Concepts and Semantic Relations in Information Science

Wolfgang G. Stock
Heinrich-Heine-University Düsseldorf, Department of Information Science, Universitätsstr. 1, D-40225 Düsseldorf, Germany. E-mail: stock@phil-fak.uni-duesseldorf.de

Concept-based information retrieval and knowledge representation are in need of a theory of concepts and semantic relations. Guidelines for the construction and maintenance of knowledge organization systems (KOS) (such as ANSI/NISO Z39.19-2005 in the U.S.A. or DIN 2331:1980 in Germany) do not consider results of concept theory and theory of relations to the full extent. They are not able to unify the currently different worlds of traditional controlled vocabularies, of the social web (tagging and folksonomies) and of the semantic web (ontologies). Concept definitions as well as semantic relations are based on epistemological theories (empiricism, rationalism, hermeneutics, pragmatism, and critical theory). A concept is determined via its intension and extension as well as by definition. We will meet the problem of vagueness by introducing prototypes. Some important definitions are concept explanations (after Aristotle) and the definition of family resemblances (in the sense of Wittgenstein). We will model concepts as frames (according to Barsalou). The most important paradigmatic relation in KOS is hierarchy, which must be arranged into different classes: Hyponymy consists of taxonomy and simple hyponymy, meronymy consists of many different part-whole-relations. For practical application purposes, the transitivity of the given relation is very important. Unspecific associative relations are of little help to our focused applications and should be replaced by generalizable and domain-specific relations. We will discuss the reflexivity, symmetry, and transitivity of paradigmatic relations as well as the appearance of specific semantic relations in the different kinds of KOS (folksonomies, nomenclatures, classification systems, thesauri, and ontologies). Finally, we will pick out KOS as a central theme of the Semantic Web.

Introduction

A knowledge organization system (KOS) is made up of concepts and semantic relations that represent a knowledge domain terminologically. In knowledge representation, we distinguish between five approaches to KOSs: nomenclatures, classification systems, thesauri, ontologies and, as a borderline case of knowledge organization systems, folksonomies (Stock & Stock, 2008). Knowledge domains are thematic areas that can be delimited, such as a scientific discipline, an economic sector, or a company's language. A knowledge organization system's goal in information practice is to support the retrieval process. We aim, for instance, to offer the user concepts for searching and browsing, to index automatically, and to expand queries automatically: We aim to solve the vocabulary problem (Furnas, Landauer, Gomez, & Dumais, 1987). Without KOS, a user will select a word A for his search, while the author of a document D uses A' to describe the same object; hence, D is not retrieved. Concept-based information retrieval goes beyond the word level and works with concepts instead. In the example, A and A' are linked to the concept C, leading to a successful search. Susan Feldman (2000) expressed the significance of this approach very vividly from the users' perspective: "Find what I mean, not what I say."

This article will deal with problems of KOSs in general as well as of all kinds of knowledge domains. How do we have to represent concepts and relations in order to make the goals of perfect concept-based information retrieval achievable? What is new in this article? We will expand the status quo of the understanding of concepts in information science with concept explanations (after Aristotle) and family resemblance (after Wittgenstein), introduce syncategoremata (after Menne), thematize vagueness (with Black), as well as prototypes (with Rosch) and model concepts as frames (according to Barsalou). In terms of the relations (which are structural invariants of attributes in frames), it is important to pay attention to transitivity in all forms of hyponymy and meronymy. Additionally, it seems necessary to forego any unspecific association relations as far as possible, in order to work with the respective specific concept relations instead.

We are concerned with a new view of concepts which will touch on known and established theories and models but also be suitable for exploiting all advantages of KOSs, and in particular of folksonomies and ontologies, for information science and practice. If we are to create something like the 'Semantic Web,' we must perforce think about the concept of the concept, as therein lies the key to any semantics (Hjørland,

Received March 8, 2010; revised May 3, 2010; accepted May 5, 2010

 • Published online 21 June 2010 in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/asi.21382

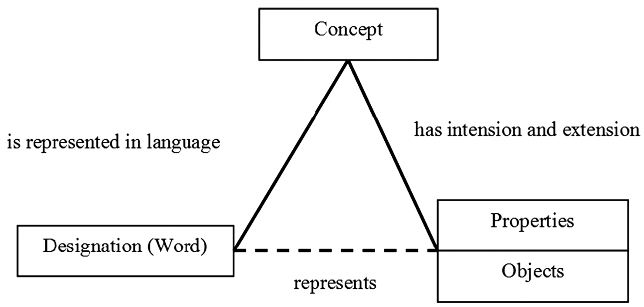

FIG. 1. The semiotic triangle in information science.

2007). Excursions into general linguistics, philosophy, cognitive science, and computer science are also useful for this purpose.

Concepts

The Semiotic Triangle

In language, we use symbols, e.g., words, in order to express a thought about an object of reference. We are confronted with the tripartite relation consisting, following Charles K. Ogden and Ivor A. Richards (1985 [1923]), of the thought (also called reference), the referent, or object of reference, and the symbol. Ogden/Richards regard the thought/reference as a psychological activity ("psychologism"; according to Schmidt, 1969, p. 30). In information science (Figure 1), 'concept' takes the place of (psychological) thought. As in Ogden and Richards' classical approach, a concept is represented in language by words. Such designations can be natural-language words, but also terms from artificial languages (e.g., notations of a classificatory KOS). The concept 'concept' is defined as a class containing certain objects as elements, where the objects have certain properties.

We will discuss the German guidelines for the construction of KOS, which are very similar to their counterparts in the United States, namely ANSI/NISO Z39.19-2005. The norms DIN 2330 (1993, p. 2) and DIN 2342/1 (1993, p. 1) understand a concept as "a unit of thought which is abstracted from a multitude of objects via analysis of the properties common to these objects." This DIN definition is not unproblematic. Initially, it would be wise to speak, instead of (the somewhat psychological-sounding) "unit of thought," of "classes" or "sets" (in the sense of set theory). Furthermore, it does not hold for each concept that all of its elements always and necessarily have "common" properties. This is not the case, for example, with concepts formed through family resemblance. Lacking common properties, we might define *vegetable* as "is cabbage vegetable or root vegetable or fruit vegetable, etc." But what does 'family resemblance' mean? Instead of *vegetable*, let us look at a concept used by Ludwig Wittgenstein (2008 [1953]) as an example of this problem, *game*. Some games may have in common that there are winners and losers, other games—not all, mind you—are entertaining, others again require skill and luck of their players, etc. Thus, the concept of the game cannot be defined via exactly one set of properties. We must admit that a concept can be determined not only through a conjunction of properties, but also, from time to time, through a disjunction of properties.

There are two approaches to forming a concept. The first goes via the objects and determines the concept's extension, the second notes the class-forming properties and thus determines its intension (Reimer, 1991, p. 17). Gottlob Frege uses the term 'Bedeutung' (meaning) for the extension, and 'Sinn' (sense) for the intension. Independently of what you call it, the central point is Frege's discovery that extension and intension need not necessarily concur. His example is the term *Venus*, which may alternatively be called *Morning Star* or *Evening Star*. Frege (1892, p. 27) states that Evening Star and Morning Star are extensionally identical, as the set of elements contained within them (both refer to Venus) are identical, but are intensionally nonidentical, as the Evening Star has the property "first star visible in the evening sky" and the Morning Star has the completely different property "last star visible in the morning sky."

The extension of a concept M is the set of objects O_1, O_2, etc., that fall under it:

$$M = df\{O_1, O_2, \ldots, O_i, \ldots\},$$

where "=df" means "equals by definition." It is logically possible to group like objects together (via classification) or to link unlike objects together (via colligation, e.g., the concept Renaissance as it is used in history) (Shaw, 2009, 2010; Hjørland, 2010a).

The intension determines the concept M via its properties f_1, f_2, etc., where most of these properties are linked via "and" (&) and a subset of properties is linked via "or" (v) (where ∀ is the universal quantifier in the sense of "holds for all"):

$$M = df\forall x\{f_1(x) \,\&\, f_2(x) \,\&\, \ldots \,\&\, [f_g(x) v f_{g'}(x) v \ldots v f_{g''}(x)]\}.$$

This definition is broad enough to include all kinds of forming concepts such as concept explanations (founding on ANDing properties) and family resemblance (founding on ORing properties). It is possible (for the "vegetable-like" properties) that the subset of properties f_1, f_2, etc., but not f_g, is a null set, and it is possible that the subset f_g, $f_{g'}$, etc., but not f_1, f_2, etc., is a null set (in the case of concept explanation). The purpose of concept formation is conceptual coherence (Spiteri, 2008), at least for a knowledge domain and for a certain time period.

Concept Theory and Epistemology

How do we arrive at concepts, anyway? This question calls for an excursion into epistemology. Birger Hjørland (2009; see also Szostak, 2010; Hjørland, 2010b) distinguishes between four different approaches to this problem: empiricism, rationalism, hermeneutics, and pragmatism. We add a fifth approach, critical theory, which can, however—as Hjørland suggests—be understood as an aspect of pragmatism (see Figure 2). Any concept formation and any semantic relation may have some foundation. Is it determined

DOI: 10.1002/asi

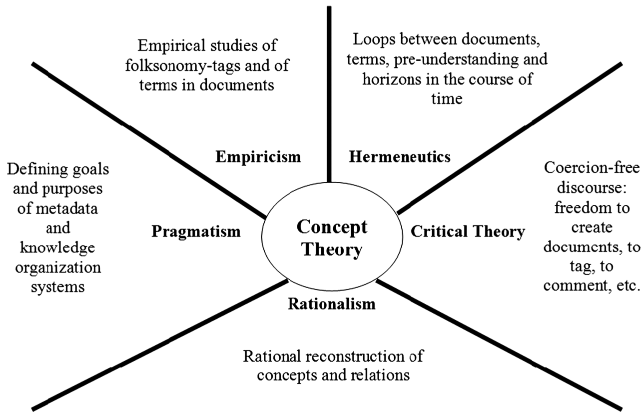

FIG. 2. Epistemological foundations of concept theory.

by empirical research (empiricism)? Is it based on logical deductions (rationalism)? Is it grounded on the historical development (hermeneutics)? Are there any purposes (pragmatics)? Is there a coercion-free discourse on the concept and on the relation (critical theory)?

Empiricism starts from observations; thus one looks for concepts in concrete, available texts which are to be analyzed. Some typical methods of information science in this context are similarity calculations between text words, but also between tags when using folksonomies and the cluster-analytical consolidation of the similarity relations (Knautz, Soubusta, & Stock, 2010). Rationalism is skeptical toward the reliability of observations and constructs a priori concepts and their properties and relations, generally by drawing from analytical and formal logic methods. In information science we can observe such an approach in formal concept analysis (Ganter & Wille, 1999; Priss, 2006). Hermeneutics (called "historicism" in Hjørland, 2009, p. 1525) captures concepts in their historical development as well as in their use in a given "world horizon." In understanding this, human's "being thrown" into the world plays an important role (Heidegger, 1962 [1927]). A text is never read without any understanding or prejudice. This is where the hermeneutic circle begins: The text as a whole provides the key for understanding its parts, while the person interpreting it needs the parts to understand the whole (Gadamer, 1975 [1960]). Prejudices play a positive role here. We move dynamically in and with the horizon, until finally the horizons blend in the understanding (Stock & Stock, 2008, p. 93). In information science, the hermeneutical approach leads to the realization that concept systems and even bibliographical records (in their content-descriptive index fields) are always dynamic and subject to change (Gust von Loh, Stock, & Stock, 2009). Pragmatism is closely associated with hermeneutics, but stresses the meaning of means and goals. Thus, for concepts one must always note what they are being used for: "The ideal of pragmatism is to define concepts by deciding which class of things best serves a given purpose and then to fixate this class in a sign" (Hjørland, 2009, p. 1527).

