Theories are Knowledge Organizing Systems (KOS)

Birger Hjørland

Royal School of Library and Information Science, University of Copenhagen, 6 Birketinget, DK-2300 Copenhagen S, Denmark, <pnl617@iva.ku.dk>

Birger Hjørland holds an M.A. in psychology and Ph.D. 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 and coordinator of computer based information services 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 (February 2015) 37 in Google Scholar and 21 in Web of Science.

Hjørland, Birger. Theories are Knowledge Organizing Systems (KOS). Knowledge Organization. 42(2), 113-128. 100 references.

Abstract: The notion “theory” is a neglected concept in the field of information science and knowledge organization (KO) as well as generally in philosophy and in many other fields, although there are exceptions from this general neglect (e.g., the so-called “theory theory” in cognitive psychology). This article introduces different conceptions of “theory” and argues that a theory is a statement or a conception, which is considered open to be questioned and which is connected with background assumptions. Theories form interconnected systems of grand, middle rank and micro theories and actions, practices and artifacts are theory-laden. The concept of knowledge organization system (KOS) is briefly introduced and discussed. A theory is a form of KOS and theories are the point of departure of any KOS. It is generally understood in KO that concepts are the units of KOSs, but the theory-dependence of concepts brings theories to the forefront in analyzing concepts and KOSs. The study of theories should therefore be given a high priority within KO concerning the construction and evaluation of KOSs.

Received: 1 January 2015; Revised: 16 February 2015; Accepted: 17 February 2015

Keywords: theory, theories, concept, knowledge, knowledge organization, information science, knowledge organization systems

1.0 Introduction

Thomas Kuhn’s well-known book The Structure of Scientific Revolutions (1st edition 1962) is broadly understood as a turning point in the philosophy of science. Although there are different views about the role and quality of this book, it contributed—according to conventional wisdom—to bringing logical positivism to a fall. The substitute to positivism (as well as logical positivism’s dual ancestors: empiricism and rationalism) is a philosophy that is closer related to historicism, pragmatism and hermeneutics (cf. Mallery, Hurwitz and Duffy, 1992). Whereas a key issue for the logical positivists was “verified facts,” a key issue for Kuhnian philosophy is the theory-laden view of observation—and hence the set of background assumptions, “paradigms” or “theories” that guides and determines observational statements. (It has been disputed whether this “positivism” is a “straw man” because Kuhn was unspecific and unclear in his criticism of positivism; see, for example, Friedman 2003; Moges 2010 and Reisch 1991).

In his later career, Kuhn developed theories of concepts, categories and taxonomic systems and he inspired new developments in cognitive psychology and cognitive science. Among these theories, the so-called “theory-theory” is important for the present paper. These developments seem important in many respects, not least in meta-sciences such as information science.

“Theory” is a much-used term in science, philosophy and everyday life language. It is therefore strange, that it is relatively seldom examined. Routledge Encyclopedia of Philosophy (1998), Stanford Encyclopedia of Philosophy (http://plato.stanford.edu/) and the Internet Encyclopedia of Philosophy (http://www.iep.utm.edu/t/), for example, contain many articles about specific theories or special issues related to...
theory, but no entry about the concept “theory” itself. There are of course exceptions from this general neglect of theories of theory; including the aforementioned “theory theory.” There are also many scattered works including Biesta, Allan and Edwards (2014), Birns (2010), Corvellec (2013), DiMaggio (1995), Dubin (1978), Dubem (1962), Elliott and Attridge (2011), Fawcett and Downs (1986), Mintzberg (2009), Mjøset (2001), Suppe (1977, 1989, 1998), and works mentioned in other places in this article. There has in addition been a debate in administrative science/information systems research including Gregor (2006), Sutton and Straw (1995), Weick (1995) and Whetten (1989). The reader may say that the present article contains many references to the concept of “theory” and thus disconfirm the claim that is a neglected concept; this literature is, however, very scattered and unsystematic and therefore the claim that “theory” is an understudied topic is maintained. In section 3 below “levels of theories,” some unresolved problems related to the study of theories will be indicated.

The basic theses of this article are:

- The notion “theory” is important for information science and knowledge organization (KO) but is still a neglected concept.
- The positivist dichotomy between observational and theoretical statements is wrong (see Putnam 1962). Even every-day concepts (such as a “blackbird” or a “hammer”) are theoretical constructions and have theoretical implications for how we think and act (although we seldom realize this and seldom know how our concepts have been constructed).
- “Theory” itself—like other concepts—is theoretically loaded. Different philosophies such as positivism and Kuhnian theory provide different meanings of the term “theory.”

In information science, the notion of theory operates on two levels:

The information producers, mediators and users studied by information scientists are acting according to their pre-understanding and perspectives, i.e., theories. The understanding and explaining of these actors and their products (i.e. “information”) must therefore involve an analysis of the theories that have influenced these actors.

When information scientists study information, information systems and information users, they always operate from a theoretical pre-understanding, from a certain theoretical perspective or position—which may be termed a “meta-theory,” a “paradigm” or a “tradition”.

Of specific relevance for KO is the claim that a theory is a knowledge organization system (KOS) and vice versa: Any KOS is, if not a theory, at least theoretically and ideologically loaded.

2.0 What is a Theory?

As a start, we shall have a look at how a standard dictionary defines “theory.” WordNet 3.1 distinguishes three meanings:

- theory (a well-substantiated explanation of some aspect of the natural world; an organized system of accepted knowledge that applies in a variety of circumstances to explain a specific set of phenomena); “theories can incorporate facts and laws and tested hypotheses,” “true in fact and theory.”
- hypothesis, possibility, theory (a tentative insight into the natural world; a concept that is not yet verified but that if true would explain certain facts or phenomena); “a scientific hypothesis that survives experimental testing becomes a scientific theory”; “he proposed a fresh theory of alkalis that later was accepted in chemical practices”
- theory (a belief that can guide behavior): “the architect has a theory that more is less”; “they killed him on the theory that dead men tell no tales.”

