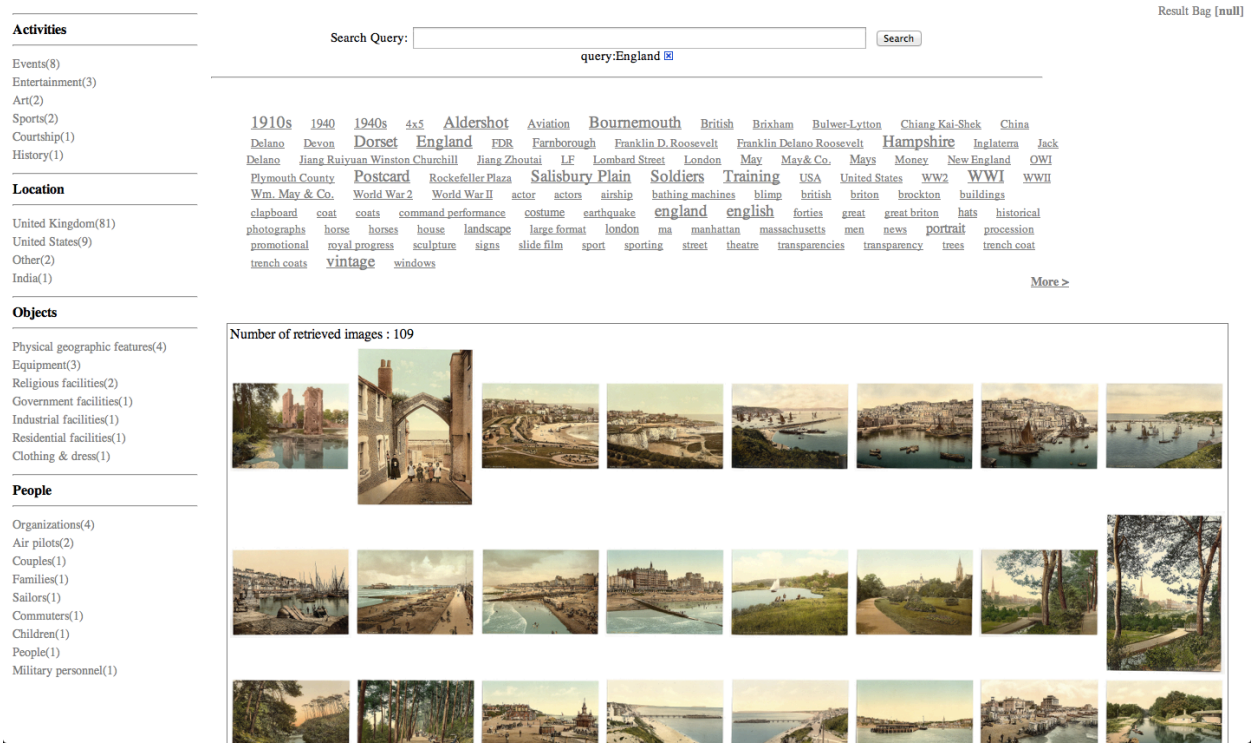

Figure 3. The dual-perspective navigation interface

A typical usage scenario of DPNF starts when a user chooses a starting point from any one of the searching and browsing mechanisms. By selecting a subject heading, a tag, or issuing a query, the user initiates the search and the system retrieves a set of images based on all textual

Information Processing & Management, In Press, 2016

content associated with images: subject headings, tags, and descriptions. Based on the returned images, the system generates and presents subject heading facets as well as the corresponding tag cloud that describes the retrieved subset of the original images.

Now the user has the freedom to explore the presented images, the four facets of subject headings, the tag cloud, or any combination therein. The current query is shown below the search text box to help users keep track of the current selection criteria. The flexibility of choosing a method for refining the search is the key concept in the DPNF design. The user can refine the search from each point of exploration in all three ways: by entering a new query, by selecting a specific tag from the tag cloud, or by clicking one subject heading from any one of the four facets. When the user adds a search query, subject heading, or tag to the current search, the system updates the image list so that only those images satisfying all selected criteria are displayed. The ranking of images is now determined by their relevance to the selected criteria. Since this re-filtering reduces the number of displayed images, it also affects the set of associated subject headings and the corresponding tag cloud. Thus, the user's facet browsing with subject headings results in updating and narrowing the set of displayed tags; in the same manner, selections in the tag cloud not only update the tag cloud but also update and narrow the displayed subject headings. Unlike earlier explored parallel but independent application of faceted browsing and tag-based navigation, DPNF offers seamless cross-navigation between two features, which achieves our goal of fully integrated dual perspective navigation support.

At any time, the user can also remove any of the search criteria by clicking the "x" beside a given search criterion displayed under the search query box. Similar to adding search criterion, removing a criterion would cause updates to the list of displayed images, and thus trigger updates to both the facets and the tag cloud. A search restarts when the user removes all the search

Information Processing & Management, In Press, 2016

criteria.

Figure 3 shows DPNF loaded with the Flickr Commons collection. The query "England" was issued as the initial query on our experimental tasks (European travel chapter). Relevant images are displayed as the search results, corresponding subject heading as facets are displayed at the left side, and the tag cloud is displayed at the top. The user's interactions with DPNF would be the same as presented above. When a user clicks on a thumbnail, the interface switches to a full image presentation window to show the image in higher resolution accompanied by the image description and associated tags and subject headings (see Figure 4).

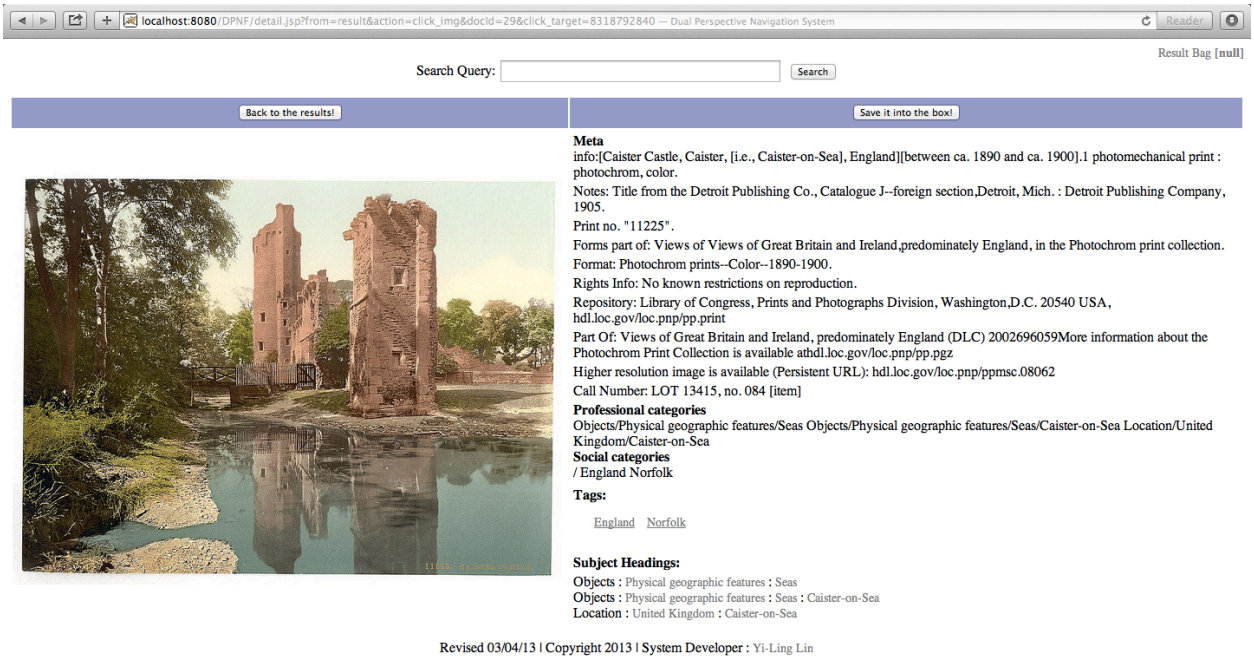

Figure 4. The detail page of the dual-perspective navigation interface

4. Research Design

To assess whether DPNF does provide an efficient, effective, and user-oriented method to support image finding, we performed a controlled user study. In this section, we introduce the study design while the next section focused on result analysis.