Critical Theory (Habermas, 1987 [1968]) stresses coercion-free discourse, the subject of which is the individual's freedom to use both words and concepts (as for example during tagging) at his or her discretion, and not under any coercion (Fuchs, 2008). For Jürgen Habermas (1998, p. 306), in an ideal social world the only force is discourse, or "the unforced force of the better argument." Concerning tagging (or concept formation), we would like to formulate, "the unforced force of the better tag (the better concept)." The media of an inclusive civil society "must empower citizens to participate in and respond to a public discourse that, in turn, must not degenerate into a colonizing mode of communication" (Habermas, 2006, p. 420). Habermas' view of "the Internet" is very pessimistic; for him weblogs and chat rooms play a "parasitical role of online communication" (Habermas, 2006, p. 423), because "Internet-based discourse communities have fragmented the public" (Geiger, 2009, p. 2). Habermas did not realize that there are two kinds of power relations in Web 2.0 communities (Jarrett, 2008). There are Web 2.0 services (and their discourse communities) where the participants actively cooperate, e.g., blogging and backtracking, to twitter and to retweet, or to work together on Wikipedia articles. This kind of relation is called *collaborative*. Here, following Surowiecki (2005, p. XIX) and Peters (2009, p. 169), it is possible to find "wisdom of crowds," but it is also possible to find "madness of crowds" and maybe parasitical activities as well. Surowiecki (2005, p. 10) defines four principles that are necessary for the successful working of wisdom of crowds: "Diversity of opinion (each person should have some private information …), independence (people's opinions are not determined by the opinions of those around them), decentralization (people are able to specialize and draw on local knowledge), and aggregation (some mechanism exists for turning private judgments into a collective decision." This second kind of power relation is called *collective* (Schmidt & Stock, 2009, p. 873); it is able to realize "collective intelligence." Isabella Peters (2009, p. 170) concludes, "for folksonomies in knowledge representation, this means that Collective Intelligence can develop best if users tag the same resource multiple times independently of each other, and thus implicitly compile a collective vocabulary that reflects their collective opinion on the resources." In this case, Habermas' pessimistic argument does not work. There is no fragmentation of the public and no parasitical use of Web 2.0 services. There is no human communication at all, but statistical aggregation done by machines. In the case of collective tagging, the "public sphere may very well operate at a micro-level via Habermasian ideal discourse, but it is at a macro-level constructed not through communication, but algorithms" (Geiger, 2009, p. 25).

Each of the five epistemological theories is relevant for the construction of concepts and relations in information science research as well as in information practice, and should always be accorded due attention in compiling and maintaining knowledge organization systems.

Concept Types

Concepts are the smallest semantic units in knowledge organization systems; they are "building blocks" (Hjørland,

DOI: 10.1002/asi

2010c) or “units of knowledge” (Dahlberg, 1986, p. 10). A KOS is a concept system in a given knowledge domain. In knowledge representation, a concept is determined via words that carry the same, or at least a similar, meaning (this being the reason for the designation “Synset,” which stands for “set of synonyms,” periodically found for concepts) (Fellbaum, 1998). In a first approach, and in unison with DIN 2342/1 (1992, p. 3), synonymy is “the relation between designations that stand for the same concept.” There is a further variant of synonymy, which expresses the relation between two concepts, and which we will address below.

Some examples for synonyms are *autumn* and *fall* or *dead* and *deceased*. A special case of synonymy is found in paraphrases, where an object is being described in a roundabout way. Sometimes it is necessary to work with paraphrases, if there is no name for the concept in question. In German, for example, there is a word for “no longer being hungry” (*satt*), but none for “no longer being thirsty” (Bertram, 2005, p. 41). This is an example for a concept without a concrete designation.

Homonymy starts from designations; it is “the relation between matching designations for different concepts” (DIN 2342/1, 1992, p. 3). An example for a homonym is *Java*. This word stands, among others, for the concepts Java (island), Java (coffee), and Java (programming language). In word-oriented retrieval systems, homonyms lead to big problems, as each homonymous—and thus polysemous—word form must be disambiguated, either automatically or in a dialog between man and machine. Varieties of homonymy are homophony, where the ambiguity lies in the way the words sound (e.g., *see* and *sea*), and homography, where the spelling is the same but the meanings are different (e.g., *lead* the verb and *lead* the metal). Homophones play an important role in information systems that work with spoken language, homographs must be noted in systems for the processing of written texts.

Many concepts have a meaning which can be understood completely without reference to other concepts, e.g., *chair*. Albert Menne (1980, p. 48) calls such complete concepts “categorematical.” In knowledge organization systems that are structured hierarchically, it is very possible that such a concept may occur on a certain hierarchical level:

. . . with filter.

This concept is syncategorematical; it is incomplete and requires other concepts in order to carry meaning (Menne, 1980, p. 46). In hierarchical KOS, the syncategoremata are explained via their broader concepts. Only now does the meaning become clear:

Cigarette
. . . with filter

or

Chimney
. . . with filter.

One of the examples concerns a filter cigarette, the other a chimney with a (soot) filter. Such an explication may take the incorporation of several hierarchy levels. As such, it is highly impractical to enter syncategoremata on their own and without any addendums in a register, for example.

Concepts are not given, like physical objects, but are actively derived from the world of objects via abstraction (Klaus, 1973, p. 214). The aspects of concept formation (in the sense of information science, not of psychology) are first and foremost clarified via definitions. In general, it can be noted that concept formation in the context of knowledge organization systems takes place in the area of tension between two contrary principles. An economical principle instructs us not to admit too many concepts into a KOS. If two concepts are more or less similar in terms of extension and intension, these will be regarded as one single “quasi-synonymous” concept. The principle of information content leads in the opposite direction. The more precise we are in distinguishing between intension and extension, the larger each individual concept’s information content will be. The concepts’ homogeneity and exactitude will draw the greatest profit from this. Lloyd K. Komatsu (1992, p. 501) illustrates this problematic situation (he uses “category” for “concept”):

> Thus, economy and informativeness trade off against each other. If categories are very general, there will be relatively few categories (increasing economy), but there will be few characteristics that one can assume different members of a category share (decreasing informativeness) and few occasions on which members of the category can be treated as identical. If categories are very specific, there will be relatively many categories (decreasing economy), but there will be many characteristics that one can assume different members of a category share (increasing informativeness) and many occasions on which members can be treated as identical.

The solution for concept formation (Komatsu, 1992, p. 502, uses “categorization”) in KOS is a compromise:

> The basic level of categorization is the level of abstraction that represents the best compromise between number and informativeness of categories.

According to the theory by Eleanor Rosch (Mervis & Rosch, 1981; Rosch, 1975a,b, 1983; Rosch & Mervis, 1975; Rosch, Mervis, Gray, Johnson, & Boyes-Braem, 1976), we must distinguish between three concept levels: the superordinate level, the basic level, and the subordinate level:

> Suppose that basic objects (e.g., chair, car) are the most inclusive level at which there are attributes common to all or most members of the category. Then total cue validities are maximized at that level of abstraction at which basic objects are categorized. That is, categories one level more abstract will be superordinate categories (e.g., furniture, vehicle) whose members share only a few attributes among each other. Categories below the basic level will be subordinate categories (e.g., kitchen chair, sports car) which are also bundles of predictable attributes and functions, but contain many attributes which overlap with other categories (for example, *kitchen*

DOI: 10.1002/asi

> *chair* shares most of its attributes with other kinds of chairs) (Rosch et al., 1976, p. 385).

Thus, many people agree that on the basic level the concept *chair* is a good compromise between *furniture*, which is too general, and *armchair, Chippendale chair*, etc., which are too specific. In a knowledge organization system for furniture, the compromise looks different, as we must differentiate much more precisely: Here we will add the concepts from the subordinate level. If, on the other hand, we construct a KOS for economic sciences, the compromise might just favor *furniture*; thus, we would restrict ourselves to a superordinate-level concept in this case.

Concepts whose extension is exactly one element are individual concepts, their designations are proper names, e.g., of people, organizations, countries, products, but also of singular historical events (e.g., German Reunification) or individual scientific laws (Second Law of Thermodynamics). All other concepts are general concepts (Dahlberg, 1974, p. 16). We want to emphasize categories as a special form of general concepts. Moving upwards through the levels of abstraction, we will at some point reach the top. At this point—please note: always in the context of a knowledge domain—no further step of abstraction can be taken. These top concepts represent the domain-specific categories. Fugmann (1999, p. 23) introduces categories via the concepts' intension. Here a concept with even fewer properties no longer makes any sense. In faceted KOS (Gnoli, 2008), the categories form the framework for the facets.

According to Fugmann (1999), the concept types can be distinguished intensionally. Categories are concepts with a minimum of properties (to form even more general concepts would mean, for the knowledge domain, the creation of empty, useless concepts). Individual concepts are concepts with a maximum of properties (the extension will stay the same even with the introduction of more properties). General concepts are all the concepts that lie between these two extremes. Their exposed position means that both individual concepts and categories can be manipulated quite easily via methods of knowledge representation, while general concepts can easily lead to problems. Although individual concepts, referring to named entities, are quite easy to define, some kinds of KOS, e.g., classification systems, do not consider them for controlled vocabulary (in classification systems, named entities have no dedicated notations) (Buizza, 2010).

Vagueness and Prototype

Individual concepts and categories can generally be exactly determined. But how about the exactitude of general concepts? We will continue with our example *chair* and follow Max Black (1937, p. 433) into his imaginary chair exhibition:

> One can imagine an exhibition in some unlikely museum of applied logic of a series of "chairs" differing in quality by at least noticeable amounts. At one end of a long line, containing perhaps thousands of exhibits, might be a Chippendale chair; at the other, a small nondescript lump of wood. Any "normal" observer inspecting the series finds extreme difficulty in "drawing the line" between chair and not-chair.

The minimal distinctions between neighboring objects should make it near impossible to draw a line between chair and not-chair. Outside the "neutral area," where we are not sure whether a concept fits or not, we have objects that clearly fall under the concept on the one side, and on the opposite side, objects that clearly do not fall under the concept. However, neither are the borders between the neutral area and its neighbors exactly definable. Such blurred borders can be experimentally demonstrated for many general concepts (Löbner, 2002, p. 45). "The knowledge approach posits that we must be willing to accept a degree of uncertainty and some fuzzy boundaries in the design of concepts, but we can still find enough areas of commonalities to make concepts coherent across a domain," Spiteri (2008, p. 9) reports.

As a solution, we might try not searching for the concept's borders at all and instead work with a "prototype" (Rosch, 1983). Such a prototype can be regarded as "the best example" for a Basic-Level concept. This model example possesses "good" properties in the sense of high recognition value.

If we determine the concept intensionally, via a prototype and its properties, the fuzzy borders are still in existence (and may cause the odd mistake in indexing these border regions), but on the plus side, we are able to work satisfactorily with general concepts in the first place. If we imagine a concept hierarchy stretching over several levels, prototypes should play a vital role, particularly on the intermediate levels, i.e., in the Basic Level after Rosch. At the upper end of the hierarchy are the (superordinate) concepts with few properties, so that with all probability, one will not be able to imagine a prototype. And the at the bottom level, the (subordinate) concepts are so specific that the concept and the prototype will coincide.

No concept (and no KOS) remains stable over time. "Our understandings of concepts change with context, environment, and even personal experience" (Spiteri, 2008, p. 9). In science and technology, those changes are due to new observations and theories or—in the sense of Thomas S. Kuhn (1962)—to scientific revolutions. A good example is the concept *planet of the solar system*. From 1930 to 2006 the extension of this concept consisted of nine elements; now there are only eight (Pluto is no longer accepted as a planet in astronomy) (Boyle, 2010).

Definition

In knowledge representation practice, concepts are often only implicitly defined—e.g., by stating their synonyms and their location in the semantic environment. It is our opinion that in knowledge organization systems the concepts used are to be exactly defined, since this is the only way to achieve clarity for both indexers and users.

