Approaches to Knowledge Organization (KO)

by Birger Hjørland
Lecture given at the University of Rome

April 20, 2007

What is Knowledge Organization?

Knowledge Organization (KO) is about activities such as document description, indexing and classification performed in libraries, in bibliographical databases, in archives and in other kinds of “memory institutions” and on the Internet.

These activities are done librarians, by archivists, by information specialists, by subject specialists, by laymen *as well as by computer algorithms*.

What is Knowledge Organization?

KO as a field of study is concerned with the nature and quality of such knowledge organizing processes (KOP) as well as the knowledge organizing systems (KOS) used to organize documents, document representations and concepts.

Examples of KOS are controlled vocabularies, authority files, classification systems, thesauri, semantic networks and ontologies.

What is Knowledge Organization?

In this presentation I shall provide an overview of approaches to KO. Often are such approaches not explicit theories, but are practices implying different views of knowledge, cognition, language, and social organization – which involves difficult questions of an interdisciplinary nature.

This is why this presentation is not just an overview of explicit theories, but an uncovering as well of such theories. You may find it surprising that I include e.g. IR and bibliometrics. This is because they represent alternative ways to accomplish the same goal as KO.

The digital challenge

In order to consider the future of KO it is important to evaluate past approaches, the heritage that they provide, and potentialities in future environments.

In the past has each library often classified/indexed its own monographs (but not considered articles/analytics, which was mostly taken up by the documentation movements & commercial services).

The digital media and the networked systems has seriously challenged this model.

The digital challenge

It is an important part of the picture that *users have always found most of their references in alternatives to the library catalog.* Libraries may be important for the physical delivery of documents. The identification of documents that the users want to borrow (or browse) may, however, be obtained by other tools.

This presentation is restricted to the consideration of systems and processes for subject access (in libraries and elsewhere). Library catalogs are today seriously challenged in this function by competing alternatives available to the users.

The digital challenge

For my own part, for example, I find many books in Amazon.com, in book reviews, in bibliographical references, in title searches in the library catalog etc. I order a lot of books in the library but very seldom books that I have identified by using the library's UDC-system. I believe that my case is typical of many scholars' way of finding books.

However, systems such as the UDC is a typical example of what is regarded knowledge organization within LIS, which is the topic for this presentation.

The digital challenge

In the 1980s *the Royal Library* in Copenhagen added Library of Congress classifications (and other data too) to its newly established OPAC. These data came from the MARC-records from LC and BL. The Royal Library continued to classify all books according to its own system in addition to the new data from MARC records.

The State Library in Århus changed in 2001 their own system to the DDC. The rationale, I believe, was that most new books purchased by the library is already classified with DDC from LC.

The digital challenge

DDC may not be a superior system, but it is economical and time consuming not to have to classify the major part of the acquired books by the library's own staff.

This is what networked library systems can do (and also to some degree did already at the time of printed catalog cards). *The full consequence might be that each book just have to be classified once and for all*.

It is interesting, or perhaps rather painful, to observe that the old Dewey system established 1876 is the most used one in libraries today in spite of all research and development in LIS and KO in more than 100 years!

Overview of approaches to KO

It is not easy to get an overview of approaches to KO. Perhaps is the reason that *KO has been dominated by different technologies* and relatively neglected as a theoretical and academic subject.

Also the circumstance that different persons (like Ranganathan, Salton, Garfield etc) have tended to be interested in only one approach, which they have considered superior without critical examination of its presumptions (and without comparison with other approaches) may be a cause.

Overview of approaches to KO

Some approaches are also unclear and the literature about them may not be clear about how a system inspired from one approach is different from a system inspired from another approach. The task to analyze the presumptions and implications of different approaches to KO has just begun.

It is also my claim that we can only have clarity about approaches to KO if we have clarity about theories of knowledge, language, cognition etc., which are difficult fields.