Information Processing & Management, In Press, 2016

4.1 Dataset

We utilized two collections of images. One collection is from the "Teenie" Harris archive belonging to the Carnegie Museum of Art in Pittsburgh, Pennsylvania, which catalogs a 40-year period of Pittsburgh history through the eyes of an African-American journalist and amateur historian. Since this collection contains professional subject headings by the Library of Congress without social tags, in order to fulfill our study setting to test both features, we selected 1,864 (out of 80,000) images from this collection to collect social tags from the Amazon Mechanical Turk. There were 986 images of the selected set featured in the exhibition (Oct 2011-April 2012) at the Carnegie Museum of Art. The remaining images were included in this study as they provide a more granular overview of the entire collection. For every image in our 1,864-image collection, we obtained user tags through Amazon's Mechanical Turk (MTurk) services (Rashtchian, Young, Hodosh, & Hockenmaier, 2010). The use of MTurk was motivated by the reported success of its application in similar contexts, which ranged from labeling images with keywords (Nowak & Rüger, 2010; Sorokin & Forsyth, 2008) to judging the relevance of search results (Grady & Lease, 2010). We gathered 5,634 unique tags created by 256 users for the 1,864 images in our collection. The tag-collecting and some simple quality control has been depicted in the third section of our previous work (Y.-L. Lin, Trattner, Brusilovsky, & He, 2014). Figure 5 shows an example of the interface used in MTurk for collecting tags. We only showed the image without any description or subject heading while collecting the social tags.

The second dataset was crawled from Flickr. It contains images uploaded to Flickr by the Library of Congress[2] prior to January 2013. It contains 15,194 images that are identified by the "Library of Congress" tag. Around 83% of images (12,541) have more than one tag. Overall,

[2] http://www.flickr.com/photos/library_of_congress/

Information Processing & Management, In Press, 2016

there are 1,216,318 tags provided by the Library of Congress and Flickr's users, among which 12,896 are unique tags. The maximum number of tags per image was 73, which is close to Flickr's limit of 75 different tags per image. Since the images belong to the Library of Congress collection, they also have assigned subject headings. We were able to retrieve subject headings by the Library of Congress for 6,923 images, which in total have 27,232 subject headings (of which 1,596 are unique). We further identified 5,281 images with both subject headings and social tags, and used these images to create the second dataset for our experiment. Table 1 shows the detailed information of the two datasets.

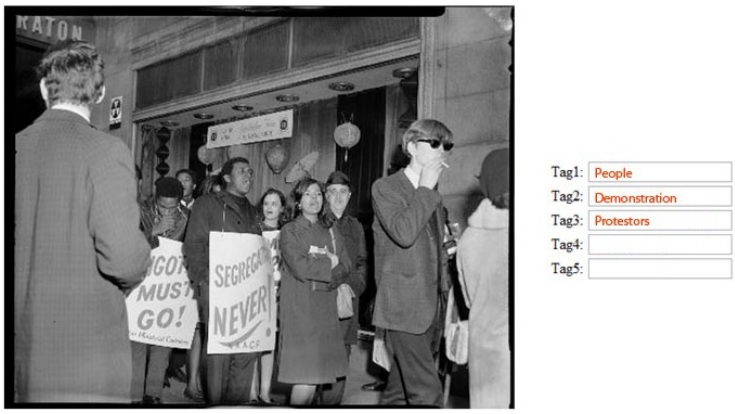

Figure 5. The tagging interface in MTurk

Table 1. Summary of two studied dataset collections (Numbers are unique)

	No. of unique SHs	Mean SHs per image	No. of unique tags	Mean tags per image	No. of images
Teenie Harris	607	6.35	5,634	17.23	1,864
Flickr	1,596	5.45	12,896	15.30	5,281

4.2 Interfaces

To examine the value of DPNF, i.e., to assess whether users will find images more efficiently and effectively when two kinds of information descriptors (subject headings from

Information Processing & Management, In Press, 2016

experts and tags from the general public) are presented, we developed two baseline interfaces based on a single kind of information descriptor.

The *tag-only interface* employs the tag cloud used in the DPNF interface and can be considered as a DPNF interface with no subject-heading component. This interface provides a search text box for a basic keyword search with query suggestions, a social tag cloud of the entire dataset at the top, and image examples at the bottom (see Figure 6-1). It facilitates users' familiarity with the high-level information *aboutness* of the entire collection and gives users the freedom to choose the starting point from any one of the searching and browsing mechanisms.

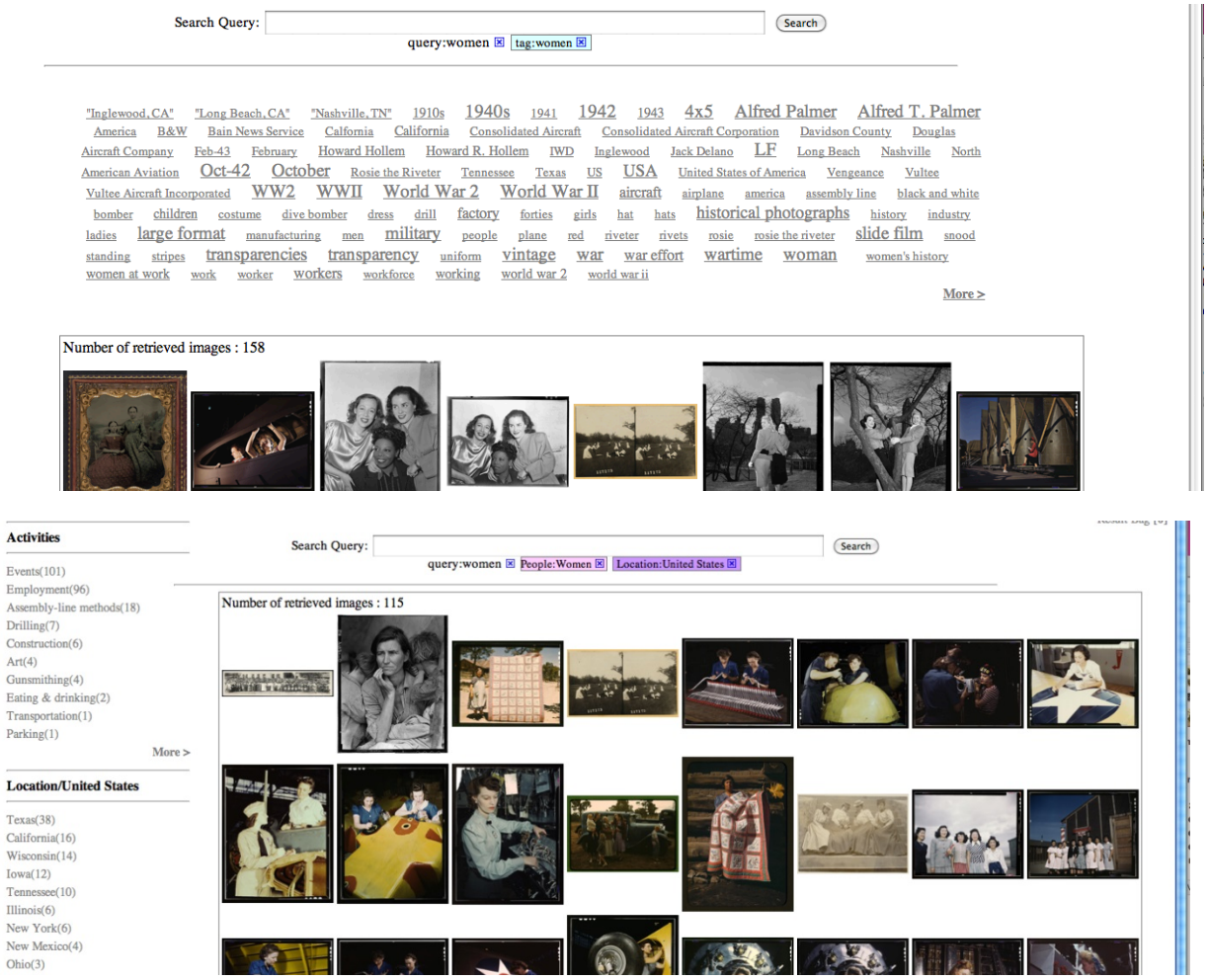

Figure 6. Examples of a tag-only interface (top) and subject-headings only interface (bottom)

Information Processing & Management, In Press, 2016

The *subject-headings only* interface uses faceted browsing to represent structural subject heading vocabulary along several dimensions. It could be considered as a DPNF interface without the tag cloud component (see Figure 6-2). The form of faceted browsing used in the subject-headings only have been shown to be helpful and preferred by users over the traditional search interface (Y. Lin et al., 2010; Yee et al., 2003), and adopted by different studies (Chan & O'Neill, 2010; Sigurbjornsson & Zwol, 2010).

