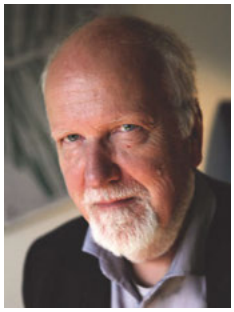
Knowledge Organization (KO) *†

Birger Hjørland

Royal School of Library and Information Science, University of Copenhagen,
6 Birketinget, DK-2300 Copenhagen S, Denmark,
<birger.hjorland@hum.ku.dk>

Birger Hjørland holds an MA in psychology and PhD in library and information science. He is Professor in knowledge organization at the Royal School of Library and Information Science in Copenhagen since 2001 and at the University College in Borås 2000-2001. He was research librarian at the Royal Library in Copenhagen 1978-1990, and taught information science at the Department of Mathematical and Applied Linguistics at the University of Copenhagen 1983-1986. He is chair of ISKO's Scientific Advisory Council and a member of the editorial boards of *Knowledge Organization*, *Journal of the Association for Information Science and Technology* and *Journal of Documentation*. His H-index is 38 in Google Scholar and 21 in Web of Science.



Hjørland, Birger. 2016. "Knowledge Organization (KO)." *Knowledge Organization* 43: 475-484. 61 references.

* Derived from the article of the same title in the ISKO Encyclopedia of Knowledge Organization, version 1.0; published 2016-08-01; article category: disciplines

† Thanks to Maria Teresa Biagetti for serving as the editor for this article and to two anonymous referees for providing valuable feedback.

1.0 Introduction

KO is a field of research, teaching and practice, which is mostly affiliated with library and information science (LIS). KO is first and foremost institutionalized in professorships at universities around the world, in teaching and research programs at research institutions and schools of higher education, in scholarly journals (for example, *Knowledge Organization*, 1993-), in national and international conferences, in national and international organizations (for example, the International Society for Knowledge Organization, ISKO, cf., Dahlberg 2010).

KO is about describing, representing, filing and organizing documents, document representations, subjects and concepts both by humans and by computer programs (cf., Hjørland 2008). For these purposes rules and standards are developed, including classification systems, lists of subject headings, thesauri and other forms of metadata. The organization of knowledge in classification systems and concept systems are core subjects in KO. The two main aspects of KO are (1) knowledge organization processes (KOP) and (2) knowledge organization systems (KOS). Knowledge organization processes (KOP) are, for example, the processes of cataloging, subject analysis, indexing and classification by humans or computers. Knowledge organization systems (KOS) are selection of concepts with indication of selected semantic relations, for example, classification systems, lists of subject headings, thesauri, ontologies and other systems of metadata.

People like Charles A. Cutter, W. C. Berwick Sayers and Ernest Cushington Richardson established the field "knowledge organization" as an academic field around 1900. Henry Bliss' 1929 book *The Organization of Knowledge and the System of the Sciences* also represents one of the main intellectual contributions establishing the field. These authors argued that book classification should be based on knowledge organization as it appears in science and scholarship.

2.0 Research traditions, approaches and basic theoretical issues in KO

Traditionally, approaches to KO are divided into human-based approaches versus machine-based approaches (cf. Anderson and Pérez-Carballo 2001a, b). There are, however, many different kinds of human approaches and many different kinds of computer-based approaches, and they are not necessarily distinct. For example, human-based approaches may be very mechanical, if humans just follow simple rules they have learned such as alphabetical arrangement, or finding best matches for book titles in a given KOS. Both humans and machines may or may not

base their classification on citations, and if they both do, they are applying a similar approach. Hjørland (2011b) therefore argued that this traditional distinction is theoretically unfruitful. Alternatively, it was suggested that human indexers as well as programmers are guided by their knowledge and theories, which—at the deepest—is connected to their (often-implicit) theories of knowledge. However, it is often difficult to reveal what kind of theoretical assumption guide KOPs. Such processes are often done intuitively and some systems have been difficult to relate to a theory. However, the following eight+ traditions in and outside KO are probably the most influential and important today.

2.0A. Approaches developed inside KO

2.1 Practicalist and intuitivist approaches

These are approaches, which give priority to practical matters such as using the same classification system for several libraries and thereby facilitating centralization of classification and indexing. From this perspective, KO should be balanced between, on the one hand, adequate and updated subject knowledge, and on the other hand, the need for stability in order to avoid reclassification. The model here is the *Dewey Decimal Classification* system (*DDC*, first edition constructed by Melvil Dewey in 1876), which today is the dominant library system worldwide. (The practicalism described in this section should not be confused with pragmatism, which has a deep intellectual foundation and commitment and which is important in the domain-analytic approach described in 2.5 below).

Another example is the journal classification in the citation databases: "The Institute of Scientific Information (ISI) itself provides a classification of journals at the level of the database that has been based on intuitive criteria (Pudovkin & Garfield 2002)." (Leydesdorff 2006, 602). In other words, no kind of research based criteria were used, just the intuition of the classifiers.

