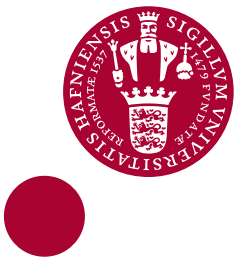

The Importance of Theories of Knowledge

Indexing and Information retrieval as an example

Hjørland, Birger

Published in:
Journal of the American Society for Information Science and Technology

DOI:
10.1002/asi.21451

Publication date:
2011

Document version
Early version, also known as pre-print

The Importance of Theories of Knowledge: Indexing and Information Retrieval as an Example

Birger Hjørland
Royal School of Library and Information Science, 6 Birketinget, DK-2300 Copenhagen S, Denmark.
E-mail: bh@iva.dk

A recent study in information science (IS), Lykke and Eslau (2010; hereafter L&E), raises important issues concerning the value of human indexing and basic theories of indexing and information retrieval, as well as the use of quantitative and qualitative approaches in IS and the underlying theories of knowledge informing the field. The present article uses L&E as the point of departure for demonstrating in what way more social and interpretative understandings may provide fruitful improvements for research in indexing, knowledge organization, and information retrieval. The artcle is motivated by the observation that philosophical contributions tend to be ignored in IS if they are not directly formed as criticisms or invitations to dialogs. It is part of the author's ongoing publication of articles about philosophical issues in IS and it is intended to be followed by analyzes of other examples of contributions to core issues in IS. Although it is formulated as a criticism of a specific paper, it should be seen as part of a general discussion of the philosophical foundation of IS and as a support to the emerging social paradigm in this field.

Introduction

The purpose of the present article is to demonstrate implications of theories of knowledge by considering a published empirical study in IS from a theory of knowledge point of view and analyzing its epistemological position and demonstrating how an alternative view contributes to the further advancement of the field. Lykke and Eslau's (2010; hereafter L&E) study is important and interesting; the authors raise the question of whether human indexing has *any* advantages compared with automated indexing and they provide empirical indication that it hasn't. The purpose of our analysis in this article is to examine this from the point of view of epistemology/theory of knowledge, thereby arguing that there are alternative ways to look at this important issue and such alternative approaches may be able to advance IS in a fruitful way.

Initial view of L&E's (2010) Paper

The L&E study (which is openly available on the Internet) investigates three search strategies used in a pharmaceutical electronic document management system (EDMS):

> Strategy 1: A metadata search strategy based on human controlled indexing.
> Strategy 2: A simple natural language strategy based on automatic indexing.
> Strategy 3: An advanced natural language strategy based on automatic indexing and with the use of a corporate thesaurus for query expansion.

The controlled indexing in strategy 1 was made by using a corporate thesaurus developed in 2001, which comprises 5,600 preferred, controlled terms (concepts) used for human indexing and a total of 16,000 terms. Indexing was performed by authors, librarians, assistants, and research staff. The same thesaurus was also used in strategy 2 (using synonyms and narrower terms, including their synonyms).

The evaluation framework was experimental (Lykke & Eslau, p. 91). Testing was performed in a database that comprised 25,384 documents and 10 "realistic" search tasks (SJ[1]1-SJ10) were developed. Relevance was measured by a 4-point scale assessment (highly relevant, fairly relevant, marginally relevant, and irrelevant). Precision and relative recall was calculated for each strategy and search task.

The study stresses that it is made in the context of enterprise retrieval systems. The findings and the issues may, however, also be relevant for traditional bibliographical databases using thesauri or other kinds of controlled vocabularies. It is, thus, assumed that the issues discussed by both L&E and the present article may have very broad

Received August 13, 2010; revised September 27, 2010; accepted September 28, 2010

 DOI: 10.1002/asi.21451

[1]SJ stands for "Search Job." They are used in Table 1–4 of L&E, pp. 92–93.

TABLE 1. Human versus automated indexing and derived versus assigned indexing.