We see here, that in a broad sense “theory” is synonym with “belief,” in a narrow sense it requires that such a belief is “a well-substantiated explanation of some aspect of the natural world” (one may add some aspect of the social-cultural world or of any “world”). To say—as in the two first senses—that a theory is something that is not yet verified implies that some knowledge claims are verified to a degree that can no longer be considered theories. This distinction between knowledge claims that are verified and those that are not fully verified runs into difficulties because of the principle known as fallibilism, according to which “any claim justified today may need to be revised or withdrawn in light of new evidence, new arguments, and new experiences” (Kompridis 2006, 180). Philosopher Karl Popper (1980, 111) wrote:

The empirical basis of objective science has thus nothing ‘absolute’ about it. Science does not rest upon solid bedrock. The bold structure of its theories rises, as it were, above a swamp. It is like a building erected on piles. The piles are driven down from above into the swamp, but not down into any natural or “given” base; and if we stop driving the piles deeper it is not because we have reached firm ground. We simply stop when we are satisfied that
the piles are firm enough to carry the structure, at least for the time being.

Fallibilism—if it is accepted—implies that no knowledge can be considered finally proved (some versions of fallibilism restrict this principle to scientific theories, cf., Rescher 1998). Another way to say this is that all knowledge shall be considered theories, although some knowledge claims are considered extremely “robust” and well established. The implications of considering fallibilism a true doctrine may be more fruitful for KO compared to considering it wrong: By considering it true, we are open to different perspectives and the only thing we seem to lose is a somewhat forced language in which even claims such as “we shall all die” are termed “theories.” Fallibilism does not insist on the falsity of our claims or that knowledge is unavailable, but rather on their openness to revision: It does not imply skepticism (the view that we fail to know anything). In any case: The evaluation of knowledge claims should be based on the quality of the arguments, including, of course, its empirical support.

The concept “theory” is thus connected with the issue of how certain knowledge claims are and whether they can be finally verified or not. Different theories of knowledge and science have different views on this issue and therefore imply different conceptions of “theory.” According to William Outhwaite (1998),:

The “standard view” in Anglo-American philosophy of science derived from the Vienna Circle’s logical empiricism and [was] consolidated by Karl Popper in the middle decades of the twentieth century …. In this conception, theories consist essentially of law-like statements verified or, in Popper’s more refined version, falsified, in a fairly direct confrontation with “the facts.”

The article continues:

Both elements of this relation came under fire from three directions. [1] Philosophers of science such as Mary Hesse, Rom Harré, Norwood Russell Hanson and Michael Scriven questioned the deductivist model of scientific theory. [2] Historians and sociologists of science, building on the pioneering work of Thomas Kuhn (1962) …, noted that scientists were much more collectivistic and conservative in their theoretical affiliations than Popper’s model suggested …. [3] And social scientists, beginning with historians and philosophers of history, pointed out that explanations by reference to general laws had little application in the social world.

Today the term “theory” is therefore understood differently in different disciplines and by different epistemologies (Heelan and Schulkin 1998, 274-5):

Philosophers of science, seeing physics as the privileged exemplar of science, took theory to be a mathematical model tested against observations, while pragmatists, seeing experimental praxis as the privileged exemplar of science, took theory to be descriptive of scientific entities as these were perceived in laboratory praxis. The biological and social sciences today tend to use “theory” in the latter sense, while the physical sciences continue to exploit the mathematical imagination in search of new theoretical models, or in their terms, simply theories. One of the problems that hermeneutical method and philosophy will address is the diversity in the meaning given to theories and the usage of the term.

Sven-Eric Liedman (2013) provided a short history of the concept of theory. He quoted Gadamer’s Warheit und Methode (1960) and wrote that the notion of theories as constructions which succeed one another was deeply akin to his way of thinking (p. 45).

In critical theory (see, for example, Geuss 1998 and Zima 2007) there has also been a rejection of a widely held view about what a “theory” is: a set of formally specified and interconnected general propositions that can be used for the successful explanation and prediction of the phenomena in some object domain (Geuss 1998).

This conception of theory, the members of the Frankfurt School argued, is extremely misleading because it directs attention away from the social context within which theories necessarily arise, are tested and are applied, and which alone they are fully comprehensible. The term “theory” should be used in the first instance to designate a form of (ideally social) activity with an especially salient cognitive component, and only derivatively for the propositions that might be formulated in the course of such activity.

Likewise, Zima (2007, 14 emphasis original) suggests the following alternative definition to the commonly held view: “Theory is an interest-guided discourse whose semantic and narrative structure is developed and reflected upon by a self-critical subject who is aware of the theory’s historical, social and linguistic origins.”

Are these critical-theoretical definitions of theory fruitful? Are they not just saying that theorists should be critical and reflective, without providing a helpful understanding on how a “theory” is different from a “discourse”? I agree
with the following quoted evaluation of critical theory’s contribution to philosophy (Geuss 1998):

In retrospect the most important contribution of critical theory to philosophy in the late twentieth century would seem to be their criticism of positivism and their demand that social theory be reflexive; that is, that theorists try to be as aware as possible of their own position, the origin of their beliefs and attitudes, and the possible consequences their theorizing might have on what they are studying.

I also agree with critical theory’s claim that the traditional understanding of “theory” as being based on neutral and objective science in the positivist sense is problematic, and therefore that a clear-cut distinction between “theory” and “ideology” cannot be maintained. This is something that critical theory shares with many anti-positivist theories, including versions of social epistemology (Fuller 1988), social semiotics (Hodge and Kress 1988), pragmatism (Johnson 2006), social constructivism (Bijker 2001), and more. Still, however, critical theory has in my opinion failed to provide a better definition of the term “theory.” We have to look elsewhere.