2.2 Consensus based approaches

Henry E. Bliss (1929; [1933]1939) found that library classification should be based on what he referred to as the scientific and educational consensus. "Topics should be collocated and placed in classes not according to the whim of the person who devises the classification system, but according to the standards set by scientists and educators" (Drobnicki 1996, 3). It was characteristic a) that Bliss consulted the scholarly literature and b) believed that one is able to detect an underlying pattern of agreement. Eugene Garfield has described Henry Bliss as "a true scholar. His goals and aspirations were different from those of Melvil Dewey, whom he certainly surpassed in intellectual ability, but by whom he was dwarfed in organizational ability and drive" (Garfield 1974, 291). Bliss' view of consensus probably reflected the positivism or modernism of his time. He wrote ([1933]1939, 37):

> The more definite the concepts, the relations, and the principles of science, philosophy, and education become, the clearer and more stable the order of the sciences and studies in relation to learning and to life; and so the scientific and educational consensus becomes more dominant and more permanent.

Kruk (1999) is among the critiques of this view and wrote "In the twentieth century knowledge is not perceived as a solid structure any more. The universal library is a utopian vision and it belongs to the same category as the universal encyclopedia and the universal language." Today Bliss' view is contrasted by a view of knowledge much more concerned with conflicting interests and perspectives (cf. the domain analytic view, 2.5). His engagement with the literature to be classified is, however, still an important principle.

Bliss' reception may reveal something about hostility that serious academic work may encounter in librarianship dominated by practicalism (Campbell 1976, 139):

> Bliss had announced his intention to develop a new general classification in the *Library Quarterly* in 1910. The announcement met with bitter hostility, not from Melvil Dewey (Bliss always said that his personal relations with Dewey were cordial …) but from some of Dewey's disciples. Bliss gradually became a rather solitary figure in the American library scene, and his later work met with apathy.

Further (Campbell 1976, 139):

> Bliss's first book, *The Organization of Knowledge and the System of the Sciences*, was published in 1929 by Henry Bolt & Co., New York, after he had failed to interest the American Library Association in it. Only three of Bliss's papers were ever published by the Association, and two of those were condensed …. The American Library Association, after negotiations lasting several years, refused to publish his second book without a generous subsidy from the author sufficient to cover all publishing costs.

Fortunately, this hostility did not hinder Bliss' recognition: "The two books … and the outline version of his scheme, *A System of Bibliographic Classification* (1935, 2nd ed., 1936) won him a reputation in many parts of the world as an original thinker of great power, and a classifi-

cationist who was not afraid to tread out new paths" (Campbell 1976, 139).

2.3 Facet-analytic approaches

The facet-analytic paradigm is probably the most distinct approach to knowledge organization developed within LIS. It is mainly attributed to S. R. Ranganathan and the British Classification Research Group, but is fundamentally based on principles of logical division developed more than two millennia ago (Mills 2004). Faceted systems differ from enumerative systems by not listing all their classes, but provide building blocks from which specific classes for each document may be formed. It still has a strong position in the field and it is the most explicit and "pure" theoretical approach to knowledge organization (KO). The strength of this approach is its logical principles and the way it provides structures in knowledge organization systems (KOSs). The main weaknesses of this approach are 1) its lack of empirical basis in its methodology (although, of course, any given facetted classification must have a basis in some empirically derived list of concepts) and 2) its speculative ordering of knowledge without basis in the development or influence of theories and socio-historical studies. It seems to be based on the problematic assumption that relations between concepts are a priori and not established by the development of models, theories and laws (see further in Hjørland 2013b).

2.4 User-based and cognitive approaches

A distinction should be made between user-friendly KOSs and user-based KOSs. Today it seems evident that KOSs should be user friendly, but this was not always the case (see Hjørland 2013c and Jensen 1973). It is not evident, however, that user-friendly systems should be produced on information from users or about users. Extremely successful systems such as Apple's iPhone, Dialog's search system and Google's PageRank, for example, are not based on the empirical studies of users. Actually, the idea that KOS should be based on user studies (rather than, for example, on literary warrant, logical division, word statistics or scholarly theories) seems to be an unsupported hypothesis. Nonetheless, it is a family of approaches that has its supporters (for further information see Hjørland 2013c).

2.5 Domain-analytic/epistemological approaches

A core principle of the domain-analytic approach is: "The starting point for understanding classification is one that any object, any document and any domain could be classified from multiple equal correct perspectives" (Mai 2011, 723). In other words (Hjørland 2002, 116):

> Different communities may be interested in the same object (e.g. a stone in the field [or a given book]) but may interpret it differently (e.g. from an archeological or geological point of view). What is informative (and thus information) depends on the point of view of the specific community.

In contrast to consensus based approaches (3.2 above), domain analysis assumes the existence of multiple perspectives: Disagreement is common and "the picture is really not one of agreement, but of conflicting schools, and the closer the neighbours the sharper the conflict" (Broadfield 1946, 69). Of course, the degree of consensus is stronger in some domains compared to others. Recently a revolution has taken place in ornithology, and it seems as if the new classification of birds has a very strong scientific basis and a high degree of consensus (see Fjeldså 2013). To examine the warrant for a classification is of course part of the domain-analytic framework. It is also important to realize that not every perspective or classification is as important as any other is. One should not subscribe to relativism due to convenience, i.e. abstain from considering strengths and weaknesses in different perspectives or paradigms.

