Classification

by Birger Hjørland

Table of contents:

Abstract:
This article presents and discusses definitions of the term *classification* and the related concepts *concept*, *categorization*, *ordering*, *taxonomy* and *typology*. It further presents and discusses theories of classification including the influences of Aristotle and Wittgenstein. It presents different views on forming classes, including logical division, numerical taxonomy, historical classification, hermeneutical and pragmatic/critical views. Finally, issues related to artificial versus natural classification and taxonomic monism versus taxonomic pluralism are briefly presented and discussed.

1. Introduction

This article is about *classification* as a basic term in an interdisciplinary perspective. Classification is a fundamental concept and activity in knowledge organization, but it is also an important concept in many other fields, including biology and philosophy. In knowledge organization and library and information science (LIS), it is mostly about classifying documents, document representations, and concepts (e.g., in thesauri), and library classification systems and ontologies are well-known kinds of knowledge organization systems (KOS). These activities and systems are based on more fundamental conceptions and theories of classifications that are presented in this article.

The ISKO Encyclopedia of Knowledge Organization (IEKO) plans to cover a very broad spectrum of articles related to classification besides the present one. We already have an article about → logical division, → notation, → automatic subject indexing, and further articles are planned about, for example, library classification, numerical taxonomy, classification of the sciences, classification in specific domains (including biology, physics and chemistry), and much more.

This article covers a very complex concept and is therefore highly compressed and abbreviated. In particular, most of the theories mentioned in section 4 deserve to be enlarged on in independent articles that it is hoped will be forthcoming at some point in the future.

2. The meaning of the word *classification*

Among the many contributors to the definition of *classification* two (Frederik Suppe and Henry E. Bliss) are here selected as outstanding. Frederick Suppe distinguished two senses of *classification*: a broad and a narrow meaning. He called the broad meaning "conceptual classification".

Classification is intrinsic to the use of language, hence to most if not all communication. Whenever we use nominative phrases we are classifying the designated subject as being importantly similar to other entities bearing the same designation; that is, we classify them together. Similarly the use of predicative phrases classifies actions or properties as being of a particular kind. We call this *conceptual* classification, since it refers to the classification involved in conceptualizing our experiences and surroundings" (Suppe 1989, 292).

Classification in the narrower meaning Suppe called "systematic classification".

A second, narrower sense of classification is the *systematic* classification involved in the design and utilization of taxonomic schemes such as the biological classification of animals and plants by genus and species (Suppe 1989, 292).

Henry E. Bliss (1929) also considered the senses of the word *classification* and wrote: "[t]his term, like other English derivatives ending in *ion*, is ambiguously used both in the predicative and in the substantive sense, now for the action and now for the act, sometimes for the process and sometimes for the product" (142). In order to remove this ambiguity, he suggested three definitions proceeding from *class* as a substantive:

1. The verb *to class* denotes likening, referring, or assigning a thing to some class, or several things to their respective classes, as may be requisite or relevant to interest involved. This verb is used not only transitively, but sometimes intransitively. Thus it may be said that olive oil *classes* as a luxury.
2. The verb *classify* means primarily to make, or conceive, a class, or classes, from a plurality of things, and secondary to arrange classes in some order or to relate them in some system according to some principle or conception, purpose or interest [...]. These three processes, classing, forming classes, and arranging classes, are so implicated that it is not easy to separate them in thought or in terminology; yet we propose here that this should be done as conductive to precision in this study. There is an important distinction between assigning a thing, or things, to some class or classes, and arranging classes in some order or system. [...]
3. A classification is a series or system of classes arranged in some order according to some principles or conception, purpose or interest, or some combination of such (Bliss 1929, 142-143).

There are many more definitions of classification than the ones given above. For a chronological sample of definitions of classification, see the Appendix.

The objects we classify may be physical objects, persons, processes, ideas, concepts, words, etc. Some of these entities, such as concepts, may be both the elements classified and a result of a (new) classification.

Table 1: Selected terms used about the units classified and the resulting groups

What may be classified	**Results of classification may be termed:**
Concepts	Categories
Documents	Clades
Elements	Classes
Entities	Concepts
Ideas (including fictional ideas)	Genera
Individuals	Groups
Items	Kinds
Objects	Sets
Phenomena	Sorts
Processes	Species
Sciences	Taxa
Things	Etc.
Etc.	

The objects to be classified have attributes with values. Attributes may, for example, be color or weight. Values may be red or heavy. Classifications are made by considering different attributes and their values. [1]

In conclusion: *Classification* is a term used both about the process to classify (which is a kind of discriminative practice; see Schmidt and Wagner 2004, 392) and about the resulting set of classes, as well as the assignment of elements to pre-established classes. The wide meaning of *classification* is the process of distinguishing and distribution kinds of "things" into different groups. All narrower meanings of *classification* are based on the wide definition but

add some extra requirements or restrictions put to the classification process and the resulting classification system — for example, the requirement that a classification should use only one criterion of division at a time, that classes should be mutually exclusive, and jointly exhaustive, are requirements demanded by some specific theories of classification, but not requirements that are common for all kinds of classification as here defined.

[top of entry]

3. Related terms

3.1 Concept/conceptualization

Concept has formerly been defined the following way:

Concepts are dynamically constructed and collectively negotiated meanings that classify the world according to interests and theories. Concepts and their development cannot be understood in isolation from the interests and theories that motivated their construction, and, in general, we should expect competing conceptions and concepts to be at play in all domains at all times" (Hjørland 2009, 1522-1523).

There is a close relationship — if not total identity — between theories of classification and theories of concept. The class of `waterfowl`, for example, includes the subclasses `ducks`, `geese`, and `swans`, in exactly the same way that the concept "waterfowl" includes the subordinate concepts of "ducks", "geese", and "swans". The different theories of how we classify birds correspond to the theories of how we conceptualize birds (see Andersen et al. 2006, 19-33). Henry Bliss also emphasized this:

It is evident that a discussion of classes involves the correlation of classes to concepts, or class-concepts. The class-concept is the *mental correlate of the class*, the mental basis both of the general idea of the class and of its name, or names (Bliss 1929, 120; italics in original).

Spiteri (2008) found that an examination of traditional similarity-based concept theories suggests that they do not provide an adequate account of conceptual coherence. Library and information science needs to explore knowledge-based approaches to concept formation, which suggest that one's knowledge of a concept includes not just a representation of its features, but also an explicit representation of the causal mechanisms that people believe link those features to form a coherent whole. Spiteri (2008) found that rather than representing a universal truth based upon unitary descriptions of concepts, classification systems represent only particular points of view. She hereby supports the argument of Hjørland and Albrechtsen (1999) and Beghtol (2003) that classification research must be situated within specific contexts and the domains in which the classification systems are designed to function, as well as Mai's (2004, 41) claim that "Any classification is relative in the sense that no classification can be argued to be a representation of the true structure of knowledge [...] a classification is merely one particular explanation of the relationships in a given field that satisfies a group of people at a certain point in time".