Weiskopf (2011) writes (with reference to Gopnik and Meltzoff 1997, 32-41):

Theories are bodies of information (or, as psychologists and linguists sometimes say, bodies of knowledge) about a particular domain. … But theories are not just any body of information held in memory. What makes theories distinctive or special? Keil (1989, 279) called this “the single most important problem for future research” in the Theory-Theory tradition. Weiskopf then specifies three categories of conditions that specifies theories from other bodies of information about a domain: structural, functional, and dynamic conditions:

– Structurally, theories are abstract, coherent, causally organized, and ontologically committed bodies of information.
– Functionally, theories must make predictions, interpret evidence in new ways, and provide explanations of phenomena in their domain.
– Theories are not static representations, but have dynamic properties. This follows from the fact that they develop in response to, and may gain in credibility or be defeated by, the empirical evidence.

How much weight should be placed on each of these conditions? Some researchers have suggested that different kinds of theories satisfy different conditions. Gregor (2006, 622-31) thus suggests that a theory may have four primary goals: Analysis and description; explanation; prediction and prescription, leading to a corresponding classification of theories. The single most important problem in theory-theory is thus unsolved.

A yet broader definition is (Ayers 2008, 373): “A theory can be described as a set of concepts and the relationships among them.” This definition does not, however, make any distinction between “a statement” and “a theory.” The statement: “There are nine professors at the Royal School of Library and Information Science” contains “a set of concepts and the relationships among them,” but such a statement is normally not a theory (unless somebody questions it). In other words: We use the term “theory” when we want to say that a given statement is open to be questioned. It is relative or context-dependent when we want to consider a statement open to be questioned: for some kinds of actions and inquiries a given statement may be taken for granted, while for other kinds of actions and inquiries it needs to be considered a theory. To consider something “theory” (as opposed to fact) is to consider it open for further inquiry: It places the user in an active position regarding evaluation of the theory as opposed the passive role implied by considering something facts.

Theories are not just about explicit statements. There is implicit knowledge and background knowledge and often what has been considered “facts” later turns out to be a theory. George Lakoff reports about a theory that has been invisible for centuries (1987, 6):

From the time of Aristotle to the later work of Wittgenstein, categories were thought be well understood and unproblematic. They were assumed to be abstract containers, with things either inside or outside the category. Things were assumed to be in the same category if and only if they had certain properties in common. And the properties they had in common were taken as defining the category.

This classical theory was not the result of empirical study. It was not even a subject of major debate. It was a philosophical position arrived at on the basis of a priori speculation. Over the centuries it simply became part of the background assumptions taken for granted in most scholarly disciplines. In fact, until very recently, the classical theory of categories was not even thought of as a theory. It was taught in most disciplines not as an empirical hypothesis but as an unquestionable, definitional truth.

I therefore propose the following definition: A theory is an explicit or implicit statement or conception that might be
questioned (and thus met with an alternative theory), which is more or less substantiated and dependent on other theories (including background assumptions). We use the term theory about a statement or conception when we want to emphasize that it might be wrong, biased, bad or insufficient for its intended use and therefore should be considered and perhaps replaced by another theory. Carroll and Campbell (1989) and Dillon (1995) described artifacts as theories because artifacts are the results of design processes influenced by theories about how software or things should be designed in order to satisfy users.

A dynamic force in the development of knowledge is the dialog between different theories. The post-Kuhnian emphasis on theory has also been characterized in this way (Lloyd 1993, 32):

Perhaps the greatest advance in understanding the nature of explanation made in the post-positivist and post-Kuhnian era is the general realization that methodologies, theories, and explanations are related to each other via extra-logical, historically variable constellations variously described as “background knowledge,” “traditions,” paradigms,” “research programmes,” “fields,” or “domains.” We can call all of these “framework concepts.”

Post-Kuhnian philosophy therefore acknowledges the importance of disagreements much more than did logical positivism.

The important contribution to the understanding of what “theory” means is its embeddedness in larger frameworks of human activity systems thus implying a cultural-historical and pragmatic perspective. As we shall see in the next section theories are related to larger frameworks such as metatheories, paradigms and philosophical positions.

### 3.0 Levels of Theories

The sociologist of science, Robert Merton (1910-2003) made a classification of three levels of theories and advocated theories of the middle-range, which were defined this way (1968, 39):

Theories that lie between the minor but necessary working hypotheses that evolve in abundance during day-to-day research and the all-inclusive systematic efforts to develop a unified theory that will explain all the observed uniformities of social behavior, social organization and social change.

The *SAGE* Encyclopedia of *Qualitative Research Methods* describes these three kinds of theories in the following way (Ayres 2008, 373):

Grand theories, sometimes referred to as conceptual frameworks or conceptual models, develop overall explanations for a discipline or body of knowledge. The concepts addressed by grand theories are highly abstract and cannot easily be operationalized into variables or used in hypotheses. Thus, grand theories are untestable. Some authors have described grand theories as normative; that is, that grand theories describe not the way a discipline is, but the way that discipline should be. Grand theories, though untestable, are often useful as organizing frameworks for knowledge development or as foundations for mid-range theory development.

Mid-range theories, which have been described as being particularly useful for practice disciplines, are more abstract and inclusive than micro theories but remain testable, although such testing may require a program of research or series of studies in which specific concepts and relationships in the theory are tested individually. Mid-range theories have been described as particularly useful for practice disciplines and have been the focus of recent theory development efforts in (for example) nursing.