Overview of approaches to KO

1. “Traditional approaches”
2. Management oriented approaches
3. Logical and facet-analytic approaches
4. Computer based approaches
5. Bibliometric approaches
6. User oriented and cognitive approaches
7. Domain analytic approaches
8. Other approaches

1. Traditional approaches

If we take systems like DDC, LCC, UDC,and Bliss 1 – all established in the late 1800s or early 1900s – as our point of departure, what approach to KO do they represent?

Obviously, they represent different approaches. However, compared to the other approaches to be introduced, they share some attributes. They are more or less enumerated systems, more or less based on scientific disciplines and on the need to classify knowledge according to the scientific world view.

1. Traditional approaches

Two system designers, Melvil Dewey and Henry Bliss, stand out as quite different persons. Dewey was more like a businessman and his system was based on principles inspired by a management & business philosophy. Bliss, on the other hand, was more like an intellectual or a scholar, trying to study the theoretical principles of KO as a field. In many ways I feel that what Bliss attempted was very much the same as what I myself attempt to do. Although I have critical views of Bliss' research he may turn out to be my favorite model in the history of our field.

1. Traditional approaches

Dewey's approach to KO shall be considered under the label "Management oriented approaches" and thus distinguished from the more "intellectual" approaches although the DDC also adopted many principles from other approaches.

LCC (Library of Congress Classification) was based on the wish to reflect the collection in one library as perfect as possible. It is very much based on literary warrant and subject specialists in the LC. Despite its concentration on one collection it has proven fruitful for many other collections.

1. Traditional approaches

In the traditional approaches was subject expertise used to design classification systems as well as to index/classify documents. The necessity of subject knowledge on behalf of both the classifier of specific books and on behalf of the construction of classification schemes was mostly taken for granted. No special method was developed.

Bliss (1 ed.) was established on the wish to establish a system which reflected the true order of Nature and Science.

1. Traditional approaches

The traditional approaches were based on intellectual aspirations to reflect a true order of reality. Besides a few principles, they did not establish a specific methodology for bibliographic classification. Classification was implicitly treated as a “neutral reading” of the true semantic relations. Besides subject knowledge, classification was often considered based on “common sense” or “intuition”.

Example: Subject categories in ISI’s citation indexes are just based on intuitive criteria (Leydesdorff, 2006, p. 602).

1. Traditional approaches

From the point of view of a theory of KO, traditional approaches often represent a difficulty:

"It is quite hard to discern any strong theoretical principles underlying LCC [Library of Congress Classification]". Broughton (2004, p. 143)

Also some formulations by S. R. Ranganathan (e.g., 1951) suggest that “traditional” systems seem to lack a theoretical foundation (in his eyes as opposed to his own approach).

1. Traditional approaches

The implication of the necessity of subject knowledge is, that librarians, in order to classify books, should know about scientific developments. This should also be reflected in their education:

"Again from the standpoint of the higher education of librarians, the teaching of systems of classification . . . would be perhaps better conducted by including courses in the systematic encyclopedia and methodology of all the sciences, that is to say, outlines which try to summarize the most recent results in the relation to one another in which they are now studied together. . . ."
(Ernest Cushing Richardson, quoted from Bliss, 1935, p. 2).

1. Traditional approaches

This important principle has been implicit in the management of research libraries and bibliographic databases such as MEDLINE, in which subjects specialists are often hired to do the work in KO. The importance of subject knowledge has not been explicit in the following approaches to KO - except in domain analysis.

1. Traditional approaches

Among the other principles, which may be attributed to the traditional approach to KO are:

- Principle of controlled vocabulary
- Cutter's rule about specificity (1876)
- Hulme's principle of literary warrant (1911)
- Principle of organizing from the general to the specific

1. Traditional approaches

Here in the early 2000s “the traditional approach” seems still to be very much used, but seems to lack distinctness as an approach to KO. How classifiers interpret the literature in order to base their decisions on “literary warrant” seems most to rely on a kind of positivism in which the sources are assumed to speak for themselves: There is one correct answer, which the classifier is supposted to be able to “read”.

Principles on how to interpret semantic relations and how such relations reflect interests and points of view were not translated to principles or theories of KO.