4.3 Search tasks

Two types of search tasks were examined in this study - lookup search and exploratory search. To study lookup search behavior, we selected nine different images and performed a known-item search by showing the user one of the images throughout the search and asking him/her to find that exact picture with the selected interface. The exploratory search scenario simulated a more complicated situation where a user has a broader information need that requires multiple searches interwoven with browsing and analysis of the retrieved information (Y. Lin et al., 2010; Marchionini, 2006). To support this scenario, we designed three exploratory search tasks. Sample images used for lookup search and a sample exploratory task are shown in Table 2. The 9 lookup images and 3 exploratory tasks are collection specific since the collections are distinct. The testing shows that there was no collection effect.

Table 2. Samples of search tasks and descriptions of the Flickr collection

Search Tasks	Search Task Descriptions
Lookup	Find the following picture
	

	(only one images was presented to the user at a time)
Exploratory	Background: You would like to add a new chapter to a travel book with some historical pictures about Europe. You are looking for images from the Library of Congress Flickr Commons collection The new chapter will include photographs of natural scenery, landmarks or buildings, and events in Europe. You want to present 4 countries. For each country, you will collect one representative picture of its natural scenery, one for its modern facilities, and one for its activities. All three pictures have to be in the same location (e.g., in the same region, state, province, or city of the country). You should gather 12 photos from this search.

4.4 Subjects

We recruited sixty-two participants from the greater Pittsburgh area. The participants were paid $12 per hour. Eight took part in a pilot study that helped us to shape the main study. The remaining 54 participants participated in the main study. The characteristics of the participants' demographics by collection are reported in Table 3. The participants were distributed evenly according to gender. The participants who scored higher than the mean (73.43%) of the working memory scores from all participants are defined as the group with the high level of working memory. Although the percentage of the native English speakers was slightly lower than non-native English speakers, we ensured that at least one third of the participants were native English speakers in both collections. Although high and low working memory was almost equally distributed (13 vs. 14 in both collections), the high-low distinction was made without consideration of the collection a user was assigned to.

Table 3. Demographics of the participants by interfaces in two collections

Characteristic	Teenie Harris (N=27)	Flickr (N=27)
Working memory		
- Low	14(52%)	14(52%)

Information Processing & Management, In Press, 2016

	- High	13(48%)	13(48%)
Native			
	- Non-native English speaker	16(59%)	17(63%)
	- Native English speaker	11(41%)	10(37%)
Gender			
	- Female	13(48%)	13(48%)
	- Male	14(52%)	14(52%)
Major			
	- Non-computer related	13(48%)	17(63%)
	- Computer related	14(52%)	10(37%)

*Values represent the number of cases (the percent of cases) in each category

4.5 Design and procedure

The study was organized as a collection-specific within-subject experiment. Since the collections are distinct, the participants were divided into two groups to test different collection separately. Each of the participants evaluated the three search interfaces during one experimental session. To compare the interfaces in both lookup search and exploratory search context, each interface was examined in both types of search tasks. The experiment was conducted in the usability lab located at the University of Pittsburgh's School of Information Sciences.

Each subject was assigned to work on three different lookup search tasks and one exploratory task. Over the course of the experiment, each subject had to perform the required tasks on each of the three interfaces. To counter the impact of fatigue and learning, the order of search tasks across the interfaces and the order of the work within each interface were rotated using the Latin square design.

The overall procedure was as follows: First, we informed the recruited participants about the objectives of the study and obtained their consent (~10 minutes). Then each participant

Information Processing & Management, In Press, 2016

completed a short background survey (~2 minutes) and took a working-memory capacity test (~3 minutes). Before performing the official tasks within each interface, each participant was trained to use the features of the targeted interface with a detailed explanation of the different requirements of each search task. The participant was given sufficient time to become familiar with the interface and the both types of search tasks (~5 minutes per interface).

Lookup task. In the course of the pilot study, we observed that the participants spent an average of 104.69 seconds (SD =71.91 seconds) to find an image. Based on this data, we rounded the average plus one standard deviation, and imposed a limit of 3 minutes (180 seconds) per image for each lookup search task in the main study.

Exploratory task. A description of the task was given to the participant before they started to work on the task. The participants were allowed to ask any questions about the task except how to search for relevant images. In our pilot study, participants spent an average of 506.29 seconds (SD=94.15 seconds) to achieve the requirements of each exploratory task. Consequently, the main study imposed a limit of 10 minutes (600 seconds) for each participant to complete each assigned exploratory search task. Participants were told to complete the task as quickly as possible within the 10-minute limit.

A *post-task* questionnaire (Appendix A) with questions about the difficulty of finding images with the corresponding interface was presented upon completion of each assigned task. After the *post-task* questionnaire (2 minutes), a NASA-TLX workload survey (Hart & Staveland, 1988) was administered to assess the participant's workload when interacting with the corresponding interface (3 minutes). After the participant completed work with all three interfaces, he or she was asked to fill in the *post-experiment* questionnaire (Appendix B) that focused on comparing the three search interfaces in terms of the participant's preference,

Information Processing & Management, In Press, 2016

perception, etc. This survey was followed by a structured interview. The experiment took approximately 120 minutes to complete.

4.6 Hypotheses

In light of the definitions of findability (Morville, 2005), we expected that the proposed DPNF contributes to the enhancement of image findability by supporting people's ability to find their way to target items in an efficient and effective manner. Effectiveness, efficiency, and satisfaction of the ISO 9241 standard for usability (1998) were adopted to assess the usefulness of DPNF. According to the ISO definition, effectiveness is the accuracy and completeness with which users achieve specified goals, efficiency is the resources required in relation to the accuracy and completeness with which users achieve the goals, and satisfaction is the positive attitudes toward the user of the product (ISO, 1998). Consequently, we defined three main research questions on effectiveness, efficiency, and subjective perception to examine the usability of DPNF. The three research questions and their corresponding list of hypotheses and measurements are shown in Table 4.

The RQ1 was defined as the question focusing on effectiveness. Referring to Hornbaek (2006), we focused on task completion (Westerman, S., Cribbin, T., Wilson, 2001) and accuracy measures (Marshall, D., Foster, J.C., Jack, 2001) as the two main effectiveness hypotheses: H1-1 and H1-2. In addition, to test whether a feature-rich interface costs more working memory to process (Findlater & McGrenere, 2007; Marchionini, 1993), we defined the H1-3 hypothesis.

The RQ2 was defined as an *efficiency* question. In information foraging theory, searchers navigate through information patches to find what they need. With diverse information descriptors provided as information scent, only the clearest indication (strongest scent) can quickly lead users closer to the information they require. To test whether our DPNF provides

Information Processing & Management, In Press, 2016

strong scent, search time and the interactions with scent were important to investigate. Therefore, we define three hypotheses to examine the question in terms of search time and interactions, including H2-1 (focusing on lookup tasks), H2-2 (focusing on exploratory tasks), and H2-3. In addition, information foragers search for a scent trail and follow it toward their desired resource. When they lose the scent, they often retrace back through the coming route until they encounter a better alternative scent or give up. If they typically have to backtrack, the circumstance might indicate that information scent is not clear enough to provide users with a good direction. Thus,

Table 4. Research questions, hypotheses, and measurements

Research Questions	Hypotheses	Metrics
RQ1: Does the dual-perspective navigation approach provide information that helps users achieve their goals in a more effective manner than the single-perspective approaches?	• H1-1: Users will successfully complete more tasks with DPNF. • H1-2: Users will make less futile searches with DPNF. • H1-3: Users with different working memory capacity will not have different search performance with DPNF.	• Task success • Number of selected pictures • Number of futile searches[3] • Interaction by working memory capacity and interface
RQ2: Does the dual-perspective navigation approach guide users to their targets with fewer required resources than the single-perspective navigation approaches?	• H2-1: Users will spend less time to find a target item with DPNF. • H2-2: Users will spend less time to complete a task with DPNF. • H2-3: Users will reach the task goal with fewer interface interactions with DPNF. • H2-4: Users will use less back tracking[4] with DPNF.	• Time to find a targeted item • Time to complete a task • Number of the actions to reach the task goal • Number of back tracking used • NASA's Task Load

[3] A futile search is defined as a search, which returns an empty list of results.

[4] Back tracking action is defined as deleting a search query or going back to the result set after examining a specific picture.