Micro theories, sometimes referred to as partial or situational theories, have the narrowest scope. Micro theories are restricted to a particular phenomenon or, as the name suggests, situation. Some scientists have equated micro theories with research hypotheses because their narrow scope makes it possible for such theories to be tested with as little as one research study.

Each of these levels has met different interests at different times. Poole (1985) is a book about theories of the middle range in information science. Morgan and Wildemuth (2009), citing Poole (1985, 42), write: “Middle-range theories are concrete enough to clearly apply to phenomena of interest to a professional field like information and library science, while simultaneously being abstract enough to apply to settings beyond the context in which they were developed.”

Wagner and Berger (1985) considered metatheoretical frameworks as important orienting strategies and Hjörland (2011) used this understanding to explain browsing behavior. Skinner (1985) presents and discusses prominent philosophers such as Gadamer, Derrida, Foucault, Habermas, Althusser and Levi-Strauss. Its title: *The return of grand theory in the human sciences* implicates that it is a reaction to a period in which grand theory has not been pursued. Leckie, Given and Buschman (2010) present prominent critical theorists for library and information...
science—and thereby also represent “grand theory.” A reviewer, Keilty (2011, 1), wrote that the book “represents an important and reasoned contribution at the advent of critical theory to metatheoretical discourse within information studies. The necessity of its intervention cannot be overstated.” He also said (p. 2):

In certain ways, however, the volume does not live up to its potential. Most of the essays are largely formulaic: introduce a theorist, summarize a few of his/her ideas, and suggest ways in which future research within information studies might engage with those ideas. With a few exceptions, the effect of such a formula is to argue persuasively for an engagement with critical theory without doing the work of that engagement in a substantive way. Most of the essays point to future potentials rather than accomplishing the tough work of such critical analysis. Where, for example, in a discussion of the importance of interpretative analysis of representation as a corrective to empirical methods, is such a sophisticated analysis of an object of representation?

This analysis is very important. It can never be a goal in itself to introduce a theory in information science or knowledge organization if it is not explicated in what way the theory offers a new perspective for the field which has clear implications for practice.

In Table 1 below, the levels of theory from the general philosophical level (or grand theory) over metatheories (or paradigms and traditions) to specific theories and finally to practice can be understood as a hierarchy of mutually dependent theories in a domain such as information science.

It is important to realize that Table 1 is based on the idea of interacting levels of theories, and that practice cannot be understood as an alternative to theory, but is always influenced by theory (Hookway 2013):

All the pragmatists, but most of all [John] Dewey, challenge the sharp dichotomy that other philosophers draw between theoretical beliefs and practical deliberations. In some sense, all inquiry is practical, concerned with transforming and evaluating the features of the situations in which we find ourselves.

It is important to realize that a given discipline like information science is dependent on general philosophical/interdisciplinary theories. A study of the general philosophical level is not in itself, however, a contribution to information science: In order to be a contribution to our field, it is necessary to provide a well-argued proposal on how information science problems may be better solved by considering a given philosophical perspective. In the end, the purpose of all theoretical work is contributing solving information problems in the real world. Therefore, the theoretical analysis must go back and forth between practical problems, theories, and philosophies. It may, for example, consider the theoretical basis for search engines and suggest how alternative philosophies may improve such engines according to some criteria of what a good search engine should provide.

A related model of relations between levels of theory is provided by Iivari, Hirschheim and Klein (2001, 189) in the field of information systems development. They also provided a four-tiered framework with certain structural similarities to the model developed in Table 1. It will be too much to present their figure and its many concepts in the present paper, but a brief summary shall be given: At the top of their model are four information systems development paradigms (related to different views of ontology, epistemology, methodology and ethics). This is followed by eleven different information systems development approaches (related to different views on goals, guiding principles, fundamental concepts and principles for the information systems development process). This again is followed by eighteen different information systems development methodologies (relations between techniques and the detailed information systems development process). Finally, twenty different information systems development techniques (with detailed concepts and notations) are displayed. This model is clearly different from Table 1 but it is mentioned here because of a certain structural analogy with Table 1 and because it seems an important goal for information science (and all other domains) to develop similar models as a way of organizing the knowledge in a given domain. This will not be an easy task because the philosophical terminology is rather unclear, and the different paradigms, approaches, methodologies and techniques are different to identify and classify in a satisfactory way. Nonetheless, it seems an important task that need to be done—and which is interdisciplinary in nature. We need more knowledge of theories—and how to identify them in a given domain both diachronically and synchronically and how to organize them in levels from general philosophy to specific approaches.

4.0 Knowledge Organization Systems

The concept “knowledge organization system” (KOS) is today a common term in KO used as a generic term for, among other terms, classification systems, thesauri, taxonomies, ontologies, etc. (cf., Smiraglia 2014, 4). The term became common with Gail Hodge’s report (2001, section 1) about digital libraries:
The term knowledge organization systems is intended to encompass all types of schemes for organizing information and promoting knowledge management. Knowledge organization systems include classification and categorization schemes that organize materials at a general level, subject headings that provide more detailed access, and authority files that control variant versions of key information such as geographic names and personal names. Knowledge organization systems also include highly structured vocabularies, such as thesauri, and less traditional schemes, such as semantic networks and ontologies. Because knowledge organization systems are mechanisms for organizing information, they are at the heart of every library, museum, and archive.

An important question is whether one kind of KOS can be reused as the core foundation for establishing another kind? Can, for example, a classification system be transformed into a thesaurus? Or can a thesaurus be transformed into an ontology? The answer to this question is of both practical and theoretical interest. It is of theoretical interest, because a precise answer to this question reveals something about what different kinds of KOS have in common, as well as how they are principally different.