	Human-based indexing	Automated indexing
Derived indexing	a) Human-derived indexing	b) Automatically derived indexing
Assigned indexing	c) Human-assigned indexing	d) Automatically assigned indexing

implications. We found the findings surprising. In general, the expanded search queries performed best on both recall and precision. The metadata searches resulted in the lowest precision, which is remarkable because the human indexers were tailored specifically to meet information tasks such as those search tasks constructed for the experiment. It is also remarkable that the expanded search retrieved more of the highly relevant documents (which is described as discouraging, Lykke & Eslau, p. 93).

The authors write (Lykke & Eslau, p. 94) that the study was a "case study" and therefore the findings do not provide any conclusive results as to whether a thesaurus is better used as support for expanded natural language searches compared with controlled metadata searching.[2]

Why choose L&E as the case for criticism? If specific criticisms can be raised against L&E, then why not choose a study in which such criticism cannot be raised? Also, why choose a paper in a festschrift rather than a journal? Or, in what way can L&E be considered representative for research in information science?

The answer is that L&E was selected because it—more than any other study I have encountered—provided an opportunity to formulate important arguments of a very general nature. From an outside view it may seem easy to select a paper for this kind of analysis, but this looks differently when you start searching. L&E is an important study with some challenging findings, which call out for closer analysis and a published study has to be considered as a contribution whether it is a journal article, a conference paper, a festschrift article, or anything else. One cannot assume that one kind of publication is preliminary and later to be replaced by another. A more relevant view is that any paper is part of an ongoing development in the field and that both authors and citers influence the development of future papers in a dialogical manner. Concerning the representativity of L&E, Saracevic (2008, p. 772) has described an important kind of study that is no longer made in IS. In this way, L&E is typical for research in IS. Also the neglect of theoretical issues in indexing and retrieval is typical. My argument below, using L&E as a representative example, is that such neglects may be fatal for the field.

[2]Is there perhaps a contradiction in claiming that the study is both experimental and a case study? At least we are not informed about what more is needed—in the eyes of the authors—before we may confidently conclude as to whether a thesaurus is better used as support for expanded natural language searches compared with controlled metadata searching. These problems regarding the generalization of the findings are addressed below in the section "Back to L&E."

Human Based Indexing Versus Automated Indexing

First, we shall consider the distinctions made by L&E (Lykke & Eslau, p. 87) between manual human indexing versus automatic indexing and between assigned and derived indexing. For clarity, we shall consider the four alternative possibilities in more detail than L&E: AQ1

a. Human-derived indexing is done, for example, when humans underline (or highlight) words and phrases in texts to be used as index terms. This can be done more or less mechanically, depending on the instructions given to the indexer as well as the indexer's understanding of the documents and the indexer's anticipation of the kinds of questions that the documents (and thus the indexing) are supposed to help answer.

b. Automatically derived indexing is done, for example, when a computer program selects all terms from texts with the possible exception of "stopwords" (listed in a "stoplist"). Again, the production of the stoplist may be more or less mechanical or based on experiences with particular genres and domains. Also, other criteria may be used to select words from texts. It is common to select terms based on statistical distributions, rather than using words that are either too common or too seldom. There are also other possibilities, and the choice of method reflects a theory of indexing (discussed below).

c. Human-assigned indexing is done, for example, if users "translate" terms in the text to synonyms listed as descriptors (or preferred terms) in thesauri or other kinds of "controlled vocabularies." This may again represent a relatively mechanical way of indexing. A much more creative way of assigning index terms is to conceptualize the text and describe the text based on this conceptualization. This is done, for example, when librarians assign genre terms to fiction (which typically do not contain genre labels).

d. Automated-assigned indexing is done, for example, in the Institute for Scientific Information (ISI)-citation databases, where "keywordplus" are search terms automatically added to records, based on the frequency of words in titles of cited papers. There are many other kinds of automated-assigned indexing made with or without the use of thesauri or other kinds of controlled vocabularies. Again, automated-assigned indexing may be done in many ways, some of which are rather crude, while others are very innovative and based on a deep knowledge of subject literatures, concepts, and relevance criteria.