Frické (2012, 33), however, is opposed to considering concepts as mental constructs, and writes that the word *concepts* "amounts roughly to "general notion" or "general idea" or even "meaning". Many describe concepts as being mental or mental constructions; however, we regard them as abstractions or abstract objects (in the standard Fregean third realm)."

In spite of this close connection between classification and concepts, the discourses on concepts and the discourses about classification seem mostly to be separated in the literature.

[top of entry]

3.2 Categorization

Elin K. Jacob found that classification and categorization are different processes:

Although systems of classification and categorization are both mechanisms for establishing order through the grouping of related phenomena, fundamental differences between them influence how that order is effected — differences that do make a difference in the information contexts established by each of these systems. While traditional classification is rigorous in that it mandates that an entity either is or is not a member of a particular class, the process of categorization is flexible and creative and draws nonbinding associations between entities — associations that are based not on a set of predetermined principles but on the simple recognition of similarities that exist across a set of entities. Classification divides a universe of entities into an arbitrary system of mutually exclusive and nonoverlapping classes that are arranged within the conceptual context established by a set of established principles. The fact that neither the context nor the composition of these classes varies is the basis for the stability of reference provided by a system of classification. In contrast, categorization divides the world of experience into groups or categories whose members bear some immediate similarity within a given context. That this context may vary — and with it the composition of the category — is the basis for both the flexibility and the power of cognitive categorization" (Jacob 2004, 527-528).

Jacob's distinction is based on the narrow meaning of classification presented in section 2. Her distinction is supported by Schmidt and Wagner (2004), who introduced some distinctions between classification and other forms of discriminatory practice:

The point we want to make is that we have to be quite specific in distinguishing different types of discriminative practice: seeing something, seeing something for what it is as opposed to something else (reflecting on what one is seeing), physically separating things in some regular way, saying that x is C ("categorizing" x as C), and classifying x as C according to an inscribed, publicly available classification system. These are radically different practices, involving radically different forms of convention, principles of abstraction, etc. (Schmidt and Wagner 2004, 45-46).

About categorization, the same authors wrote:

Categorization, by contrast [to seeing and recognizing], is a linguistic operation of ascribing a category or concept to a particular phenomenon by the means of signs. Merely talking about phenomena, however, is not necessarily categorizing them, although talking involves the application of concepts. To categorize is to make a conceptual proposition ("red is a color"). In categorizing what you see as trees and birds you emphasize certain aspects of the world while abstracting from others, for instance that the trees and birds may all have green colors or that clouds and leaves may all be moved by the wind. An act of categorization cuts the world into pieces in that it emphasizes certain features at the expense of others ("x belongs to category C").
(In themselves acts of separating objects are not acts of categorization, as they are not

necessarily linguistic operations. Peeling onions or removing dirt from one's body by means of soap and water are acts of separation but not acts of categorization, although they may be subjected to acts of categorization, for instance when one is instructing children in how to do it. Similarly, when sorting the garbage (putting paper in this container, potato peels in that container) one may, or may not, be following instructions involving categorizations)" (Schmidt and Wagner 2004, 391-392).

These two sources pose a question about the broad definition of classification mentioned in section 2 of the present article. Nevertheless, this broad definition is widely used in the literature and it will introduce problems to restrict the term *classification* to the narrow definition. Therefore, the choice made here is to consider classification as synonymous with categorization but to maintain the distinction between classification in a wide and a narrow sense.

[top of entry]

3.3 Ordering

WordNet 3.1 provides two senses of the noun *ordering*:

- ordering, order, ordination (logical or comprehensible arrangement of separate elements): "we shall consider these questions in the inverse order of their presentation"
- order, ordering (the act of putting things in a sequential arrangement): "there were mistakes in the ordering of items on the list"

Some authors do not consider historicist classifications (like cladistics systems) as following the concept of classifications:

For several centuries all ordering systems were thought to be classifications and the two terms were treated virtually as synonyms. Eventually, however, it was realized that classification means making classes and that ordering systems that are not based on classes, such as sequential listing or cladifications (Mayr 1995), are not classifications. Hence, ordering systems denotes the general concept that includes classification as one of its subdivisions. (Mayr and Bock 2002, 172)

Instead, Mayr and Bock suggest that cladistics systems should be considered as ordering systems in a broader category. However, this terminology is not generally used, and it is deviant from the suggestions made in the present article.

Ordering depends on conceptual classification but it is broader than systematic classification. Books can be ordered by, for example, by size, language, or publication date, or alphabetically by author or title.

For further information about order and ordering systems, see Meinhardt et al. (1984), Mayr (1995), Mayr and Bock (2002) and Schmidt and Wagner (2004).

[top of entry]

3.4 Taxonomy

The term *taxonomy* originates from Greek *taxis* (meaning "order" or "arrangement") and *nomos* (meaning "law" or "science"). It was first used in 1813 by the French naturalist Augustin-Pyramus de Candolle (Candolle 1813). It became widely used in relation to biological classification but has since spread to other domains. Some authors consider it synonymous with *classification*, whereas others make distinctions between the two terms. In the following quote, the two terms are considered synonymous:

[T]axonomy has acquired a wide range of meanings no longer restricted to the classical understanding of biology. Taxonomy is now applying its early sense of organizing things in accord with particular principles ("taxis": arrangement; "nomos": law) to a broader range of domains after several centuries of limitation to biology and other natural sciences. In the 1990s, taxonomy was redefined as any semantically significant, systematic organization of content or as the process of developing such organization. This definition sometimes includes any collection whose individual elements have been assigned to various nodes of a classification system. Thus, taxonomy is sometimes considered the process of matching collection items with predefined labels, and sometimes it is the creation and arrangement, as well as the resulting product, of the classification system itself (Grove 2010, 5139; references omitted).

An example of a distinction between classification and taxonomy is:

'Taxonomy' is generally considered a synonym of 'systematic' and is traditionally divided into classification (the orderly arrangement of organisms into taxonomic groups on the basis of similarity), nomenclature (the labeling of the units), and identification (the process of determining whether an unknown belongs to one of the units identified). (Vandamme 2015, 255).

Adams and Adams (1991, 202-3):

The term taxonomy, like other words relating to classification, has no generally accepted, precise definition. For many systematists, especially in the biological sciences, it is synonymous with classification itself [references here omitted]. Our usage, however, is more restricted. We apply the term only to classificatory systems having an explicit hierarchic feature; that is, systems in which basic types are either clustered into larger groups or split into smaller ones, or both.
Most of the time, taxonomic ordering is simply the classification of classes: the clustering of basic types into larger and more inclusive units on the basis of some but not all of their attributes […]
The classic example of a taxonomy is the Linnean system of biological classification, in which species are clustered into genera, genera into families, and so on up to the level of biological classes.