2. Management oriented approaches

Melvil Dewey's system, the DDC, is today the most successful of all systems developed within (LIS). Its success stands, however, in sharp contrast to its lack of basis in research.

As already stated was Dewey's goal was to establish a manageable system that could be used in (and sold to) many libraries and where practicalities were the most important consideration. His system has been called an empty semiotic shell and there is a particular strong connection between his principles and the criticism raised against library classification for it's devoid of substantive intellectual content.

2. Management oriented approaches

For example, Dewey wanted women employed in libraries because they demanded lower salaries. He preferred lower salaries for higher qualifications (although he also wanted a qualified staff). Dewey also believed that librarianship should not concern itself with what was inside the books ! (cf Tredinnick, 2006, p. 39-40). Another principle was to limit classification as reflecting reality to the easy or obvious and then to do something “practical” (such as alphabetical subdivisions) (cf. Miksa, 1998).

Most important, I believe, is the attempt to standardize knowledge organization.

2. Management oriented approaches

Standardization may sound fruitful. Consider, however, that developments in knowledge is dynamic. If knowledge is developing dynamically then any attempt to standardize classification either implies that there is a large amount of arbitrarity in classification that can be standardized or it imply that the correspondence to the external world is (partly) given up, that library classifications become more or less independent systems rather than systems optimized to reflect the structure of collections, subject domains, or user communities for which they are used.

2. Management oriented approaches

DDC may be understood not as a tool to optimize users retrieval of information, but first and foremost as a tool to optimize library administration, e.g. by being able to hire staff educated in the use of the system and by shared cataloging among many libraries. It is probably not accidental that LIS was termed library economy in 1876 (and so termed in DDC until (and including) 14. edition, 1942).

It is of course of benefit for users to have well managed libraries - and a well managed classification too.

2. Management oriented approaches

"[Dewey's) legacy, however, is mixed. On the one hand the scheme has over the decades saved millions of dollars and countless hours of time. Because it has become so widely accepted throughout the world it has permitted one person to classify one title for the hundreds of thousands of libraries using the decimal system. In addition, the system itself has become familiar to millions of people who can feel relatively confident that their knowledge of the system used in one library will serve them well in another.

2. Management oriented approaches

On the other hand the doctrine of Anglo-Saxonism Dewey wove so tightly into his system has over the years resisted the introduction of new threads with more culturally pluralistic origins." (Wiegand, 1998).

2. Management oriented approaches

From the point of view of LIS and KO as an academic field (and librarians as involved in intellectual activities), a management and business perspective like that of DDC is, however, problematic. Its success has been very great up to know. One of the cost of lack of intellectual involvement may be, however, that this systems (and libraries at large) may be less competitive compared to other systems developed outside LIS (such as Google).

3. Logical & facet analytic approaches

Some approaches are characterized by using logic rather than empirical research. Facet analysis is the most influential approach in LIS based primarily based on logic.

W. C. Berwick Sayers (1881-1960), a teacher of Ranganathan, was able to structure the methodology of library classification in terms of canons and axioms and my thus bee seen as belonging to this approach.

Today represents “Formal concept analysis” an approach, which is also mainly logical. Here we shall concentrate on the so-called facet-analytic approach.

3. Facet-analytic approaches

Facet analysis is probably the most distinct approach in the history of KO. It is more methodic compared to “traditional approaches” and it is more genuine LIS compared to computer based approaches. The great name in this tradition is of course S. R. Ranganathan, who was educated in mathematics, which influenced this approach very much. Important dates in the history of this tradition was the publication of the first edition of *Colon Classification* in 1933 and the establishment of the British *Classification Research Group* (CRG) in 1952.

3. Facet analytic approaches

Faceted classification is also called analytico-synthetic classification, named after the two main processes involved in the composition of a call number: Analysis, breaking down each subject into its basic concepts and synthesis, combining the relevant units and concepts to describe the subject matter of the document. A title: The police in Denmark during World War II" is analyzed in categories:

Time (1940-1945),

Place: (Denmark) and

"Personality": (Police)

3. Facet analytic approaches

Each of those categories (plus a few more) have their own independent part of the classification system. In subject description in classification codes or call numbers is the book's subject synthesized from classes from each of the categories. (The class name for time, place and personality, among others).

