

June Abbas

State University of New York at Buffalo, Buffalo, New York, USA

In the margins: Reflections on scribbles, knowledge organization, and access

Extended abstract for the position paper submitted to the North American Symposium on Knowledge Organization (NASKO), Toronto, Ontario, 2007. Full paper to appear in *Knowledge Organization* vol. 34 no. 2 (2007).

Extended abstract: A favored text, dog-eared and yellowed from use, yet still useful, brings back insights that we try to impart to our students when we teach knowledge organization, organization and control of recorded information courses, whichever words we have chosen to label them. Scribbled in the margins¹ are notes to self, keywords, subject headings? “tags”? to remind us of why this particular passage was relevant to us. These scribbles include notes about the thoughts, subjects, eloquent linguistics that we wish to remember, and to access at a later time, maybe even our thoughts that occurred as we read the words. Should someone pick up this same text and read the passages and also the notes, would one necessarily draw the same conclusions, or would one have yet other insights into the author’s meanings, the scribbles, the words?² Wilson (1968) reminds us that “What a text says is not necessarily what it reveals or what it allows us to conclude. . . . but what is not said may interest us more than what is said” (p. 18). How then do we access the facts, truths, assertions, that the text conveys, or doesn’t convey, or the different truths, assertions, that occur to another when they read the text? Our knowledge organization structures provide access points to follow. Classification schemes, controlled vocabularies, ontologies, taxonomies, and the like, have been used to access various levels of subject content within the texts.³ How then, do we access the “meaning”, the conclusions, insights others’ make while reading the words, the scribbles in the margins?

This is an old argument. Knowledge organization structures are not static. We struggle to update classification schemes and conduct research to determine if they work. Controlled vocabularies have been criticized as being out of date, containing arcane, discriminatory, Anglo-centric terminology (Olson, 2002). We have conducted studies that show that users don’t understand how to use subject headings (Markey, 1984; Drabenstott and Vizine-Goetz, 1990), or that the words they choose for searching do not match subject headings (Taylor, 1984; Carlyle, 1989; Doyen and Wheeler, 1989; Lester, 1989; Abbas, 2001). So what have we done with the knowledge we gained from this research? Has it changed our way of thinking about knowledge organization and subject access?

On the surface, it seems the Web has taken much of knowledge organization out of our hands. Users can access this vast depository of texts by entering a few words into a search box,

and they do. Studies have shown us that most web searchers are not concerned with thinking up precise, well defined Boolean search strings. They enter a few key (relevant to them) words and click a button. They then sift through the multitude of hits and find at least one or more that satisfice their information need. In online collaborative sharing communities, such as Flickr (<http://www.flickr.com>), del.icio.us (<http://www.delicious.com>), and LibraryThing (<http://www.librarything.com>), users can organize images, cluster bookmarks, and catalog their own personal libraries, etc. using words that are relevant to them. They are not using our knowledge organization devices. They are creating their own as they use/view others' tags. Vander Wahl (2006) has been credited with coining the term "folksonomy", or the resulting cluster of terms that emerges when a community describes texts. Folksonomies are then used for subject representation. Other proponents of this concept/or the process of enabling subject access using user-defined descriptors are: Hastings (1995); O'Connor (1996); Bates (1998); O'Connor, O'Connor, and Abbas (1999); Abbas (2001, 2005) to name a few. The Web gave us an environment to test the efficacy of using user-defined descriptors for subject (as well as physical) access. We might then assume that collaborative sharing communities are in effect, "scribbling in the margins" when they tag their images, their bookmarks, their libraries.⁴ They are choosing a few words or phrases to represent the "meaning" of the text to them. They are then re-using these words as their own controlled vocabularies. Others are sometimes invited to provide their own tags, thereby providing *their* meaning for the object. Tag clouds (the resulting structures built as a result of tagging objects) then become visual representations of meaning to at least this one user, microcommunities, and to a larger society of users. Tag clouds become mechanisms not just for representation, but for retrieval.

Blair (1990) provides a further context for examining social representation and access issues. He posits that the language we use to represent both our information needs and to index texts is learned in a social context or community. Blair explains the theory of "language games", as first developed by the early twenty century philosopher Ludwig Wittgenstein and the process in which we learn language and meaning. We do not acquire language purely by learning the word and its definition, but instead learn its use and appropriateness within the context of our "forms of life" or everyday experiences. Furthermore, we have to possess some prior understanding of the form of life or the language game context we are engaged in before the words can have meaning. Users of online sharing communities are engaging within the social context of a particular community. Each person who contributes tags is engaging in "language games" as they go through their daily "forms of life" or experiences. Where this practice may differ from Wittgenstein's conception, is that there are few limits on what is accepted or unaccepted practice. Users can tag using their own constructions, experiences, meanings, with the only limits imposed being of technological nature.

So, where does this leave us? Where do we go from here? We have a rich source that is untapped. Our OPACs gather users' search terms and search sessions. Websites also track and collect this same information about access. Online collaborative sharing sites are developing folksonomies. Each of these sources can tell us volumes about how our users access information. These sources provide us with a glimpse into user's perceptions and cognitive processes as they scribble in the margins. At the very least, these sources provide us with the

terms used, and with further study, may potentially provide contextual meaning. What we need to consider now is *how* we can use these sources to adapt, augment, revitalize our knowledge organization structures. There are efforts underway to do just this. Museum and library communities, for example, are exploring the usefulness, as well as logistics, of gathering and incorporating users' tags into their websites, online exhibits, and WebPACs (Trant, 2006; Spiteri, 2006; Sweda, 2006). Digital libraries that have been developed for youth are also exploring the idea of using user-defined descriptors as subject headings (Abbas, 2001, 2005; Reuter and Druin, 2004).