Hedden (2016) also uses the term *taxonomy* in a very broad sense, not just about classifications (hierarchical or non-hierarchical) but as a synonym for any kind of knowledge organization system (KOS). However, one may ask, if the term *taxonomy* is not used with a specific meaning in relation to classification and KOS, why then use it at all? Hlava (2014) is another book that uses the word "taxonomies" in the title, but does not distinguish between "taxonomy" and "classification". In the glossary is written: "Taxonomy. As defined in ANSI/NISO Z39.19-2010R, 'A collection of controlled vocabulary terms organized into a hierarchical structure. Each term in a taxonomy is in one or more parent/child

(broader/narrower) relationships to other terms in the taxonomy'" (Hlava 2014, Vol. 1, 50). This way of using terminology is confusing; instead of trying to clarify terms and to aim at consensus in the use of terms for given concepts, it does the opposite.

ISO 25964-2: Information and Documentation-Thesauri and interoperability with other vocabularies, Part 2: Interoperability with other vocabularies (International Organization for Standardization 2013) contains a lengthy section 19 "Taxonomies" (p. 59-66), in which taxonomies are just defined (60) as "monohierarchical as well as polyhierachical classification of any subject matter" (i.e., a hierarchical classification). However, the same standard writes in section 3.10 (3) "Classes are the basic units of which a classification scheme is constructed. In a taxonomy, although this is a type of classification scheme, they are generally known as categories". Aside from demanding that a taxonomy is hierarchical, no defining characteristics are given in this standard.

Dahlberg's Classification System for the Literature on Knowledge Organization has class 5 "On Special Objects Classification Systems (Taxonomies)", as opposed to class 6 "On Special Subjects Classification Systems" (Dahlberg 1993, 219). Similarly, according to Aida Slavic the difference between classifications and taxonomies is based on the distinction between aspect classification (or "disciplinary classification") on one side and entity classification (or "phenomenon classification") on the other.

Knowledge classification can be, and often is, TAXONOMIC (sometimes called "entity classification") like the classification of zoology, classification of plants, or classification of chemical elements (which means that they are going to list one concept in one place only in the classification structure).
Bibliographic classifications i.e. those one has to use to describe real documents ARE NOT and CAN NOT be taxonomic. They are by all means ASPECT or disciplinary classifications. This means that they will list one concept in all disciplines and fields where that concept might be studied: e.g. "water" will have to appear under chemistry, physics, in geology, medicine, sport etc.
This is of critical importance for information retrieval as aspect classification helps to establish the context in which one concept or phenomenon might be studied within the document" (Slavic 2000; electronic source, no pages).

This quote from Slavic is, however, contradicted by other uses of the terminology. Bibliographical classifications may be phenomenon classifications (for example, the system by James Duff Brown (1862-1914); cf. Beghtol 2004; Gnoli 2016) and disciplinary-based library classification systems like the Dewey Decimal Classification (DDC) are also sometimes termed "taxonomies" (see Waltinger et al. 2011).

Carl E. Landweh et al. found that a taxonomy is based on a theory.

A taxonomy is not simply a neutral structure for categorizing specimens. It implicitly embodies a theory of the universe from which those specimens are drawn. It defines what data are to be recorded and how like and unlike specimens are to be distinguished. In creating a taxonomy of computer program security flaws, we are in this way creating a theory of such flaws, and if we seek answers to particular questions from a collection of flaw instances, we must organize the taxonomy accordingly (Landweh et al. 1994, 214).

However, classifications, too, are based on theories (and an atheoretical classification or taxonomy may be considered an oxymoron; see Hjørland 2016b). Therefore, the theoretical basis cannot be used as a criterion for distinguishing classification and taxonomy.

Marradi suggested the following distinctions:

A taxonomy obtains when several fundamenta divisionis [criteria of division] are considered in succession, rather than simultaneously, by an intensional cl. [classification]. The order in which fundamenta are considered is highly relevant: the taxonomy obtained by using property X to classify a genus and then property Y to classify its species is by no means the same as that obtained by considering property Y first and property X afterwards (Marradi 1990, 146).

Marradi suggested the following differences between classifications, typologies, and taxonomies as products:

1. Classification schemes
   When only one fundamentum divisionis is considered, a classification scheme is produced — usually by an intensional classification. The extensions of each class must be mutually exclusive, and jointly exhaustive. Classes need not be at the same level of generality, and may be ordered (Marradi 1990, 129).
2. Typologies
   When several fundamenta are jointly considered, a typology is produced. This may be done through either intensional or extensional classification (Marradi 1990, 129).
3. Taxonomies
   When several fundamenta are considered in succession through a series of intensional classifications, a taxonomy is produced. Specific concepts/terms (such as taxon, rank, clade) are needed to deal with taxonomies (Marradi 1990, 129).

It is not difficult to find examples of the use of the terms *classification*, *typology*, and *taxonomy* in disagreement with Marradi's definitions. Here, it will not be discussed whether or not it is a good idea to use his definitions prescriptively.

[top of entry]

3.5 Typology

Typology is derived from the two Greek words *typo* meaning "type" and *logos* meaning "word." The word *typology* literally means the study of types (subdivisions of particular kinds of things). We saw above (3.4) that according to Marradi (1990) typologies are kinds of classifications in which more than one fundamental criterion of division is simultaneously taken into account. Another definition was provided by Kenneth D. Bailey:

Typology is another term for a classification. Two characteristics distinguish typologies from generic classifications. A typology is generally *multidimensional* and *conceptual*. Typologies generally are characterized by labels or names in their cells. (Bailey 1994, 4; italics in original)

Bailey exemplifies:

As a hypothetical example, let us use two dimensions to construct a classification. These dimensions are intelligence (dichotomized as intelligent/unintelligent) and motivation (dichotomized as motivated/unmotivated). Combining these two dimensions creates a fourfold typology; as shown in Table 1.1. These four categories can be defined as *cells* in the table. In this case, they are *types*, or *type concepts*. A motivated and intelligent person can be labeled as successful; an intelligent but unmotivated person is likely to be an underachiever; while a motivated but unintelligent person is an overachiever; and one who lacks both intelligence and motivation is likely doomed to failure. (Bailey 1994, 4; italics in original).

A hypothetical fourfold typology (after Bailey 1994, 5; Table 1.1):

	Motivated	**Unmotivated**
Intelligent	Success, 1	Underachiever, 2
Unintelligent	Overachiever, 3	Failure, 4

The term *typology* is used in many fields. For example are Carl G. Jung's psychological types famous (Jung 1971). In Library and Information Science (LIS) is *typology* used, for example about document typologies. Web of Science, for example, distinguishes between article, book review, letter, review, proceeding paper and other types of documents.