Ranganathan proposed 5 categories, PMEST:

Personality, Matter, Energy, Space and Time

3. Facet analytic approaches

- Personality is the distinguishing characteristic of a subject
- Matter is the physical material of which a subject may be composed
- Energy is any action that occurs with respect to the subject
- Space is the geographic component of the location of a subject.
- Time is the period associated with a subject.

The Classification Research Group expanded the number of categories.

3. Facet analytic approaches

When we consider the methodological principles developed within this approach, they are impressing (see for example the introduction to the Bliss 2 system). Those principle have been taught for generations in schools of LIS and have brought a feeling of rigor and sound scientific principles.

Still are textbooks such as *Classification made simple* (Hunter, 2002) and *Essential classification* (Broughton, 2004) based on principles developed in this tradition.

3. Facet analytic approaches

Facet analysis is not properly based on empirical research. It is a rationalist approach much more than an empiricist approach.

In practice are classifications of course constructed on the basis of some library collections or the terminology of some disciplines. Its methodology does not, however, describe the problems in selecting this empirical basis, nor is system evaluation based on empirical studies within this tradition. The emphasis is on the logical analysis on a given set of terms and concepts. And evaluations tend to focus on whether explicit rules have consequently been applied in a given case.

3. Facet analytic approaches

Miksa (1998, 71-73) raises tree kinds of negative aspects of Ranganathan's work:

1. Ranganathan helped to promote the "atomization" of subjects.
2. Neither Ranganathan himself nor those who have adopted his approach to subjects have ever critically analyzed the analogy at the base of his approach. "In the end, there is strong indication that Ranganathan's use of faceted structure of subjects may well have represented his need to find more order and regularity, in the realm of subjects, than actually exist" (Miksa, 1998, p. 73).
3. Ranganathan vigorously pursued the goal of finding *one* best subject classification system.

3. Facet analytic approaches

Facet analysis is still used, also in the design of web-pages (see La Barre, 2006). It should also be said that the revisions of traditional systems such as DDC increasingly use facet-analytic principles.

However, today the interest in this approach is relatively low. None classification researcher (except Brian Vickery) from the traditional approaches or from the facet analytic approaches are visible in bibliometric maps of LIS such at the one made by White & McCain (1998).

3. Facet analytic approaches

The reason may well be research results obtained from computer based experiments such as Cranfield and TREC (or implicit feelings that points of view as those formulated by Miksa) may be true.

It is important however, that LIS-researchers examine the basic assumptions in different approaches so that the further development of our field are based on research rather than on intuitions and fads. Facet analytic approaches should carefully be compared with other approaches.

4. Computer based approaches

From the 1950s have computers been important in many levels, including as genuine, algorithmic-based approaches to KO.

Other approaches are semi-automatic or machine-aided, such as the technique “text categorization“, which is based on manually predetermined categories.

Computers also influenced KO by providing for example citation databases which allows new forms of analysis and KO.

As stated in the introduction has computers and networks seriously challenged how classification work should be organized.

4. Computer based approaches

In 1957-1962 the Cranfield I experiments claimed that traditional forms of classification and indexing (e.g. the UDC) were rather inefficient compared to retrieval based on simple “UNITERMS" or alphabetical subject headings (cf., Ellis, 1996, 3-6):

UNITERM	82,0% recall
Alphabetical subject headings	81,5% recall
UDC	75,6% recall
Facet classification scheme	73,8% recall

4. Computer based approaches

These results were seriously challenged and many experiments and developments have taken place since then. The main tendency may, however, be interpreted as empiricist in more than one way:

- Emphasizes on empirical testing (Based on relevance measures: Recall & Precision)
- Emphasis on “free text searching” rather than on human indexing or abstracting. I.e. an assumption that a document reveals its own subject matter in an objective way independent of use contexts.
- An attempt to uncover patterns bottom-up rather than top-down.

