



1-2013

# An exploratory study of user-centered indexing of published biomedical images

Sujin Kim

*University of Kentucky*, [sujinkim@uky.edu](mailto:sujinkim@uky.edu)

**[Click here to let us know how access to this document benefits you.](#)**

Follow this and additional works at: [https://uknowledge.uky.edu/pathology\\_facpub](https://uknowledge.uky.edu/pathology_facpub)



Part of the [Pathology Commons](#)

---

## Repository Citation

Kim, Sujin, "An exploratory study of user-centered indexing of published biomedical images" (2013). *Pathology and Laboratory Medicine Faculty Publications*. 1.

[https://uknowledge.uky.edu/pathology\\_facpub/1](https://uknowledge.uky.edu/pathology_facpub/1)

This Article is brought to you for free and open access by the Pathology and Laboratory Medicine at UKnowledge. It has been accepted for inclusion in Pathology and Laboratory Medicine Faculty Publications by an authorized administrator of UKnowledge. For more information, please contact [UKnowledge@lsv.uky.edu](mailto:UKnowledge@lsv.uky.edu).

---

**An exploratory study of user-centered indexing of published biomedical images**

**Notes/Citation Information**

Published in *Journal of the Medical Library Association*, v. 101, no. 1, p. 73-76.

© 2013, Authors.

**Digital Object Identifier (DOI)**

<http://dx.doi.org/10.3163/1536-5050.101.1.011>

# BRIEF COMMUNICATION

## An exploratory study of user-centered indexing of published biomedical images\*

Sujin Kim, PhD

See end of article for author's affiliation.

DOI: <http://dx.doi.org/10.3163/1536-5050.101.1.011>

### BACKGROUND

User-centered image indexing—often reported in research on collaborative tagging, social classification, folksonomy, or personal tagging—has received a considerable amount of attention [1–7]. The general themes in more recent studies on this topic include user-centered tagging behavior by types of images, pros and cons of user-created tags as compared to controlled index terms; assessment of the value added by user-generated tags, and comparison of automatic indexing versus human indexing in the context of web digital image collections such as Flickr. For instance, Golbeck's finding restates the importance of indexer experience, order, and type of images [8]. Rorissa has found a significant difference in the number of terms assigned when using Flickr tags or index terms on the same image collection, which might suggest a difference in level of indexing by professional indexers and Flickr taggers [9]. Studies focusing on users and their tagging experiences and user-generated tags suggest ideas to be implemented as part of a personalized, customizable tagging system. Additionally, Stvilia and her colleagues have found that tagger age and image familiarity are negatively related, while indexing and tagging experience were positively associated [10].

A major question for biomedical image indexing is whether the results of the aforementioned studies, all of which dealt with general image collections, are applicable to images in the medical domain. In spite of the importance of visual material in medical education and the prevalence of digitized images in formal medical practice and education, medical students have few opportunities to annotate biomedical images. End-user training could improve the quality of image indexing and so improve retrieval. In a pilot assessment of image indexing and retrieval quality by medical students, this study compared concept completion and retrieval effectiveness of indexing terms generated by medical students on thirty-nine histology images selected from the PubMed Central (PMC) database. Indexing instruc-

tion was only given to an intervention group to test its impact on the quality of end-user image indexing.

### METHODS

#### Medical students and indexing instruction

Emails were sent through the University of Kentucky's medical student email discussion list soliciting participants, and personal emails were sent to medical school faculty soliciting their support in recruiting participants. An exemption certificate was issued prior to recruitment by the University of Kentucky's Internal Review Board. All of the indexing worksheets along with study information and instruction were individually emailed to those students who agreed to participate in the study. A total of sixteen indexers agreed to participate. However, some participants were dropped as their results were incomplete. The final sample included twelve first- and second-year medical students.

The medical students were divided into 2 groups: Instruction Group (n=6) and No Instruction Group (n=6). The indexing worksheet given to both groups contained general study information, a list of hyperlinks to 39 randomly selected figure images, captions, and MEDLINE records. The students were told to look at the study images along with the associated captions to generate indexing keywords for the given images. If the captions were not clear, they were asked to review the MEDLINE records. Within a one-month period, all indexing results were returned for data analysis. The National Library of Medicine (NLM) training manual for subject analysis was used to develop a 7-minute-long, web-based instructional video <<http://128.163.165.41/IndexingTutorial.html>> [11]. This video includes a short introduction to microscopic images, concept of subject analysis, and how to assign topical keywords by introducing Medical Subject Headings (MeSH), Subheadings, Check Tags, and key concepts for microscopic images. A link to the video was given to half the students, the Instruction Group, to test whether or not the instruction affected indexing results.