Ingetraut Dahlberg has expressed the view that KO is part of the metasciences (Dahlberg 2014; here cited from Dodebei 2014):

> I consider Knowledge Organization as a subdiscipline of Science of Science with application fields not only in the Information Sciences but also for all subject fields (domains) needing Taxonomies (classification systems of objects) and other fields like Statistics, Commodities, Utilities, Weapons, Patents, Museology etc. According to Science Theory, every domain has its own area of objects and of methods and processes, next to other relationships.

Also, Hjørland (2011b) claims the importance of theory of knowledge for indexing and information retrieval. Today medical doctors often rely on systematic reviews based on the paradigm termed evidence based medicine (EBM, or interdisciplinarily: evidence based practice, EBP). By implication, indexing and retrieval have to adapt to the criteria for what counts as knowledge in this paradigm. The same is of course the case in other fields and in the case of conflicting paradigms. In general: Criteria for organizing knowledge are to be found in the subject fields, their theories and paradigms. It is therefore important with Dahlberg to consider KO as a science of science.

From the domain-analytic perspective, the term KO better reflects the connection to the metasciences than does the term information organization, IO. KO points to the related fields of the history, philosophy and soci-

ology of knowledge (among other fields). This is one argument that KO should be considered the preferred term (see further in Hjørland 2012b).

A model of a domain-analytic study is Ørom (2003) who identified different "paradigms" in art studies and compared them with major library classification systems.

2.0B. Approaches developed outside KO (but representing competing approaches, which are necessary to consider)

2.6 Bibliometric approaches

Bibliometrics (with altmetrics, informetrics, scientometrics and webometrics) is an interdisciplinary field with strong affiliations to LIS. This field developed techniques for producing bibliometric maps based on co-citation analysis, bibliographic coupling or direct citation. Such maps may serve information retrieval and is a form of competing or supplementary approach to knowledge organization, although the fields of KO and bibliometrics have so far not had much mutual contact. Among the main bibliometric researchers are names such as Eugene Garfield, Henry Small and Howard D. White. Bibliometric methods are sometimes considered "objective," but Hjørland (2013a and 2016b) argues that this is not the case, and considers the strong and weak sides of this approach to KO.

2.7 IR approaches

Information retrieval (IR) is today a term mainly related to computer science. Formerly it had strong relations to information science, but the field has largely immigrated to computer science. Among the basic assumptions and techniques in this approach is the study of the statistical relation between terms, documents and collections of documents. Among the main IR researchers are names such as Gerald Salton, Karen Spärck Jones, Stephen Robertson and C. J. "Keith" van Rijsbergen. Again, if the purpose of a KOS is to help users identifying relevant documents, then IR is a family of competing approaches compared to approaches studied by the KO community. As such, it is a very successful family of approaches. Robertson (2008) stated, "statistical approaches won, simply. They were overwhelmingly more successful [compared to other approaches such as thesauri]." This issue is further addressed in, for example, Hjørland (2016a)

2.8 Other approaches

Many other approaches exist. Here just two will be mentioned. Heinrich Herre (2013) discussed an ontological approach that provides formal specifications and harmonized definitions of concepts used to represent knowledge of specific domains. It make use of the onto-axiomatic method, of graduated conceptualizations, of levels of reality, and of top-level-supported methods for ontology-development.

Jack Andersen is today a main representative of a genre approach to knowledge organization. He writes (2015, 14-15):

> As Bazerman (2012) reminds us, while recognizing the social importance of effective search engines and other systems of structuring knowledge and inscribing writing, we still need to understand the activity contexts of those producing and using knowledge and information because no matter how fragmentary, how automatic, and how fast information comes to a user, the very user (herself/himself placed in an activity contexts [sic]) must ultimately make sense of the information found and that sense cannot be made without understanding the various of activity (and the practices) producing that information.

We have now presented an overview of approaches to KO and competing approaches from outside KO. It is obvious that these as well as other approaches need careful considerations and that important strategic decisions are involved in this theory choice. The future of the field of KO is dependent on whether the research, teaching and practice in the future provide helpful systems and services for given user groups or whether existing systems like Google already provide satisfying results. A core issue is therefore to evaluate the relative strength and weaknesses of different approaches. As already stated, Hjørland (2015a) argued that for serious purposes such as medical decisions classical databases are still needed and that KO need to be further developed to make searches more efficient.

3.0 KO on different technological platforms

Ideally, KO should be understood as a knowledge base that can be applied to all technological platforms. However, its development has often been technology-driven. Therefore, an overview of KO in different platforms therefore is provided in this section.

3.1a KO in physical libraries

KOSs in libraries are mainly classification systems and indexing systems such as the *Dewey Decimal Classification* (*DDC*) and the *Library of Congress Subject Headings* (*LCSH*).

Library classification systems may be developed for the double function of shelving physical documents and

as a tool for information retrieval (IR) and for browsing in printed catalogs (from the 1980s in OPACs, online public access catalogs). The function as shelving tools puts major restrictions on classifications because such systems must arrange all documents in a linear sequence. This double function of classification systems may be an economic and management advance in some contexts but it implies that the function of classifications as IR tool is based on restrictions that are unnecessary from the retrieval perspective.