As we have seen each of the four possibilities allow for very different kinds of indexing to take place, *each possibility may utilize more mechanical procedures or more interpretative (and creative) procedures.* Human indexing may be performed in a very mechanical way and automated indexing may be performed in a very creative way, and there is a wide spectrum of indexing strategies between the most mechanical and the most interpretative. The way

human beings index documents is dependent, among other things, on

- how the indexers have been instructed to do the indexing and
- the indexer's knowledge of the domain, including the documents to be indexed, the vocabulary in the fields, the questions raised, and the criteria used to decide what is relevant and important.

The way programmers make automatic indexing systems is, in a similar way, based on

- which technologies the programmers have learned about and
- the programmer's knowledge of the characteristics of the domain and the relevance of the given documents to potential queries.

By implication, the four alternatives do not represent four theories of indexing (or four theoretically based approaches to indexing). The categories the L&E study uses are the very categories that are assumed to constitute the main [oppositional] approaches to indexing. Although this is not directly claimed by L&E, these are the categories used, and no alternative indexing theories are discussed. I want to argue that this is not a fruitful way of putting things: If a human indexer does a very mechanical job (for example, transferring the title-words of the documents into index terms), then that human being is by principle behaving like a computer. Because a dominant ideal of indexing in IS has often been connected with the establishing of rules for indexing, theories of human indexing have in reality been very machine-like. Alternatively, *we need to distinguish theories of indexing that cross these four approaches*.[3] Such theories are presented in the following section.

Theories of Indexing

In Hjørland (1997) and later works, I proposed a classification of indexing approaches based on the theories' epistemological assumptions, as follows.

Rationalist theories of indexing (such as Ranganathan's theory) suggest that subjects are constructed logically from a fundamental set of categories. The basic method of subject analysis is then "analytic-synthetic," to isolate a set of basic categories (=analysis) and then to construct the subject of any given document by combining those categories according to some rules (=synthesis). Also, the applications of other rules such as logical division are by principle part of the rationalist view.

Cognitive views suggest that people index and search documents in a specific way because they have a certain cognitive or mental structure, which cognitive studies may uncover and somehow provide a basis for indexing, i.e., that criteria for similarity, and thus indexing, is somehow hardwired into our brain or cognitive structure. In other words: the rules for indexing are parts of our cognitive structures, which in this view (as opposed to the historicist and pragmatic views) are connected to universal, biological given structures. Cognitive views of indexing seem, however, to be theoretically unclear and problematic.[4]

Empiricist theories of indexing are based on the idea that similar (informational) objects share a large number of properties. Objects may be classified according to those properties, but this should be based on neutral criteria, not on the selection of properties from theoretical points of view because this introduces a kind of subjective criteria, which is not approved by empiricism. Numerical statistical procedures are based on empiricist philosophy. Also, the search for consensus among indexers is an approach that may be interpreted as based on empiricism: the correct indexing is the one that indexers agree on, empirical studies of inter-indexer agreement are believed to reveal correct indexing (which is a problematic assumption because, as argued by Cooper (1969), indexing—as done may the majority of indexers—may be consistently bad).[5]

Historicist and hermeneutical theories of indexing suggest that the subject of a given document is relative to a given discourse or domain and is why the indexing should reflect the need of a particular discourse or domain. According to hermeneutics, a document is always written and interpreted from a particular horizon.[6] The same is the case with systems of knowledge organization and with all users searching such systems. Any question put to such a system is put from a particular horizon. All those horizons may be more or less in consensus or in conflict. To index a document is to try to contribute to the retrieval of "relevant" documents by knowing about those different horizons.