4. Computer based approaches

Julian Warner has characterized the traditional computer approaches as “query transformation” :

"Two antithetical, if not always clearly distinguished, traditions can be detected in information retrieval system design and evaluation. The idea of query transformation, understood as the automatic transformation of a query into a set of relevant records, has been dominant in information retrieval theory. A contrasting principle of selection power has been valued in ordinary discourse, librarianship, and, to some extent, in practical system design and use". (Warner, 2002).

4. Computer based approaches

I believe it is not interesting any longer to ask if computers can organize knowledge. Computer systems such as Google is used all the time by all of us. The question is: Are there kinds of KO that can be done better by human beings? And: Is it worth the money?

Today the question is asked whether systems such as Google can or will replace the academic library. Are more traditional forms of KO still needed?

4. Computer based approaches

I believe that Warner's notion of selection power is relevant and is something that more traditional forms of KO may have in the luggage.

Take the recent discussion between Karen Sparck Jones and Birger Hjørland in *Journal of Documentation*.

Sparck Jones (2005) expressed that view that traditional forms of classification may be obsolete and replaced by, for example systems with relevance feedback mechanisms.

4. Computer based approaches

This may be a good illustration of Warner's point. Relevance feedback is an advanced form of query transformation. Can it really replace pre-classification or "selection power"?

I believe not. I have made an example about geography. I do not believe that people, who do not know, for example, Swedish geography, can make adequate feedback about whether a document is about a Swedish place or not. Classification made previously is needed during information searching.

4. Computer based approaches

We should not make too strong a difference between computer based approaches and human based approaches. Often humans function rather mechanically: For example, if a certain word is used in the title of a document is the document classified in a class containing the same word. This is an algorithmic procedure whether performed by humans or machines. So, we have to specify what kind of analysis (intellectual analysis or mechanical analysis) is at play: We have to consider theoretical approaches more deeply.

4. Computer based approaches

Human indexing is often rather inconsistent and it has been claimed that "The effectiveness of manual TC [text categorization] is not 100% anyway (Cleverdon 1984) and, more importantly, it is unlikely to be improved substantially by the progress of research." (Sebastiani, 2002, p. 41).

I believe the last part of the sentence is wrong and provocative. I believe we can learn people how to improve indexing and that this can be improved substantially by the progress of research.

4. Computer based approaches

I also believe that a better understanding of knowledge, language, genres and human practices may help to improve computers. Also a top-down approach which can help both humans and computers to identify the best "selection power" or criteria.

I have more to say about this under the domain-analytic approach.

5. Bibliometric approaches

Papers usually cite other papers allowing users to retrieve references by searching cited and citing references in citation databases.

This has been used to provide, for example, ”Atlas of science” or bibliometric maps, which are important, but rather different kinds of KOS.

Scientific papers is a kind of self organizing system in which authors provide a kind of subject representation through references.

http://www.db.dk/bh/lifeboat_ko/Bibliometric_MAP_LIS.PDF

http://www.db.dk/bh/lifeboat_ko/Bibliometric_MAP_LIS.PDF

5. Bibliometric approaches

Advantages

- Citations are provided by experts.
- The number of references reflect the indexing depth and specificity (average about 10 references)
- A highly dynamic form of subject representation
- References are distributed in papers, allowing the utilization of paper structure in the contextual interpretation of citations.

Disadvantages

- The relation between citations and subject relatedness is indirect and somewhat unclear.
- Does not provide clear logical structure with mutually exclusive and collectively exhaustive classes.
- Namedropping and other forms of imprecise citations causes noise.

5. Bibliometric approaches

Personally I believe that bibliometric KO is one kind of KO that is rather different from traditional forms. No doubt is bibliometric maps also useful as tools for retrieving documents. Bibliometrics is a kind of social KO, while traditional forms mostly are a mixture of social and cognitive KO.

If we, for example, consider a traditional geographical map (or its representation in for example UDC) as a kind of KO. Is it reasonable to believe that it can be produced by bibliometric techniques? I do not believe so.