	• H2-5: Users will expend less mental effort when using DPNF.	Index questionnaire
RQ3: Does the dual-perspective navigation approach make users feel confident in their image finding ability and create a positive perception of the approach?	• H3-1: Users will be more confident in performing a search task with DPNF. • H3-2: Users will be more satisfied with DPNF.	• Post-questionnaire • Structured interview

we explore this issue by examining another hypothesis -- H2-4. Furthermore, crafting an interface with optimal levels of information scent can reduce the mental effort that users have to expend to find their desired resource. Therefore, the workload of using interfaces is also an important factor to explore, so we proposed another hypothesis in order to test it -- H2-5.

Usability studies (Gutwin, 2002; McGrenere et al., 2002; Rui, Gupta, & Cadiz, 2001) have found that if participants are confident with the information scent provided by a system, they tend to believe they are still on the pathway to their goal so they will keep using the system, which ultimately creates a more positive perception of the system for the participants. Therefore, we defined RQ3 as the *perception* research question. Accordingly, we defined hypothesis H3-1 and H3-2 to test it. We designed post-task (Appendix A), post-experiment (Appendix B) questionnaires and a structured interview to assess the differences among the three search interfaces based on three dimensions, preference, satisfaction, and perception of users.

4.7 Statistic method

The generalized estimating equation (GEE) was applied to model and analyze the data (Liang, Zeger, & Apr, 2007). GEE can specify the repeated measures on two variables -- interface and search_type -- in an appropriate manner that allows us to define the distribution and

Information Processing & Management, In Press, 2016

link function to model different types of outcome variables such as linear, Gamma with log link, ordinal logistic, Poisson log-linear, binary logistic, etc. It provides many correlation structures and produces model-based and empirical estimates, as well as the proportional odds model (Pedhazur, 1982) to help us to easily interpret the associations found in the data.

5. Analysis of Results

The experimental conditions included interfaces, collections, search tasks, interface order, and task order (for lookup tasks only). The system log was analyzed to investigate users' performance while the participants' responses on the questionnaires were used to evaluate their subjective perceptions from the participants. This section is organized into three parts based on the general construct of usability with objective and subjective factors.

5.1 Effectiveness

This section focuses on the *effectiveness* research question, "Does the dual-perspective navigation approach provide information that helps users achieve their goals in a more effective manner than single-perspective approaches?" First, we measured the participants' task success with the following steps:

1) For each lookup task, a participant had to find the required picture within three minutes. The variable, task_success, was coded as zero for those who didn't find the assigned picture within three minutes and as one for those who did find the assigned picture within the time limit. This study required each participant to perform three lookup tasks per interface.
2) For the exploratory task, a participant had to find twelve pictures within ten minutes to fulfill the requirements for this task. The variable was coded as zero for those who didn't

Information Processing & Management, In Press, 2016

find all twelve pictures within ten minutes, and one for those who completed the task successfully within the ten minutes allowed.

Table 5. The numbers of non-success and success observations in interfaces and search types

Search type	Subject Heading		Tag		Dual	
	Non-success	Success	Non-success	Success	Non-success	Success
Lookup (N=162)	39	123	38	124	40	122
Exploratory (N=54)	5	49	9	45	0	54

Table 5 shows the number of successful and failed cases for the two types of search tasks with three interfaces. As the data show, users have different performance rates between the two search types. There is approximately 25 % failure rate in the lookup task, whereas the failure rate in the exploratory task ranges from 0 to approximately 10%.

A binominal distribution with log link function is applied with GEE to evaluate the association between success and two within-subject variables -- interface and search type. Since 54 participants were all successful while using the dual-perspective navigation interface for the exploratory tasks (Table 5), it is invalid to apply the binary logistic to model success with zero observation in this condition. The GEE analysis of task success among interface was limited to the cases with the lookup search tasks. In addition, these two types of search tasks are designed quite differently with different time limitations (three minutes for each of the three lookup tasks and 10 minutes for the exploratory task) and requirements (one picture for each of the three lookup tasks and 12 pictures for the exploratory task) for each interface. Therefore, the cases are split based on the search type for the following performance analyses using GEE.

To predict task success in the lookup tasks, we controlled for experimental conditions, subject demographics, and the interaction effect between interface and working memory. The

Information Processing & Management, In Press, 2016

model has significant effects on collection [χ (1) ^{2}=25.757, *p<0.001*] and working memory [χ(1)2=5.516, *p=0.019*], but no significant effect on interfaces. The complete results are shown in Table 6. When comparing low working memory and high working memory for number of successes, low working memory participants had 3.295 times the number of successes that the high working memory participants had (OR=3.295, *p=0.002*). The participants with low working memory capacity are more successful in completing the lookup task. The result indicates that working memory capacity affects the search performance differently from what we usually expect (i.e., users with high working memory capacity might have higher success in cognitive tasks). Except for collection, the rest of the variables have no effect on success. With the collections, images were found 3 times more frequently in the Teenie Harris collection than in the Flickr collection.

Table 6. Significant effect influencing lookup task success

Parameter		S.E.	Wald χ^2	Sig	Exp(B)
collection	Teenie Harris	.2165	25.757	<.000	3.000
working memory	Low	.3777	9.971	.002	3.295

To test the difference in task success rates among the interfaces in the exploratory task, we transformed the data and adopted a nonparametric test (Cochran's Q Test) for the binary response analysis. To compare the pattern across the interfaces, we obtained all of the pairwise comparisons among the three interfaces. We applied a Bonferroni correction for *p<0.05* rule to protect against Alpha inflation. To do this, a Type I error across the pairwise comparisons was adjusted to be less than a 5% chance, which was accomplished by dividing .05 by the number of comparisons. The percentage of successes that occurred in each interface was 90.74% with the subject heading-only interface, 83.33% with the tag-only interface, and 100% with the dual

Information Processing & Management, In Press, 2016

interface [Q(2)=10.167, *p=0.006*]. The pairwise comparisons using a Bonferroni correction of *p=.0167* (.05/3) indicates that the participants with the dual-perspective navigation interface had more successes than with the tag-only interface [Q(1)=9.00, *p=0.004]*, but there was no significant difference between the subject heading-only and the dual-perspective interfaces, nor between the subject heading-only and the tag-only interfaces.

In addition to using the variable task_success to test differences while completing the exploratory tasks among interfaces, we also adopted the number of pictures selected during the exploratory search as another outcome variable for an exploratory task's success. GEE Poisson was applied to predict the number of pictures selected since the Poisson distribution can interpret the distribution of this counting outcome variable. Table 7 shows that there is a significant interface effect [$\chi(2)^2$=14.643, *p=0.001*]. Compared to the DPNF interface, the participants selected 2% fewer pictures (OR=0.984, *p=0.014*) with the subject heading-only interface, and 3.8% fewer pictures (OR=0.962, *p=0.013*) with the tag-only interface.

Table 7. Significant effect influencing success based on the number of selected pictures for the exploratory task

Parameter		S.E.	Wald χ^2	Sig	Exp(B)
interface	Subject Headings	.0065	6.091	.014	.984
	Tags	.0158	6.164	.013	.962

Another effectiveness metric -- futile search -- was defined to measure how frequently empty lists were returned to users while searching using DPNF. For the exploratory search tasks, the participants failed to get returned results (i.e., their query returned an empty list) significantly more frequently (*p=0.038)* with the tag-only interface (mean=1.65, S.E.=0.256) than with DPNF (mean=0.93, S.E.=0.142). With the lookup tasks, there was no effect found.

Information Processing & Management, In Press, 2016

In summary, from the perspective of effectiveness, we found a significant interface effect for exploratory search tasks. The participants had higher success rate when they performed an exploratory search task with DPNF compared to with the tag-only interface. DPNF also outperformed both the subject heading-only and tag-only interfaces in terms of the number of pictures selected, which is an important productivity measure for the exploratory search. DPNF interface also significantly reduced the number of unsuccessful (futile) searches in comparison to the tag-only interface. No significant impact of the interface was found for the lookup search. We found, however, that participants' performance in the lookup search was affected by the collection used and by individual levels of working memory.

5.2 Efficiency

This section addresses the efficiency research question, "Does the dual-perspective navigation approach guide users to their targets with fewer required resources than the single perspective approaches?"

To assess the hypotheses, "H2-1: Users will spend less time to find a targeted item with DPNF" and "H2-2: Users will spend less time to complete a task with DPNF," the search time is examined with different search types. Table 8 shows the descriptive statistics of search time in two search types with three interfaces.