[top of entry]

3.6 Division

→ Logical division is sometimes seen as a concept closely related to classification, but is here understood as one among other approaches to classification, more precisely, logical division is an a-priori, rationalist approach to classification. Parry and Hacker (1991, 136-7):

Logical division and classification are closely related. In logical division, one starts with a class (genus) and divide it into subclasses (but not into individuals). In classification, one starts with individuals or classes, and groups them into classes on the basis of properties they share. Like many abstract words ending in "-tion", "division" and "classification" are used both for a process and for the product resulting from the process — in this case, a mental or verbal activity, namely a hierarchical arrangement of concepts or terms. In the sense of such an arrangement, every table that results from logical division may also be regarded as if it were the result of classification.
As a process of knowledge, division is characteristically an a priori process based on meaning. Classification, however, is primarily inductive, and may require empirical knowledge. In less technical language, this means that a table of pure division can be constructed solely from the meaning of terms in the table; no recourse to experience is necessary. I know, for example, that squares are a subclass of parallelograms, not from experience but from the meanings of the terms "square" and "parallelogram". On the other hand, if I engage in the process of classifying plants, I would have to observe the similarities and dissimilarities of actual plants. I could not construct a useful classification of plants unless I had observed the plants I wished to classify.
In the natural sciences, classification is far more important than division, since the principal analysis and relationships are based on empirical data. Division plays an important tole in the formal sciences of mathematics and logic, and in the empirical areas to which these

disciplines can be applied. But the formal scientist often uses methods partly inductive to develop the knowledge that he will later expound deductively.

[top of entry]

4. Theories of classification [2]

In this section, the following theories are presented.

- 4.1: The "classical view" of classes as defined by sets of elements with necessary and sufficient attributes versus the views proposed by, derived from, or related to Ludwig Wittgenstein's late philosophy. As opposed to necessary and sufficient attributes, these views consider that classes and concepts are graded structures. This section briefly presents the prototype theory suggested by Eleanor Rosch, as well as theories developed by Thomas Kuhn and Michael Billig.
- 4.2: The way of forming classes (e.g., by logical division, by measuring similarity among elements, by collecting elements with a common ancestry, or by collocating tools to support human activities) (the epistemology of classification). [3]
- 4.3: The view that there is one correct or best classification versus the view that there are different classifications for different purposes (the metaphysics of classification).

[top of entry]

4.1 The "classical view" versus "prototype theory" [4]

Aristotle developed a theory of classification in which all elements in a given class share at least one characteristic with all other members. Classes should be designed so membership of a class is given by a set of necessary and sufficient characteristics. For example, according to Aristotle's *per genus et differentiam* definition, man is a rational animal. This definition first considers a class or concept consisting of all animals (including humans). It then claims the essential difference between humans (men) and all other animals is that humans are rational. In this way, the class of `animals` is divided into two non-overlapping classes: `rational animals` (humans) and `non-rational animals` (all other animals). In order for some organism to belong to the class of humans, it is *necessary* that it is rational: all elements in the class *must* have this characteristic. It is also a *sufficient* condition: if an organism has the attributes of all other animals plus the attribute of being rational, it must be human (it is quite a different task to find out *if* a given organism is rational or not). This has also been called monothetic classification [5] or the feature theory of classification (for further information about the classical view see, for example, Parry and Hacker 1991; Moss 1964; Frické 2016).

In the middle of the twentieth century, a new theory of classification was proposed, which considered itself an alternative to the feature theory that had existed for millennia. Wittgenstein (1953) claimed that not all concepts consist of elements which have a set of necessary or sufficient characteristics. He used the metaphor of family resemblance for a series of overlapping similarities, where no one feature is common to all the elements in the concept. In a family, some members may be alike in one respect (e.g., the shape of the nose) while other members may be alike in other respects. Wittgenstein's famous example is games, where he claimed that no characteristic common to all kind of games exists. Experimental psychologist Eleanor Rosch (1978), inspired by Ludwig Wittgenstein's later philosophy, first

defined prototype theory [6]. Given the Roschian theory, some elements are better representatives than others. For example, if the class or concept is bird, in the classical theory this concept may be defined by attributes such as feathers, beak, and the ability to fly, and every bird is as representative as any other. In the prototype theory, on the other hand, a blackbird is considered a good example (at the least by Westerners), while a penguin is considered a bad example. Instead of being defined by necessary or sufficient characteristics, classes are determined by the overall likeness to a prototype — hence the name of the theory.

A now well-established division between two theories of concepts and classification is therefore classical or Aristotelian classification on the one side and prototype theory (or polythetic classification) on the other. George Lakoff wrote about these two theories:

From the time of Aristotle to the later work of Wittgenstein, categories [and classification] were thought to be well understood and unproblematic. They were assumed to be abstract containers, with things either inside or outside the category. Things were assumed to be in the same category if and only if they had certain properties in common. And the properties they had in common were taken as defining the category.
This classical theory was not the result of empirical study. It was not even the subject of major debate. It was a philosophical position arrived at on the basis of a priori speculation. Over the centuries it simply became part of the background assumptions taken for granted in most scholarly disciplines. In fact, until very recently, the classical theory of categories was not even thought of as a *theory*. It was taught in most disciplines not as an empirical hypothesis but as an unquestionable, definitional truth" [7] (Lakoff 1987, 6; emphasis in original).

Geoffrey C. Bowker provided another description:

An Aristotelian classification works according to a set of binary characteristics, which the object being classified either presents or does not present. At each level of classification, enough binary features are adduced to place any member of a given population into one, and only one class. [...] Aristotelian models [...] have traditionally informed formal classification theory in a broad range of sciences, including biological systematics, geology, and physics. Rosch's (1978) prototype theory argues that, in daily life, our classifications tend to be much fuzzier than we might at first think. We do not deal with a set of binary characteristics when we decide that this thing we are sitting on is a chair. Indeed, it is possible to name a population of objects that people would in general agree to call chairs that have no two binary features in common" (Bowker 1998, 256).

There are still many people subscribing to the classical view, and Wittgenstein's view — and thereby, indirectly, prototype theory — is criticized by Needham (1975), Sutcliffe (1993) and Margolis (1994). See Fox (2011) for a recent discussion of prototype theory in knowledge organization.