While many (big) libraries developed tailored classifications, some systems were used by many libraries and may be considered kinds of standards. Among the best known library classification systems are the *DDC* (first published in 1876, 23nd edition published in 2011), the Library of Congress *Classification* (LC*C*) 1901- (regularly updated) and the Universal Decimal Classification (UDC) first published 1905-1907 (latest "full edition" 2005).

From a research perspective, we may ask what kind of theory underlies such KOSs? It could be said that *DDC* emphasizes practicalities, efficient management, and standards rather than a scholarly, theoretical approach. It is the world's most widely used library classification system, but is not optimal to any particular collection or target group and does not – according to, among others, James Blake (2011, 469-70)—reflect current scientific knowledge. Although Blake found that "such 'outdated' classifications may still do their job well" (470) this seems to reflect a lack of ambition in providing up-to date information and to prioritize library management issues rather than advanced IR requirements. *DDC* is probably the system, which has meant most for the institutionalization and ideology of LIS and KO.

LC*C* was developed on the basis of the collections of Library of Congress thus reflecting this specific collection. Major principles of this system are its basis in "literary warrant" and enumeration of classes (as opposed to faceted systems). Vanda Broughton (2004, 143) wrote about it: "It is quite hard to discern any strong theoretical principles underlying LC*C*." Some formulations by S. R. Ranganathan (e.g. 1951) also suggest that such "traditional" systems seem to lack a theoretical foundation (in his eyes as opposed to his own approach). The LC*C* and UDC reflected in the past much better current scholarly knowledge compared to *DDC* (but the UDC scheme in particular has not generally been updated, cf. Hjørland 2007a). When it is said that such systems lack a theoretical foundation it can be argued that their implicit principles are:

> 1. that they should reflect current subject knowledge: that their theoretical basis should be found in the epistemological assumptions on which they reflect the subject fields covered;
>
> 2. the principle of literary warrant first formulated by Hulme (1911) which means that they are based on the literature they classify (in particular, the LC*C* is based on classifying the books in the Library of Congress, but because the size of the collections it has turned out to be fruitful for many other large research libraries).

Faceted library classification systems developed in the first half of the twentieth century as opposed to enumerative systems. LC*C* is a model of an enumerative system, in which all the classes are listed (and the systems therefore is comprehensive, LC*C* fills up about 41 volumes). Faceted systems, on the other hand, do not list all their classes, but provide building blocks from which specific classes for each document may be formed (Ranganathan was inspired by the Meccano toy). While UDC may be considered a forerunner partly based on facet analytic principles, the most well-known systems in this tradition is the *Colon Classification* (*CC*) developed by S. R. Ranganathan in 1933 and the *Bliss Bibliographic Classification*, 2nd ed. (*BBC*2) developed by Jack Mills, Vanda Broughton and others from 1977 (still in progress). While these systems represent a progress in research and development, their practical influence have been disappointing—although their principles have gradually influenced other systems, including the *DDC*.

BBC2, *CC*, *DDC*, LC*C* and UDC are universal systems, covering all fields of knowledge, although some (e.g. *BBC2*, LC*C* and UDC) may be considered sets of domain-specific systems, which as a whole form a universal system. Universal systems are less important for special libraries and for scholarly subject retrieval compared with special systems designed for subject bibliographies such as MEDLINE or PsycINFO. When online bibliographic databases developed from about 1963 (cf., Hahn 1998), the development of domain-specific thesauri for online searching became a research front in KO. Some researchers, for example, Szostak, Gnoli & López-Huertas (2016), argue, however, that universal systems are important for interdisciplinary research. Although research is still done on library classification and indexing systems, this area have lost importance compared with research on other kinds of KOSs better adapted to and used in online retrieval systems.

The most used basis for organizing in universal systems has been division (or collocation) by scholarly disciplines. *DDC*, for example, states that (Dewey 1979, p. xxxi):

> A work on water may be classed with many disciplines, such as metaphysics, religion, economics, commerce, physics, chemistry, geology, oceanography, meteorology, and history. No other feature of

> the *DDC* is more basic than this: that it scatters subjects by discipline.

The alternative principle, collocation by phenomena has also sometimes been preferred and used and have supporters (see, for example, Ahlers Møller 1981; Beghtol 2004; Brown 1914; Szostak, Gnoli & López-Huertas 2016).

During the 1980s, library catalogs became available as online public access catalogs (OPACs). This allowed users to search the catalog from remote terminals, e.g. from the users' homes. OPACs also provide better search possibilities, but to a wide extend continued to use the same kinds of KOSs as developed in the age of the card catalog.

3.1b KO in archives

Archival science is an independent field with its own journals, conferences, textbooks and encyclopedias (e.g. Fox and Wilkerson 1998; Duranti and Franks 2015). Knowledge organization in archives should, however, also be considered a part of KO as defined in the beginning of this article. Archives may contain official records, business records, images, letters, diploma etc. The most important specific principle of organization for this domain is the principle of provenance (Wikipedia 2016):

> Provenance is a fundamental principle of archival science, referring to the individuals, groups, or organizations that originally created or received the items in a collection, and to the items' subsequent chain of custody. According to archival theory and the principle of provenance, records which originate from a common source (or fonds) should be kept together—either physically, or, where that is not practicable, intellectually in the way in which they are catalogued and arranged in finding aids—in accordance with what is sometimes termed the principle of archival integrity or respect des fonds. Conversely, records of different provenance should be separated. In archival practice, proof of provenance is provided by the operation of control systems that document the history of records kept in archives, including details of amendments made to them. The authority of an archival document or set of documents of which the provenance is uncertain (because of gaps in the recorded chain of custody) will be considered to be severely compromised.