6. User-oriented and cognitive approaches

User-oriented approaches are approaches mainly based on information about or from users. For example, the examining of logging files, interviews and so on.

The most important contribution from user-oriented approaches is the emphasis that in the end it is the users' needs that are should be in the focus of all LIS-services and products. Libraries are no ends in themselves and should not reflect patronizing attitudes.

6. User-oriented and cognitive approaches

Why are studies of users important?

In some cases, say music, the research establishment may tend to disregard the kind of music favored by many people. In such cases may expert knowledge be of less use compared to users knowledge and preferences. (Or rather “users” are the experts in this case).

6. User-oriented and cognitive approaches

A new tendency is that users themselves organize knowledge in "**folksonomies**", which are collaboratively generated, open-ended labeling systems that enables Internet users to categorize content such as Web pages, online photographs, and Web links. (Wikipedia, 2006).

Not to be confused with **folk taxonomies**, cultural practices that has been widely documented in anthropological work. Folk taxonomies are culturally supplied, intergenerationally transmitted, and relatively stable classification systems that people in a given culture use to make sense of the entire world around them (not just the Internet). (Wikipedia, 2006).

6. User-oriented and cognitive approaches

Historical periods in biological systematics (after Mishler, 2000)

1) Pre-history. Folk classifications

*2) Ancient Greeks through Linneaeus: Essentialism

*3) Natural system. Overall resemblance; "importance".

4) Darwin. Evolutionary language added (Only a superficial effect for a long time, cf. 6)

5) Numerical Phenetics. Computers added. (Only a superficial effect)

*6) Phylogenetic systematics (Cladistics). [A late Darwian approach]

[*7) Systematics based on DNA-analysis]

6. User-oriented and cognitive approaches

In the previous survey of periods in biological systematics is Folk classifications referred to as the pre-history of biological systematics.

Should KOS in library and information science be based on folk-classification and user studies or on scientific principles?

I believe that much interest in “user studies” in our field is based on attempts to short-circuit the needed subject knowledge.

Hjørland & Nissen Pedersen (2005):

1. Classification is the ordering of objects (or processes or ideas, whatsoever) into classes on the basis of some properties.
2. The properties of objects are not just "given" but are only available to us on the basis of some descriptions and pre-understandings of those objects. It is not possible to enumerate all properties of an object in a description.

7. Domain analytic approaches

3. Description (or every other kind of representation) of objects is both a reflection of the thing described and of the subject doing the description. Descriptions are more or less purposeful and theory-laden. Pharmacologists, for example, in their description of chemicals, emphasize the medical effects of chemicals, whereas "pure" chemists emphasis other things such as their structural properties.

4. The selection of the properties of the objects to be classified must reflect the purpose of the classification. There is thus no "neutral" or "objective" way to select properties for classification.

7. Domain analytic approaches

Example: Whether to classify by form or color. The figures below may be classified according to color or shape. None of those properties are "objectively" more important than the other:

■ □ ▲

5. The (false) belief that there exist objective criteria for classification may be termed "empiricism" (or "positivism"), while the belief that classifications are always reflecting a purpose may be termed “pragmatism". Our paper is thus an argument for the pragmatist way of understanding.

7. Domain analytic approaches

6. We saw that different domains (chemistry and pharmacology) may need different descriptions and classification of objects to serve their specific purpose in the social division of labor in society. The criteria for classification are thus generally domain-specific. Different domains develop specific languages (LSPs) that are useful to describe, differentiate and classify objects in their respective domain.

7. Domain analytic approaches

7. In every domain exist different theories, approaches, interests or "paradigms", which also tend to describe and classify the objects according to their views and goals.

(The documentation for this claim is established in the descriptions of the different domains in the *Epistemological Lifeboat* http://www.db.dk/jni/lifeboat/home.htm).