Pragmatic and critical theories of indexing are in agreement with the historicist point of view that subjects are relative to specific discourses but emphasizes that subject analysis should support given goals and values and should consider the consequences of indexing. These theories emphasize that indexing cannot be neutral and that it is a wrong goal to try to index in a neutral way. Indexing is an act (and computer-based indexing is acting according to the programmer's intentions). Acts serve human goals. Libraries and information services also serve human goals, and this is why their indexing should be done in a way that supports these

[3] It should be recognized, however, that L&E (pp. 92–93) describes some important domain-specific principles on indexing.

[4] It is difficult to understand what "a cognitive view of indexing" means. As expressed by Andersen (2004, p. 143): "It is, however, difficult to see what a cognitive approach to indexing offers and, if it offers something, what is cognitive about it." He also quotes Mai (2000, pp. 123–124): "[Farrow's cognitive model of indexing] adds no further knowledge or instruction to the process." The literature is also very unclear about the differences between e.g., cognitive approaches, user-oriented approaches, and other approaches; therefore, the viewpoint is unclear and difficult to test and it is difficult for IS to advance properly.

[5] This is easy to understand if we assume that indexers are influenced by how they have been taught, and if we consider the different views of indexing in the literature of IS, which is supposed to form the basis for teaching indexing: If unfruitful theories have dominated, then we should assume that indexing implication has been consistently bad.

[6] Gadamer (1989) is probably the best description of how texts are understood.

DOI: 10.1002/asi

goals as much as possible. The core of indexing is, as stated by Rowley and Farrow (2000), to evaluate a paper's *contribution to knowledge* and index it accordingly;[7] or, with the words of Hjørland (1997), to index its *informative potentials*.[8] What is known as "request oriented indexing"[9] is very much in accordance with pragmatic and critical views of indexing.

Basically, I therefore claim that theories of indexing are related to theories of knowledge and that the consideration of such theories may provide important insights about how to improve indexing. More specifically, I argue that the pragmatic/critical view is the most fruitful and that the application of this view is able to improve both human indexing and automated indexing.

Back to L&E

A well-known distinction in the literature about research methodology is between qualitative versus quantitative research approaches. L&E is in the quantitative pole of this continuum. There is an almost total absence of qualitative data, which might help us to interpret the results. The 10 "realistic" search tasks (SJ1-SJ10) were thus not included in the paper (for example, as an appendix). We have only the authors' words that these tasks were realistic and we have no possible way to check this claim.[10] Similarly, we have only quantitative data about which documents were classified as respectively highly relevant, fairly relevant, marginally relevant, and irrelevant; we do not have qualitative information about these categories of documents or about the relevance criteria used to make this classification (cf. Hjørland, 2010). Most important: We have no information that can help answer whether the indexing or the search queries could be improved in ways that might alter the results significantly.

My main argument against the methodology (in the present study as well as in the dominant trend in IS today) is that it just helps us to choose one among three *given* alternatives without providing information about how to improve each alternative (or other alternatives). In other words: *The study is based on a methodology that does not help us improve indexing because it does not help us understand the underlying qualitative problems in the indexing process.*

There is, however, one exception that should be recognized and that may simultaneously illuminate my point. In search scenario 7, the search system Verity K2 did not recognize search term with slashes as a prefix, and thus documents indexed this way could not be retrieved. This is, of course, relatively easily to change in the database (or perhaps in the search system), and it is an example of how this study contributed to providing knowledge about how indexing could be improved. It can be viewed as an example of "failure analysis".[11] With this exception, the study does not, however, help us to improve indexing and searching, and it also fails to provide information that enables us to generalize below the present context.

L&E classified the documents into four degrees of relevance.[12] This classification is, in itself, a controversial aspect of the study, and an aspect which very much involves the theory of knowledge. Even considering the work task situation relevance may be based on different criteria, e.g., "topicality" research methodology (as in evidence-based medicine) on other criteria (see Hjørland, 2010). This is not, however, an issue that will be further addressed here. *The point of the departure for us is that L&E did establish a classification of the documents into four degrees of relevance in relation to the SJs, and that the indexing/searching provided by humans failed to identify the documents considered "relevant" by L&E to the same degree as the search in the full text documents did*. If we consider the relevance "given," then the question is to explain why the indexing and retrieval failed to retrieve all the relevant documents and why the indexing and retrieval did retrieve nonrelevant documents.