Archives normally collect unique objects in contrast to libraries collecting single copies of published works of which many more copies typically exist. For further information about KO in archives, see also Sweeney (2010).

3.1c KO in museums

Museology or museum studies is like archival science an independently organized field. Museums have—like archives and libraries—developed systems for organizing their objects and the knowledge they transmit (cf., Neilson 2010). Like archives, museums normally collect unique objects.

ICONCLASS is an example of a subject-specific international classification system for iconographic research and the documentation of images. It contains definitions of objects, persons, events, situations and abstract ideas that can be the subject of an image. It consists of a classification system with approximately 28,000 definitions, an alphabetical index, and a bibliography with 40,000 references to books and articles of iconographical and cultural historical interest (http://www.iconclass.nl/).

Ørom (2003) points out that the organizing principles of museum exhibitions may reflect a worldview or a scholarly paradigm that is not only reflected in the organization of objects in museums, but as well in the literature and in the classification systems in libraries. In other words, Ørom demonstrated a common theoretical basis of KOSs.

3.2 KO in classical bibliographic databases

The electronic scholarly databases such as MEDLINE, PsycINFO and Science Citation Index developed earlier than OPACs and in many ways represent IR-systems that are more advanced. Such "classical databases" are records databases (cf., Voss 2013, 79) in which each document is represented by a bibliographical record consisting of many separate fields, thus providing well-defined data. Hjørland (2015a) made the following points:

- A given record contains a mixture of fields derived from the document it represents, as well as information added by the database producer (it may contain the whole document in addition to value-added information). It may also contain information imported from third parties, as well as user-added information in the case of social tagging and related technologies
- Some fields contain controlled vocabularies developed by information professionals (e.g., descriptors and classification codes). Other fields contain "natural language" (i.e., the authors' language for special purposes). Many databases today include citation indexing, that is, the possibility of searching bibliographical references in each document represented.
- It is important to realize that the efficiency of given fields for optimal search strategies is relative from domain to domain (the value of searching document ti-

tles, for example, varies according to how titles are used in different domains. In the social sciences, for example, the use of metaphors may thus limit the value of title searches). See further in Hjørland and Kyllesbech Nielsen (2001).
- By implication, the experienced searcher should know not just about the database systems and the bibliographic records (or full-text records) but also about the concepts and genres of the primary literature. This aspect connects information science with fields such as scholarly communication, written composition, genre studies, and language for special purposes. Whereas KO in the narrow sense is about the design of bibliographical records and systems of controlled vocabularies, KO in the broader sense is about how knowledge is organized in different domains and how this can be used for IR. This broader perspective on KO becomes increasingly important in the context of full-text databases and the Internet.
- The specific requirement of indexing for Boolean searches is to represent different facets of the document that are used in the search process. Each facet is constructed by combining terms with the Boolean operator "or," and the facets are combined by the Boolean operator "and." This is known as "building blocks search strategy" (cf., Harter 1986, 242), which in some respects looks like the facet analytic approach mentioned above, but faceted classifications are seldom or never used for this purpose. Perhaps the reason is that whereas the tradition of facet analysis is mainly logical or speculative, it is important to anticipate what facets need to be combined during a search. For example, in evidence based practice (EBP) the methodological facet is important (e.g., a descriptor for randomized clinical trials; which, by the way, has only been considered in MEDLINE since about 1994, after the breakthrough of the EBP paradigm). See further in Hjørland (2011a). The construction of facets should therefore be based on studies of researchers' criteria of relevance (which are most explicated in EBP).

The most important developments in the classical databases from the perspective of KO were:

> 1. The study of the relative importance of "natural language" and "controlled vocabularies" (cf., Svenonius 1986)
> 2. The realization that many different "subject access points" (SAP) supplements each other and that no system can guarantee a full retrieval of relevant records without noise (cf. Hjørland & Kyllesbech Nielsen 2001)
> 3. Emphasis on the development of domain specific thesauri for IR.
> 4. The development of citation indexes and derivatives such as bibliometric maps based on co-citation analysis and related techniques as unique and fundamentally different kinds of KOSs (see Hjørland 2013a).

Classical bibliographical databases have in general lost importance compared with Internet search engines. It is today an open question whether, for example, the traditional thesaurus still has a role to fill in modern information retrieval (see Dextre Clarke and Vernau 2016). Hjørland (2015a) argued however, that for serious scholarly purposes it is important that users or intermediaries are able to control the search process. For such tasks, classical databases seem to be the most advanced tools.

3.3 KO on the Internet

The Internet and its search engines have revolutionized the way people search and find information. Compared to classical databases which require professional information specialists or information competent end-users, search engines are (or seems to be) very easy to use. In addition, search engines have a broad and comprehensive coverage of many kinds of documents. The Internet has become the most important medium for organizing and searching information and documents. The field information architecture developed as a new field, which is concerned with organizing knowledge on the WWW (see e.g., Rosenfeld and Morville 1998). Its medium is new, but its basic principles are part of the field of KO as defined in the beginning of this article.