Table 1. Information science’s theories and traditions.

| General philosophical level | How are great thinkers such as Aristotle (384—322 BCE), René Descartes (1596—1650), Charles Darwin (1809—1882), Karl Marx (1818—1883), Charles Sanders Peirce (1839—1914), Ludwig Wittgenstein (1889—1951), Michel Foucault (1926—1984), Thomas Kuhn (1922—1996), Hans-Georg Gadamer (1900—2002) and Jürgen Habermas (born 1929) relevant to information science? |
| Meta-level/paradigms: Information sciences’ paradigms and traditions | Information science approaches include facet analysis, user-based views and cognitive views, bibliometric views, systems-oriented views, domain-oriented views, critical approaches etc. (see, for example, Bates 2005; Ellis 1992a+b; Fisher et al. 2005; Fuchs 2011; Hjørland 2013a+b+c+d+e); Sonnenwald (in press). |
| Theory level | The theory level is the level of the specific assumptions, which may guide practitioners’ decisions. For example: “users’ utilization of a library is inversely correlated with distance to the library”; “users’ preferences are based on their individual personalities”; “users’ preferences are formed by market forces”; “the most cited documents are the best documents”; “the most cited documents reflect the dominant ideology” etc. Each theory is related to a metatheory, which is in turn related to the general philosophical level. |
| Application level (practical activities done by information specialists) | Helping users search for documents, information, knowledge, and art. Designing and evaluating search systems, classifications, ontologies, and so on. Cataloguing, classifying, indexing, and annotating documents. Building and managing collections/Cultural Resource Management. … |

The relationship between topic maps and traditional classification schemes might be that topic maps are not so much an extension of the traditional schemes as on a higher level. That is, thesauri extend taxonomies, by adding more built-in relationships and properties. Topic maps do not add to a fixed vocabulary, but provide a more flexible model with an open vocabulary. A consequence of this is that topic maps 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.

Hjørland (2007) understood KOS in a broad and in a narrow sense. In the narrow sense KOS is a synonym for semantic tools, which is understood as selections of concepts and an indication of some of their semantic relations. Different KOS displays or emphasize different semantic relations, which is closely related to the idea of a “Semantische Treppe” (English: “semantic staircase,” sometimes called “semantic spectrum”) as suggested by...
Blumauer and Pellegrini (2006, 16). This staircase represents a ranking of KOS according to semantic richness in which the ontology represents the highest semantic richness of all KOS. Figure 1 displays this staircase from the English translation provided by Olensky (2010).

Figure 1 can briefly be explained in the following way:

- **Glossary.** Although there are many kinds of dictionaries with many kinds of information about words, glossaries and subject dictionaries may define a term by referring to a synonym or “definition by species and genus” (e.g. defining “man” as “thinking animal”). The dominating semantic relation in glossaries is the generic relation (man is a kind of animal), but other kinds of semantic relations may also be applied.

- **“Folksonomy”** means a user generated taxonomy, typically less formal and accurate compared to traditional taxonomies. Folksonomies are examples of uncontrolled vocabularies.

- **Taxonomy** (or classification) is typically a hierarchical system in which generic relations primarily organize terms. For example, a taxonomy may list classes representing all species under the class “animal.”

- **Thesaurus** is a KOS that represents concepts in a domain. It typically relates synonyms to a given concept, distinguishes homonyms, and for each concept specifies its broader concepts, narrower concepts and related concepts. Broader and narrower concepts may be generic or partitive related. “Related concept” is a collected term covering all kinds of semantic relations considered important except hierarchical relations.

- **Topic Map** is a special kind of ontology. Ontologies are KOS in which the kinds of semantic relations are unlimited. They are produced for making logical inferences by computers and puts therefore high demands on the formal specifications. An ontology may, for example, specify which drugs have a specific side effect. There is a fundamental ambiguity in the use of the term “ontology” in information science: On the one hand, it is used (as in Figure 1) as one kind of KOS with some specific requirements (see Soergel et al. 2004). On the other hand are other kinds of KOS also sometimes considered ontologies.

Considering the aforementioned views of KOS, it is surprising that Kless et al. (2015) suggest: Ontologies cannot be simply seen as the “better thesaurus” as it is suggested by simple spectrums of “formality” …. Ontologies and thesauri have to be treated as two orthogonal kinds of models with different characteristics that serve different purposes” (p. 17) and: “an ontology is not a good thesaurus (p. 1).”

Kless et al. also question the widespread view that ontologies are a kind of controlled vocabulary (p. 1). A controlled vocabulary is a restricted and authoritative list of words or terms used for describing, indexing or classifying documents or information. It is controlled because (1) only terms from the list may be used for indexing (for the system applying this controlled vocabulary) (also the revision process of the vocabulary is controlled) (2) the meaning of the terms are controlled (e.g. when two terms should be considered synonyms or not). Therefore, all terms in a controlled vocabulary should have an unambiguous, non-redundant definition. Because ontologies ful-

---

**Figure 1:** The semantic staircase (after Olensky 2010, section 2.3.3).
fil these requirements, they should be considered a kind of controlled vocabularies.

A main argument considering the concepts in Figure 1 as strongly related and as subclasses of the concept of KOS is that their construction and evaluation require the same basic kind of knowledge. For example, the common blackbird (Turdus merula) consist of 16 kinds, which were considered different subspecies of the same species in, for example, Peters (1931-1987) Check-list of Birds of the World. New research based on DNA-analysis have revealed, however, that they cannot be considered one species, but simply are similar adaptations in which the male birds developed a black plumage contrasting the yellow bill as an effective means of demonstrating dominance within their territory (Fjeldså 2013, 141). When this new understanding of blackbirds becomes common, we may expect that it will come to dominate the biological literature (and thus providing “literary warrant”) as well as lay peoples understanding (and thus providing “user warrant”). We should expect that the semantic relations related to these birds become integrated in all kinds of KOS, including those represented in Figure 1. (An important decision in KOS is when to adapt a new system; it has to have certain stability first. Mayr & Bock (1994) argued that the time was still not ripe for a new taxonomy of birds, whereas Fjeldså (2013, 141) is indicating that the time now seems to be ripe for introducing the new classification of birds, although many questions are not yet settled).