Table 8. Descriptive (mean ± SE) of search time and total actions by search type and interface

Search Type	Measure	SH	Tag	Dual
Lookup	Successful cases	123	124	122
	Search time	63.80±4.20	56.79±3.78	64.00±4.11
	Total actions	12.93±.67	11.40±.60	12.84±.73
Exploratory	Successful cases	49	45	54
	Search time	354±15.75	402.98±15.96	354.94±13.47

Information Processing & Management, In Press, 2016

Total actions	80.45±2.94	83.64±3.54	78.74±2.53

With the lookup task, there is only a main effect of collection [$\chi(1)^2$=13.182, *p<0.001*] that was significant. When comparing the Flickr and Teenie Harris collections on time spent to reach the target item, we see that the participants working with the Teenie Harris collection spent only 74.2% of the time that the participants spent on the Flickr collection (OR=0.742, *p<.0.001*). Because the lookup task asks participants to find the required picture in the collection, the total number of pictures in the collection might influence participants' performance in the lookup task. There is no other significant main effect, as well as no significant interaction effect for the lookup task.

With the exploratory task, the result indicates that there is a main effect for the interface [$\chi(2)^2$=6.364, *p<0.042*] when controlling for experimental conditions, subject demographics, and the interaction between interface and working memory. The results of the pairwise comparisons with the sequential Bonferroni adjustment indicates that the participants spent significantly less time (*p=0.020)* with the subject heading-only interface (mean=352.87, S.E.=16.21) and (*p=0.050)* the dual-perspective navigation interface (mean=352.10, S.E.=13.03) than with the tag-only interface (mean=395.70, S.E.=15.24) to complete a task.

Table 9. Significant effect influencing average time for finding a required picture in the exploratory task

Search type	Parameter		S.E.	Wald χ^2	sig	Exp(B)
Exploratory	Interface	Subject Heading	.0788	5.684	.017	1.207
		Tag	.0847	10.364	.001	1.313

Due to the time limit of the exploratory task, participants obtained a different number of images within the same search time (up to 10 minutes -- the time limit in the experimental

Information Processing & Management, In Press, 2016

setting). For example, one participant might successfully obtain 12 pictures in the required 10 minutes while a second participant might only obtain six pictures within that time limit. Their performance should be considered differently. To address this, we used an average time per picture selected as another dependent variable to be able to predict more accurately how fast a participant is in the exploratory search. There is a main effect of interface [$\chi(2)^2$=19.694, *p<0.001*]. When comparing the DPNF interface to the other two interfaces on average time spent (see Table 9), the participants using the subject heading-only interface spent 20.7% longer (OR=1.207, *p=0.017*), and the participants using the tag-only interface spent 31.3% longer (OR=1.313, *p=0.001*).

To test the hypothesis, "H2-3: Users will reach the task goal with fewer interface interactions with DPNF," the count of the overall interactions -- total actions -- is used as the outcome variable. Table 8 shows the descriptive statistics of total actions in two search types with three interfaces. Since the distribution of the count variable could be described with a Poisson distribution, we applied the Poisson distribution with log link in GEE to predict participants' interactions with the interfaces given the independent variables from the experimental conditions, subject demographics, and the interaction between interface and working memory.

For the lookup task, the results show that there is a main effect found in collection [$\chi(1)^2$=11.393, *p=.001*]. When comparing the collections for number of actions, the participants who worked in the Teenie Harris collection took 80.6% of the actions that the participants who worked in the Flickr collection took (OR=. 806, *p=0.001*). No other effects were found.

For the exploratory task, no effect was found. However, when we included failure cases into the analysis (with all cases), we found an effect of interface [$\chi(2)^2$=7.441, *p=0.024*], which

Information Processing & Management, In Press, 2016

indicates that users who failed to achieve the task goal were strongly affected by the interface type. In total, users had 19% more interactions with the subject-heading only interface (OR=1.190, *p=0.005*) and 25.2% more interactions with the tag-only interface (OR=1.252, *p=0.001*) than with the dual-perspective navigation interface.

To investigate whether users back tracked less frequently with DPNF for H2-4, the number of back tracking actions were examined with the negative binominal log link in GEE. For the lookup search task, there was no significant effect found based on successful cases. When failure cases were included into the analysis, we found an effect of collection [$\chi(1)^2$=7.219, *p*=0.*007]*. The participants working in the Teenie Harris collection used 30.2% of the back tracking actions that the participants working in the Flickr collection used (OR=0.302, *p=0.007*). The reason might be that the number of total images contained in the Teenie Harris collection is around two thirds smaller than the number of images in the Flickr collection. In the lookup task, each participant was required to find an exact picture, and this type of task seems to be easier to complete in the small collection.

For the exploratory search task, an effect of collection [$\chi(1)^2$=5.654, *p=0.017*] was found. The participants needed 13.6% (OR=1.136, *p=0.017*) more back tracking actions to complete the exploratory task in the Teenie Harris collection than in the Flickr collection. The rest of the effects are not significant. When we included all cases, the results reveal that the participants applied less back tracking during search with DPNF than with the other interfaces. Compared to DPNF, the participants used 31% (OR=1.310, *p=0.004*) more back tracking actions with the subject heading-only interface and 30.4% (OR=1.304, *p=0.003*) more with the tag-only interface respectively. Users who failed to achieve the task goal used less back tracking with the

Information Processing & Management, In Press, 2016

DPNF interface than with the other interfaces. This shows that DPNF has relatively strong information scent to lead users to the target resources.

To examine the hypothesis **"**H2-5: Users will have less mental effort when using DPNF", we used NASA TLX to collect users' workload evaluations after they experienced each type of tasks with each interface. The outcome of workload is a continuous dependent variable that was calculated by NASA TLX. Since the Gama distribution can well interpret continuous probability distributions, GEE Gamma was applied to predict workload in the system given the experimental conditions, subject demographics, performance, and the interaction between interface and working memory. The result does not provide any evidence that the DPNF interface significantly reduces mental effort for either the lookup tasks or the exploratory tasks. It is interesting, however that the mental effort analysis reveals a significant effect of native language on the workload. The participants who are not native English speakers had 23% (OR=1.23, $p=0.012$; OR=1.233, $p=0.013$) more workload to complete either the lookup or exploratory task than the participants who are native English speakers. This might be related to the language used in both kinds of image descriptors. Both tags and subject headers describe the *aboutness* of the image in English, so the participants with better English ability might expend less mental effort to consume the information provided by the systems. While it doesn't support our original hypothesis, it provides evidence about the validity of the workload measures.

In summary, from the *efficiency* prospect, the DPNF approach guided users to their targets with less search time, fewer total interactions, and fewer back tracking actions than both of the single perspective approaches when the participants performed the exploratory search task. In addition, we observed that participants had different behavior in different collections for different search tasks. For the lookup search, participants who worked in the Teenie Harris

Information Processing & Management, In Press, 2016

collection used less search time, fewer total interactions, and fewer back tracking actions than those working in the Flickr collection. For the exploratory search, the results show an opposite collection effect (more discussions will be provided in section 6.2). We did not observed significant differences between the interfaces in terms of mental workload, although we did find that the mental workload is significantly affected by the native language of participants.

5.3 Subjective perception

This section focuses on the subjective perception research question, "RQ3: Does the dual-perspective navigation approach make users feel confident in their image finding ability and create a positive perception of the approach?" to understand whether users are more satisfied with DPNF (H3-2) and more confident in performing a search task with DPNF (H3-1). We collected data for this section by distributing questionnaires (see Appendices A and B). Table 10 shows the descriptive statistics of satisfaction and confidence with three interfaces.

Table 10. Descriptive (mean ± SE) of confidence and satisfaction by interface

Measure	SH	Tag	Dual
Confidence	4.06±.10	4.06±.09	4.19±.09
Satisfaction	4.00±.09	3.20±.14	4.67±.07

A set of questions related to confidence was asked in the *post-task* questionnaire. The participants were asked to rate their confidence on a scale from 1 (not at all) to 5 (very much). In addition, the study included two more confidence questions via a *post-experiment* questionnaire designed to inquire about which interface participants felt most confident using after their experience with the three different interfaces, and why. We found no significant effects on confidence in the post-task rating data, and only a marginal effect of interface order. However, the *post-experiment* questionnaire, which recorded participants' perceptions after they gained

Information Processing & Management, In Press, 2016

practical experience with both tasks and all three types of interfaces provided evidence in favor of DPNF. When asked the question, "Which one of the interfaces would you feel most confident using for other search tasks?" 85% of the participants chose the DPNF interface, 11% chose the subject heading-only interface, and only 3.7% chose the tag-only interface. The reasons, according to participants' comments, can be categorized into three groups. First, the DPNF interface provided them with diverse information/options to use with different needs. Second, the participants felt that the DPNF interface is more powerful for transitioning between search methods easily, i.e., between professional categories and social tags. Lastly, the subject headings and tags can complement and support each other seamlessly to assist the participants' search.