In parallel with the development of the Internet, a new kind of KOS termed "ontologies" became important from the 1990s. Dagobert Soergel wrote (1999, 1119):

> Classification has long been used in library and information systems to provide guidance to the user in clarifying her information need and to structure search results for browsing, functions largely ignored by the text retrieval community but now receiving increasing attention in the context of helping users to cope with the vast amount of information on the Web. Fairly recently, other fields, such as AI, natural language processing, and software engineering, have discovered the need for classification, leading to the rise of what these fields call ontologies .... But a classification by any other name is still a classification (1120).

Soergel's main point of view is thus that ontologies basically are classification systems and that they represent a "reinvention of classification" by new research communities with little communication and mutual learning in relation to the field of KO. Ontologies may, however, be considered more general and more abstract forms of KOS. All traditional forms such as classification systems and thesauri may just be understood as restricted kinds of ontologies. Lars Marius Garshol wrote (2004): "Topic maps [ontology-based systems] can actually represent taxonomies, thesauri, faceted classification, synonym rings, and authority files, simply by using the fixed vocabularies of these classifications as a topic map vocabulary."

Given this perspective, it seems less important for KO to investigate specific forms of KOSs such as classification systems or thesauri. It seems important to make the abstraction to systems of concepts and their semantic relations (Hjørland 2007b) and to understand each specific kind of KOS as based on principles that are general for all KOSs. Hjørland (2016a) thus argued that thesauri would benefit from adopting some of the principles used in ontologies.

From the point of view of KO as a field of research, teaching and practice, we may ask: What are the implications for us? Is classification still needed after Google? (cf. Hjørland 2012a). Does the traditional thesaurus have a place in modern information retrieval? (cf. Dextre Clarke and Vernau 2016). To answer these questions we have to examine the potential value of different approaches and on this basis estimate how we may continue contributing making documents findable, cf. section 2: Research traditions, approaches and basic theoretical issues in KO.

4.0 Other names and other fields

The term KO is also used in other fields such as cognitive psychology and information management (and is thus a homonym). This entry is about KO as related to LIS (KO in a narrow sense). KO in a broader sense is concerned with:

1. How knowledge is organized in society, e.g. in scholarly disciplines and in the social division of labor. This is a social KO perspective and is in particular relevant for disciplinary classification. An example is Oleson & Voss (1979) *The Organization of Knowledge in Modern America, 1860-1920*. This dimension is covered by fields such as the "sociology of knowledge" and the "social history of knowledge," among others.
2. How knowledge is organized in scholarly theories such as biological taxonomies. This may be termed "intellectual classification" (as opposed to social KO). An example is Fjeldså (2013) about the classification of birds. This dimension is covered by the single sciences and in fields such as philosophy and science studies.

This differentiation of the social and intellectual organization of knowledge is here taken from Whitley ([1984]2000). There are of course mutual interactions between these social KO and intellectual KO. KO in the narrow sense is dependent on KO in the broader sense (i.e. subject knowledge about intellectual classification, for example, the classification of documents about birds reflects how birds themselves are classified).

As described elsewhere in this article there is a tendency that different aspects of KO isolate themselves using separate names such as "information architecture." One of the basic claims in this entry is, however, that the phenomena listed in the beginning of this article need to be considered independent of the specific media on which they are used and independent of the specific traditions and methodologies in which they have been investigated.

5.0 Conclusion

KO may be understood in narrow as well as in broad senses. The narrow senses are, for example, the KOS and KOPs taking place within LIS. The broad senses are, for example, the conceptual systems, social fields and activity systems existing or taking place in all spheres of society. For us in the KO community within LIS, the purpose of studying and teaching KO is to develop better information services, whatever that means. Different approaches and theories exists in and outside KO and it is strategically important that our teaching and research in KO is based on well-considered and well-informed choices. The broad kind of KOSs (e.g. activity systems and scientific theories) are important because they form the background knowledge needed in order to organize knowledge in the narrow LIS sense (see Hjørland 2015b).

References

Ahlers Møller, Bente. 1981. "Subject Analysis in the Library. A Comparative Study." *International Classification* 8:23-27.

Andersen, Jack. 2015. "Re-describing Knowledge Organization." In *Genre Theory in Information Studies*, ed. Jack Andersen. Bingley, UK: Emerald Group Publishing Limited, 13-42.

Anderson, James D. and José Pérez-Carballo. 2001a. "The Nature of Indexing: How Humans and Machines Analyze Messages and Texts for Retrieval. Part I: Re-

search, and the Nature of Human Indexing." *Information Processing & Management* 37:231-54.

Anderson, James D. and José Pérez-Carballo. 2001b. " The Nature of Indexing: How Humans and Machines Analyze Messages and Texts for Retrieval. Part II: Machine Indexing, and the Allocation of Human versus Machine Effort." *Information Processing & Management* 37:255-77.

Bazerman, Charles. 2012. "The Order of Documents, the Order of Activity, and the Order of Information." *Archival Science* 12:377-88.