Definitions must match several criteria in order to be used correctly (Dubislav, 1981, p. 130; Pawłowski, pp. 31–43).

DOI: 10.1002/asi

Circularity, i.e., the definition of a concept with the help of the same concept, which—as a mediate circle by now—can now be found across several definition steps, is to be avoided. To define an unknown concept via another, equally unknown concept (ignotum per ignotum) is of little help. The inadequacy of definitions shows itself in their being either too narrow (when objects that should fall under the concept are excluded) or too wide (when they include objects that belong elsewhere). In many cases, negative definitions (*a point is that which has no extension*) are unusable, as they are often too wide (Menne, 1980, p. 32). A definition should not display any of the concept's superfluous properties (Menne, 1980, p. 33). Of course, the definition must be precise (and thus not use any figurative truisms, for example) and cannot contain any contradictions (such as *blind viewer*). Persuasive definitions, i.e., concept demarcations aiming for (or with the side effect of) emotional reactions (e.g., to paraphrase Buddha, Pariah is a man who lets himself be seduced by anger and hate, a hypocrite, full of deceit and flaws…; Pawłowski, 1980, p. 250), are unusable in knowledge representation. The most important goal is the definition's usefulness in the respective knowledge domain (Pawłowski, 1980, p. 88). In keeping with our knowledge of vagueness, we strive not to force every single object under one and the same concept, but sometimes define the prototype instead.

From the multitude of different sorts of definition (such as definition as abbreviation, explication, nominal, and real definition), concept explanation and definition via family resemblance are particularly important for knowledge representation.

Concept explanation starts from the idea that concepts are made up of partial concepts:

$$\text{Concept} = \text{df Partial Concept}_1, \text{Partial Concept}_2, \ldots$$

Here one can work in two directions. Concept synthesis starts from the partial concepts, while concept analysis starts from the concept. The classical variant dates from Aristotle and explains a concept by stating genus and differentia. Aristotle works out criteria for differentiating concepts from one another and structuring them in a hierarchy. The recognition of objects' being different is worked out in two steps; initially, via their commonalities—what Aristotle, in his "Metaphysics," calls "genus"—and then the differences defining objects as specific "types" within the genus (Aristotle, 1057b 34 et seq.). Thus, a concept explanation necessarily involves stating the genus and differentiating the types. It is important to always find the nearest genus, without skipping a hierarchy level.

What determines the differentiation of a genus' types? Aristotle distinguishes between two aspects—the arbitrary nature of an object (i.e., that horses have tails while humans do not) on the one hand, and the fundamental, specific properties that make a difference on the other (e.g., that humans are able to reason, while horses are not). In the Middle Ages this thesis of Aristotle's was summed up in the following, easy to remember phrase (Menne, 1980, p. 28): "Definitio fit per genus proximum et differentiam specificam." Hence, concept explanation works with the following partial concepts:

Partial Concept$_1$: Genus (concept from the directly superordinate genus),

Partial Concept$_2$: Differentia specifica (fundamental difference to the sister concepts).

The properties that differentiate a concept from its sister terms (the concepts that belong to the same genus) must always display a specific, and not an arbitrary property (accidens). A classical definition according to this definition type is:

Homo est animal rationale.

Homo is the concept to be defined, *animal* the genus concept, and *rational* the specific property separating man from other creatures. It would be a mistake to define mankind via *living creature* and *hair not blond*, since (notwithstanding jokes about blondes) the color of one's hair is an arbitrary, not a fundamental property.

The fact that over the course of concept explanations, over several levels from the top down, new properties are always being added means that the concepts are becoming ever more specific; in the opposite direction, they are getting more general (as properties are shed on the way up). This also means that on a concept ladder properties are "inherited" by those concepts further down. Concept explanation is of particular importance for KOS, as their specifications necessarily embed the concepts in a hierarchical structure.

In concept explanation, it is assumed that an object wholly contains its specific properties if it belongs to the respective class; the properties are joined together via a logical AND. This does not hold for the vegetable-like concepts, where we can only distinguish a family resemblance between the objects. Here the properties are joined via an OR (Pawłowski, 1980, p. 199). If we link concept explanation with the definition according to family resemblance, we must work with a disjunction of properties on certain hierarchical levels. Here, too, we are looking for a genus concept, e.g., for Wittgenstein's *game*. The family members of *game*, such as board game, card game, game of chance, etc., may very well have a few properties in common, but not all. Concepts are always getting more specific from the top down and more general from the bottom up; however, there are no hereditary properties from the top down. On those hierarchy levels that define via family resemblance, the concepts pass on some of their properties, but not all.

Let us assume, for instance, that the genus of *game* is *leisure activity*. We must now state some properties of games in order to differentiate them from other leisure activities (such as *meditating*). We define:

Partial Concept$_1$/Genus: Leisure Activity

Partial Concept$_2$/Differentia specifica: Game of Chance v Card Game v Board Game v . . .

DOI: 10.1002/asi

If we now move down on the concept ladder, it will become clear that *game* does not pass on all of its properties, but only ever subsets (as a game of chance needs not be a card game). On the lower levels, in turn, there does not need to be family resemblance, but only "normal" (conjunctive) concept explanation. For each level it must be checked whether family resemblance has been used to define disjunctively or "normally." We have to note that not all concept hierarchies allow for the heredity of properties; there is no automatism. This is a very important result for the construction of ontologies.

Frames

How can a concept be represented? One successful approach works with frames (Minsky, 1975). Frames have proven themselves in cognitive science, in computer science (Reimer, 1991, p. 159), and in linguistics. In Lawrence W. Barsalou's (1992, p. 29) conception, frames contain three fundamental components:

- sets of attributes and values (Petersen, 2007),
- structural invariants,
- rule-bound connections.

Among the different frame conceptions, we prefer Barsalou's version, as it takes into consideration rule-bound connections. We can use this option in order to automatically perform calculations on the application side of a concept system (Stock, 2009, pp. 418–419).

The core of each frame allocates properties (e.g., Transportation, Location, Activity) to a concept (e.g., Vacation), and values (say, Kauai or Las Vegas) to the properties, where both properties and values are expressed via concepts. After Minsky (1975), the concept is allocated such attributes that describe a stereotypical situation. There are structural invariants between the concepts within a frame, to be expressed via relations (Barsalou, 1992, pp. 35–36):

> Structural invariants capture a wide variety of relational concepts, including spatial relations (e.g., between *seat* and *back* in the frame for *chair*), temporal relations (e.g., between *eating* and *paying* in the frame for *dining out*), causal relations (e.g., between *fertilization* and *birth* in the frame for *reproduction*), and intentional relations (e.g., between *motive* and *attack* in the frame for *murder*).

The concepts within the frame are not independent but form manifold connections bound by certain rules. In Barsalou's *Vacation*-frame, there are, for example, positive (the faster one drives, the higher the travel cost) and negative connections (the faster one drives, the sooner one will arrive) between the transport attributes. We regard the value for the location *Kauai* on the attribute level, and the value *surfing* on the activity level. It is clear that the first value makes the second one possible (one can surf around Kauai, and not, for example, in Las Vegas). A formulation in a terminological logic (also called description logic; Nardi & Brachman, 2003) and the separation of general concepts (in a TBox) and individual concepts (in the ABox) allow us to introduce the option of automatic reasoning to a concept system, in the sense of ontologies. If some of the values are numbers, these can be used as the basis of automatic calculations.

Barsalou (1992, p. 43) sees (at least in theory) no limits for the use of frames in knowledge representation. Some groundwork, however, must be performed for the automatized system:

> Before a computational system can build the frames described here, it needs a powerful processing environment capable of performing many difficult tasks. This processing environment must notice new aspects of a category to form new attributes. It must detect values of these attributes to form attribute-value sets. It must integrate cooccurring attributes into frames. It must update attribute-value sets with experience. It must detect structural invariants between attributes. It must detect and update constraints. It must build frames recursively for the components of existing frames.

Where the definition as concept explanation at least leads to one relation (the hierarchy), the frame approach leads to a multitude of relations between concepts and, furthermore, to rule-bound connections. As concept systems absolutely require relations, frames—as concept representatives—ideally consolidate such methods of knowledge representation. This last quote of Barsalou's should inspire some thought, though, on how not to allow the mass of relations and rules to become too large. After all, the groundwork and updates mentioned above must be put into practice, which represents a huge effort. Additionally, it is feared that as the number of different relations increases, the extent of the knowledge domain in whose context one can work will grow ever smaller. To wit, there is according to Daniele Nardi and Ronald J. Brachmann (2003, p. 10) a reverse connection between the language's expressiveness and automatic reasoning:

> [T]here is a tradeoff between the expressiveness of a representation language and the difficulty of reasoning over the representation built using that language. In other words, the more expressive the language, the harder the reasoning.

KOS designers should thus keep the number of specific relations as small as possible, without for all that losing sight of the respective knowledge domain's specifics.

Semantic Relations

Syntagmatic and Paradigmatic Relations

Concepts do not exist in independence of each other, but are interlinked. We can make out such relations in the definitions (e.g., via concept explanation) and in the frames. We will call relations between concepts "semantic relations" (Khoo & Na, 2006; Storey, 1993). This is only a part of the relations of interest for knowledge representation. Bibliographical relations (Green, 2001, p. 7) register relations that describe documents formally (e.g., "has author," "appeared in

DOI: 10.1002/asi

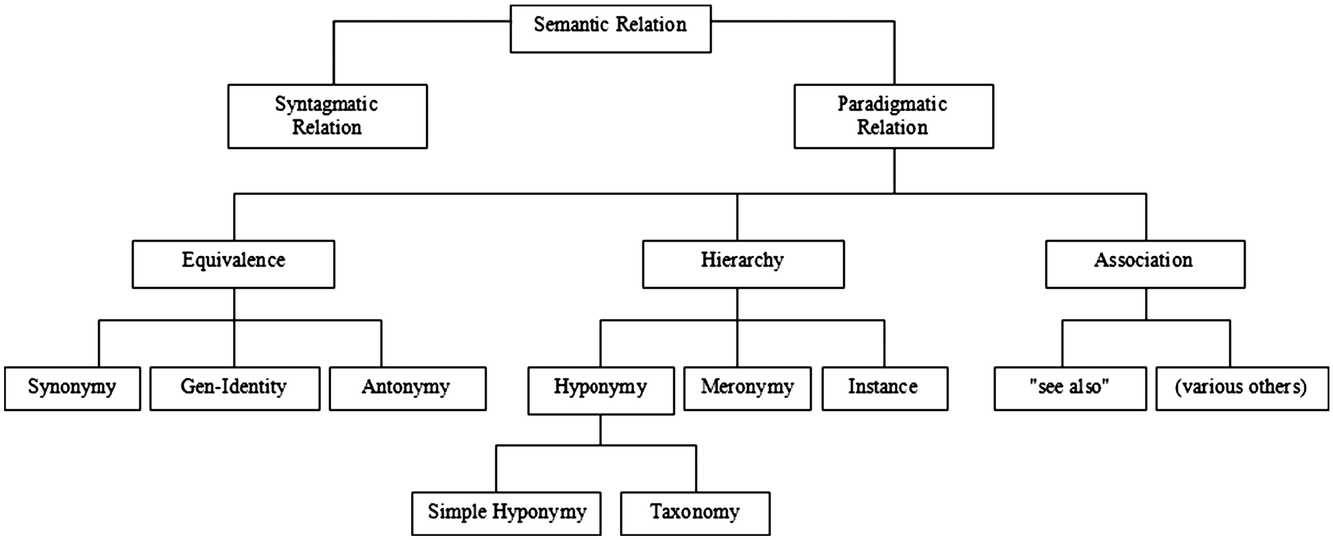

FIG. 3. Semantic relations.

source," "has publishing date"). Relations also exist between documents, e.g., insofar as scientific documents cite and are cited, or Web documents have links. We will concentrate exclusively on semantic relations here.