Thomas Kuhn [8] is well known for his book *The structure of scientific revolutions* (1962) in which he made the concepts of "scientific paradigm" and "paradigm shift" widespread — including in common language. It is less well known that his later research focused on concepts, classification, and scientific taxonomies and inspired a new theory of concepts called "theory theory". The best introduction to this work is probably that by Andersen et al. (2006), who wrote:

According to this theory, the basic conceptual structure of science is a classification system that divides objects into groups according to similarity relations. The grouping is not determined by identifying necessary and sufficient conditions, but by learning to identify similarities and dissimilarities between the objects. It was one of Kuhn's central claims that one learns such concepts by being guided through a series of encounters with objects that highlight the relations of similarity and dissimilarity currently accepted by a particular community of concept users. [...].
Kuhn's standard example of a learning process of this sort is a child learning the concepts "duck", "goose", and "swan" (Kuhn 1974). In this example, an adult familiar with the classification of waterfowl guides a child ("Johnny") through a series of ostensive acts until he learns to distinguish ducks, geese, and swans. Johnny is shown various instances of all three concepts, being told for each instance whether it is a duck, a goose, or a swan (Andersen et al. 2006, 20-21).

The most important aspects of Kuhn's theory are:

- People learn concepts (or classifications) according to how these concepts are understood in a given society by being confronted with exemplars and similarities as well as dissimilarities compared with other concepts (e.g., by parents and teachers).
- Two people can correctly identify the same concepts/classes even if they use different characteristics to make the correct identification.
- Dissimilarity plays as important a role as similarity in classification. Similarity alone is not enough (see Andersen et al. 2006, 24ff).
- A given concept/classification is based on a paradigm. For example:
    - Paradigm one: Ptolemaic astronomers might learn the concepts "star" and "planet" by having the Sun, the Moon, and Mars pointed out as instances of the concept "planet" and some fixed stars as instances of the concept "star".
    - Paradigm two: Copernicans might learn the concepts "star", "planet", and "satellites" by having Mars and Jupiter pointed out as instances of the concept "planet", the Moon as an instance of the concept "satellite", and the Sun and some fixed stars as instances of the concept "star". Thus, the concepts "star", "planet", and "satellite" got a new meaning and astronomy got a new classification of celestial bodies.

    The difference before Copernicus, and later, say, Newton, is striking: after a paradigm shift we learn to distinguish concepts in new ways (see Andersen et al. 1996).

A contemporary example: Ornithologists have recently discovered that the blackbird, which so far has been considered one species, should be considered different species which have becoming similar by adapting to the same ecological niche. In the future, we may therefore expect that not just experts, but also amateur ornithologists, schoolchildren, and the rest of us learn to classify birds in a new way.

Michael Billig, a social psychologist inspired by rhetoric, proposed another new theory that is probably in harmony with Kuhn's view in important respects. Billig considers that thinking is like a quiet internal argument. Therefore, psychological and rhetorical theories are closely linked, and the psychology of classification/categorization can learn much from rhetoric. Billig (1996) describes modern cognitive psychology's tendency to consider categorization a fundamental cognitive process in both animals and humans. Modern cognitive psychology tends to view "[t]he individual as an active processor of information" in which "the effect of a

stimulus depends on how it is categorized and interpreted by the perceiver" (quotations from Eiser 1980, 8). This basic psychological process is often attributed a biological status by cognitive psychologists (this criticism is also raised against Eleanor Rosch). Billig finds that the implication of this view is that humans are tied to prejudiced and bureaucratic modes of thinking. Billig does not consider it wrong that categorization is an important process, but from his studies in rhetoric he argues that there must be two fundamental processes: categorization and particularization, the latter being a reverse process in which something is not just considered an element of a class or a category, but is considered something special. Billig further demonstrates with many examples how humans are able not just to categorize and particularize but also to discuss and consider the way things are categorized. Arguments about categories and particulars are important elements in human communication and thinking, and they are often related to wide-ranging theoretical and ideological issues and conflicts. By including particularization as a basic psychological process, Billig is able to make room for people, who are not just prejudiced and bureaucratic but also open-minded and flexible. In relation to research in knowledge organization, Billig's research raises serious problems for the cognitive view that tries to base classifications on the study of the human mind.

The basic lessons from these new theories of classification may be summarized this way:

1. Humans do not classify in a given way according to inborn or "given" characteristics, but according to the human activities and goals that we have (which may require different classifications).
2. Instead of the classical model of sets of mutually exclusive and jointly exhaustive criteria, we may need alternative models. Andersen et al. (2006) found that the so-called "dynamic frames" represent the best way to represent classes. It has not been examined, however, whether this form of knowledge representation is applicable or fruitful in relation to the construction of knowledge organization systems (KOS) in LIS.

[top of entry]

4.2 The methodology of forming classes, the epistemology of classification

The fundamental elements of any classification are its theoretical commitments, basic units and the criteria for ordering these basic units into a classification. (Hull 1998)

4.2a Elaine Svenonius (2004) proposed that three epistemological theories are important for knowledge organization (or, as she preferred, knowledge representation):

- α Operationalism
- β The picture theory of meaning
- γ The contextual or instrumental theory of meaning

These views may deserve their own entry in this encyclopedia. They are outlined below.

Concerning α, Svenonius raised the criticism that all operational definitions lack validity and that operationalism represents a form of logical positivism.

Concerning β, what Svenonius termed "the referential or picture theory of meaning", she found that this, too, "derives from an empiricist view of knowledge" (2004, 574). The author summarized the basic problems with this theory:

First, the picture theory assumes a universal form of language in which the meaning of propositions picturing the world are prescribed, relatively fixed, and generally understood. The objection here is that pictures can be differently interpreted. A cup is half full or half empty. A picture of a duck from another viewpoint could be a picture of a rabbit; a picture of a block could be interpreted as a triangular prism.
Secondly, the picture theory implies fixity of reference. But the meanings of words are not necessarily fixed in the sense of referring to sets of homogeneous objects in the real world or clearly delineated mental concepts. Many words have fluid boundaries. (A chair with three legs is still a chair.) Fluidity is necessary if words are to function in a variety of different contexts. The picture theory falls down particularly in the case of abstract words whose referents are mental constructs and function words, such as adverbial particles and prepositions.
A third problem with the picture theory is that it represents knowledge of the world as the conjunction of knowledge of independent microworlds. To regard the totality of knowledge as a simple aggregation is simplistic [...]" (Svenonius 2004, 578; note omitted).

Concerning γ, Svenonius's last theory was the contextual or instrumental theory of meaning. The basic tenet of the instrumental theory of meaning is that we know what a word means when we know how to use it. Svenonius found that this way of thinking led to adoption of the methods of numerical taxonomy. However, numerical taxonomy may be considered a form of empiricism rather than of the contextual or instrumental theory of meaning as developed by pragmatic philosophers. If this understanding is true, we may conclude that Svenonius has not really suggested an alternative to empiricism and logical positivism. Such alternatives are presented below (4.2c). Before this, however, another important view will be presented.