The paper by L&E does not allow us to say whether the failure to identify the relevant (and only the relevant) document was because of bad indexing, of bad search profiles, or, most likely, a combination of indexing and retrieval. Because *the underlying issue is the same* in both indexing and retrieval, this is not an issue we need to consider much further in this article: To establish why the indexers failed to discriminate the relevant documents properly involves the same theory as to establish why the searchers failed to discriminate the relevant documents. We may therefore concentrate of the indexing alone.

Let us imagine, as a thought experiment, that the most relevant documents in one of the SJ's were all findings based on either randomized controlled trials (RCT) or other kinds

[7] "In order to achieve good consistent indexing, the indexer must have a thorough appreciation of the structure of the subject and the nature of the contribution that the document is making to the advancement of knowledge" (Rowley & Farrow, 2000, p. 99).

[8] Excellent examples of domain-analytic studies of indexing and classification systems are provided in the domain of art by Ørom (2003) and in music by Abrahamsen (2003).

[9] Request-oriented indexing is indexing in which the anticipated request from users influences how documents are being indexed. The indexer asks himself: "Under which descriptors should this entity be found?" and "think of all the possible queries and decide for which ones the entity at hand is relevant" (Soergel, 1985, p. 230). Request-oriented indexing may be indexing that is targeted towards a particular audience or user group. For example, a library or a database for feminist studies may index documents differently to a historical library. It is probably better, however, to understand request-oriented indexing as policy-based indexing: The indexing is done according to some ideals and reflects the purpose of the library or database doing the indexing. In this way, it is not necessarily a kind of indexing based on user studies.

[10] Of course the absence of these search tasks may be caused by a lack of available space in the paper, and perhaps published in a later, more comprehensive report of the experiment.

[11] Saracevic wrote: "A lot can be learned from failure analyses, particularly about human performance. Regrettably, failure tests are no longer conducted, mostly because they are complex, very time consuming, and CANNOT be done by a computer" (Saracevic, 2008, p. 772).

[12] L&E (p. 91) writes about this: "To avoid subjective judgments, we asked the relevance assessors to assess the documents retrieved according to the work task situation and the indicative request." Still, however, different assessors tend to provide different judgments, which is why it is problematic to claim that subjectivity is absent.

of controlled experiments. If this has not been indicated by the indexer in the controlled metadata, then this indexing is not fruitful in discriminating relevant versus nonrelevant documents. Although it is very probable that this information appears somewhere in the full-text document, it is obvious why the controlled metadata failed when competing with full-text retrieval. In other words: If the SJs are *given* and if the relevance criteria are *given*, THEN the question of indexing and retrieval is just a question of discriminating the different kinds of relevance in indexing and in retrieval. What Saracevic (2008) termed "failure analysis" should be applied to evaluate the indexing to improve it.

The study reported by L&E failed to retrieve the relevant documents based on the controlled metadata better than or equal to full-text searching. Clearly the indexing failed to provide the necessary discrimination. For me, there is only one logical explanation: *The indexers were not instructed properly to discriminate between the documents according to the relevance criteria employed in this study*.[13]

The title of this article emphasized that it is about "enterprise settings." In the article it is emphasized that "the indexing policy, the metadata scheme, the indexing checklist, and the corporate thesaurus are tailored specifically to meet information tasks as the ones investigated" (Lykke & Eslau, 2010, p. 92). Further:

> This finding [that metadata searches resulted in the lowest precision] is remarkable, as human indexers should be better at weighting the significance of subjects, and be more able to distinguish between important and peripheral compared with computers that base significance on term frequency. This is especially true in context of enterprise retrieval, where retrieval is embedded in and targeted to specific information tasks. Compared with retrieval system with a more general, broader scope it should be easier for a human indexer to interpret and relate document content to work domain characteristics. (Lykke & Eslau, 2010, p. 93)

As already said, the results of this article suggest that something has gone extremely wrong in the indexing in the setting in which this study took place.[14] I think it is wrong to try to answer the question of whether or not human beings can compete with computers. Rather than suggesting that human indexers in all settings (or in all enterprise settings, or in other classes of settings) can or cannot compete with computers, quite different questions should be asked. The questions I raise are as follows:

- Given the poor performance of human indexers, would improvements in their education, their instructions, and their tools provide a significantly different picture?
- How has knowledge from IS contributed to the qualifications, the instructions, and the tools provided in this setting?

Theories of indexing in IS are mostly based on universalist, objectivist, and cognitivist assumptions that have perhaps been problematic, and may explain at least a part of the failures of the indexing. It should be recognized that the present study, to some degree, reflects domain-specific, pragmatic, and request-oriented views of indexing. However, we still need to know whether this kind of understanding has penetrated to all persons and all practical procedures in thesaurus construction and indexing in this setting.

This study does not help to answer these questions and I see this as a problem inherent in its positivist philosophy. We need to know more about what is specific for this domain, its terminology, relevance criteria, documents, and genres. A lot of assumptions need to be examined, among them: In what ways (if any) are "enterprise settings" different from disciplinary domains such as medicine, chemistry and pharmacology? Also, in disciplines, should indexing be tailored to the specific needs and tasks of the domain? As I have written:

> A stone on a field could contain different information for different people (or from one situation to another). It is not possible for information systems to map *all* the stone's possible information for every individual. Nor is any *one* mapping the one "true" mapping. But people have different educational backgrounds and play different roles in the division of labor in society. A stone in a field represents typical one kind of information for the geologist, another for the archaeologist. The information from the stone can be mapped into different collective knowledge structures produced by e.g. geology and archaeology. Information can be identified, described, represented in information systems for different domains of knowledge. (Hjørland, 1997, p. 111, emphasis in original)

Although I find that it is fruitful that L&E connects indexing to the specific interests of a domain, there seems to be a need for research illuminating what a domain is, whether an enterprise should be considered a domain—in the perspective of being able to generalize findings. As things are reported in the present paper—and in former papers by the authors—there seems to be no arguments for claiming, on the one hand, that findings can be generalized to other enterprises and, on the other hand, that enterprises are more like each other than a pharmacological enterprise is like the discipline of pharmacology.

Conclusion

The paper by L&E does not consider its own theoretical position and this is itself an important point of criticism from the point of view of the theory of knowledge.

L&E was published in a festschrift to Peter Ingwersen who is known for his commitment to the label "the cognitive view." However, the study considered in the present article cannot, in my opinion, be understood as "cognitive" but must rather be understood as "physical" in the sense of Ellis (1992, 1996).[15]

[13] Alternatively, of course, the indexers may have used fruitful relevance criteria that this study failed to acknowledge.

[14] I am not saying that indexing is better in other settings: Bad indexing in practice may reflect a problematic understanding and teaching of theories of indexing in IS.

[15] I have the impression that "the cognitive view" in many cases collapses to the "system-driven" approach, which it was originally meant to substitute. This is, however, a point which will not be further addressed in this paper.

DOI: 10.1002/asi

L&E shows certain connections to the domain-analytic view developed by the present author such as the recognition that indexing needs to be tailored to the needs of a specific domain. In other ways, it differs, for example, in its positivist assumption about general laws about the efficiency of specific strategies. The article raises important and interesting problems and seems to confirm that the physical approach by Ellis (1996, pp. 177–180), also termed "the archetypal approach," i.e., Cranfield-type experiments, in information retrieval is still dominating.