Beghtol, Clare. 2004. "Exploring New Approaches to the Organization of Knowledge: The Subject Classification of James Duff Brown." *Library Trends* 52, no. 4:702-18.

Blake, James. 2011. "Some Issues in the Classification of Zoology." *Knowledge Organization* 38:463-72.

Bliss, Henry Evelyn. 1929. *The Organization of Knowledge and the System of the Sciences*. New York: Henry Holt and Company.

Bliss, Henry Evelyn. (1933) 1939. *The Organization of Knowledge in Libraries and the Subject-Approach to Books*. 2nd ed. New York: H. W. Wilson. Citations refer to the 2nd ed.

Broadfield, A. 1946. *The Philosophy of Classification*. London: Grafton.

Broughton, Vanda. 2004. *Essential Classification*. London: Facet Publishing.

Brown, James Duff. 1914. *Subject Classification, with Tables, Indexes, etc., for the Sub-division of Subjects*. 2nd rev. ed. London: Grafton.

Campbell, D. John. 1976. "A Short Biography of Henry Evelyn Bliss (1870-1955)." *Journal of Documentation* 32:134–45.

Dahlberg, Ingetraut. 2010. "International Society for Knowledge Organization (ISKO)." In *Encyclopedia of Library and Information Sciences*, 3rd ed., ed. Marcia J. Bates & Mary Niles Maack. Boca Raton, Florida: CRC Press, IV:2941-49.

Dewey, Melvil. 1979. *Dewey Decimal Classification and Relative Index*. 19th ed., v. 1). Albany, NJ: Forest Press.

Dextre Clarke, Stelle and Judi Vernau, eds. 2016. "Special Issue: The Great Debate: 'This House Believes that the Traditional Thesaurus has no Place in Modern Information Retrieval.'" *Knowledge Organization* 43:135-214.

Dodebei, Vera. 2014. "13ª ISKO International Conference, 19-22 Maio 2014, Cracóvia, Polônia: Relatório da participação da ISKO-Brasil, Maio de 2014." http://isko-brasil.org.br/wp-content/uploads/2014/06/relat_iskoCrac%C3%B3via2014.pdf

Drobnicki, John A. 1996. *Bliss: The Man and the Classification*. New York: City University of New York (CUNY): CUNY Academic Works. http://academicworks.cuny.edu/cgi/viewcontent.cgi?article=1014&context=yc_pubs

Duranti, Luciana and Patricia C. Franks. 2015. *Encyclopedia of Archival Science*. Lanham, Md.: Rowman & Littlefield.

Fjeldså, Jon. 2013. "Avian Classification in Flux." In *Handbook of the Birds of the World*, special volume 17: *New Species and Global Index*, ed. J. del Hoyo et al. Barcelona: Lynx Edicions, 77-146.

Fox, Michael J. and Peter L. Wilkerson, Peter L. 1998. *Introduction to Archival Organization and Description*. Los Angeles, CA: Getty Information Institute.

Garfield, Eugene. 1974. "The 'Other' Immortal, A Memorable Day with Henry E. Bliss." *Wilson Library Bulletin* 49, no. 4: 288-92. http://www.garfield.library.upenn.edu/essays/v2p250y1974-76.pdf

Garshol, Lars Marius. 2004. "Metadata? Thesauri? Taxonomies? Topic maps! Making sense of it all." *Journal of Information Science* 30:378-91. http://www.ontopia.net/topicmaps/materials/tm-vs-thesauri.html

Hahn, Trudi Bellardo. 1998. "Text Retrieval Online: Historical Perspectives on Web Search Engines." *Bulletin of the American Society for Information Science* 24, no. 4:7-10. http://www.asis.org/Bulletin/Apr-98/hahn.html

Harter, Stephen P. 1986. *Online Information Retrieval. Concepts, Principles, and Techniques*. New York: Academic Press.

Herre, Heinrich. 2013. "Formal Ontology and the Foundation of Knowledge Organization." *Knowledge organization* 40:332-39.

Hjørland, Birger. 2002. "Principia Informatica: Foundational Theory of Information and Principles of Information Services." In *Emerging Frameworks and Methods. Proceedings of the Fourth International Conference on Conceptions of Library and Information Science* (CoLIS4), ed. Harry Bruce, Raya Fidel, Peter Ingwersen and Pertti Vakkari. Greenwood Village, Col.: Libraries Unlimited, 109-21.

Hjørland, Birger. 2007a. "Arguments for 'The Bibliographical Paradigm.' Some Thoughts Inspired by the New English Edition of the UDC." *Information Research* 12, no 4: paper colis06. http://InformationR.net/ir/12-4/colis/colis06.html

Hjørland, Birger. 2007b. "Semantics and Knowledge Organization." *Annual Review of Information Science and Technology* 41:367-405.

Hjørland, Birger. 2008. "What is Knowledge Organization (KO)?" *Knowledge Organization* 35:86-101.

Hjørland, Birger. 2011a. Evidence-Based Practice. An Analysis Based on the Philosophy of Science. *Journal of the American Society for Information Science and Technology*, 62:1301-10.

Hjørland, Birger. 2011b. "The Importance of Theory of Knowledge: Indexing and Information Retrieval as an Example." *Journal of the American Society for Information Science and Technology* 62:72-77.