In information science we distinguish between paradigmatic and syntagmatic relations where semantic relations are concerned (Peters & Weller, 2008a; Stock, 2007, p. 451). This differentiation goes back to Ferdinand de Saussure (2005 [1916]) (de Saussure uses "associative" instead of "paradigmatic"). In the context of knowledge representation the paradigmatic relations form "tight" relations, which have been established (or laid down) in a certain KOS. They are valid independently of documents (i.e., "in absentia" of any concrete occurrence in documents). Syntagmatic relations exist between concepts in specific documents; they are thus always "in praesentia." The issue here is that of co-occurrence, be it in the continuous text of the document (or of a text window), in the selected keywords (Wersig, 1974, p. 253), or in tags when a service applies a folksonomy (Peters, 2009). The syntagmatic relation is the only semantic relation that occurs in folksonomies (Peters, 2009; Peters & Stock, 2007, 2008). In the sense of a bottom-up approach of building KOS, folksonomies provide empirical material for potential controlled vocabularies, as well as material for paradigmatic relations, even though the latter is only "hidden" in the folksonomies (Peters & Weller 2008a, p. 104) and must be intellectually revealed via analysis of tag co-occurrences. Isabella Peters and Katrin Weller regard the (automatic and intellectual) processing of tags and their relations in folksonomies as the task of "tag gardening" with the goal of emergent semantics (Peters & Weller, 2008b).

Paradigmatic relations, however, always express the type of relation. From the multitude of possible paradigmatic relations, knowledge representation tries to work out those that are generalizable, i.e., those that can be used meaningfully in all or many use cases. Figure 3 provides us with an overview of semantic relations.

Order and S – R – T

Relations can be differentiated via the amount of their argument fields. Two-sided relations connect two concepts, three-sided relations three, etc. It is always possible to simplify the multisided relations via a series of two-sided relations. *To heal*, for instance, is a three-sided relation between a person, a disease, and a medication. This would result in three two-sided relations: person – disease, disease – medication, and medication – person. We will assume, in this article, that the relations in question are two-sided.

The goal is to create a concept system in a certain knowledge domain which will then serve as a knowledge organization system. KOS can be characterized via three fundamental properties (where x, y, z are concepts and ρ a relation in each case). Reflexivity in concept systems asks how a concept adheres to itself with regard to a relation. Symmetry occurs when a relation between A and B also exists in the opposite direction, between B and A. If a relation exists between two concepts A and B, and also between B and C, and then again between A and C, we speak of transitivity. We will demonstrate this on several examples:

R	Reflexivity	$x \rho x$
		"…is identical to…"
	Irreflexivity	$-(x \rho x)$
		"…is the cause of…"
S	Symmetry	$(x \rho y) \rightarrow (y \rho x)$
		"…is equal to…"
	Asymmetry	$(x \rho y) \rightarrow -(y \rho x)$
		"…is unhappily in love with…"
T	Transitivity	$[(x \rho y) \& (y \rho z)] \rightarrow (x \rho z)$
		"…is greater than…"
	Intransivity	$[(x \rho y) \& (y \rho z)] \rightarrow -(x \rho z)$
		"…is similar to…"

An order in a strictly mathematical sense is irreflexive (−R), asymmetrical (−S), and transitive (T) (Menne, 1980,

DOI: 10.1002/asi

p. 92). An order that has as its only relation *is more expensive than*, for example, has these properties: A certain property, say a lemon, is not more expensive than a lemon (i.e., −R); if a product (our lemon) is more expensive than another product (an apple), then the apple is not more expensive than the lemon but cheaper (−S); if, finally, a lemon is more expensive than an apple and an apple more expensive than a cherry, then a lemon, too, is more expensive than a cherry (T).

For asymmetrical relations, we speak of an inverse relation if it addresses the reversal of the initial relation. If in (x ρ y) ρ is the relation *is hyponym of*, then the inverse relation ρ' in (y ρ' x) is *is hyperonym of*.

Insofar as a KOS has synonymy, which of course is always, symmetrical (if x is synonymous to y, then y is synonymous to x), it will never be an order in the mathematical sense. An open question is whether all hierarchical relations in KOSs are transitive as a matter of principle. We can easily find counterexamples in a first, naïve approach to the problem. Let us assume, for instance, that the liver of Professor X is a part of X and Professor X is a part of University Y, then transitivity dictates that the liver of Professor X is a part of University Y, which is obviously nonsense. But attention! Was that even the same relation? The liver is an organ; a professor is part of an organization. Only because we simplified and started from a general part-whole relation does not mean that transitivity applies. Intransitivity may thus mean, on the one hand, that the concept system (wrongly) summarizes different relations as one single relation, or on the other hand, that the relation is indeed intransitive.

Why is transitivity in particular so important for information retrieval? Central applications are query expansion (automatically or manually processed in a dialog between a user and a system) (Stock, 2007, p. 480) or (in ontologies) automatic inference. If someone, for example, were to search for stud farms in the Rhein-Erft district of North-Rhine Westphalia, they would formulate:

Stud Farm AND Rhein-Erft District

The most important farms are in Quadrath-Ichendorf, which is a part of Bergheim, which in turn is in the Rhein-Erft district. If we expand the second argument of the search request downwards, proportionately to the geographical structure, we will arrive, in the second step, at the formulation that will finally provide the search results:

Stud Farm AND (Rhein-Erft District OR Bergheim
OR …OR Quadrath-Ichendorf)

Query expansion can also lead to results by moving upwards in a hierarchical chain. Let us say that a motorist is confronted with the problem of finding a repair shop for his car (a Ford, for instance) in an unfamiliar area. He formulates on his mobile device:

Repair Shop AND Ford AND ([Location], e.g., determined
via GPS).

The retrieval system allocates the location to the smallest geographical unit and first takes one step upwards in the hierarchical chain, and at the same time back down, to the sister terms. If there are no results, it's one hierarchy level up and again to the sister terms, and so forth until the desired document has been located.

A query expansion by exactly one step can be performed at every time. If we imagine the KOS as a graph, we can thus always and without a problem incorporate those concepts into the search request that are linked to the initial concept via a path length of one. (Whether this is always successful in practice is moot. The incorporation of hyperonyms into a search argument in particular can expand the search results enormously and thus negatively affect precision.) If we want to expand via path lengths greater than one, we must make sure that there is transitivity, as otherwise there would be no conclusive semantic relation to the initial concept.

Equivalence

Two designations are synonymous if they denote the same concept. Absolute synonyms, which extend to all variants of meaning and all (descriptive, social, and expressive) references are rare; an example is *autumn* and *fall*. Abbreviations (*TV/television),* spelling variants (*grey/gray),* inverted word order (*the sweet night air/the night air sweet)*, and shortened versions (*The Met/The Metropolitan Opera*) are totally synonymous as well. Closely related to total synonymy are common terms from foreign languages (*rucksack/backpack*) and divergent linguistic usage (*media of mass communication/mass media*).

After Löbner (2002, p. 46), most synonymy relations are of a partial nature: They do not designate the exact same term but stand for (more or less) closely related concepts. Differences may be located in either extension or intension. Löbner's (2003, p. 117) example *geflügelte Jahresendpuppe* (literally, *winged end-of-year doll*, in the German Democratic Republic's official lingo) may be extensionally identical to *Weihnachtsengel* (*Christmas angel*), but is not intensionally so. As opposed to true synonymy, which is a relation between designations and a concept, partial synonymy is a relation between concepts.

In information practice, most KOSs treat absolute and partial synonyms, and furthermore, depending on the purpose, similar terms (as 'quasi-synonyms') as one and the same term. If two terms are linked as synonyms in a concept system, they are (right until the system is changed) always a unit and cannot be considered in isolation. If the concept system is applied to full-text retrieval systems, the search request will be expanded by all the fixed synonyms of the initial search term. Synonymy is reflexive, symmetrical, and transitive.

Certain objects are "gen-identical" (Menne, 1980, p. 68). This is a weak form of identity, which disregards certain temporal aspects. A human being at different ages (Person X as a child, adult, and old man/woman) is thus gen-identical. A possible option in concept systems is to summarize terms for gen-identical objects as quasi-synonyms. There is, however, also the possibility of regarding the respective terms individually and linking them subsequently.

If gen-identical objects are described by different terms at different times, these terms will be placed in the desired context via chronological relations. These relations are called "chronologically earlier" and—as the inversion—"chronologically later." As an example, let us consider the city located where the Neva flows into the Baltic Sea:

> Between 1703 and 1914: Saint Petersburg
> 1914–1924: Petrograd
> 1924–1981: Leningrad
> Afterwards: Saint Petersburg again.

Any neighboring concepts are chronologically linked:

> *Saint Petersburg* [*Tsar Era*] is chronologically earlier than *Petrograd*.
> *Petrograd* is chronologically earlier than *Leningrad*.

The chronological relation is irreflexive, asymmetrical, and transitive.

Two concepts are antonyms if they are mutually exclusive. Such opposite concepts are, for example, *love – hate, genius – insanity,* and *dead – alive*. We must distinguish between two variants: Contradictory antonyms know exactly two shadings, with nothing in between. Someone is pregnant or isn't pregnant—tertium non datur. Contrary antonyms allow for other values between the extremes; between love and hate, for example, lies indifference. For contradictory antonyms it is possible, in retrieval, to incorporate the respective opposite concept—linked with a negating term such as "not" or "un-"—into a query. Whether or not contrary antonyms can meaningfully be used in knowledge representation and information retrieval is an open question at this point. Antonymy is irreflexive, symmetrical, and intransitive.

Hierarchy

The most important relation of concept systems, the supporting framework so to speak, is hierarchy. Emile Durkheim (1995 [1912]) assumes that hierarchy is a fundamental relation used by all men to put the world in order. As human societies are always structured hierarchically, hierarchy is—according to Durkheim—experienced in everyday life and, from there, projected onto our concepts of "the world."

If we do not wish to further refine the hierarchy relation of a KOS, we will have a "mixed-hierarchical concept system" (DIN 2331:1980, p. 6). It is called "mixed" because it summarizes several sorts of hierarchical relation. This approach is a very simple and naïve world view. We distinguish between three variants of hierarchy: hyponymy, meronymy, and instance.

Hyponym-Hyperonym Relation

The abstraction relation is a hierarchical relation that is subdivided from a logical perspective. "Hyperonym" is the term in the chain located precisely one hierarchy level higher than an initial term; "hyponym" is a term located on the lower hierarchy level. "Sister terms" (first-degree parataxis) share the same hyperonym. Concepts in hierarchical relations form hierarchical chains or concept ladders. In the context of the definition, each respective narrower term is created via concept explanation or—as appropriate—via family resemblance. If there is no definition via family resemblance, the hyponym will inherit all properties of the hyperonym. In the case of family resemblance, it will only inherit a partial quantity of the hyperonym's properties. Additionally, it will have at least one further fundamental property that sets it apart from its sister terms. For all elements of the hyponym's extension, the rule applies that they are also always elements of the hyperonym. The logical subordination of the abstraction relation always leads to an implication of the following kind (Löbner, 2002, p. 85; Storey, 1993, p. 460):

> If x is an A, then x is a B
> if A is a hyponym of B.

If it is true that *bluetit* is a hyponym of *tit*, then the following implication is also true:

> If it is true that: x is a bluetit, then it is true that: x is a tit.

The abstraction relation can always be expressed as an "IS-A" relation (Khoo & Na, 2006, p. 174). In the example

> Bird – Songbird – Tit – Blue Tit

(defined in each case without resorting to family resemblance), it is true that:

> The bluetit IS A tit.
> The tit IS A songbird.
> The songbird IS A bird.

Properties are added to the intension on the journey upwards: A *songbird* is a *bird* that sings. The *bluetit* is a *tit* with blue plumage. Mind you: The properties must each be noted in the term entry (keyword entry, descriptor entry, etc.) via specific relations; otherwise any (automatically implementable) heredity would be completely impossible.