On what basis did Kless et al. (2015) argue that thesauri and ontologies are “two orthogonal kinds of models”? Their methodology consists of comparing existing thesauri and thesauri standards to existing ontologies and the OWL description logic semantics. The explicitly wrote: “the comparative performance of thesauri and ontologies in specific application contexts are not the subject of this article” (p.4). They are well aware of the problems of comparing existing thesauri and ontologies: “Our comparison had to overcome the problem that the quality of real-life thesauri and ontologies varies considerably, but also that there are different understandings of what ontologies are” (p.3).

Kless et al. described differences and similarities between typical thesauri and ontologies in theory and practice and listed 12 differences:

Difference 1: Concepts in thesauri are a bundle of different ontological entities
Difference 2: For the labeling of classes in ontologies, versus the use of terms in thesauri, the precision of concept descriptions is valued over literary or user warrant.
Difference 3: Unlike thesauri, ontologies allow specifying the meaning of relata through membership conditions.

Difference 4: Meaning-defining ontologies convey the meaning of relata more precisely than do thesauri.
Difference 5: The hierarchical relationship and the associative relationship in a thesaurus are a bundle of ontologically different relationships.
Difference 6: In contrast to the hierarchical (is-a) relationship in ontologies, the hierarchical relationship in thesauri is not necessarily transitive.
Difference 7: The is-a relationship in a meaning-defining ontology tends to be more consistent than does the generic relationship in a thesaurus.
Difference 8: The structure of an ontology can be checked for consistency more comprehensively than can the structure of a thesaurus, and, in contrast to a thesaurus, the consequences of the modeling decisions in ontologies can be seen by inferring class subsumptions automatically using a reasoner.
Difference 9: Top-level ontologies provide a more rigorous basis for building redundancy-free hierarchies in ontologies than do the “intuitive” application and interpretation of categories permitted by the thesaurus standard ISO 25964-1:2011.
Difference 10: There may be little, if any, overlap between use of the whole-part and associative relationships in thesauri and use of their semantically equivalent counterparts in ontologies.
Difference 11: Apart from the is-a and the instance-of relationships, relationships in ontologies express membership conditions while thesaurus relationships serve mainly navigational and information retrieval purposes.
Difference 12: Relationships in ontologies are always directed and—apart from the is-a and the instance-of relationships—do not imply that their inverse is true while relationships in thesauri are always bidirectional (i.e., reciprocal).

While these differences seem to reflect typical thesauri and ontologies and the standards used for their construction, it does not follow that thesauri would not improve, if these characteristics from ontologies were adapted. The question is why thesauri are limited to the relatively few kinds of semantic relations (and therefore tend to bundle different relationships)? As far as I know, there has never been put forward arguments or research demonstrating the functionality of such a bundling. The set of relations used in thesauri have to my knowledge never been theoretically motivated! (They may be intuitively motivated by the need of searchers in online databases to increase “recall” and “precision” but this function has never been properly examined and for me it seems unlikely that a broader set of specified semantic relations should not provide better results).
If we consider the differences #1, #4, #5, #7, #8 and #9 it seems unlikely that thesauri would not be improved by taking over these characteristics from ontologies. Difference #2 seems to be the one with the most debatable status. How should priorities be made between logical criteria, empirical criteria, user-based criteria, pragmatic criteria etc.? Current thesauri and their standards may not be based on the most fruitful criteria—as argued by Hjørland (2015). Thesauri as well as their theoretical basis are open for improvement, therefore the claim that “an ontology is not a good thesaurus” has not been demonstrated. If such a demonstration should have been provided, it should have been based on the methods, which was explicitly excluded Kless et al.: “the comparative performance of thesauri and ontologies in specific application contexts are not the subject of this article”. I therefore tend to support ISKO UK AGM (2015) proposition “that the traditional thesaurus has no place in modern information retrieval” and suggest that we in KO focus less on specific kinds of KOS (such as thesauri) and more on general principles of KOS.

Kless et al. (2015, p. 17) are right in their criticism that KOS cannot be understood just as “simple spectrums of ‘formality’” because they vary in many different ways. All of them may, however, be understood as (1) a selection of concepts representing a domain and (2) a specification of some semantic relations between those concepts. The criteria for (1) selecting the concepts (2) selecting the semantic relations and (3) determining the semantic relations between the concepts is different across the whole spectrum of KOS. A dictionary, for example, may claim to be based on cognitive theory as WordNet (Miller 1998, p. 43) or on a historiist theory (as claimed about the Oxford English Dictionary by Miller 1998, p. 43) or it may be based on critical theory (as discussed by Benson 2001).

There are many other approaches for developing and evaluating KOS. Regarding classification schemes one approach is facet-analysis, which is based on establishing logical categories and on logical subdivisions (presented and discussed by Hjørland 2013b). We shall not present further approaches for developing KOS in this article but just state that any KOS is always influenced by some kind of theoretical assumptions and that such assumptions often are forming traditions or “paradigms” in KO.