#### Figure images and textual description sources

The study used histology/microscopic figure images and their accompanying textual descriptions, which were located by searching the PubMed Central (PMC) database. To focus on a certain disease area, the search was limited to only free, full-text articles on breast neoplasms. The search statement used was:

```
"((((("breast neoplasms"[MeSH Terms] OR breast neoplasms [Figure/Table Caption] OR breast neoplasms[Body - All Words] OR breast neoplasms[Title] OR breast neoplasms[Abstract]) AND figure[All Fields]) AND "breast neoplasms"[MeSH Major Topic]) AND "pathology"[Subheading]) AND ("1997/01/01"[PubDate] : "2008/08/31"[PubDate])
```

\* This project was supported in part by the Donald A. B. Lindberg Research Fellowship (2009) from the Medical Library Association (MLA). In addition, this publication was made possible by grant number P20RR-16481 from the National Center for Research Resources (NCRR), a component of the National Institutes of Health (NIH). Its contents are solely the responsibility of the author and do not necessarily represent the official views of NCRR or NIH.

A total of 353 valid figure images† out of 1,586 initially retrieved PMC records were identified, and only 39 images were selectively chosen for the indexing task given to the students. MeSH descriptors, the full-text of individual articles, and the accompanying caption of figure images were provided to both indexer groups.

### Measures of concept completion and retrieval effectiveness

The study used 2 measures, concept completion and retrieval effectiveness, to assess indexing by medical students. The measures were chosen based on results of previous studies that emphasized the importance of index term frequency when assessing coverage of indexed material [1–8]. Concept completion measured completeness of coverage of identified indexable units in individual images. To assess completeness of coverage, the terms assigned by student indexers were compared to a gold standard agreed upon by 2 domain experts. These experts—who had backgrounds in breast pathology, molecular studies, tissue bank, and imaging informatics—reviewed captions and images independently, and their results were combined into one gold standard. Concept completion scores were assigned between 0, representing no indexable concepts were identified, and 10, representing instances in which all the concepts identified by the 2 experts were present. Two study staff decided whether the same term was being indexed as those in the gold standard provided by the domain experts. If the 2 staff disagreed, the average score was used. These scores represent the percentage of terms assigned (e.g., 5 represent 50% of the experts' terms). Additionally, the indexing terms used by the medical students were mapped into Unified Medical Language System (UMLS) vocabularies through the MetaMap transfer engine to identify the terms' semantic types.

Retrieval effectiveness was defined as the ability to retrieve what was wanted and to avoid what was not wanted. To measure effectiveness, the topical keywords assigned by the students were entered into the PMC Image search engine. The unique term entries from each indexing set were searched in the PMC Images database‡ by using "OR" operators to identify correct histology images in a full study data set (n=353). Invalid images were manually determined

by 2 staff members with a background in pathology images and image indexing.

## RESULTS

### Concept completion

The Instruction Group (mean=6.94) generated fewer terms per image than the No Instruction Group (mean=7.88). However, this difference was not statistically significant ( $P>0.005$ ). The concept completion scores for the 2 groups were almost identical (No Instruction=5.76 vs. Instruction=5.79) ( $P>0.005$ ). For the aggregated index terms, the total number of index entries for the Instruction Group (n=3,925) was slightly less than that of the No Instruction Group (n=4,212). Slightly more duplicate entries (No Instruction=1,879 vs. Instruction=1,871) were assigned by the medical students without instruction. One thousand two hundred two of the assigned words (62%) were the same in both groups. The most popular words in common related to anatomy (e.g., breast neoplasms).

For the MetaMap result, the study found that the Instruction Group (n=1,633, 87.3%) generated more mapped terms than the No Instruction Group (n=1,473, 78.41%). This was a very high match score considering the poor mapping performance reported by other automatic indexing studies [12–14]. The No Instruction Group (n=115, 86.47%) mapped more MetaMap semantic types than the Instruction Group (n=100, 75.19%). It is interesting to note that the Instruction Group generated more of the semantic types such as Laboratory Procedure; Amino Acid, Peptide, or Protein Receptor; and Medical Device, Research Device, that were emphasized in the training video as a part of core components for histology images. Although the ranked order of the identified semantic types, as mapped, was different between two groups, the sets of semantic types identified from both groups were similar to each other (Table 1).