Hjørland, Birger. 2012a. "Is Classification Necessary after Google?" *Journal of Documentation* 68:299-317.

Hjørland, Birger. 2012b. "Knowledge Organization = Information Organization?" In *Categories, Contexts and Relations in Knowledge Organization: Proceedings of the Twelfth International ISKO Conference, 6 - 9 August 2012, Mysore, India*, ed. A. Neelameghan and K.S. Raghavan. Advances in Knowledge Organization v. 13. Würzburg: Ergon-Verlag, 8-14.

Hjørland, Birger. 2013a. "Citation Analysis: A Social and Dynamic Approach to Knowledge Organization." *Information Processing & Management* 49:1313–25.

Hjørland, Birger. 2013b. "Facet Analysis: The Logical Approach to Knowledge Organization." *Information Processing & Management* 49:545–57.

Hjørland, Birger. 2013c. "User-based and Cognitive Approaches to Knowledge Organization: A Theoretical Analysis of the Research Literature." *Knowledge Organization* 40:11–27.

Hjørland, Birger. 2015a. "Classical Databases and Knowledge Organization: A Case for Boolean Retrieval and Human Decision-Making During Searches." *Journal of the Association for Information Science and Technology* 66:1559-75. doi: 10.1002/asi.23250

Hjørland, Birger. 2015b. "Theories are Knowledge Organizing Systems (KOS)." *Knowledge Organization* 42:113-28.

Hjørland, Birger. 2016a. "Does the Traditional Thesaurus Have a Place in Modern Information Retrieval?" *Knowledge Organization* 43:145-59.

Hjørland, Birger. 2016b. "Informetrics Needs a Foundation in the Theory of Science." In *Theories of Informetrics and Scholarly Communication*, ed. Cassidy Sugimoto. Berlin: Walter de Gruyter, 20-46.

Hjørland, Birger and Lykke Kyllesbech Nielsen. 2001. "Subject Access Points in Electronic Retrieval." *Annual Review of Information Science and Technology* 35:249-98.

Hulme, E. Wyndham. 1911. "Principles of Book Classification." *Library Association Record* 13:354-58; 389-94; 444-49.

Jensen, Povl Johannes. 1973. *Catalogue and Scholarship: D. G. Moldenhawers Catalogue in the Royal Library of Copenhagen*. Copenhagen: The Royal Library.

Leydesdorff, Loet. 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* 57:601-13.

Mai, Jens-Erik. 2011. "The Modernity of Classification." *Journal of Documentation* 67:710-30.

Mills, Jack. 2004. "Faceted Classification and Logical Division in Information Retrieval." *Library Trends* 52, no 3:541-70.

Neilson, Dixie. 2010. "Museum Registration and Documentation." In *Encyclopedia of Library and Information Sciences,* 3rd ed., ed. Marcia J. Bates and Mary Niles Maack. Boca Raton, Florida: CRC Press, V:3739-53.

Oleson, Alexandra and John Voss, eds. 1979. *The Organization of Knowledge in Modern America, 1860-1920*. Baltimore: Johns Hopkins University Press.

Pudovkin, Alexander I. and Eugene Garfield. 2002. "Algorithmic Procedure for Finding Semantically Related Journals." *Journal of the American Society for Information Science and Technology* 53:1113-19.

Ranganathan, Shiyali Ramamrita. 1951. *Philosophy of Library Classification*. Copenhagen: E. Munksgaard.

Robertson, Stephen. 2008. "The State of Information Retrieval." ISKO-UK. Presentation and audio recording freely available at http://www.iskouk.org/content/state-information-retrieval-researchers-view

Rosenfeld, Louis and Peter Morville. 1998. *Information Architecture for the World Wide Web*, 1st ed. Cambridge Sebastopol, CA: O'Reilly.

Soergel, Dagobert. 1999. "The Rise of Ontologies or the Reinvention of Classification." *Journal of the American Society for Information Science* 50:1119-20.

Svenonius, Elaine. 1986. "Unanswered Questions in the Design of Controlled Vocabularies." *Journal of the American Society for Information Science* 37:331-40.

Sweeney, Shelley. 2010. "Provenance of Archival Materials." In *Encyclopedia of Library and Information Sciences,* 3rd ed., ed. Marcia J. Bates and Mary Niles Maack. Boca Raton, Florida: CRC Press, VI:4315-23.

Szostak, Rick, Claudio Gnoli and María López-Huertas. 2016. *Interdisciplinary Knowledge Organization*. Cham: Springer.

Voss, Jakob. 2013. *Describing Data Patterns: A General Deconstruction of Metadata Standards*. Berlin: Humboldt University (Dissertation). http://edoc.hu-berlin.de/

Whitley, Richard R. (1984) 2000. *The Intellectual and Social Organization of the Sciences*. Oxford: Oxford University Press. 2nd ed. with a new introduction. Citations refer to the 2nd ed.

Wikipedia, the free encyclopedia. Provenance. http://en.wikipedia.org/wiki/Provenance

Ørom, Anders. 2003. "Knowledge Organization in the Domain of Art Studies: History, Transition and Conceptual Changes." *Knowledge Organization* 30:128-43.