8. Any given classification will always be a reflection of a certain view or approach to the objects being classified. Ørom (2003), for example, shows how different library classifications are reflecting different views of the Arts. Ereshefsky (2000) argues that Linnaean classification is based on criteria that are pre-Darwinian and thus problematic. Sometimes, however, a given classification seems to be immune to criticism. This may be the case with *the Periodical System* of Chemistry and Physics. Such immunity is caused by a strong consensus in the underlying theory.

7. Domain analytic approaches

9. A given literature to be classified is always - more or less - a merging of different domains and approaches/theories/views. Such different views may be explicit or implicit. If they are implicit they can be uncovered by theoretical and philosophical analysis.

10. Classification systems that do not consider the different goals and interest reflected in the literature of a given domain are "positivist".

The criteria for classification should be based on an understanding of the specific goals, values and interest at play. They are not to be established a priory, but by "literary warrant": by examining the literature. (This cannot either be done in a "neutral" or "objective" way, but may be done more or less qualified by considering the different arguments).

7. Domain analytic approaches

In her reply Sparck Jones (2005) acknowledge the pragmatic point of view. Her final suggestion is, however:

"At the same time, one of the most important techniques developed in retrieval research and very prominent in recent work, namely relevance feedback, raises a more fundamental question. This is whether classification in the conventional, explicit sense, is really needed for retrieval in many, or most, cases, or whether classification in the general (i.e. default) retrieval context has a quite other interpretation.

7. Domain analytic approaches

Relevance feedback simply exploits term distribution information along with relevance judgements on viewed documents in order to modify queries.

In doing this it is forming and using an implicit term classification for a particular user situation. As classification the process is indirect and minimal." (Sparck Jones, 2005).

7. Domain analytic approaches

We have already considered the arguments against pure relevance feed-back.

The domain-analytic view imply that if LIS-professionals learn about domains, especially about paradigms and sociological issues in domains can better classifications be constructed and documents better indexed. (Besides, LIS-professionals need this knowledge for interacting with users, interpreting information needs etc.).

7. Domain analytic approaches

Tredinnick. (2006): *Digital information contexts: Theoretical approaches to understanding digital information* explores different frameworks. Among the implications of post-modernism Tredinnick presents the view that information professionals have to become scholar-librarians and active participants in the discourses that surround particular collections. This is a change that requires fundamental changes in the professions dominating beliefs.

This comes, I believe, very close to ”domain analysis”

8. Other approaches

There are many other approaches to KO, for example, based on genres, semiotics, document composition, “information architectures” etc. Much is of course inspired by new Internet technologies.

Many of those approaches shares with domain-analysis a believe in human interpretation as a supplement to computer based forms of KO.

Conclusion

What is the future of KO?

I believe there will be a pressure on KO performed by libraries. People will use the best systems (they are just one click away). There is not much need for librarians who classify the same books almost the same ways in multiple libraries. There will not be much room for routine work (or more mechanical work).

If KO within LIS is going to survive, I believe that excellence is needed. Rather than routine workers are LIS-professionals going to be a kind of researchers.

Conclusion

I believe that indexing and abstracting provided by LIS-professionals are better than author-based indexing and abstracting (Or at least may add value to document descriptions).

LIS professionals consider a given document in the context of a collection or a subject literature.

They may also serve as critical assessors on behalf of their users and potential users.

Conclusion

If users have a special interest, they may search the bibliographical database (including OPAC) that has the best coverage and subject description in relation to their needs.

One Danish library (a military library) have begun to write abstracts of their books. That is probably one way to make a difference.

Rather than trying to cover all genres, media, levels and subjects, I believe libraries should cooperate, each developing competence within a narrow specialty.

Conclusion

This must also be the case of public libraries. If a user is interested in say Mozart, he would like to browse in fine descriptions of works by and on Mozart. He would not care much if this system is made in one public library or another. It would be better to have one public library contributing something to the best KO that to have all public libraries making lesser systems.

Conclusion

I believe the best way to prepare students for the future in KO is to provide a critical review of approaches available and cooperate in the discipline to establish the best possible theoretical ground for evaluating existing technologies and providing excellent KO.