### Retrieval effectiveness

The terms assigned by the Instruction Group found more of the study images (n=310) than the terms assigned by the No Instruction Group (n=111). The difference was almost triple ( $P<0.05$ ). However, approximately 70% of the all retrieved images were incorrect (not histology/microscopic images such as graphs, tables, gel images, etc.). It appears likely that the Instruction Group assigned the least general terms and that reduced the recall of correct histology images.

## DISCUSSION AND CONCLUSION

The main purpose of this study was to assess image indexing quality of keywords that medical students assigned to images. The major findings from this study were that concept completion was only 50% compared to experts and that effectiveness was limited by the large number of incorrect images retrieved. The instruction group used fewer terms, but these terms produced

† Graphs or tables or non-histology/microscopic images were out of scope of this study, so those images and the captions were removed. After collecting an initial set of figure images and accompanying captions, the study manually screened microscopic images for the analysis.

‡ The newly launched National Center for Biotechnology Information (NCBI) Images database <<http://www.ncbi.nlm.nih.gov/images>> allows one to search published PubMed Central (PMC) Figure Images by basic search parameters such as All Fields, Author, Figure Caption, Filter, Image Height and Width, Text Words, and so on.

**Table 1**  
Semantic types of mapping results using the National Library of Medicine's MetaMap 2010 by indexer groups

Students without instruction	Frequency	Students with instruction	Frequency
Neoplastic Process	134	Gene or Genome	84
Gene or Genome	108	Neoplastic Process	83
Functional Concept	93	Body Part, Organ, or Organ Component	35
Biomedical Occupation or Discipline	78	Laboratory Procedure	29
Body Part, Organ, or Organ Component	68	Qualitative Concept	23
Qualitative Concept	62	Amino Acid, Peptide, or Protein, Biologically Active Substance	22
Laboratory Procedure	53	Amino Acid, Peptide, or Protein, Receptor	16
Pathologic Function	41	Cell Function	16
Amino Acid, Peptide, or Protein, Biologically Active Substance	38	Mammal	15
Organism Function	31	Amino Acid, Peptide, or Protein, Immunologic Factor	12
Amino Acid, Peptide, or Protein, Receptor	24	Finding	12
Finding	23	Genetic Function	11
Amino Acid, Peptide, or Protein, Enzyme	22	Functional Concept	10
Biologic Function	22	Therapeutic or Preventive Procedure	10
Amino Acid, Peptide, or Protein, Immunologic Factor	21	Indicator, Reagent, or Diagnostic Aid, Organic Chemical	9
Mammal	21	Biomedical Occupation or Discipline	8
Cell	20	Cell	8
Cell Function	20	Quantitative Concept	8

better retrieval results in the limited PMC search. The identified sets of the semantic types found in this study can be used in further development of retrieval systems as a guide to essential indexing components for histology images. In publishing, as a standard part of article submission (or in pathology practice), journals could require that images (figures) reporting molecular or histology findings about cell compounds, antibody, staining, tissues, and magnification along with quantitatively measured study data include information in the image legend for better indexing and retrieval.

The medical students were quite impressed with the video, which indicated an interest in learning how to analyze the contents (subject matter) of biomedical images. This interest suggests offering brief training on how to index biomedical images (as well as biomedical contexts) early in medical education. Considering the fast growth of social networks, it will be important to train end users in how to assign subject terms so that tagged content (whether images or not) becomes retrievable information on the web. Throwing "stuff" into "never-discoverable" web spaces will not advance health care in any manner. With the advancement of biomedical imaging technologies, the value of images will be improved by adding effective and relevant descriptions combined with thematic descriptors. In this sense, studies on mapping user-given tags and evaluating the retrieval effectiveness of images relevant to concept completion can improve the use of biomedical images in both practice and research. More value would certainly be added to the practices currently used by medical librarians.