Since approximately 1990 the cognitive view has increasingly been challenged by researchers such as Jack Andersen, Bernd Frohmann, Birger Hjørland, Jens-Erik Mai, Sanna Talja, and others, suggesting more social and interpretative approaches to IS. The present article provides some specific examples on how such a social-interpretative understanding may provide a better basis for improving indexing in different contexts. We need, first, qualitative information about relevance criteria, indexing, and query formulation. Without such information IS cannot advance.

The optimistic prognosis is as follows: When human indexers learn about automatic indexing with the documents they are indexing, they probably get a clear idea of what is not necessary to do (because it is done by the computer) and to be able to concentrate on value-added indexing that requires human judgment and interpretation. By implication it is necessary that all indexers are experienced searchers in the domain.

Acknowledgments

Thanks to two anonyme reviewers for valuable feedback that improved this article.

References

Abrahamsen, K.T. (2003). Indexing of musical genres. An epistemological perspective. Knowledge Organization, 30(3/4), 144–169.

Andersen, J. (2004). Analyzing the role of knowledge organization in scholarly communication: An inquiry into the intellectual foundation of knowledge organization. PhD dissertation. Copenhagen: Department of Information Studies, Royal School of Library and Information Science. Retrieved July 20, 2010, from http://biblis.db.dk/Archimages/78.PDF

Cooper, W.S. (1969). Is inter-indexer consistency a hobgoblin? American Documentation, 20(3), 268–278.

Ellis, D. (1992). The physical and cognitive paradigms in information retrieval research. Journal of Documentation, 48(1), 45–64.

Ellis, D. (1996). Progress and problems in information retrieval. London: Library Association.

Gadamer, H.-G. (1989). Truth and method (2nd rev. ed.). Trans. J. Weinsheimer and D.G.Marshall. New York: Crossroad.

Hjørland, B. (1997). Information seeking and subject representation. An activity-theoretical approach to information science. Westport & London: Greenwood Press.

Hjørland, B. (2008). What is knowledge organization (KO)? Knowledge organization. International Journal Devoted to Concept Theory, Classification, Indexing and Knowledge Representation, 35(2/3), 86–101. AQ2

Hjørland, B. (2010). The foundation of the concept of relevance. Journal of the American Society for Information Science and Technology, 61(2), 217–237.

Lykke, M., & Eslau, A.G. (2010). Using thesauri in enterprise settings: Indexing or query expansion? In B. Larsen, J.W. Schneider & F. Åström (Eds.), The Janus faced scholar. A festschrift in honor of Peter Ingwersen (pp. 87-97). Copenhagen: Royal School of Library and Information Science (This special volume of the ISSI e-newsletter, Vol. 06-S, June 2010). Retrieved July 20, 2010, from http://www.issi-society.info/peteringwersen/pif_online.pdf

Mai, J.-E. (2000). The subject indexing process: An investigation of problems in knowledge representation. Unpublished doctoral dissertation. The University of Texas at Austin. Retrieved July 20, 2010 from http://individual.utoronto.ca/jemai/Papers/2000_PhDdiss.pdf

Ørom, A. (2003). Knowledge organization in the domain of art studies—History, transition and conceptual changes. Knowledge Organization, 30(3/4), 128–143.

Rowley, J.E., & Farrow, J. (2000). Organizing knowledge: An introduction to managing access to information (3rd. ed.). Aldershot: Gower Publishing Company.

Saracevic, T. (2008). Effects of inconsistent relevance judgments on information retrieval test results: A historical perspective. Library Trends, 56(4), 763–783. Retrieved June 12, 2009, from http //www.scils.rutgers.edu/≈tefko/LibraryTrends2008.pdf

Soergel, D. (1985). Organizing information: Principles of data base and retrieval systems. Orlando, FL: Academic Press.

Author Queries

AQ1: “than” appears to be a typo. Please revise this for greater clarity.
AQ2: This was not cited. Please either cite it or delete it.