If we define via family resemblance, the situation is slightly different. In the example:

> Leisure Activity – Game – Game of Chance

it is true, as above, that:

> A game of chance IS A game.
> A game IS A leisure activity.

As we have delimited *game* via family resemblance, *game of chance* does not inherit all properties of *game* (e.g., not necessarily *board game, card game*), but only a few. The hyponym's additional property (is a game of chance) is in this

DOI: 10.1002/asi

case already present as a part of the hyperonym's terms, which are linked via OR. The clarification is performed by excluding the other family members linked via OR (for instance, like: is precisely a game requiring luck).

It is tempting to assume that there is reciprocity between the extension and intension of concepts in a hierarchical chain: To increase the content (i.e., to add further properties on the way up) would go hand in hand with a decrease of the number of objects that fall under the concept. There are certainly more birds in the world than there are songbirds. Such a reciprocal relation can be found in many cases, but it has no general validity. It is never the case for individual concepts, as we could always add further properties to those without changing the extension. The intention of *Karl May,* for instance, is already clearly defined by *author, born in Saxonia, invented Winnetou*; adding *has business relations with the Münchmeyer publishing house* would not change the extension in the slightest. We can even find counterexamples for general concepts, i.e., concepts that display an increase in extension as their content is augmented. The classical example is by Bolzano (1973 [1837]). Walter Dubislav (1981, p. 121) gives a lecture on this case:

> Let us use with *Bolzano* the concept of a "speaker of all European languages" and then augment the concept by adding the property "living" to the concept "speaker of all living European languages." We can notice that the intension of the first concept has been increased, but that the extension of the new concept thus emerging contains the extension of the former as a partial quantity.

Of course, we have to assume that there are more speakers of all living European languages than speakers of all European languages which include dead languages such as Gothic, Latin, or Ancient Greek.

We can make out two variant forms of the abstraction relation: taxonomy, and nontaxonomical, "simple" hyponymy. In a taxonomy, the IS-A relation can be strengthened into IS-A-KIND-OF (Cruse, 2002, p. 12). A taxonomy does not just divide a larger class into smaller classes, as is the case for simple hyponymy. Let us consider two examples:

? A queen IS A KIND OF woman.
(better: A queen IS A woman).
? A stallion IS A KIND OF horse.
(better: A stallion IS A horse).

In both cases, the variant IS A KIND OF is unrewarding; here, we have simple hyponymy. If we instead regard the following examples:

A cold blood IS A KIND OF horse.
A stetson IS A KIND OF hat,

we can observe that here the formulation makes sense, as there is indeed a taxonomical relation in these cases. A taxonomy fulfills certain conditions, according to Cruse (2002, p. 13):

> Taxonomy exists to articulate a domain in the most effective way. This requires "good" categories, which are (a) internally cohesive, (b) externally distinctive, and (c) maximally informative.

In taxonomies, the hyponym, or "taxonym," and the hyperonym are fundamentally regarded from the same perspective. *Stallion* is not a taxonym of *horse*, because *stallion* is regarded from the perspective of gender and *horse* is not. In the cases of *cold blood* and *horse*, though, the perspectives are identical; both are regarded from a biological point of view. The hyponymy-hyperonymy relation is irreflexive, asymmetrical, and transitive.

Meronym-Holonym Relation

If the abstraction relation represents a logical perspective on concepts, the part-whole relation starts from an objective perspective (Khoo & Na, 2006, p. 176). Concepts of wholeness, "holonyms," are divided into concepts of their parts, "meronyms." If in an abstraction relation it is not just any properties that are used for the definition but precisely the characteristics that make up its essence, then the part-whole relation likewise does not use any random parts but the "fundamental" parts of the wholeness in question. The meronym-holonym relation has several names. Apart from "part-whole relation" or "part-of relation," we also speak of "partitive relations" (as in DIN 2331:1980, p. 3). A system based on this relation is called "mereology" (Simons, 1987).

In individual cases, it is possible that meronymy and hyponymy coincide. Let us consider the pair of concepts:

Industry – Chemical Industry.

Chemical industry is as much a part of industry in general as it is a special kind of industry.

Meronymy is expressed by "PART OF." This relation does not exactly represent a concept relation but is made up of a bundle of different partitive relations. If one wants—in order to simplify, for example—to summarize the different part-whole relations into a single relation, transitivity will be damaged in many cases. Winston, Chaffin, and Herrmann (1987, pp. 442–444) compiled a list of (faulty) combinations. An example may prove intransitivity:

Simpson's finger is part of Simpson.
Simpson is part of the Philosophy Department.
? Simpson's finger is part of the Philosophy Department.

The sentence marked with a question mark is a false conclusion. We can (as a "lazy solution") do without the transitivity of the respective specific meronymy relations in information retrieval. In so doing, we would deprive ourselves of the option of query expansion over more than one hierarchy level. But we do not even need to make the effort of differentiating between the single partitive relations.

DOI: 10.1002/asi

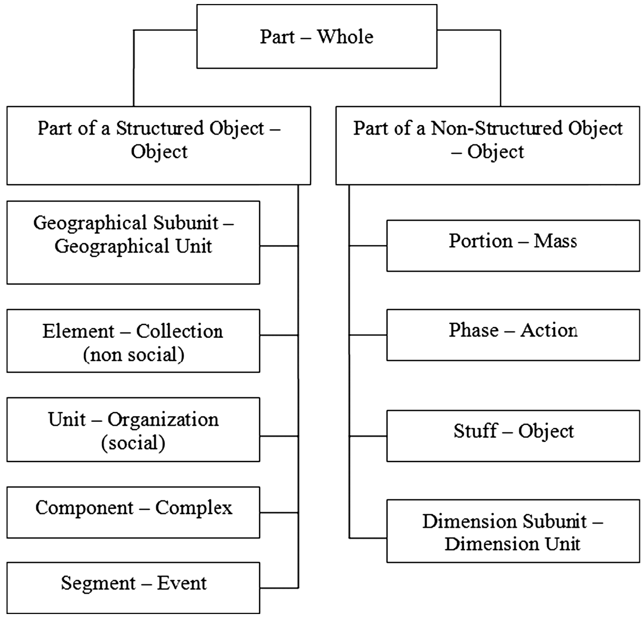

FIG. 4. Specific meronym-holonym relations.

The elaborated solution distinguishes the specific meronymy relations and analyzes them for transitivity, thus providing the option of query expansion at any time and over as many levels as needed.

We will follow the approach, now classical, of Winston et al. (1987) and specify the part-whole relation into meaningful kinds. Winston et al. distinguish six different meronymy relations, which we will extend to nine via further subdivision (Figure 4) (Weller & Stock, 2008).

Insofar as wholenesses have a structure, this structure can be divided into certain parts (Gerstl & Pribbenow, 1996; Pribbenow, 2002). The five part-whole relations displayed on the left of Figure 4 distinguish themselves by having had wholenesses structurally subdivided. Geographical data allow for a subdivision according to administrative divisions, provided we structure a given geographical unit into its subunits. *Northrhine-Westphalia* is a part of *Germany*; the locality *Kerpen-Sindorf* is a part of *Kerpen*. (Nonsocial) uniform collections can be divided into their elements. A *forest* consists of *trees*; a *ship* is part of a *fleet*. A similar aspect of division is at hand if we divide (uniform) organizations into their units, such as a *university* into its *departments*. Johansson (2004) notes that in damaging uniformity there is not necessarily transitivity. Let us assume that there is an association Y, of which other associations $X_1, \ldots, X_n$ (and only associations) are members. Let person A be a member of X_1. In case of transitivity, this would mean that A, via membership in X_1, is also a part of Y. But according to Y's statutes, this is absolutely impossible. The one case is about membership of persons, the other about membership of associations, which means that the principle of uniformity has been damaged in the example. A contiguous complex, such as a *house*, can be subdivided into its components, e.g., the *roof* or the *cellar*. Meronymy, for an event (say a *circus performance*) and a specific segment (e.g., the *trapeze act*), is similarly formed (temporally speaking in this case) (Storey, 1993, p. 464).

The second group of meronyms works independently of structures (on the right-hand side in Figure 4). A wholeness can be divided into random portions, such as a *cup* (after we drop it to the floor) into *shards* or—less destructively—a *bread* into servable *slices*. A continuous activity (e.g., *shopping*) can be divided into single phases (e.g., *paying*). One of the central important meronymy relations is the relation of an object to its stuff, such as the *aluminum* parts of an *airplane* or the *wooden parts* of my *desk*. If we have a homogenous unit, we can divide it into subunits. Examples are *wine* (*in a barrel*) and *1 liter of wine* or *meter – decimeter*.

All described meronymy-holonymy relations are irreflexive, asymmetrical, and transitive, insofar as they have been defined and applied in a "homogeneous" way.

We have already discussed the fact that in the hierarchical chain of a hyponym-hyperonym relation the concepts pass on their properties (in most cases) from the top down. The same goes for their meronyms: We can speak of meronym heredity in the abstraction relation (Weller & Stock, 2008, p. 168). If concept A is a partial concept (e.g., a *motor*) of the wholeness B (a *car*), and C is a hyponym of B (let's say: an *ambulance*), then the hyponym C also has the part A (i.e., an *ambulance* therefore has a *motor*).

Instance

In extensional definition the concept in question is defined by enumerating those elements for which it applies. In general, the question of whether the elements are general or individual concepts is left unanswered. In the context of the instance relation, it is demanded that the element always be an individual concept. The element is thus always a "named entity" (Stock, 2007, p. 254).

Whether this element-class relation is regarded in the context of hyponymy or meronymy is irrelevant for the instance relation. An entity can be expressed both via "is a" and via "is part of." In the sense of an abstraction relation, we can say that:

Persil IS A detergent.

Cologne IS A university city on the Rhine.

Likewise, we can formulate:

Silwa (our car) IS PART OF our motor pool.

Angela Merkel IS PART OF the CDU.

Instances can have hyponyms of their own. Thus, in the last example, *CDU* is an instance of the concept *German political party*. And obviously our *Silwa* has parts, such as chassis or motor. The instance relation is reflexive, asymmetrical, and intransitive.

Further Specific Relations

There is a wealth of other semantic relations in concept systems, which we will, in an initial approach, summarize

DOI: 10.1002/asi

under the umbrella term "association." The association relation as such therefore does not exist, there are merely various different relations. Common to them all is that they—to put it negatively—do not form (quasi-)synonyms or hierarchies and are—positively speaking—of use for knowledge organization systems.

In a simple case, which leaves every specification open, the association relation plays the role of a "see also" link. The terms are related to each other according to practical considerations, such as the link between products and their respective industries in a business administration KOS (e.g., *body care product* SEE ALSO *body care product industry* and vice versa). The unspecific "see also" relation is irreflexive, symmetrical, and intransitive.

For other, now specific, association relations, we will begin with a few examples. Schmitz-Esser (2000, pp. 79–80) suggests the relations of usefulness and harmfulness for a specific KOS (of the world fair "Expo 2000"). Here it is shown that such term relations have persuasive "secondary stresses." In the example:

> Wind-up radio IS USEFUL FOR communication in remote areas.

there is no implicit valuation. This is different for:

> Overfishing IS USEFUL FOR the fish meal industry.
> Poppy cultivation IS USEFUL FOR the drug trade.

A satisfactory solution might be found in building on the basic values of a given society ("useful for whom?") in case of usefulness and harmfulness (Schmitz-Esser, 2000, p. 79) and thus reject the two latter examples as irreconcilable with the respective moral values or—from the point of view of a drug cartel—keep the last example as adequate.