4.2b Alberto Marradi distinguished the following senses of classification as an operation (intellectual and otherwise):

- α Intensional classification (or subdivision or downwards classification) [9]
  "[T]he subdivision of the extension of a concept (*genus*) into several extensions corresponding to as many concepts of lower generality (*species*). The former and all the latter concepts have the same intension except for one aspect (*fundamentum divisionis*): on that account each species concept is a different partial articulation of the genus concept" (Marradi 1990, 129; emphasis in original).
- β Extensional classification (or numerical taxonomy or upwards classification) [10]
  "[T]he grouping of the objects/events of a set into several subsets according to the perceived similarities of their states on one or more properties" (Marradi 1990, 129).
  "[B]otanist Michel Adanson stated that "all parts and qualities, or properties and faculties of plants ... barring not even one" ought to be considered before attempting a classification (1763, clvi). Along with this idea, Adanson operated extensional [classification] and produced taxonomies based on the rate of equal states on the total of properties considered between any two plants (1763, vol. I)" (Marradi 1990, 136).
  "The predominant criterion is to maximize homogeneity within classes and heterogeneity between classes" (Marradi 1990, 135).
  "Other labels have been proposed for the operation, over and above the old ones — "classification" and "taxonomy". Among them "numerical taxonomy" (Sokal and

Sneath 1963), "class formation" (Capecchi and Möller 1968), "cluster analysis", etc." (Marradi 1990, 136).

- γ Classing
  "[T]he assignment of objects/events to classes defined by the first operation [subdivision] (or of new objects/events to groups created by the second operation with other objects/events)" (Marradi 1990, 129).

Marradi's two first options are further discussed in the next section. His third option is not related to a new fundamental method and is not further discussed in this article, but this issue is partly dealt with in the entry on Subject. Forthcoming entries on subject analysis, indexing, etc. will also cover this topic.

4.2c Birger Hjørland has suggested that there are four basic theories and approaches to classification [11]: α rationalism; β empiricism; γ historicism, and δ pragmatism/critical theory. All four will be presented and discussed below. The first two (rationalism and empiricism) are related to theories already presented.

- α Rationalism. Hjørland considers than the first of Marradi's operations, intensional classification or subdivision, corresponds to what he in different writings has related to rationalism:

  > Rationalist theories of indexing (such as Ranganathan's theory) suggest that subjects are constructed logically from a fundamental set of categories. The basic method of subject analysis is then "analytic-synthetic", to isolate a set of basic categories (=analysis) and then to construct the subject of any given document by combining those categories according to some rules (=synthesis). The application of rules such as logical division is by principle part of the rationalist view (Hjørland 2011, 74).

  According to Marradi, logical division dominated for centuries until challenged (or supplemented) by extensional (empiricist) classification.

  > In our opinion, this belated development [of empiricist classification] depends on the fact that, in order to be somehow formalized from a spontaneous activity into a respectable intellectual operation within a scientific discipline, extensional cl. [classification] had to wait for the development of another intellectual tool, viz. the idea of orderly recording the states of a vector of objects on a vector of properties—in other words, for the intellectual forefather of what is presently known as the data matrix (Marradi 1990, 135-136).

  What are the major benefits and drawbacks of logical division as a method of classifying? Frické (2016, 547) stated: "Logical division produces classifications with admirable qualities. Everything has a place in a leaf, its own unique place, and the classification schedule embodies the maximum amount of general information about the items being classified". Its weaknesses have been known for a long time: "Aristotle had argued that logical division was an inappropriate tool for the classification of organized beings" (Stevens 1998; electronic source, no pages). One limit of this method is that it seems better suited to some kinds of object (e.g., formal objects, such as mathematical objects) compared to other kinds of object (e.g., "organized beings"), but this is an open issue today. A modern criticism is its

relationship with essentialism. If the basis of division (fundamentum divisionis: classification principle) is not to be arbitrary, it has to be deduced from what are considered essential criteria, but this idea is heavily criticized today (Wilkins 2013; Frické 2016).

- β Empiricism. The second of Marradi's operations, extensional classification, has, according to Marradi, also been termed "numerical taxonomy" and "cluster analysis", among others. It corresponds to what Farradane (1950) termed "inductive classification" and Parrochia (2016) "phenomenal classifications", and has been related by Hjørland to empiricism:

> Empiricist theories of indexing are based on the idea that similar (informational) objects share a large number of properties. Objects may be classified according to those properties, but this should be based on neutral criteria, not on the selection of properties from theoretical points of view because this introduces a kind of subjective criteria, which is not approved by empiricism. Numerical statistical procedures are based on empiricist philosophy (Hjørland 2011, 74).

This means that overall likeliness, sometimes termed "phenetics" (e.g., statistical measures of similarity based on a great number of attributes), is the basis of this method. This seems at first to be properly scientific but, on further examination, it turns out to be based on a number of problematic assumptions. The first question is about the number of properties needed.

> A controversial question is how many properties are to be considered. Parsimony and elegance would advise to keep that number down; on the other hand, it may be remarked that "increasing the number of variables increases the probability of correct classification" (May 1982, 43). Since the concept of "correct [classification]" is rather questionable [...], a better argument might be that, with a formalized approach, there is no way to consider the information on the objects'/events' properties unless those properties are included in the matrix. Therefore, leaving properties out of the matrix entails a loss of information of unknown amount (Marradi 1990, 137).

If we assume that any object has an unlimited number of properties, then it is impossible to consider them all, even in theory. By implication, any empirical classification is biased in ways that cannot be controlled.

Another problem is that the descriptions of objects, on which this method is based, cannot be atheoretical or objective (cf. Hjørland 2016a, 2016b) — or, as formulated by Gitelman (2013), "Raw data is an oxymoron". In other words, the data used are always theoretically biased in ways we often cannot recognize or control.

The third problem, closely related to the second, concerns the concept of similarity. Classification has often been defined as bringing like things together (and thus separating unlike things). "Likeness" is a concept that may also be expressed by other terms such as "similarity", "sameness", "resemblance", or "equivalence". The problem is that things cannot be similar in an objective way. Any object is similar to another object in some ways and dissimilar in other ways. For any three objects, two different

classifications can be constructed which fulfils the demand of bringing the like objects together. Consider Figure 1 below, wherein the items may be classified according to color or shape. None of those properties is objectively more important than the other. For some purposes, the two squares are most alike and should be classified together. For other purposes, the two black figures (a square and a triangle) are most alike and should be classified together (see also Popper 1959, 441).

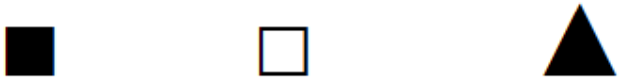

Figure 1: Classification criteria. Which two are similar?

