Book Reviews

Edited by Joseph T. Tennis

Book Review Editor

Brian C. O'Connor, Jodi Kearns, and Richard L. Anderson. *Doing Things with Information: Beyond Indexing and Abstracting*. Westport, Connecticut: Libraries Unlimited, 2008. xix, 241 pages.ISBN: 987-1-59158-577-0.

The authors state that this book emerged from a proposal to do a second edition of *Explorations in Indexing and Abstracting* (O'Connor 1996); much of its content is the result of the authors' reaction to the reviews of this first edition and their realization for "the necessity to address some more fundamental questions" (p. ix). The scope, goals, and objectives of this book are not stated as clearly, perhaps, as one would like: The authors discuss a number of points they wish to address, but what is lacking is a coherent statement of the purpose of this book. The final chapter of the book provides a clearer outline of this purpose, namely to present readers with new ways "to think about messages in all sorts of media and how they might be discovered, analyzed, synthesized, and generated" and to "[bring] together philosophical, scientific, and engineering notions into a fundamental model for just how we might understand doing this with information" (p. 225). Much of the preface is concerned with addressing reviewers' comments of *Explorations in Indexing and Abstracting;* while the authors justify this approach by saying that it presents a snapshot of the thinking by the three authors during the early period of constructing this book, this approach does tend to muddy the clarity of the scope of the book. Readers must interpret the subtitle of this book with caution; as the authors state in the preface, this book is not specifically about indexing and abstracting, nor is it in any way a primer for these two activities. Once again, this caveat is made clear in the concluding chapter, rather than stated explicitly in the preface. The authors are influenced heavily by Patrick Wilson's vision of interpreting and understanding information (Wilson 1977, 1983). We are told that when we design representation systems (e.g., indexing and abstracting), we need to "exhaust all possible representations of each document, considering relevance to the smallest granularities of both document meanings and structures" (p. xix). Until we have done this, we have not done our jobs as indexers, abstractors, or cataloguers, which the authors call surrogate engineers: An ambitious task indeed for any surrogate engineer.

The book is organized as follows:

Preface
1. Background concepts and models
2. Considerations of representation
3. Representation, function, and utility
4. Failures of representation: Indeterminacy and depth
5. Aboutness and user-generated descriptors
6. Responses to indeterminacy
7. Doing things with word-based documents
8. Functional application of information measurement
9. Functional ontology construction
10. Creek pebbles: A summary metaphor and touchstone for exploration

The first nine chapters of the book consist of discussions and analyses of different methods by which can be modelled the relationship amongst the user, the document, and the environment in which they exist. The user is presented as a person with a question. The authors prefer this phrase to the perhaps more common "information need." The document is the potential source of information to address this question. The authors argue that document representation systems often fail to address the person's questions because they do not adequately take into account the varying and individual nature of the relationship between a user and a document. These systems need to consider:

– How purpose influences mode of representation (e.g., the author's stance on a topic),
– No representation without a common code (i.e., users often do not understand the language used

in the systems to describe subject content, such as Library of Congress Subject Headings (LCSH), and the rules used to design this language
– Synchronic and diachronic attributes. Library and Information Science (LIS) is good at representing the diachronic (or unchanging) attributes of a document, such as title and author, but not at representing the synchronic attributes, e.g., the individual's interpretation of a document, validity of the data, etc.

These three considerations are discussed throughout the first eight chapters. The authors do a good job of presenting interdisciplinary perspectives of these considerations, drawing from philosophy, engineering, information science, behaviourism, and so forth. The authors' focus on the centrality of the user in the design of document representation systems is not new.

Calvin Mooers' (1950) Law that "an information retrieval system will tend not to be used whenever it is more painful and troublesome for a customer to have information than for him not to have it," emphasized the need to design indexing terms that address the needs of users for the document. In her analysis of the two prevailing approaches to indexing, Fidel (1994) points to (a) the document-oriented approach (e.g., Borko and Bernier 1978; Rowley 1988), whereby indexing represents the content of a document in a prescribed indexing language, and (b) the user-oriented approach, whereby indexing reflects the requests and needs of the users for which the documents may be relevant (e.g., Lancaster 1991; Soergel 1995). The fact that indexing languages such as LCSH are not always readily understood by searchers has been a topic of discussion in LIS for some time (e.g., Carlyle 1989; Greenberg 2006; Yee 1991). While these two preceding points are understandably important to an understanding of the relationship between a user and a document and, by extension, the document's representation, the authors do not offer much that is new to this discussion. It is in their discussion of synchronic and diachronic attributes in indexing languages that the authors cover ground that is just emerging in LIS; for example, Tennis (2007) posits a model that focuses on largely diachronic attributes in indexing languages, although he does not make the same point as the authors about the unchanging attributes of these languages. The discussion pertaining to diachronic and synchronic attributes is an illustration of an aspect of the book that readers may find frustrating: The authors point to inherent problems or inadequacies with document representation systems, but little in the way of suggested solutions. Although the book does not purport to be an indexing manual, the discussion of at least some tangible solutions would certainly enhance the authors' arguments.