The questions about satisfaction were asked in the *post-experiment* questionnaire after experiencing all three interfaces. The participants rated their satisfaction from a scale of 1 (not at all) to 5 (very much). The assessment of *satisfaction* shows that the participants were significantly more satisfied with the DPNF interface than with the other two interfaces. When the participants used the subject heading-only interface and the tag-only interface, they had only 12.5% (OR=0.125, $p<0.001$) and 21% (OR=.0.021, $p<0.001$) of the satisfaction comparing to when they used the DPNF interface (see Table 11). Further support for this assessment of users' satisfaction across the three interfaces was found in the preference analysis. Most of the participants (77.8%) preferred the DPNF interface, 18.5% preferred the subject heading-only interface, and only 3.7% (one participant) preferred the tag-only interface. Although they had different preferences for different tasks, the vast majority of the participants still preferred to use the dual-perspective navigation interface.

Table 11. Significant effect influencing satisfaction and confidence

Parameter	S.E.	Wald χ^2	sig	Exp(B)

Information Processing & Management, In Press, 2016

Satisfaction	Interface	SH	.3797	29.928	<.000	.125
	Interface	Tag	.4675	69.016	<.000	.021

Overall, the result of subjective perception analysis demonstrated that the user satisfaction was affected by the type of interface. The participants were significantly more satisfied with the DPNF interface than with the other two interfaces. Although the main effect of interface was not found on confidence, DPNF was selected as the interface in which the participants were most confident for future use, and as the top choice to be recommended to cultural heritage institutions.

6. Further Analysis and Discussion

Following the presentation of main "bottom line" results, this section attempts to examine user performance on a more granular level and to discuss the difference between the two search types.

6.1 Fine-grain Action Analysis

To better understand differences in user behavior across tasks and interfaces, we performed a fine-grain log-based analysis of users' interactions, mainly focusing on navigation and search actions (Table 12 and Table 13). The analysis exposed that in the exploratory search task, the participants made 78.7% (OR=1.787, $p<0.001$) more search actions and 56% (OR=. 441, $p=0.005$) less navigation actions with the tag-only interface than with the DPNF interface. Similarly, in the context of lookup search, the number of navigation actions performed by participants using the tag-only interface was only 39.2% of the number of navigation actions performed by participants working with the DPNF interface (OR=0.392, $p<0.001$). It shows that the tag-only interface might not be able to provide sufficient information scent that is critical in guiding users to their target items. Consequently, users of the tag-based interface used navigation

Information Processing & Management, In Press, 2016

mechanisms much less frequently and replaced them with search activities. In contrast, the DPNF interface provided more features and richer information scent to guide users to the target items, so they tended to use navigation mechanisms more frequently.

To understand if there is any difference in the use of subject headings or tags between single perspective (subject heading-only or tag-only) and dual perspectives (DPNF), we split the navigation actions into subject-heading navigation and tag navigation then performed further analyses between the subject-heading only and DPNF and between tag-only and the dual approach. To complete an exploratory task, the participants applied significantly more tag navigations (*p=0.003)* with the tag-only interface (mean=2.87, S.E.=0.487) than with the DPNF interface (mean=1.63, S.E.=0.615). However, we didn't find any significant difference between the subject heading-only and DPNF.

Table 12. Descriptive (mean ± SE) of actions by interface in lookup task

	SH		Tag		Dual	
Measure	All	Successful	All	Successful	All	Successful
Cases	162	123	162	124	162	122
Navigation	2.48±.24	1.56±.18	1.25±.14	.83±.13	2.71±.23	1.80±.20
SH Navigation	2.48±.24	1.56±.18	-	-	1.65±.17	1.10±.14
Tag Navigation	-	-	1.25±.14	.83±.13	1.06±.16	.70±.12
Search	3.72±.25	2.77±.20	4.22±.27	2.94±.20	3.63±.24	2.75±.22
Futile Search	.51±.08	.31±.07	.63±.09	.37±.09	.54±.07	.39±.07
Back Tracking	.14±.04	.04±.02	.12±.04	.08±.04	.08±.03	.02±.01

Table 13. Descriptive (mean ± SE) of actions by interface in exploratory task

	SH		Tag		Dual	
Measure	All	Successful	All	Successful	All	Successful
Cases	54	49	54	45	54	54

Information Processing & Management, In Press, 2016

Navigation	6.22±.79	6.45 ±.87	4.46±.67	3.67±.63	6.67±.73	6.67±.73
SH Navigation	6.22±.79	6.45 ±.87	-	-	4.70±.64	4.70±.64
Tag Navigation	-	-	4.46±.67	3.67±.63	1.96±.40	1.96±.40
Search	5.76±.61	4.96±.53	9.46±.83	8.36±.76	5.20±.62	5.20±.62
Futile Search	1.37±.22	1.09±.17	1.70±.24	1.53±.22	1.04±.19	1.04±.19
Back Tracking	20.78±.98	19.84±.94	21.93±1.34	20.29±1.43	18.61±.87	18.61±.87

6.2 Lookup Search vs. Exploratory Search

In addition to for the main interface effect reported in the section 5, we also examined several effects to predict a variety of measures. The significant effect of collection was found mainly in lookup search on task success, search time, and total action. An interesting finding is that this effect had different directions between lookup and exploratory search on search action, futile search, and back tracking action.

Another interesting result from the performance analysis that deserves further examination is that a significant difference between interfaces was observed for exploratory tasks but not for lookup tasks. The difference might be because the participants could have had different search strategies to complete the assigned tasks that might have been the result of both the different time constraints in the experimental setting (three minutes for the lookup search and 10 minutes for the exploratory search) and the different requirements for these two types of search tasks.

For the lookup search, we observed that the participants usually tried strategies that might be familiar to them through their daily search experience. They often started with issuing several queries and used the navigation tools if they failed to locate the right images after a few searches. This strategy might be a result of the tighter time constraints for the lookup task, causing participants to avoid using less familiar navigation features in order to complete the task within the time allowed. Although the pilot study results indicate that three minutes should have been

Information Processing & Management, In Press, 2016

sufficient to perform each lookup task, some participants mentioned in the *post-experiment* questionnaire and interview that they felt time pressure while they performed the lookup tasks.

For the exploratory search tasks, users were returned an empty list of results (a clear sign of failed search) significantly more frequently ($p=0.038$) with tag-only interface (mean=1.65, S.E.=0.256) than with the DPNF interface (mean=0.93, S.E.=0.142). This might be correlated with a considerable increase of search actions at the expense of navigation. Compared to the DPNF interface, users of the tag-only interface performed 78.7% more search actions (OR=1.787, $p<0.001$) in exploratory tasks. The observation, that while in the exploratory search context, users of the tag-only interface searched more and received more empty results than did users of other interfaces, may indicate that tags of the tag-only interface were not as supportive as needed to support the role of information scent.

7. Conclusions

The work presented in this paper aimed to enhance image findability (Morville, 2005) by suggesting and examining the dual-perspective navigation framework (DPNF) for image search. We expected that DPNF could provide more comprehensive aboutness of an item, which can increase the information scent and therefore provide better support to users to their desired items in an efficient and effective manner. Effectiveness, efficiency, and satisfaction of the ISO 9241 standard for usability (1998) were adopted to evaluate the usability of DPNF. We summarize the important results for the three research questions as the following:

RQ1: Participants successfully completed more exploratory tasks with the DPNF interface than with the tag-only interface. While performing the exploratory search, participants with the DPNF interface made fewer unsuccessful (futile) searches than participants with the tag-only interface.

Information Processing & Management, In Press, 2016

RQ2: Participants spent significantly less time, fewer interface interactions, and less back tracking to complete an exploratory task with the DPNF interface than with the subject-heading only and tag-only interfaces. In addition, the DPNF interface did not cause any extra workload for participants as compared to using the other single perspective interfaces.

RQ3: Participants were more satisfied with the DPNF interface than with the others. DPNF interface was selected as the interface that the participants were most confident using in the future, and as the top choice to be recommended to cultural heritage institutions.

This study compared user performance and feedback for three types of image finding interfaces in the context of two types of search tasks - lookup and exploratory search. The results demonstrated that the DPNF interface outperformed the subject heading-only and tag-only interfaces. Both objective performance analysis and subjective perception analysis produced significant findings.