By implication, empirical criteria for classification are not enough and we need some guidance on how to determine which criteria should be used when determining similarity. If we take a biological example, scientists have long recognized that modern-day birds and reptiles share a common ancestor. Both groups lay shelled eggs and have scales (in birds, confined to the legs), nucleated red blood cells, and a number of skeletal similarities. Different methods and paradigms in biological taxonomy thus arrive at different results: Methods based on historical development consider birds and reptiles as related groups (birds may be considered a kind of reptile), while numerical taxonomy, based on a quantitative study of many structural similarities, considers birds and reptiles to be less related (birds are not reptiles). The reason is that many animals develop similar structures because they adapt to the same ecological niches, but may have very different phylogenetic backgrounds and are by biologists considered different groups. (Considering the reverse, in some spiders the male and the female are very different, and if numerical taxonomy is used, the male and the female might be classified as different species, which obviously is problematic.) Therefore, modern biological taxonomy is dominated by quite a different philosophy and method: cladism. We shall return to this below. Consider, however, that Marradi saw these two methods (logical division and phenetics) as exhaustive, but that Hjørland (2009) presented two additional methods of classification related respectively to historicism and pragmatism/critical theory. The biological examples should provide sufficient argument for the first of these, and others have argued in a similar way. [12]

- γ Historicist approaches to classification. To say that two elements belong to the same class (or "clade") if they share a common ancestor is clearly different from defining membership of a class by similarity (sets of characteristics as arranged by logical division or numerical taxonomy). Today, this is the dominant approach in biological systematics. This approach (termed "cladistics", "phylogenetic classification", "historical classification", → "genealogical classification", or "genetic classification") is based on the historical or evolutionary development of the classified objects (Hennig 1966; Hjørland 2003, 107; Gnoli 2006). It is not only used in biology, but also, for example, for classification of languages and musical instruments — and should be considered one of four general approaches to classification. Hjørland (2013, 2016a) suggested genealogical classification may also be used in bibliometrics and information retrieval as an alternative to classifying documents according to similarity.

To define membership of classes, clades, or species by common ancestor is different from, but related to, an evolutionary ordering of classes (see Dousa 2009 for early discussion of evolutionary order in library classification).

Although cladism seems to dominate biological taxonomy today, it has also met with skepticism:

> [I]t is at the same time becoming clearer that there is very likely no such ideal classification. There is no reason why a classification that reflects the origins of the things classified should coincide exactly with one aimed at the ecological relations of those things, and it is increasingly perceived that these can and do diverge (Dupré 2002, chs. 3-4). This possibility becomes even clearer in view of the difficulties that are emerging in the project of evolutionary-based classification. Speciation was once seen as an all or nothing affair leading to complete isolation of one group from another. It is now clear that for micro-organisms, in particular, there is very little such isolation, and genetic material moves in many ways from one kind of organism to another. In fact it has become common to conceive of the genome of an ecosystem (the soil of an area, or a body of water) rather than the privatized genome of an individual organism (e.g. Venter et al. 2004). The classical picture of speciation applies quite well to some of the most complex multi-celled organisms, such as mammals and birds, though much less well to plants. An important movement in biology is to transcend the anthropocentrism that takes the peculiarities of our own corner of the living world as the model for all (Dupré 2006, 31).

Dupré's reservations are not about the validity of the cladistics approach as such, but about the idea of one ideal classification. It seems clear that genetic classification is a distinct approach, with some major benefits — to know about the origin of things is to know things in a deeper way than just to know about sets of attributes (which, in the historical perspective, often looks superficial).

There is also a subjective side of classification, and this subjectivity has developed historically. The classifying subject is influenced by his or her culture, paradigms, and tradition. This may be termed, for example, hermeneutics, historicism, or social epistemology. This can be fruitfully illustrated in the classification of colors:

> [W]hen the colour vocabularies of various languages are considered and compared, the researcher finds that there are many different ways in which humans categorize and "label" colours, resulting in an amazing array of misunderstandings. Monoglot individuals invariably believe that their own colour system is clear and obvious, and they are often mystified when confronted with an alternative system. So the first step which the reader has to take when entering the world of colour semantics is probably the most difficult of all; s/he must restrict his or her own colour system to normal, everyday speech, and learn to set it aside when considering foreign or historical colour descriptions. The aim is to dispose of any preconceptions about how colour "should" be classified and described, so as to gain insights into the workings of other languages and cultures, and into the nature of colour itself (Biggam 2015, 1).

Formerly, Hjørland wrote:

> [H]ermeneutical theories of indexing suggest that the subject of a given document is relative to a given discourse or domain and is why the indexing should reflect the need of a particular discourse or domain. According to hermeneutics, a document is always written and interpreted from a particular horizon [note omitted]. The same is the case with systems of knowledge organization and with all users searching such systems. Any question put to such a system is put from a particular horizon. All those horizons may be more or less in consensus or in conflict. To index a document is to try to contribute to the retrieval of "relevant" documents by knowing about those different horizons (Hjørland 2011, 74).

Historicism is therefore, as we have seen, an approach that may be applied to both the object and the subject in classification. If both the object and the subject are considered, we may speak of a united historicist theory. B. M. Kedrow presents such a united historicist view of classification:

> Historicism as a key to any natural classification.
> Of crucial importance for the analysis of the problem at hand is the historical approach to its consideration and solution, in other words, the principle of historicism. This refers both to the development history of the objects studied by the sciences as well as to the evolution of scientific knowledge itself.
> In fact, any artificial classification of things or knowledge of things is mainly characterized by the lack of historicism. In that case, the relationships between things and between things and the knowledge of them will not reveal the classification as a result necessarily incurred in the course of development but rather as random and superficial relations that sometimes are also fixed by man himself in order to understand the given material.
> In contrast, the truly natural classification reflects the real connections between objects as they have evolved in their development or in the development of knowledge of them. This was, for example, the case with the classification of the chemical elements. This truly natural system could only be made after the relationships between the elements had been established as real interactions, which had found static relations and given tabular form. All previously established systems of elements proved to constantly be one-sided, in essence artificially, in a sense, arbitrary, because none of them were based on the principle of historicism (Kedrow 1975, vol. 1, 4-5; translated from German by BH. Concerning natural versus artificial classification, see section 4.3a).

A more recent view was expressed by Fulvio Mazzocchi:

> Hermeneutics and postpositivist epistemology emphasized, respectively, the historicity of understanding and the incommensurability [13] of alternative scientific paradigms. Postmodernist theories argued for the breakdown of "grand narratives", indicating the need to embrace pluralistic views. What is basically common to many of these approaches is the refusal of the belief that an absolute vantage point can be reached. There is no ultimate criterion for univocally distinguishing accidental from distinctive features: the fixing of

such a distinctiveness always depends on a given perspective (Mazzocchi 2017, 373).