More needs to be considered. More needs to be learned. What do we know about social classification, tagging, and its meaning and use for users? Potential areas of exploration include:

- What does tagging mean to users? Is it a way to describe a text, a scribble in the margins, or a search term? Are these potential uses different to users?
- What are users' motivations for tagging (personal findability or organization; communal or familial sharing; meaning making)?
- Can we apply Wittgenstein's "Language Games" theory to what is happening in online sharing communities? Can this inform knowledge organization theory and practice?
- What can we learn from collaborative classification, folksonomy development? How can we incorporate this learning into classification scheme and controlled vocabulary development? Should we try to make tags more consistent and follow knowledge organization conventions or do we just watch and learn? Can we/should we apply traditional controlled vocabulary constraints on user-defined descriptors?

Notes

1 The author is in no manner condoning the practice of writing in the margins of library or other's books, but keeps this practice only at a local level.

2 The reader is reminded of the impact on scientific discovery accomplished by reading someone else's notes in the margins. Johannes Kepler's work on elliptical orbits was influenced by notes he read in the margins of a second-hand copy of Copernicus' *De revolutionibus*. (Gingerich, 2004).

3 Texts for this discussion could include any information bearing object, regardless of format, but to maintain the "argument" being developed, the word "text" will be used.

4 Use of the term scribbling in either context should in no way indicate a quick, easy process void of thought or consideration. Some tags may be created quickly, but others are only applied after much deliberation, examination of existing tags, or even by using the tag clouds, or other social classification structures of the community.

References

Abbas, J. (2001). *Smoothing the information seeking path: Removing representational*

Abbas, J. (2007). In the margins: Reflections on scribbles, knowledge organization, and access. (extended abstract) *Proceedings of the North American Symposium on Knowledge Organization*. Vol 1. Available: <http://dlist.sir.arizona.edu/1914>

- obstacles in the middle school digital library environment*. University of North Texas, Denton, TX : Unpublished doctoral dissertation.
- Abbas, J. (2005). Out of the mouths of middle school children: I. Developing user-defined controlled vocabularies for subject access in a digital library. *Journal of the American Society for Information Science and Technology*, 56(14): 1512-1524.
- Bates, M.J. (1998). Indexing and access for digital libraries and the Internet: Human, database, and domain factors. *Journal of the American Society for Information Science*, v. 49(13): 1185-1205.
- Blair, D. (1990). *Language and representation in information retrieval*. Amsterdam, The Netherlands: Elsevier Science.
- Carlyle, A. (1989). Matching LCSH and user vocabulary in the library catalog. *Cataloging and Classification Quarterly*, 10(1/2): 37-63.
- Doyen, S. and Wheeler, D. (1989). Use of a controlled vocabulary index in information retrieval tasks. In Salvendy, G. and M. Smith (Eds.), *Designing and Using Human-Computer Interfaces and Knowledge Based Systems*. Amsterdam, The Netherlands: Elsevier Science.
- Drabenstott, K. and Vizine-Goetz, D. (1990). Search trees for subject searching in online catalogs. *Library Hi Tech*, 8(3): 7-20.
- Gingerich, O. (2004). *The book nobody read: Chasing the revolutions of Nicolaus Copernicus*. New York: Penguin Books.
- Hastings, S. (1995). An exploratory study of intellectual access to digitized art images. In *Proceedings of the Fifty-ninth Annual Meeting of the American Society for Information Science* (pp. 3-8). Medford, NJ: Information Today.
- Markey, K. (1984). *Subject searching in library catalogs: Before and after the introduction of online catalogs*. Dublin, OH: Online Computer Library Center.
- O'Connor, B. (1996). *Explorations in indexing and abstracting: Pointing, virtue and power*. Englewood, CO: Libraries Unlimited, Inc.
- O'Connor, B., O'Connor, M. and Abbas, J. (1999). User reactions as access mechanism: An exploration based on captions for images. *Journal of the American Society for Information Science*, 50(8): 681-697.
- Olson, H. (2002). *The power to name: Locating the limits of subject representation in libraries*. Dordrecht [The Netherlands]: Kluwer Academic.
- Reuter, K. and Druin, A. (2004). Bringing together children and books: An initial descriptive study of children's book searching and selection behavior in a digital library. *Proceedings of the 67th ASIST Annual Meeting*, (pp. 339-348). Medford, NJ: Information Today.
- Spiteri, L. *The use of collaborative tagging in public library catalogues*. Poster presented at the 17th Annual ASIS&T SIG/CR Classification Research Workshop Saturday, November 4, 2006, Austin, TX.
- Sweda, J. (2006). *Using social bookmarks in an academic setting: PennTags*. Poster presented at the 17th Annual ASIS&T SIG/CR Classification Research Workshop Saturday, November 4, 2006, Austin, TX.

Abbas, J. (2007). In the margins: Reflections on scribbles, knowledge organization, and access. (extended abstract) *Proceedings of the North American Symposium on Knowledge Organization*. Vol 1. Available: <http://dlist.sir.arizona.edu/1914>

- Taylor, A. (1984). Authority files in online catalogs: An investigation of their value. *Cataloging & Classification Quarterly*, 4(3): 1-17.
- Trant, J. (2006). *Social classification in art museums: steve.museum*. Panel paper presented at the 17th Annual ASIS&T SIG/CR Classification Research Workshop Saturday, November 4, 2006, Austin, TX.
- Vander Wal, T. (2006, September 30). *Folksonomy to improve IA*. Presented at the OZ IA, Sydney, Australia. Retrieved May 14, 2007, from http://s3.amazonaws.com/2006presentations/OZIA/Folksonomy_for_IA.pdf
- Wilson, P. (1968). *Two kinds of power: An essay on bibliographic control*. Los Angeles, CA: University of California Press.