By truly integrating both experts' and general users' descriptors to represent more comprehensive aboutness of an item, DPNF reinforces the strengths of both the metadata-based and tag-based approaches without changing their nature or forcing users to choose one means over the other. The findings of this study can assist interface designers working on more efficient exploratory search interfaces. Although this study explicitly focuses on image search, the results may be applicable to a wide variety of other domains. The lack of textual content in image systems makes images particularly hard to locate using traditional search methods. Our study shows that the integration of folksonomies (i.e. social tags) and controlled vocabularies (i.e. subject headings) supports more effective and efficient exploratory search in the image search context.

Information Processing & Management, In Press, 2016

Although DPNF offered two types of descriptors to guide users to find images, we cannot claim that this study uncovered how different information descriptors guide individual users in finding target images. The data hint that each participant is likely to have a different background as well as a different preferred search strategy to perform a particular search task in the manner to which they are accustomed. In our future studies, we are interesting in exploring on a deeper level how participants consumed different information descriptors during their search process, and how that might relate to their background and individual differences. Further study with eye tracking augmentation may be helpful in learning more about the interaction between users' search behaviors and various types of information descriptors.

References

Al-Khalifa, H. S., & Davis, H. C. (2007). Exploring the value of folksonomies for creating semantic metadata. *International Journal on Semantic Web and Information Systems IJSWIS*, *3*(1), 12–38. Retrieved from http://eprints.ecs.soton.ac.uk/13555/

Armitage, L. H., & Enser, P. G. B. (1997). Analysis of user need in image archives. *Journal of Information Science*, *23*(4), 287–299. http://doi.org/10.1177/016555159702300403

Bateman, S., Gutwin, C., & Nacenta, M. (2008). Seeing things in the clouds : The effect of visual features on tag cloud selections. In *the 9th ACM conference on Hypertext and hypermedia* (pp. 193–202).

Begelman, G., Keller, P., & Smadja, F. (2006). Automated tag custering: Improving search and exploration in the tag space. In *Collaborative Web Tagging Workshop at WWW2006 Edinburgh Scotland* (pp. 15–33). Citeseer. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.120.5736&rep=rep1&type=pdf

Bischoff, K., Firan, C. S., Nejdl, W., & Paiu, R. (2008). Can All Tags be Used for Search ? In *Proccedings of the CIKM conference of Information and Knowledge Management* (pp. 203–212). Napa Valley, California, USA: ACM.

Candan, K. S., Di Caro, L., & Sapino, M. L. (2008). Creating tag hierarchies for effective navigation in social media. In *Proceeding of the 2008 ACM workshop on Search in social media - SSM '08* (p. 75). New York, New York, USA: ACM Press. http://doi.org/10.1145/1458583.1458597

Capra, R., & Marchionini, G. (2008). The relation browser tool for faceted exploratory search. In *Proceedings of the 8th ACM/IEEE-CS joint conference on Digital libraries* (pp. 420–420).

Pittsburgh, Pennsylvania, USA.

Chan, L. M., & O'Neill, E. T. (2010). *FAST : Faceted application of subject terminology*. Libraries Unlimited.

Fairthorne, R. A. (1969). Content analysis, specification and control. *Annual Review of Information Science and Technology*, *4*, 73–109.

Findlater, L., & McGrenere, J. (2007). Evaluating reduced-functionality interface according to feature findability and awareness. In C. Baranauskas, P. Palanque, J. Abascal, & S. Barbosa (Eds.), *Human-Computer Interaction-Interact 2007* (Vol. 9, pp. 592–605). Springer Berlin / Heidelberg.

Furner, J. (2007). User tagging of library resources: Towards a framework for system evaluation. In *World Library and Information Congress: 73rd IFLA General Conference and Council* (pp. 1–10). Durban, South Africa.

Grady, C., & Lease, M. (2010). Crowdsourcing document relevance assessment with Mechanical Turk, 172–179. Retrieved from http://dl.acm.org/citation.cfm?id=1866696.1866723

Gutwin, C. (2002). Improved focus targeting in interactive fisheye views. In *Proceedings of ACM Conference on Human Factors in Computing Systems* (pp. 267–274). ACM Press, New York, NY.

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In P. A. Hancock & N. Meshkati (Eds.), *Human Mental Workload* (Vol. 1, pp. 139–183). North-Holland.

Hassan-Montero, Y., & Herrero-Solana, V. (2006). Improving tag-clouds as visual information retrieval interfaces. In *International Conference on Multidisciplinary Information Sciences and Technologies*.

Hayman, S., & Lothian, N. (2007). Taxonomy directed fikjsibinues integrating user tagging and controlled vocabularies for Australian education networks. *World Library and Information Congress: 73rd IFLA General Conference and Council*, (August), 1–27.

Hearst, M. A. (2006). Clustering versus faceted categories for information exploration. *Communications of the ACM-Supporting Exploratory Search*, *49*(4), 59–61. http://doi.org/10.1145/1121949.1121983

Helic, D., & Strohmaier, M. (2011). Building directories for social tagging systems. In *Proceedings of the 20th ACM international conference on Information and knowledge management - CIKM '11* (p. 525). New York, New York, USA: ACM Press. http://doi.org/10.1145/2063576.2063655

Hjørland, B. (1992). The concept of "subject" in information science. *Journal of Documentation*, *48*(2), 172–200.

Hollink, L., Schreiber, a. T., Wielinga, B. J., & Worring, M. (2004). Classification of user image descriptions. *International Journal of Human Computer Studies*, *61*(5), 601–626. http://doi.org/10.1016/j.ijhcs.2004.03.002

Hornbæk, K. (2006). Current practice in measuring usability: Challenges to usability studies and research. *International Journal of Human-Computer Studies*, *64*(2), 79–102.

Information Processing & Management, In Press, 2016

http://doi.org/10.1016/j.ijhcs.2005.06.002

Hotho, A., J, R., Schmitz, C., & Stumme, G. (2006). Information Retrieval in Folksonomies : Search and Ranking. *Data Engineering*, *4011*, 411–426. http://doi.org/10.1007/11762256

ISO. (1998). *Ergonomic requirements for office work with visual display terminals (VDTs)-Part 11: guidance on usaibility (ISO 9241-11:1998)*.

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

Kammerer, Y., Nairn, R., Pirolli, P., & Chi, E. H. (2009). Signpost from the Masses : Learning Effects in an Exploratory Social Tag Search Browser. In *CHI*. Boston, MA.

Karlson, A. K., Robertson, G., Robbins, D. C., Czerwinski, M., & Smith, G. (2006). FaThumb : A Facet-based Interface for Mobile Search. In *CHI'06 Proceedings of the SIGCHI conference on Human Factors in computing systems* (pp. 711–720). Montréal, Québec, Canada: ACM, New York. http://doi.org/10.1145/1124772.1124878

Knautz, K., Soubusta, S., & Stock, W. G. (2010). Tag Clusters as Information Retrieval Interfaces. In *2010 43rd Hawaii International Conference on System Sciences* (pp. 1–10). Ieee. http://doi.org/10.1109/HICSS.2010.360

Koraljka, G., Lykke, M., & Tudhope, D. (2014). Enhancing social tagging with automated keywords from the Dewey Decimal Classification. *Journal of Documentation*, *70*(5), 801–828.

Laniado, D., Eynard, D., Colombetti, M., & Milano, P. (2007). Using WordNet to turn a folksonomy into a hierarchy of concepts. *Semantic Web Applications and Perspectives 2007*, (Section 2), 192–201. Retrieved from http://ftp.informatik.rwth-aachen.de/Publications/CEUR-WS/Vol-314/

Lee, D. H., & Schleyer, T. (2012). Social Tagging Is No Substitute For Controlled Indexing : A Comparison Of Medical Subject Headings And Citeulike Tags. *Journal of the American Society for Information Science and Technology*, *63*(9), 1747–1757.

Liang, K., Zeger, S. L., & Apr, N. (2007). Longitudinal Data Analysis Using Generalized Linear Models. *Biometrika*, *73*(1), 13–22.

Lin, Y., Ahn, J.-W., Brusilovsky, P., He, D., & Real, W. (2010). Imagesieve: Exploratory search of museum archives with named entity-based faceted browsing. *Proceedings of the American Society for Information Science and Technology*, *47*(1), 1–10. http://doi.org/10.1002/meet.14504701217

Lin, Y.-L., Trattner, C., Brusilovsky, P., & He, D. (2014). The impact of image descriptions on user tagging behavior: A study of the nature and functionality of crowdsourced tags. *Journal of the Association for Information Science and Technology*.