The basic claim in the present article is that any theory has implication for how its domain is organized and vice versa: Any specific KOS reflects in some way the theoretical understanding of its author. For example, when Carl Linnaeus (1707-1778) created his famous biological taxonomy, Charles Darwin’s theory of evolution was not yet known. A classification of living species based on a God’s given number of species cannot form the basis of a classification that satisfies the demands for a classification reflecting the biological evolution (see Ereshefsky, 2000 for a criticism of the Linnaean hierarchy). Another example is the periodic system of chemistry and physics. According to the Stowe’s “Physicist’s Periodic Table,” (Channon 2011), the basic classification parameters are the three quantum numbers: n (shell), s (spin) and m (orientation). Stowe’s system is therefore an indication that the theory of quantum mechanics has had implications for classification of the chemical elements (about the periodic table and its relation to theory see also Shapere 1977, 534-42). The same principle is also valid for the humanities; consider, for example, Ørom’s (2003) classification of arts based on different theories of art. (See also Doyt and Glick 1994).

The relation between scientific kinds and folk categories are discussed by Khalidi (2013, 55-65), who finds that “folk categories can be expected either to coincide with or to be superseded by scientific categories when the purpose for which they are introduced are roughly the same. When they are not, we should not expect them either to coincide or be superseded, but perhaps to coexist alongside scientific categories (p. 59). Thus, folk categories may deserve a place in knowledge organizing systems if they serve other purposes than scientific ones. On the other hand, Khalidi concludes that “not all purposes are created equal” (p. 62) and “I [Khalidi] privilege epistemic purposes over other purposes and I therefore accord a special status to those classifications that are introduced primarily to serve those purposes” (p. 63). Khalidi exemplify with aquarium fish, which in his view is a category that reflect human aesthetic preferences and does not mark a division between two kinds of fish, nor was it intended to do so. We shall not go deeper into this debate her, but just establish that folk categories should also be considered “theories”, and that these categories may or may not serve the same goals as scientific categories.

5.0 Relations Between Theories and Concepts

It is generally recognized in knowledge organization that concepts are the building blocks of KOS (e.g., Dahlberg 2006; Hjørland 2007; Smiraglia 2014). Although a few researchers have explicitly denied this thesis (e.g., Smith, Ceusters and Temmerman 2005), their criticism seems not well founded, and here our point of departure is that the building blocks of KOS are concepts. Concepts are also the building blocks of theories, as stated by Shoemaker et al. (2004, 15):

Concepts are the building blocks of theories—the things being studied, compared, and related to one another. A concept is an abstraction that describes a portion of reality. It is a general name for specific
instances of the phenomenon described. For example, the concept education (a generalization) describes the aggregate of people’s specific learning experiences. The concept mass media use (a generalization) describes the aggregate of individuals’ specific reading, viewing, and listening behaviors with the mass media.

Knowledge and theories are, however, not simply established from sets of pre-existing concepts. The concepts themselves are co-constructed with theories (and the very notion “concept” itself is theory-laden, see Hjørland 2009). This relation between theories and concepts is perhaps one of the most important implications of the Kuhnian revolution in the philosophy of science (Oberheim and Hoyningen-Huene 2013, 6-7). (Concerning the definitions and theories of “mass” see also Jammer 1961, 2000):

For example, the meanings of the terms “temperature,” “mass,” “chemical element” and “chemical compound” depend on which theories are used to interpret them. Conceptual changes also result in the exclusion of some old elements of the extension of a concept, while new elements come to be subsumed by it, so that the same term comes to refer to different things. For example, the term “Planet” referred to the sun but not the earth in the Ptolemaic theory, whereas it refers to the earth and not the sun in the Copernican theory. Incommensurable theories use some of the same terms, but with different meanings, to refer to different sets of things.

Weiskopf (2011) described these two relations between theories and concepts as two varieties of the theory-theory that differ on the nature of the relationship between concepts and theories: (1) the concepts in theories view, in which concepts are the constituents of theories. This view is scarcely controversial (2) the concepts as theories view, in which concepts themselves are understood as miniature theories of a particular domain:

On the concepts as theories view .... Concepts themselves are identified with miniature theories of a particular domain. For instance, Keil (1989, 281) proposes that “Most concepts are partial theories themselves in that they embody explanations of the relations between their constituents, of their origins, and of their relations to other clusters of features.” So the concept electron would itself be made up of various theoretical postulates concerning electrons, their relationship to other particles, their causal propensities which explain phenomena in various domains of physics, and so on. Concepts are not terms in theories, they are themselves theories.

Instead of considering “the concepts in theories view” and “the concepts as theories view” as two varieties of the theory-theory, they may be considered parts of the view that concepts and theories are co-constructed in a—more or less—iterative process. Alan F. Chalmers (1999, 105) wrote: “Newton could not define mass or force in terms of previously available concepts. It was necessary for him to transcend the limits of the old conceptual framework by developing a new one.” And (106): “the typical history of a concept … involves the initial emergence of the concept as a vague idea, followed by its gradual clarification as the theory in which it plays a part takes a more precise and concrete form.”

Andersen, Barker and Chen describes how concepts are learned (Andersen, Barker and Chen 1996, 349-50; see also Andersen, Barker and Chen 2006, 19-30):

One learns a concept by being guided through a series of encounters with objects that highlight the relations of similarity and dissimilarity currently accepted by a particular community of concept users. Teaching and learning depend upon examining similar or dissimilar features of a range of objects .... Kuhn’s standard example is the child learning to distinguish ducks, geese, and swans. In the learning process the child is shown various instances of all three categories, being told for each instance whether it is a duck, a goose or a swan. Also, the child is encouraged to try to point out instances of the categories. At the beginning of this process the child will make mistakes, for example mistaking a goose for a swan. In such cases the child will be told the correct category for the instance pointed out, perhaps by drawing attention to some feature that distinguishes this bird from swans. In other cases the child ascribes the instance pointed out to the correct category, and is told so. After a number of these encounters the child has acquired the ability to identify ducks, geese, and swans to the satisfaction of the instructor.