The authors' discussion of two specific problems with document representation systems, namely indeterminacy and depth, once again do not cover new ground. Indexers and cataloguers have been struggling for many years with maintaining an ideal balance between the recall (affected by the number of indexing terms assigned) and precision (affected by the specificity of the indexing terms assigned) of searches (e.g., Khosh-Khui 1987; Sparck Jones 1972; Svenonius 1971). The authors' suggestions for determining optimal level of indexing depth and specificity have guided indexers for several years, namely:

– extract whichever elements are useful to the patron
– extract however many elements that are necessary for the patron
– employ whatever form is consistent with patron abilities and requirements.

A noticeable omission in this book is the scant attention the authors pay to social tagging and folksonomies. User-generated descriptors are mentioned in Chapter 5, but are limited primarily to information retrieval systems, rather than to social tagging and bookmarking applications. Given that this book was written within the past two years, and the explosion of social tagging research within LIS, this omission seems especially surprising.

Chapter 9 presents the culmination of the authors' discussion in the presentation of their "Functional Ontology Construction" (FOC) model, which they suggest could be used to map the relationship between the user, the document, and the environment in which they exist. One would have hoped that Chapter 9 would enable the reader to "connect the dots" of the arguments and discussion of the preceding eight chapters; this is not the case, however. Most of Chapter 9 discusses the authors' rationale for basing their model on pragmatism and B.F. Skinner's Radical Behaviourism (1953); only the last three pages of the chapter focus specifically on the FOC model. Since Skinner has had a long foothold in LIS, particularly with respect to research in the area of relevance judgments (Saracevic 2006), perhaps less emphasis could have been placed on the rationale for the FOC, especially since Chapters 1-8 set the stage for this model. More time could have been spent on the model itself. Chapter 9 high-

lights a structural problem that occurs throughout most of the chapters: While the authors provide useful examples to highlight various concepts and theories, they often do so before, rather than after the concept is introduced. As a reader, I often had to read through a number of rather long examples before the actual concept under discussion was presented. I think it would be more helpful to first present the reader with a clear discussion of a concept and to then present the examples, especially as in many cases, more attention was paid to the examples, than to the concepts.

This book provides a good overview of the relationship between the document and the user; in this regard, it reinforces the importance of the client-centred approach to the design of document representation systems. In the final chapter, the authors state: "We have offered examples of new ways to think about messages in all sorts of media and how they might be discovered, analyzed, synthesized, and generated. We brought together philosophical, scientific, and engineering notions into a fundamental model for just how we might understand doing this with information" (p. 225). The authors have certainly succeeded in highlighting the complex processes, nature, and implications of document representation systems, although, as has been seen, the novelty of some of their discussions and suggestions is sometimes limited. With further explanation, the FOC model may serve as a useful way to understand how to build document representation systems to better meet user needs.

References

Borko, Harold, and Bernier, Charles L. 1978. *Indexing concepts and methods*. New York: Academic Press.

Carlyle, Allyson. 1989. Matching LCSH and user vocabulary in the library catalog. *Cataloging & classification quarterly* 10(1/2): 37-63.

Fidel, Raya. 1994. User-centered indexing. *Journal of the American Society for Information Science* 45: 572-76.

Greenberg, Jane. 2006. User comprehension and searching with information retrieval thesauri. *Cataloging and classification quarterly* 37(3/4): 203-20.

Khosh-Khui, Sam A. 1987. Effects of subject specificity. Part I: Specificity of LC subject headings and depth of subject analysis in monographic records. *Technical services quarterly* 4(2): 59-67.

Lancaster, F. Wilfrid. 1991. *Indexing and abstracting in theory and practice*. London: The Library Association.

Mooers, Calvin N. 1950. The theory of digital handling of non-numerical information and its implications to machine economics. *Proceedings of the Meeting of the Association for Computing Machinery*.

O'Connor, Brian C. 1996. *Explorations in indexing and abstracting: Pointing, virtue, and power.* Englewood, CO: Libraries Unlimited.

Rowley, Jennifer E. 1988. *Abstracting and indexing,* 2nd ed. London: Clive Bingley.

Saracevic, Tefko. 2006. Relevance: A review of the literature and a framework for thinking on the notion in Information Science. Part II. *Advances in librarianship* 30: 1-69.

Skinner, Burrhus F. 1953. *Science and human behavior.* New York: Macmillan.

Soergel, Dagobert. 1985. *Organizing information: Principles of database and retrieval systems*. Orlando, FL: Academic Press.

Sparck,Karen J. 1972. A statistical interpretation of term specificity and its application to retrieval, *Journal of documentation* 28: 11–20.

Svenonius, Elaine. 1971. *The effect of indexing specificity on retrieval performance*. Chicago: University of Chicago. Graduate Library School.

Tennis, Joseph T. (2007). *Diachronic and synchronic indexing: Modeling conceptual change in indexing languages*. Retrieved July 9, 2010, from http://www.caisacsi.ca/proceedings/2007/tennis_2007.pdf

Wilson, Patrick G. 1977. *Public knowledge, private ignorance: Toward a Library and Information policy*. Westport, CT: Greenwood Press.

Wilson, Patrick G. 1983. *Second-hand knowledge: An inquiry into cognitive authority*. Westport, CT: Greenwood Press.

Yee, Martha M. 1991. Improving LCSH for use in online catalogs. *Journal of the American Society for Information Science* 42: 78-98.

Louise F. Spiteri, Associate Professor
School of Information Management
Dalhousie University.
Halifax. Nova Scotia, B3H 3J5. Canada.
Louise.Spiteri@dal.ca