Liu, H., Little, S., & Ruger, S. (2011). Multimedia: behaviour, interfaces and interaction. In I. Ruthven & D. Kelly (Eds.), *Interactive Information Seeking, Behaviour and Retrieval* (pp. 221–234). London : Facet Publishing.

Macgregor, G., & Mcculloch, E. (2006). Collaborative tagging as a knowledge organization and resource discovery tool. *Library Review*, *55*(5), 291–300.

Information Processing & Management, In Press, 2016

Marchionini, G. (1993). Finding facts vs. browsing knowledg in hypertext systems. *Computer*, *2*(11), 70–80.

Marchionini, G. (2006). Exploratory search: From finding to understanding. *Communications of the ACM*, *49*(4), 41–46. http://doi.org/10.1145/1121949.1121979

Marchionini, G., & Brunk, B. (2003). Toward a general relation browser: A GUI for information architects. *Journal of Digital Information*, *4*(1). Retrieved from https://journals.tdl.org/jodi/article/viewArticle/97

Maron, M. E., & Studies, I. (1977). On Indexing , Retrieval and the Meaning of About. *Journal of the American Society for Information Science*, *28*(1), 38–43.

Marshall, D., Foster, J.C., Jack, M. A. (2001). User performance and attitude towards schemes for alphanumeric data entry using restricted input devices. *Behaviour & Information Technology*, *20*(3), 167–188.

McGrenere, J., Baecker, R. M., & Booth, K. S. (2002). An evaluation of a multiple interface design solution for bloated software. In *Proceedings of the SIGCHI conference on Human factors in computing systems Changing our world, changing ourselves - CHI '02* (pp. 163–170). New York, New York, USA: ACM Press. http://doi.org/10.1145/503403.503406

Millen, D. R., Feinberg, J., & Kerr, B. (2006). Dogear: Social Bookmarking in the Enterprise. In *CHI'06 Proceedings of the SIGCHI conference on Human Factors in computing systems*.

Morville, P. (2005). *Ambient Findability: What We Find Changes Who We Become*. Retrieved from http://dl.acm.org/citation.cfm?id=1121644

Nielson, J. (2003). Information foraging: Why Google makes people leave your site faster. Retrieved from http://www.useit.com/alertbox/20030630.html

Nowak, S., Liorente, A., & Motta, E. (2010). The Effect of Semantic Relatedness Measures on Multi-label Classification Evaluation. In *Proceedings of the ACM International Conference on Image and Video Retrieval* (pp. 303–310). Xi'an, China: ACM.

Nowak, S., & Rüger, S. (2010). How reliable are annotations via crowdsourcing. In *Proceedings of the international conference on Multimedia information retrieval - MIR '10* (p. 557). New York, New York, USA: ACM Press. http://doi.org/10.1145/1743384.1743478

Pedhazur, E. J. (1982). *Multiple regression and behavioral science*. *Explanation and Prediction* (Vol. 2). New York, NY: Holt, Rinehart and Winston.

Peters, I. (2006). Against folksonomies - Indexing blogs and podcasts for corporate knowledge management. In *Online Information* (pp. 93–97). London, Great Britain.

Peters, I. (2009). *Folksonomies: Indexing and Retrieval in Web 2.0.*

Pirolli, P. (2007). *Information Foraging Theory Adaptive Interaction with Information*. Cary, NC, USA: Oxford University Press.

Plangprasopchok, A., Lerman, K., & Getoor, L. (2010). Growing a tree in the forest: Constructing folksonomies by integrating structured metadata. In *Proceeding of the 16th International conference on Knowledge discovery and data mining* (pp. 949–958). ACM. Retrieved from http://dl.acm.org/citation.cfm?id=1835924

Information Processing & Management, In Press, 2016

Rashtchian, C., Young, P., Hodosh, M., & Hockenmaier, J. (2010). Collecting Image Annotations Using Amazon ' s Mechanical Turk. In *Proceedings of the NAACL HLT 2010 Workshop on Creating Speech and Language Data with Amazon's Mechanical Turk* (pp. 139–147). Retrieved from http://aclweb.org/anthology/W/W10/W10-0721.pdf

Rivadeneira, A. W., Gruen, D. M., Muller, M. J., & Millen, D. R. (2007). Getting our head in the clouds: toward evaluation studies of tagclouds. In *Proceedings of the ACM Conference on Human Factors in Computing Systems* (pp. 995–998). New York, NY, USA. Retrieved from http://portal.acm.org/citation.cfm?id=1240775

Rolla, P. J. (2009). User tags versus subject headings: can user-supplied data improve subject access to library collections? *Library Resources & Techinical Services*, *53*(3), 174.

Rui, Y., Gupta, A., & Cadiz, J. (2001). Viewing meeting captured by an omnidirectional camera. In *Proceedings of ACM Conference on Human Factors in Computing Systems.* (pp. 450–457). ACM Press.

Seifert, C., Kump, B., Kienreich, W., Granitzer, G., & Granitzer, M. (2008). On the Beauty and Usability of Tag Clouds. In *Information Visualisation, 2008. IV '08. 12th International Conference* (pp. 17–25). http://doi.org/10.1109/IV.2008.89

Sigurbjornsson, B., & Zwol, R. Van. (2010). *TAGEXPLORER : Faceted browsing of flickr photos*.

Smith, G. (2008). Navigation and Visualization. In *Tagging: People-Powered Metadata for the Social Web* (pp. 95–116). New Riders.

Sorokin, A., & Forsyth, D. (2008). Utility Data Annotation with Amazon Mechanical Turk. In *Porceedings of the 2008 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops* (pp. 1–8). IEEE. http://doi.org/10.1109/CVPRW.2008.4562953

Spool, J. M., Perfetti, C., & Brittan, D. (2004). *Designing for the scent of information*.

Steele, T. (2009). The new cooperative cataloging. *Library Hi Tech*, *27*(1), 68–77. http://doi.org/10.1108/07378830910942928

Syn, S. Y., & Spring, M. B. (2013). Finding subject terms for classificatory metadata from user-generated social tags. *Journal of the American Society for Information Science and Technology*, *64*(5), 964–980.

Trattner, C., Körner, C., & Helic, D. (2011). Enhancing the Navigability of Social Tagging Systems with Tag Taxonomies. In *Proceedings of the 11th International Conference on Knowledge Management and Knowledge Technologies - i-KNOW '11* (p. 1). Graz, Austria. http://doi.org/10.1145/2024288.2024310

Trattner, C., Lin, Y., Parra, D., & Brusilovsky, P. (2012). Evaluating Tag-Based Information Access in Image Collections. In *Proceedings of the 23rd ACM conference on Hypertext and social media* (pp. 113–122). Milwaukee, Wisconsin, USA.: ACM.

Venetis, P., Koutrika, G., & Garcia-Molina, H. (2011). On the Selection of Tags for Tag Clouds. In *Proceedings of the fourth ACM international conference on Web search and data mining - WSDM '11* (pp. 835–844). Hong Kong, China. http://doi.org/10.1145/1935826.1935855

Information Processing & Management, In Press, 2016

Westerman, S., Cribbin, T., Wilson, R. (2001). Virtual information space navigation: evaluating the use of head tracking. *Behaviour and Information Technology*, *20*(6), 419–426.

White, R. W., & Marchionini, G. (2007). Examining the effectiveness of real-time query expansion. *Information Processing & Management*, *43*(3), 685–704. http://doi.org/10.1016/j.ipm.2006.06.005

White, R. W., & Roth, R. A. (2009). *Exploratory Search: Beyond the Query-Response Paradigm*. Morgan & Claypool Publishers.

Wu, D., He, D., Qiu, J., Lin, R., & Liu, Y. (2012). Comparing social tags with subject headings on annotating books: A study comparing the information science domain in English and Chinese. *Journal of Information Science*, *39*(2), 169–187. http://doi.org/10.1177/0165551512451808

Yee, K.-P., Swearingen, K., Li, K., & Hearst, M. (2003). Faceted metadata for image search and browsing. In *CHI '03 Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 401–408). Ft. Lauderdale, Florida, USA.: ACM, New York. http://doi.org/10.1145/642611.642681

Zubiaga, A., García-Plaza, A. P., Fresno, V., & Martínez, R. (2009). Content-based clustering for tag cloud visualization. In *2009 International Conference on Advances in Social Network Analysis and Mining* (pp. 316–319). IEEE. http://doi.org/10.1109/ASONAM.2009.19