I have tried to demonstrate that approaches are associated with theories of knowledge such as positivism, rationalism, empiricism, pragmatism and post-modernism.

I believe that debate and further study of the approaches mentioned (as well as new ones) is urgent.

References

Bliss, H. E. (1929). *The organization of knowledge and the system of the sciences*. With an introduction by John Dewey. New York: Henry Holt and Co.

Broughton, V. (2004). *Essential classification*. London : Facet Publishing.

References

Ereshefsky, M. (2000). The Poverty of the Linnaean Hierarchy: A Philosophical Study of Biological Taxonomy. Cambridge: Cambridge University Press.

Ellis, D. (1996). *Progress and Problems in Information Retrieval*. London: Library Association Publishing.

References

Hjørland, B. & Nissen Pedersen, K. (2005). A substantive theory of classification for information retrieval. *Journal of Documentation, 61*(5), 582-597. http://www.db.dk/bh/Core%20Concepts%20in%20LIS/Hjorland%20&%20Nissen.pdf

Hunter, E. J. (2002). *Classification made simple*. 2nd ed. Aldershot : Ashgate Publishing Limited.

References

La Barre, K. (2006). *The use of faceted analytico-synthetic theory as revealed in the practice of website construction and design*. Ph.D.-dissertation submitted at the school of LIS at Indiana University.

Leydesdorff, L. (2006). Can scientific journals be classified in terms of aggregated journal-journal citation relations using the *Journal Citation Reports*? Journal of the American Society for Information Science and Technology, 57(5), 601-613.

References

Miksa, F. (1998). *The DDC, the Universe of Knowledge, and the Post-Modern Library*. Albany, NY: Forest Press.

Mishler, B. D. (2000). Deep Phylogenetic Relationships among "Plants" and Their Implications for Classification. *Taxon*, *49*(4), 661-683.

References

Ranganathan, S. R. (1951). *Philosophy of Library Classification*. Copenhagen: E. Munksgaard.

References

Schneider, J. & Borlund, P. (2004). Introduction to bibliometrics for construction and maintenance of thesauri: methodical considerations. *Journal of Documentation*, 60(5), 524-549.

Sebastiani, F. (2002). Machine learning in automated text categorization. *ACM Computing Surveys, 34*(1), 1-47. http://www.math.unipd.it/~fabseb60/Publications/ACMCS02.pdf

References

Sparck Jones, K. (2005). Revisiting classification for retrieval. *Journal of Documentation, 61*(5), 598-601. [Reply to Hjørland & Nissen Pedersen, 2005]

http://www.db.dk/bh/Core%20Concepts%20in%20LIS/Sparck%20Jones_reply%20to%20Hjorland%20&%20Nissen.pdf

Tredinnick, L. (2006). *Digital information contexts: Theoretical approaches to understanding digital information*. Oxford: Chandos Publishing.

References

Warner, J. (2002). Forms of labour in information systems. *Information Research, 7*(4) http://informationr.net/ir/7-4/paper135.html

White, H. D., & McCain, K. W. (1998). Visualizing a discipline: An author co-citation analysis of information science, 1972-1995. *Journal of the American Society for Information Science, 49*(4), 327-355.

References

Wiegand, W. A. (1998). The "Amherst method". The origins of the Dewey Decimal Classification System. *Libraries & Culture*, 33(2), 175-194. Available at: http://www.gslis.utexas.edu/~landc/fulltext/LandC_33_2_Wiegand.pdf

Ørom, A. (2003). Knowledge Organization in the domain of Art Studies - History, Transition and Conceptual Changes. *Knowledge Organization, 30*(3/4), 128-143.

References

Åström, F (2002) Visualizing Library and Information Science concept spaces through keyword and citation based maps and clusters. In: Bruce, Fidel, Ingwersen & Vakkari (Eds.). *Emerging frameworks and methods: Proceedings of the fourth international conference on conceptions of Library and Information Science (CoLIS4)*, pp 185-197. Greenwood Village: Libraries unlimited. Two figures: Bibliometric_MAP_LIS.PDF; Bibliometric_LIS_2.PDF