## REFERENCES

- Matusiak K. Towards user-centered indexing in digital image collections. *OCLC Syst Serv.* 2006;22(4):283–98.
- Rorissa A. User-generated descriptions of individual images versus labels of groups of images: a comparison using basic level theory. *Inf Process Manage.* 2008 Sep;44(5):1741–53.
- Muller H, Clough P, Hersh W, Deselaers T, Lehmann T, Geissbuhler A. Evaluation axes for medical image retrieval systems: the imageCLEF experience. In: *Proceedings of the 13th Annual ACM International Conference on Multimedia (MULTIMEDIA '05)*. New York, NY: Association for Computing Machinery. p. 1014–022.
- Chung E, Yoon J. Categorical and specificity differences between user-supplied tags and search query terms for images: an analysis of Flickr tags and web image search queries. *Inf Res [Internet]*. 2009 Sep;14(3), paper 408 [cited 11 Jul 2012]. <<http://InformationR.net/ir/14-3/paper408.html>>.
- Golder SA, Huberman BA. Usage patterns of collaborative tagging systems. *J Inf.* 2006;32(2):198–208.
- Morrison PJ. Tagging and searching: search retrieval effectiveness of folksonomies on the World Wide Web. *Inf Processing Manage.* 2008;44:1562–79.
- Rorissa A, Iyer H. Theories of cognition and image categorization: what category labels reveal about basic level theory. *J Am Soc Inf Sci Technol.* 2008 Dec;59(9):1383–92.
- Golbeck J, Koepfler J, Emmerling E. An experimental study of social tagging behavior and image content. *J Am Soc Inf Sci Technol.* 2011;62(9):1750–60.
- Rorissa A. A comparative study of Flickr tags and index terms in a general image collection. *J Am Soc Inf Sci Technol.* 2010;61(11):2230–42.
- Stvilia B, Jørgensen C, Wu S. Establishing the value of socially created metadata to image indexing. *Lib Inf Sci Res.* 2012;34(2):99–109.
- National Library of Medicine, Bibliographic Service Division. MEDLINE indexing: online training course [Internet]. The Library [cited 23 Jan 2009]. <[http://www.nlm.nih.gov/bsd/indexing/training/USE\\_010.htm](http://www.nlm.nih.gov/bsd/indexing/training/USE_010.htm)>.
- Pratt W, Yetisgen-Yildiz M. A study of biomedical concept identification: MetaMap vs. people. *AMIA Annu Symp Proc.* 2003;529–33.
- Osborne JD, Lin S, Zhu L, Kibbe WA. Mining biomedical data using MetaMap Transfer (MMtx) and the

Unified Medical Language System (UMLS). *Methods Mol Biol.* 2007;408:153–69.

14. Ben Abacha A, Zweigenbaum P. Automatic extraction of semantic relations between medical entities: a rule based approach. *J Biomed Semantics.* 2011 Oct 6;2(suppl 5):S4.

#### AUTHOR'S AFFILIATION

**Sujin Kim, PhD**, sujinkim@uky.edu, Associate Professor, Division of Biomedical Informatics, College of

Public Health, School of Library and Information Science, College of Communication and Information Studies and Department of Pathology and Laboratory Medicine, College of Medicine, University of Kentucky, 230 G Multidisciplinary Building, Lexington, KY 40536

*Received January 2012; accepted July 2012*

## The Leading eBook Platform for Health Science Collections. Stop Searching. Start Finding.

The R2 Digital Library is a market-leading eBook platform for health science collections featuring a comprehensive selection of medical, nursing and allied health eBooks with an intuitive interface optimized for your library.

### Specialized

Wide selection from leading STM publishers

Extensive selection of Doody's® Core Titles

High yield and focused resources

### Customized

Purchase only the resources you need

Access titles for the life of the edition

Suitable for collections of all sizes and specialties

### Optimized

Integrated drug information

Extensive image library  
Enhanced discoverability

Clean, intuitive interface



If you haven't seen the R2 Digital Library lately, don't wait. Visit us today at [www.r2library.com](http://www.r2library.com).

**800-345-6425**  
**Rittenhouse.com**  
**R2Library.com**

**Rittenhouse**  
BOOK DISTRIBUTORS

DISTRIBUTING THE LEADING SCIENTIFIC, TECHNICAL AND MEDICAL PUBLISHERS

**R2** DIGITAL  
LIBRARY



511 Fehleley Drive | King of Prussia, PA 19406