One of the knowledge domains that has enjoyed particular attention in the context of ontological knowledge representation is biology. "Gene Ontology" (GO) has even achieved an almost exemplary significance (Ashburner et al., 2000). But: This is not an ontology at all, it is a thesaurus (to be precise: three partial thesauri for biological processes, molecular functions, and cellular components), as this concept system only uses the PART OF and IS A relations, i.e., only meronymy and hyponymy. Smith, Williams, and Schulze-Kremer (2003, p. 609) have consequently established:

> The Gene Ontology, in spite of its name, is not an ontology as the latter term is commonly used either by information scientists or by philosophers. It is, as the GO Consortium puts it, a 'controlled vocabulary.'

The Gene Ontology is a good starting point, though, for explaining which relations—apart from hierarchical ones—are actually needed in biomedicine (Smith et al., 2005). For us, the Smith et al.'s approach is an example of how to specify the formerly unspecific association relation into different concrete relations. Hierarchical relations are a supporting framework in ontologies as well (Smith et al., 2005):

> *Is_a* and *part_of* have established themselves as foundational to current ontologies. They have a central role in almost all domain ontologies. . .

Since ontologies aim towards the use of relations in terminological logic, we have no use for an unspecific association relation in this context. The aim is to develop attributes that are as exact as possible—as well as, subsequently, exact values—that are characteristic for the respective knowledge domain. For the area of genetics, Smith et al. differentiate between components C ("continuants" as a generalization of the "cellular components" of the original GO) and processes P ("processes" as a generalization of "biological processes"). The following eight relations are fundamental for Smith et al. (2005), apart from hierarchical relations:

Relation	Example
C *located_in* C_i	66s pre-ribosome *located_in* nucleolus chlorophyll *located_in* thylakoid
C *contained_in* C_i	cytosol *contained_in* cell compartment space synaptic vesicle *contained_in* neuron
C *adjacent_to* C_i	intron *adjacent_to* exon cell wall *adjacent_to* cytoplasm
C *transformation_of* C_i	fetus *transformation_of* embryo mature mRNA *transformation_of* pre-mRNA
C *derives_from* C_i	plasma cells *derives_from* lymphocyte mammal *derives_from* gamete
P *preceded_by* P_i	translation *preceded_by* transcription digestion *preceded_by* ingestion
P *has_participant* P_i	photosynthesis *has_participant* chlorophyll cell division *has_participant* chromosome
P *has_agent* C	transcription *has_agent* RNA polymerase translation *has_agent* ribosome.

In the case of the relation *derives_from*, we recognize, here in the terminological field of genetics, the chronological relation of gen-identity (which we would otherwise define generally). Smith et al. (2005) distinguish three simple forms of derivation.

> [F]irst, the succession of one single continuant by another single continuant across a temporal threshold (for example, this blastocyst derives from this zygote); second, the fusion of two or more continuants into one continuant (for example, the zygote derives from this sperm and from this ovum); and third, the fission of an earlier single continuant to create a plurality of later continuants (for example, these promyelocytes derive from this myeloblast). In all cases we have two continuants c and c_1 which are such that c begins to exist at the same instant of time at which c_1 ceases to exist, and at least a significant portion of the matter of c_1 is inherited by its successor c.

DOI: 10.1002/asi

Whether a specification of the association relation will lead to a multitude of semantic relations that are generalizable (i.e., usable in all or at least most KOS) is an as-yet unsolved research problem. Certainly, it is clear that we always need a relation has_property for all concepts of a KOS. This solution is very general: it would be more appropriate to specify the kind of property (such as "has melting point" in a KOS on materials, or "has subsidiary company" in an enterprise KOS).

A necessary condition for the formulation of a new semantic relation is that it is significant in a knowledge domain. It is not wise to use specific, only "private" relations (e.g., resulting from individual differences in the interpretation of texts; Morris, 2010) in a KOS, but relations that are common knowledge to the majority of representatives of the knowledge domain.

TABLE 1. Reflexivity, symmetry, and transitivity of paradigmatic relations.

	Reflexivity	Symmetry	Transitivity
Equivalence			
– Synonymy	R	S	T
– Gen-Identity	−R	−S	T
– Antonymy	−R	S	−T
Hierarchy			
– Hyponymy			
– simple Hyponymy	−R	−S	T
– Taxonomy	−R	−S	T
– Meronymy (unspecific)	−R	−S	?
– specific Meronymies	−R	−S	T
– Instance	R	−S	−T
Specific Relations			
– "see also"	−R	S	−T
– further Relations	Depending on the relation		

Relations Between Relations

Relations can be in relation to each other (Horrocks & Sattler, 1999). In the above, we introduced meronymy and formed structure-disassembling meronymy as its specification, and within the latter, the component-complex relation, for example. There is a hierarchy relation between the three above relations. Such relations between relations can be used to derive conclusions. If, for example, we introduced to our concept system:

Roof is a component of *house*,

then it is equally true that

roof is a structural part of *house* and
roof is a part of *house*,

generally formulated:

A is a component of $B \rightarrow A$ is a structural part of $B \rightarrow A$ *is a part of* B.

Relations and Knowledge Organization Systems

We define KOSs via their cardinality for expressing concepts and relations. The three "classical" methods in information science and practice—nomenclature, classification, thesaurus—are supplemented by folksonomies and ontologies (Stock & Stock, 2008). Folksonomies represent a borderline case of KOS, as they do not have a single paradigmatic relation. Nomenclatures (keyword systems) distinguish themselves mainly by using the equivalence relation and ignoring all forms of hierarchical relation. In classification systems, the (unspecifically designed) hierarchy relation is added. Thesauri also work with hierarchy; some use the unspecific hierarchy relation, others differentiate via hyponymy and (unspecific) meronymy (with the problem [see Table 1] of not being able to guarantee transitivity). In thesauri, a generally unspecifically designed association relation ("see also") is necessarily added. Ontologies make use of all the paradigmatic relations mentioned above (Hovy, 2002). They are modeled in formal languages, where terminological logic is also accorded its due consideration. Compared to other KOS, ontologies categorically contain instances. Most ontologies work with (precisely defined) further relations. The fact that ontologies directly represent knowledge (and not merely the documents containing the knowledge) lets the syntagmatic relations disappear in this case. If we take a look at Table 2 or Figure 5, the KOS are arranged from left to right, according to their expressiveness. Each KOS can be "enriched" to a certain degree and lifted to a higher level via relations of the system to its right: A nomenclature can become a classification, for example, if (apart from the step from keyword to notation) all concepts are brought into a hierarchical relation; a thesaurus can become an ontology if the hierarchy relations are precisely differentiated and if further specific relations are introduced. An ontology can become—and now we are taking a step to the left—a method of indexing if it introduces the syntagmatic relation, i.e., if it allows its concepts to be allocated to documents, while retaining all its relations. Thus, the advantages of the ontology, with its cardinal relation framework, flow together with the advantages of document indexing and complement each other.

If there is more than one category in a KOS it is always possible to construct a faceted KOS (Gnoli, 2008). The categories mutate into the facets. There are faceted nomenclatures (Stock & Stock, 2008, pp. 281–284), faceted classification systems (Broughton, 2006; Gnoli & Mei, 2006; Slavic, 2008; Vickery, 2008), faceted thesauri (Spiteri, 1999; Tudhope & Binding, 2008), and even faceted folksonomies (Spiteri, 2010) as well.

On the Way to the Semantic Web?

With ontologies, we have arrived in the core area of the Semantic Web (Berners-Lee, Hendler, & Lassila, 2001;

DOI: 10.1002/asi

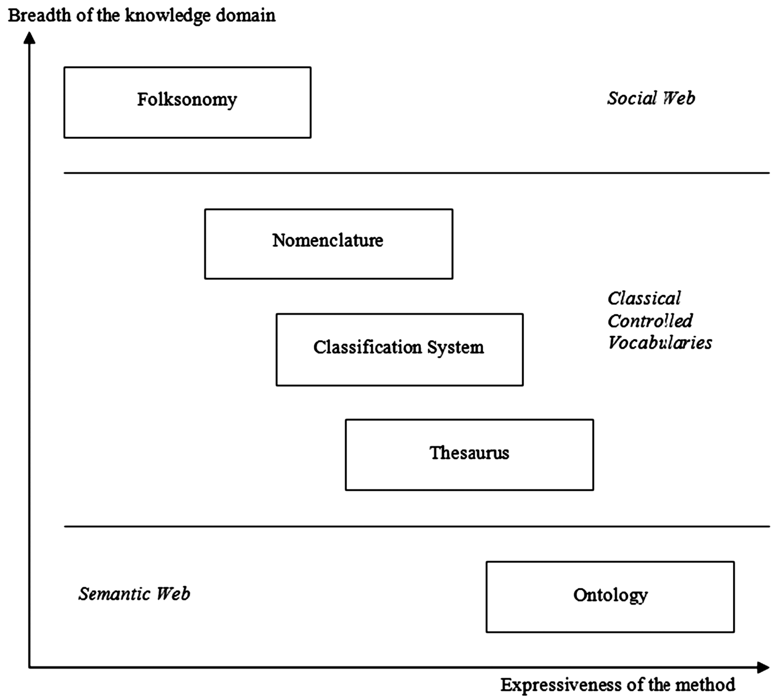

FIG. 5. Expressiveness of KOS methods and the breadth of their knowledge domains. *Source*. Stock & Stock, 2008, p. 42 (modified).

TABLE 2. Knowledge organization systems and the relations they use.

	Folksonomy	Nomenclature	Classification	Thesaurus	Ontology
Term	Tag	Keyword	Notation	Descriptor	Concept
Equivalence	–	Yes	Yes	Yes	Yes
– Synonymy	–	Yes	Yes	Yes	Yes
– Gen–Identity	–	Yes	–	–	Yes
– Antonymy	–	–	–	–	Yes
Hierarchy	–	–	Yes	Yes	Yes
– Hyponymy	–	–	–	Yes	Yes
– simple Hyponymy	–	–	–	–	Yes
– Taxonomy	–	–	–	–	Yes
– Meronymy (unspecific)	–	–	–	Yes	–
– specific Meronymies	–	–	–	–	Yes
– Instance	–	–	–	As req.	Yes
Specific Relations	–	–	–	Yes	Yes
– “see also”	–	As req.	As req.	Yes	Yes
– further Relations	–	–	–	–	Yes
Syntagmatic relation	Yes	Yes	Yes	Yes	No

Shadbolt, Hall, & Berners-Lee, 2006). Shadbolt et al. formulate the Semantic Web’s claim as follows (2006, p. 96):

> The Semantic Web is a Web of actionable information—information derived from data through a semantic theory for interpreting the symbols. The semantic theory provides an account of “meaning” in which the logical connection of terms establishes interoperability between systems.

Discussions of the Semantic Web are at first highly technically oriented: They are about RDF (resource description framework), URI (universal resource identifiers), the suitable ontology language (such as OWL, the Web ontology language), the rules of automatical inference sketched above, and ontology editors such as Protégé (Noy, Fergerson, & Musen, 2000; Noy et al., 2001). Both the concept-theoretical background and the methods for creating suitable KOS sometimes have been left open. Current attempts at a solution are discussed by Shadbolt et al. in the form of two approaches that are each based on the cooperation of participating experts. Ontologies (as described here) that are separately constructed and maintained are suited to well-structured knowledge domains (Shadbolt et al, 2006, p. 99):

> In some areas, the costs—no matter how large—will be easy to recoup. For example, an ontology will be a powerful and essential tool in well-structured areas such as scientific

DOI: 10.1002/asi

> applications.. . . In fact, given the Web's fractual nature, those costs might decrease as an ontology's user base increase. If we assume that ontology building costs are spread across user communities, the number of ontology engineers required increases as the log of the community' size.

This approach can only be used if, first, the knowledge domain is small and overseeable and, second, the members of the respective community of scientists are willing to contribute to the construction and maintenance of the ontology. Such an approach seems not to work at all beyond the borders of disciplines. The second approach, to consolidate the Semantic Web, proceeds via tagging and folksonomies (Shadbolt et al., 2006, p. 100):

> Tagging on a Web scale is certainly an interesting development. It provides a potential source of metadata. The folksonomies that emerge are a variant on keyword searches. They're an interesting emergent attempt at information retrieval.