This understanding of concepts can be stated in other words: We learn a concept (such as a swan) by growing up in a community in which this concept is understood in a certain way. That way of understanding the concept is not universal and may change over time (cf. the above-mentioned change in the conception of blackbirds). The concept (e.g. swan or blackbird) is in other words dependent on a biological taxonomic theory. The former
concept of blackbirds reflected an understanding of systematics in which similarity played a bigger role (i.e., the phyletic view) whereas the new view of blackbirds should be replaced by different concepts reflects a classification based on the phylogenetic principle (i.e., cladism). In this way even a concept like a “swan” or “blackbird” theory-dependent (although different theories often have large overlaps in their classifications).

Andersen, Barker and Chen (2006, 42-64) also presented a system for representing concepts by means of dynamic frames, which may be worth further examination in the KO-community. The concept “bird,” for example, may be defined by among other attributes the values of the attributes “beak,” “neck,” “color,” “size,” “foot” and “gait.” This system is also used to relate concepts in different scientific theories or paradigms. They write (45):

The recursive nature of frames reflects the seeming paradox that the frame, as a whole, represents a concept, but its elements, or nodes, are themselves concepts. This is not an atomistic form of analysis; there may be no ground floor or ultimate conceptual repertoire at which the chain of frames terminates. Similarly, there may be no single, unique way of drawing a frame for any given concept .... Philosophers who expect the universe to divide into a single unique set of natural kinds may be displeased with this.

Different persons might learn to identify given concepts (e.g. swans) by using different characteristics. It is not the case—as believed in traditional concept theory—that there is one set of necessary and jointly sufficient characteristics which the competent language user knows. Andersen (2002, 99) points out that: “family resemblance concepts form hierarchical structures in which a general concept decomposes into more specific concepts that may again decompose into yet more specific concepts, and so forth—in other words taxonomies.”

Andersen thus suggests that classification may be explained systematically from a family resemblance point of view and furthermore (99) argues that the family resemblance account allows for taxonomies being dynamic entities, which may undergo change. Further, theories or models “provide the causal and explanatory links that hold individual concepts together and establish taxonomic relations to other concepts” (Andersen 2002, 102).

A problem with theory-theory can be to identify the theories in which a given concept forms part (Weiskopf 2011): “For the Theory-Theory, the problem seems to be that there are too few theories. We have concepts such as car, computer, gin, lemur, and nightstick. Perhaps for some of these we have theories, at least of a highly sketchy nature. But it is less clear that we have these for other concepts.”

We presented the concept “aquarium fish” above and saw that it might be connected to a theory of human aesthetic preferences. It might also, however, be connected to more practical issues of which kinds of fish are practical to keep in aquariums and which kinds of cultural communications have favored some species of fish at the expense of other species. In other words, we need to understand the development of this hobby as a domain. Theory-theory needs to be connected with activity-theory, which study cultural-historical activity systems.

We can observe that there are different theoretical positions (including Kuhn's theory, activity theory, pragmatism, semiotics and hermeneutics), that all consider our concepts as theory-laden. They may therefore be parts of a fruitful theoretical framework for KO.

6.0 Conclusion

A theory implies a set of concepts and their relations (i.e., “ontological commitment”). A theory has implications for what concepts, observations etc. are relevant and for what should be considered less relevant. For example, there was no detailed, generally accepted theory of light before Newton proposed his particle theory. “The rival theorists of the pre-science period disagreed not only over fundamental theoretical assumptions but also over the kinds of observational phenomena that were relevant” (Chalmers 1999, p. 111). Another example is a theory of art implies a set of concepts and their relations, and how knowledge about art should be classified (cf., Ørom 2003). A theory is a KOS (it if is explicated) or has implications for construing KOS. (The main difference between theories and traditional KOS is that causal relations tend to dominate in theories while generic relations tend to dominate in classifications and thesauri). Vice versa: Given a KOS it is possible—at least by principle—to say what kinds of theoretical assumptions that have governed its construction. Of course, the influence it is seldom one consistent theory, but more like a “bricolage” (cf., Ørom 2003).

Information science is very much about constructing, using and evaluating knowledge organizing systems (KOSs). The elements in KOSs are concepts, but concepts are theory-laden. It is sometimes difficult to identify the theories, which are involved in specifying the meaning of a given term. Such an identification involves serious scholarly work but is necessary if we want to clarify which different interests are at play and if we want to contribute to make information services serve consciously chosen goals. For any given domain, we have to identify the major theories and interests at play and to uncover the corresponding concepts and their semantic
relations in order to construe, use and evaluate information systems.

References


Blumauer, Andreas and Tassilo Pellegrini. 2006. “Seman
tic Web und Semantische Technologien.” In *Semant

cfacts as Psychological Theories: The Case of Human
Computer Interaction.” *Behaviour and Information Technology* 8, no. 4: 247-56.


Channon, Martin. 2011. “The Stowe Table as the Defini
tive Periodic System.” *Knowledge Organization* 38, no. 4: 321-27.


Dillon, Andrew. 1995. “Artifacts as Theories: Conver


Elliot, Jane and Derek Attridge. 2011. *Theory after ‘The-


digms in Information Retrieval Research.” *Journal of Documentation* 48, no. 1: 45-64.


ton-Century-Crofts.


Friedman, Michael. 2003. “Kuhn and Logical Empiri
cism.” In *Thomas Kuhn*, edited by Thomas Nickles, 19-
44. Cambridge: Cambridge University Press.


ISKO UK AGM. 2015. The Great Debate: This House believes that the traditional thesaurus has no place in modern information retrieval. 19 February 2015 London. ISKO, UK-Chapter. Annual General Meeting. http://www.iskouk.org/content/great-debate#EventProgramme


B. Hjørland. Theories are Knowledge Organizing Systems (KOS)