Folksonomies, however, exclusively have syntagmatic relations. If one wants to render such an approach usable for the Semantic Web (now as the "Social Semantic Web;" Weller, 2010), focused work in "Tag Gardening" would seem to be required (Peters & Weller, 2008b). Equally open, in our view, is the question of indexing documents in the Semantic Web. Who will perform this work—the author, the users, or—as automatic indexing—a system?

For Weller (2010), principally using only ontologies for the Semantic Web is an excessive approach. Ontologies are far too complicated for John Q. Web User in terms of structure and use. Weller's suggestion boils down to a simplification of the concept system: Not all aspects of ontologies have to be realized, the desired effects might also be achieved by a less expressive method (e.g., a thesaurus). The easier the concept systems are to use for the individual, the greater the probability of many users participating in the collaborative construction of the Semantic Web will be.

Semantic-rich KOSs, notably thesauri and ontologies, only work well in small knowledge domains. Therefore, we have to consider the problem of semantic interoperability between different KOS (Gödert 2010a,b), i.e., the formulation of relations of (quasi-)synonymy, hierarchy, etc. (as means of semantic crosswalks) between singular concepts and compounds beyond the borders of single KOS (Stock & Stock, 2008, p. 291) to form a "universal" KOS which is needed for the Semantic Web.

The vision of a universal Semantic Web on the basis of ontologies is being set very narrow boundaries with the approaches introduced by Shadbolt et al. (2006). The ruins of a similar vision of summarizing world knowledge—Otlet's and La Fontaine's "Mundaneum" (Torres-Vargas, 2005)—can today be admired in a museum in Mons, Belgium. We do not propose that a Semantic Web is principally impossible; we merely wish to stress that research on concepts and relations is very useful on the way to the Semantic Web. Both the construction of all sorts of KOS as well as tag gardening are genuine domains of information science.

Conclusion

Let us summarize the important results of our theoretical journey through the jungle of concept theory, theory of semantic relations, and KOS.

(1) It is a truism, but we want to mention it in the first place: Concept-based information retrieval is only possible if we are able to construct and to maintain adequate KOS.

(2) Concepts are defined by their extension (objects) and by their intension (properties). It is possible to group similar objects (via classification) or unlike objects (via colligation) together, depending on the purpose of the KOS.

(3) There are five epistemological theories on concepts in the background of information science: empiricism, rationalism, hermeneutics, critical theory, and pragmatism. None of them should be forgotten in activities concerning KOS.

(4) Concepts are (a) categories, (b) general concepts, and (c) individual concepts. Categories and individual concepts are more or less easy to define, but general concepts tend to be problematic. Such concepts (such as *chair*) have fuzzy borders and should be defined by prototypes. Every concept in a KOS has to be defined exactly (by extension, intension, or both). In a concept entry, all properties (if applicable, objects as well) must be listed completely and in a formal way. It is not possible to work with the inheritance of properties if we do not define those properties.

(5) In information science, we mainly work with two kinds of definition, namely, concept explanation and family resemblance. Concepts, which are defined via family resemblance (*vegetable* or *game*), do not pass all properties down to their narrower terms. This result is important for the design of ontologies.

(6) Concepts can be presented as frames with sets of attributes and values, structural invariants (relations), and rule-bound connections.

(7) Concept systems are made up of concepts and semantic relations between them. Semantic relations are either syntagmatic relations (co-occurrences of terms in documents) or paradigmatic relations (tight relations in KOS). There are three kinds of paradigmatic relations: equivalence, hierarchy, and further specific relations.

(8) Especially for hierarchic relations, transitivity plays an important role for query expansion. Without proven transitivity, it is not possible to expand a search argument with concepts from hierarchical levels with distances greater than one.

(9) Equivalence has three manifestations: synonymy, gen-identity, and antonymy. Absolute synonymy, which is very sparse, is a relation between a concept and different words. All other kinds of synonymy, often called "quasi-synonymy," are relations between different concepts. Gen-identity describes an object in the course of time. Contradictory antonyms are useful in information retrieval, but only with constructions like "not" or "un-."

(10) Hierarchy is the most important relation in every KOS. It consists of the (logic-oriented) hyponym-hyperonym-relation (with two subspecies, simple hyponymy and taxonomy), the (object-oriented) meronym-holonym-relation (with a lot of subspecies), and the instance relation

DOI: 10.1002/asi

(relation between a concept and an individual concept as one of its elements). KOS designers have to pay attention to the transitivity of all relations.

(11) There are many further relations, such as usefulness or harmfulness or—in the context of genetics—has_agent or adjacent_to. It is possible to integrate all these relations into only a single association relation (as in thesauri), but it is more expressive to work with the specific relations, as they are necessary in a given knowledge domain.

(12) We can order types of KOS regarding their expressiveness (quantity and quality of concepts and semantic relations): from folksonomies via nomenclatures, classification systems, thesauri up to ontologies.

(13) Our approach is on concepts and semantic relations in general. There are no statements about special aspects or problems in single knowledge domains. So it is an object of further research to analyze concept formation and KOS in science, humanities, arts, everyday life, or the use of KOS in corporate knowledge management.

(14) All types of KOS—and not only ontologies—are able to form the terminological backbone of the semantic web. The construction of the semantic web is not only a technical task, but calls for tasks such as construction of KOS and of (automated or manual) indexing of documents.

(15) Current guidelines for the construction and maintenance of KOS (such as ANSI/NISO Z39.19-2005 or DIN 2331:1980) do not consider results of concept theory and theory of relations to the full extent. They are not able to unify the worlds of traditional controlled vocabularies (nomenclatures, classification systems, and thesauri), of the social web (tagging and folksonomies) and of the semantic web (ontologies). Here, discussions and further research are necessary.

Acknowledgments

This article is an intensely modified and updated version of a German-language paper (Stock, 2009). It benefits from discussions with my colleagues Isabella Peters and Katrin Weller from the Heinrich-Heine-University Düsseldorf and with the coauthor of our book, *Wissensrepräsentation*, Mechtild Stock. I thank Paul Becker for the translation of this piece from German to English. Finally, I thank my Danish colleague Birger Hjørland: He gave the impetus to translate and to revise the original article.

References

Aristotle. (1998). Metaphysics. London: Penguin Books.

Ashburner, M., Ball, C.A., Blake, J.A., Botstein, D., Butler, H., Cherry, J.M., et al. (2000). Gene ontology: Tool for the unification of biology. The Gene Ontology Consortium. Nature Genetics, 25, 25–29.

Barsalou, L.W. (1992). Frames, concepts, and conceptual fields. In E. Kittay & A. Lehrer (Eds.), Frames, fields and contrasts. New essays in semantic and lexical organization (pp. 21–74). Hillsdale, NJ: Lawrence Erlbaum.

Berners-Lee, T., Hendler, J., & Lassila, O. (2001). The semantic Web. Scientific American, 284(5), 28–37.

Bertram, J. (2005). Einführung in die inhaltliche Erschließung. Grundlagen – Methoden – Instrumente [Introduction to indexing. Foundations – methods – tools]. Würzburg, Germany: Ergon.

Black, M. (1937). Vagueness. Philosophy of Science, 4, 427–455.

Bolzano, B. (1973). Theory of science. Dordrecht, Netherlands: Reidel. (Original work published 1837).

Boyle, A. (2010). The case for Pluto. How a little planet made a big difference. Hoboken, NJ: John Wiley & Sons.

Broughton, V. (2006). The need for a faceted classification as the basis of all methods of information retrieval. Aslib Proceedings, 58(1/2), 49–72.

Buizza, P. (2010, February). Indexing concepts and/or named entities. Paper presented at the International Society for Knowledge Organization Conference: Paradigms and conceptual systems in knowledge organization, Rome, Italy.

Cruse, D.A. (2002). Hyponymy and its varieties. In R. Green, C.A. Bean, & S.H. Myaeng (Eds.), The semantics of relationships (S. 3–21). Dordrecht, the Netherlands: Kluwer.

Dahlberg, I. (1974). Zur Theorie des Begriffs [On concept theory]. International Classification, 1(1), 12–19.

Dahlberg, I. (1986). Die gegenstandsbezogene, analytische Begriffstheorie und ihre Definitionsarten [The object-oriented analytical concept theory and its kinds of definition]. In B. Ganter, R. Wille, & K.E. Wolff (Eds.), Beiträge zur Begriffsanalyse (pp. 9–22). Mannheim, Germany, Wien, Austria, Zürich, Switzerland: BI Wissenschaftsverlag.

DIN 2330:1993. Begriffe und Benennungen. Allgemeine Grundsätze [Concepts and denotations. Basic principles]. Berlin, Germany: Beuth.

DIN 2331:1980. Begriffssysteme und ihre Darstellung [Concept systems and their presentation]. Berlin, Germany: Beuth.

DIN 2342/1:1992. Begriffe der Terminologielehre. Grundbegriffe [Concepts of terminology. Basic concepts]. Berlin, Germany: Beuth.

Dubislav, W. (1981). Die Definition [The definition], 4th Ed. Hamburg, Germany: Meiner.

Durkheim, E. (1995). The elementary forms of religious life. New York: Free Press. (First published 1912).

Feldman, S. (2000). Find what I mean, not what I say. Meaning based search tools. Online, 24(3), 49–56.

Fellbaum, C. (Ed.). (1998). WordNet. An electronic lexical database. Cambridge, MA: MIT Press.

Frege, G. (1892). Über Sinn und Bedeutung [On sense and meaning]. Zeitschrift für Philosophie und philosophische Kritik (Neue Folge), 100, 25–50.

Fuchs, C. (2008). Internet and society. Social theory in the information age. New York: Routledge.

Fugmann, R. (1999). Inhaltserschließung durch Indexieren: Prinzipien und Praxis [Content exploitation by indexing. Principles and practice]. Frankfurt, Germany: DGI.

Furnas, G.W., Landauer, T.K., Gomez, L.M., & Dumais, S.T. (1987). The vocabulary problem in human-system communication. Communications of the ACM, 30(11), 964–971.

Gadamer, H.G. (1975). Truth and method. London, United Kingdom: Sheed & Ward. (First published 1960).

Ganter, B., & Wille, R. (1999). Formal concept analysis. Mathematical foundations. Berlin, Germany: Springer.

Geiger, R.S. (2009). Does Habermas understand the internet? The algorithmic construction of the blogo/public sphere. Gnovis. A Journal of Communication, Culture, and Technology, 10(1), 1–29.

Gerstl, P., & Pribbenow, S. (1996). A conceptual theory of part-whole relations and its applications. Data & Knowledge Engineering, 20, 305–322.

Gnoli, C. (2008). Facets. A fruitful notion in many domains. Axiomathes, 18, 127–130.

Gnoli, C., & Mei, H. (2006). Freely faceted classification for Web-based information retrieval. New Review of Hypermedia & Multimedia, 12(1), 63–81.

Green, R. (2001). Relationships in the organization of knowledge: Theoretical background. In C.A. Bean & R. Green (Eds.), Relationships in the organization of knowledge (pp. 3–18). Boston: Kluwer.

Gödert, W. (2010a). Semantische Wissensrepräsentation und Interoperabilität. Teil 1: Interoperabilität als Weg zur Wissensexploration [Semantic knowledge representation and interoperability. Part 1: Interoperability as a

DOI: 10.1002/asi

means to knowledge exploration]. Information – Wissenschaft und Praxis, 61(1), 5–18.

Gödert, W. (2010b). Semantische Wissensrepräsentation und Interoperabilität. Teil 2: Ein formales Modell semantischer Interoperabilität [Semantic knowledge representation and interoperability. Part 2: A formal model of semantic interoperability]. Information – Wissenschaft und Praxis, 61(1), 19–28.

Gust von Loh., S., Stock, M., & Stock, W.G. (2009, November). Knowledge organization systems and bibliographical records in the state of flux. Hermeneutical foundations of organizational information culture. Paper presented at the Annual Meeting of the American Society for Information Science and Technology (ASIS& T 2009), Vancouver, Canada.

Habermas, J. (1987). Knowledge and human interest. Cambridge, United Kingdom: Polity Press. (First published 1968).

Habermas, J. (1998). Between facts and norms. Cambridge, MA: MIT Press.

Habermas, J. (2006). Political communication in media society. Does democracy still enjoy an epistemic dimension? The impact of normative theory on empirical research. Communication Theory, 16, 411–426.

Heidegger, M. (1962). Being and time. San Francisco: Harper. (First published 1927).

Hjørland, B. (2007). Semantics and knowledge organization. Annual Review of Information Science and Technology, 41, 367–406.

Hjørland, B. (2009). Concept theory. Journal of the American Society for Information Science and Technology, 60(8), 1519–1536.

Hjørland, B. (2010a). Concepts. Classes and colligation. Bulletin of the American Society for Information Science and Technology, 36(3), 2–3.

Hjørland, B. (2010b). Answer to Professor Szostak (concept theory). Journal of the American Society for Information Science and Technology, 61(5), 1078–1080.

Hjørland, B. (2010c, February). Concepts, paradigms and knowledge organization. Paper presented at the International Society for Knowledge Organization Conference: Paradigms and Conceptual Systems in Knowledge Organization. Rome, Italy.

Horrocks, I., & Sattler, U. (1999). A description logic with transitive and inverse roles and role hierarchies. Journal of Logic and Computation, 9(3), 385–410.

Hovy, E. (2002). Comparing sets of semantic relations in ontologies. In R. Green, C.A. Bean, & S.H. Myaeng (Eds.). The semantics of relationships (pp. 91–110). Dordrecht, the Netherlands: Kluwer.

Jarrett, K. (2008). Interactivity is evil! A critical investigation of Web 2.0. First Monday, 13(3).

Johansson, I. (2004). On the transitivity of the parthood relations. In H. Hochberg & K. Mulligan (Eds.), Relations and predicates (pp. 161–181). Frankfurt, Germany: Ontos.

Khoo, C.S.G., & Na, J.C. (2006). Semantic relations in information science. Annual Review of Information Science and Technology, 40, 157–228.

Klaus, G. (1973). Moderne Logik [Modern logics], 7th Ed. Berlin, Germany: Deutscher Verlag der Wissenschaften.

Knautz, K., Soubusta, S., & Stock, W.G. (2010, January). Tag clusters as information retrieval interfaces. Paper presented at the 43th Annual Hawaii International Conference on System Sciences (HICSS-43), Kauai, HI. Retrieved May 19, 2010, from http://www.computer.org/portal/web/csdl/doi/10.1109/HICSS.2010.360

Komatsu, L.K. (1992). Recent views of conceptual structure. Psychological Bulletin, 112(3), 500–526.

Kuhn, T.S. (1962). The structure of scientific revolutions. Chicago: University of Chicago Press.

Löbner, S. (2002). Understanding semantics. London: Arnold.

Löbner, S. (2003). Semantik. Eine Einführung [Semantics. An introduction]. Berlin, Germany: Walter de Gruyter.

Menne, A. (1980). Einführung in die Methodologie [Introduction to methodology]. Darmstadt, Germany: Wissenschaftliche Buchgesellschaft.

Mervis, C.B., & Rosch, E. (1981). Categorization of natural objects. Annual Review of Psychology, 32, 89–115.

Minsky, M. (1975). A framework for representing knowledge. In P.H. Winston (Ed.), The psychology of computer vision (pp. 211–277). New York: McGraw-Hill.

Morris, J. (2010). Individual differences in the interpretation of text. Implications for information science. Journal of the American Society for Information Science and Technology, 61(1), 141–149.

Nardy, D., & Brachman, R.J. (2003). An introduction to description logics. In F. Baader, D. Calvanese, D. McGuinness, D. Nardi, & P. Patel-Schneider (Eds.), The description logic handbook. Theory, implementation and applications (pp. 1–40). Cambridge, MA: Cambridge University Press.

National Information Standards Organization. (2005). Guidelines for the Construction, Format, and Management of Monolingual Controlled Vocabularies (ANSI/NISO Z39.19-2005). Baltimore: Author.

Noy, N.F., Fergerson, R.W., & Musen, M.A. (2000). The knowledge model of Protégé-2000: Combining interoperability and flexibility. Lecture Notes in Computer Science, 1937, 69–82.

Noy, N.F., Sintek, M., Decker, S., Crubezy, M., Fergerson, R.W., & Musen, M.A. (2001). Creating Semantic Web contents with Protégé-2000. IEEE Intelligent Systems, 16(2), 60–71.

Ogden, C.K., & Richards, I.A. (1985). The meaning of meaning. London: ARK Paperbacks. (First published 1923).

Pawłowski, T. (1980). Begriffsbildung und Definition [Concept formation and definition]. Berlin, Germany: Walter de Gruyter.

Peters, I. (2009). Folksonomies: Indexing and retrieval in Web 2.0. Berlin, Germany: De Gruyter Saur.

Peters, I., & Stock, W.G. (2007). Folksonomies and information retrieval. In Proceedings of the 70th ASIS&T Annual Meeting (CD-ROM) (pp. 1510–1542). Medford, NJ: Information Today.

Peters, I., & Stock, W.G. (2008). Folksonomies in Wissensrepräsentation und Information Retrieval [Folksonomies in knowledge representation and information retrieval]. Information – Wissenschaft und Praxis, 59(2), 77–90.

Peters, I., & Weller, K. (2008a). Paradigmatic and syntagmatic relations in knowledge organization systems. Information – Wissenschaft und Praxis, 59(2), 100–107.

Peters, I., & Weller, K. (2008b). Tag gardening for folksonomy enrichment and maintenance. Webology, 5(3), article 58.

Petersen, W. (2007). Representation of concepts as frames. The Baltic International Yearbook of Cognition, Logic and Communication, 2, 151–170.

Pribbenow, S. (2002). Meronymic relationships: From classical mereology to complex part-whole relations. In R. Green, C.A. Bean, & S.H. Myaeng (Eds.), The semantics of relationships (pp. 35–50). Dordrecht, the Netherlands: Kluwer.

Priss, U. (2006). Formal concept analysis in information science. Annual Review of Information Science and Technology, 40, 521–543.

Reimer, U. (1991). Einführung in die Wissensrepräsentation [Introduction to knowledge representation]. Stuttgart, Germany: Teubner.

Rosch, E. (1975a). Cognitive representations of semantic categories. Journal of Experimental Psychology—General, 104(3), 192–233.

Rosch, E. (1975b). Cognitive reference points. Cognitive Psychology, 7(4), 532–547.

Rosch, E. (1983). Prototype classification and logical classification: The two systems. In E.K. Scholnick (Ed.), New trends in conceptual representation. Challenges to Piaget's theory? (pp. 73–86). Hillsdale, NJ: Lawrence Erlbaum.

Rosch, E., & Mervis, C.B. (1975). Family resemblances: Studies in the internal structure of categories. Cognitive Psychology, 7(4), 573–605.

Rosch, E., Mervis, C.B., Gray, W.D., Johnson, D.M., & Boyes-Braem, P. (1976). Basic objects in natural categories. Cognitive Psychology, 8(3), 382–439.

Saussure, F. de (2005). Course in general linguistics. New York: McGraw-Hill. (First published 1916).

Schmidt, S.J. (1969). Bedeutung und Begriff [Meaning and concept]. Braunschweig, Germany: Vieweg.

Schmidt, S., & Stock, W.G. (2009). Collective indexing of emotions in images. A study in emotional information retrieval. Journal of the American Society for Information Science and Technology, 60(5), 863–876.

Schmitz-Esser, W. (2000). EXPO-INFO 2000. Visuelles Besucherinformationssystem für Weltausstellungen [Visual Visitor Information System for World Exhibitions]. Berlin, Germany: Springer.

Shadbolt, N., Hall, W., & Berners-Lee, T. (2006). The semantic web revisited. IEEE Intelligent Systems, 21(3), 96–101.

Shaw, R. (2009). From facts to judgments. Theorizing history for information science. Bulletin of the American Society for Information Science and Technology, 36(2), 13–18.

Shaw, R. (2010). The author's response. Bulletin of the American Society for Information Science and Technology, 36(3), 3–4.

Simons, P. (1987). Parts. A study in ontology. Oxford, United Kingdom: Clarendon.

Slavic, A. (2008). Faceted classification. Management and use. Axiomathes, 18, 257–271.

Smith, B., Williams, J., & Schulze-Kremer, S. (2003). The ontology of the gene ontology. In the Proceedings of the American Medical Informatics Association (AMIA '03) (pp. 609–613). Bethesda, MD: American Medical Informatics Association. Retrieved May 19, 2010, from http://ontology.buffalo.edu/medo/Gene_Ontology.pdf

Smith, B., Ceusters, W., Klagges, B., Köhler, J., Kumar, A., Lomax, J., et al. (2005). Relations in biomedical ontologies. Genome Biology 6(5), Article R46.

Spiteri, L.F. (1999). The essential elements of faceted thesauri. Cataloging & Classification Quarterly, 28(4), 31–52.

Spiteri, L.F. (2008, October). Concept theory and the role of conceptual coherence in assessments of similarity. Paper presented at the 71st Annual Meeting of the American Society for Information Science and Technology: People Transforming Information – Information Transforming People, Columbus, OH.

Spiteri, L.F. (2010). Incorporating facets into social tagging applications. An analysis of current trends. Cataloging & Classification Quarterly, 48(1), 94–109.

Stock, W.G. (2007). Information retrieval. Informationen suchen und finden [Information retrieval. Searching and finding information]. München, Germany: Oldenbourg.

Stock, W.G. (2009). Begriffe und semantische Relationen in der Wissensrepräsentation [Concepts and semantic relations in knowledge representation]. Information – Wissenschaft und Praxis, 60(8), 403–420.

Stock, W.G., & Stock, M. (2008). Wissensrepräsentation. Informationen auswerten und bereitstellen [Knowledge representation. Indexing and providing information]. München, Germany: Oldenbourg.

Storey, V.C. (1993). Understanding semantic relationships. VLDB Journal, 2, 455–488.

Surowiecki, J. (2005). The wisdom of crowds. Why the many are smarter than the few and how collective wisdom shapes business, economies, societies, and nations. New York: Anchor Books.

Szostak, R. (2010). Comment on Hjørland's concept theory. Journal of the American Society for Information Science and Technology, 61(5), 1076–1077.

Torres-Vargas, G.A. (2005). World brain and *Mundaneum*. The ideas of Wells and Otlet concerning universal access. VINE. The Journal of Information & Knowledge Management Systems, 35(3), 156–165.

Tudhope, D., & Binding, C. (2008). Faceted thesauri. Axiomathes, 18, 211–222.

Vickery, B. (2008). Faceted classification for the Web. Axiomathes, 18, 145–160.

Weller, K. (2010). Knowledge representation on the social semantic Web. Berlin, Germany: De Gruyter Saur.

Weller, K., & Stock, W.G. (2008). Transitive meronymy. Automatic concept-based query expansion using weighted transitive part-whole-relations. Information – Wissenschaft und Praxis, 59(3), 165–170.

Wersig, G. (1974). Information – Kommunikation – Dokumentation [Information – communication – documentation]. Darmstadt, Germany: Wissenschaftliche Buchgesellschaft.

Winston, M.E., Chaffin, R., & Herrmann, D. (1987). A taxonomy of part-whole relations. Cognitive Science, 11, 417–444.

Wittgenstein, L. (2008). Philosophical investigations, 3rd Ed. Oxford, UK: Blackwell. (First published 1953).

DOI: 10.1002/asi