The most concrete implication of the united historicist view of classification is that theories become important in the explanation for our categories compared to similarity or other criteria (cf., Murphy and Medin 1985; Hjørland and Nissen Pedersen 2005). Different theories or paradigms imply different classifications; therefore to provide design principles for classifications is to negotiate the different theoretical influences on the domain to be classified. Whereas empiricists and positivists tends to "let the data speak for themselves", the hermeneutics-oriented researcher tends to apply a broad orientation which is able to uncover the theoretical influences that have produced the data and their interpretations and classifications in a given domain. This issue brings us to the last of the basic approaches to classification: pragmatic and critical theories.

- δ Pragmatic and critical approaches to classification. The pragmatic and critical approach to classification is based on considering the goals, values, interests, policies, and consequences of classification. There may be many different values at play in forming classifications [14]. Classifications based on this approach are constructed in order to support explicit interests. From this perspective, a classification can never be neutral, but will always tend to support certain goals and interests at the expense of other interests. Nobes and Stadler (2013) examined "how the classifiers themselves and the characteristics that they choose can affect classification" [15] and they observed "that objectivity is neither possible nor desirable in classification. Despite the arbitrariness, some classifications can be more reasonable or more useful than others" (573). Different interests and kinds of subjectivity may not be explicit (or they may be in conflict with their stated goals) and it is therefore an important task to uncover the hidden assumptions in classification principles, classification criteria, and in all kinds of knowledge organization systems (KOS) and information retrieval (IR). As formerly stated:

  > Pragmatic and critical theories of indexing are in agreement with the historicist point of view that subjects are relative to specific discourses but emphasize that subject analysis should support given goals and values and should consider the consequences of indexing. These theories emphasize that indexing cannot be neutral and that it is a wrong goal to try to index in a neutral way. Indexing is an act (and computer-based indexing is acting according to the programmer's intentions). Acts serve human goals. Libraries and information services [and classifications] also serve human goals, and this is why their indexing should be done in a way that supports these (Hjørland 2011, 74).

  Jesse Shera expressed the pragmatic approach very clearly:

  > The pragmatic approach to classification through meaningful units of knowledge must be based on recognition of the obvious truth that any single unit may be meaningful in any number of different relationships depending on the immediate purpose. *Thus it is the external relations, the environment, of the concept that are all-important in the act of classifying.* A tree is an organism to the botanist, an esthetic entity to the landscape architect, a manifestation of Divine benevolence to the theologian, a source of potential income to the

> lumberman. Pragmatic classification, then, denies the existence of the "essence" of tree, for each of these relationships owes its existence to different properties of the tree. Relationship is not a universal, but a specific fact unique to the things related, and just as these relations reveal the nature of the relata, so the relata determine the character of the relationship. (Shera 1951, 83-84; italics in original).

In recent years, there has been a focus on ethical issues in knowledge organization (e.g., Adler and Tennis 2013), as well as on the consequences of classification (e.g., Bowker and Star 1999) and feminist approaches to knowledge organization (e.g., Fox and Olson 2012). Together with research, uncovering hidden assumptions in classification and arguing about the paradox of atheoretical classification (e.g., Hjørland 2016b), the pragmatic/critical approach is exemplified.

One may wonder, however, if pragmatic/critical classification is scientific or able to function as the theoretical basis for classification research and practice. The first impression might be that this is a decline in scientific progress. For example, to classify animals in relation to human interests as domestic animals, pets, and pests seems primitive compared with biological classification in which no such interests and goals seem to be used. The answer to this argument is that there are levels of pragmatic classification and that basic science can be interpreted from the perspective of pragmatic philosophy. Pragmatism may be more or less short term or long term, and the pragmatic value of cladism should be understood from the perspective of long-term interests. Critical theory claims that, in the end, our scientific theories should be evaluated from the perspective of human practice. One of the reasons for considering pragmatism/critical theory as a serious approach is that the other approaches (rationalism, empiricism, and historicism) have problems that are not less significant. We cannot have a science of classification without considering criteria for classification as the fundamental problem. The claim of pragmatism and critical theory is that, in the end, such criteria must serve human values and interests (see Pihlström 2009 for scholarly argumentation). It is extremely important to realize, however, that truth is always the goal in science and scholarship. Worst of all is the manipulation of research in order to serve some specific interests. Like the historicist approach, the pragmatic/critical approach bases design principles for classification on the negotiation of the different theoretical influences on the domain to be classified, but it provides some additional criteria for theory analysis and evaluation, such as the social conditions under which knowledge is being produced.

[top of entry]

4.3 Some metaphysical issues of classification: is there one correct classification?

Has the world one unique structure ("taxonomic monism"), or is there more than one structural entity and process ("taxonomic pluralism")? Are the structures of the world mind-independent (realism), or are they artefacts projected into the world (idealism)? Can our classifications be natural, or are they always artificial? These are core issues in the metaphysics of classification. As stated by Anjan Chakravartty:

The preeminent question of the metaphysics of classification is that of whether the world is itself naturally subdivided into kinds of things. Are kinds out there, so to speak, or are they

rather artefacts of convention, existing only insofar as classificatory practices are brought to bear by creatures such as ourselves? (Chakravartty 2011, 157)

We shall here present the following issues: 4.3a Artificial versus natural classification, 4.3b Order versus disorder of things and unity versus disunity of knowledge

4.3a Artificial versus natural classification. Natural classification can be expressed by Plato's metaphor of "carving nature at its joints" (Plato c.370 BC, *Phaedrus* 265e; see also Campbell et al. 2012). Carl Linnaeus is, however, often recognized as the first scholar to clearly have differentiated "artificial" and "natural" classifications:

As far as I can see, Linnaeus was the first to draw a clear terminological distinction between artificial and natural systems, and this was praised as one of his main achievements by later naturalists and philosophers" (Müller-Wille 2007, 550; cf. Müller-Wille 2013, 311).

Although Linnaeus considered natural classification the ideal, he recognized that his own system (at least partly) represented an artificial classification.

Linnaeus realized that natural orders could not be defined. Even the most "natural", such as the Umbelliferae, the carrot family, lacked features that were unique to and constant within them. Until these were found, natural groups were "like a bell without a clapper"; in modern parlance, they were polythetic (Stevens 1998).

However, the meaning of ‘natural classification' has often been considered unclear:

In the middle of the seventeenth century, many, like Aristotle 2000 years before, believed in a nature that could be represented as some version of the scala naturae, a linear sequence of organisms arranged according to ideas of "highness" and "lowness", in which man was above all organisms (and often not part of nature), and angels and ultimately god [sic] might be above him. There were many other ways of representing nature, and as the geologist Francis Bather observed in 1927, "not a single naturalist had a clear idea of what he meant by 'natural'. All he knew was that the other fellow's classification was unnatural" (Bather 1927). In early usage, natural history itself for the most part had no historical element; "history" meant "story" or "description" (Stevens 2016, 494).

John Stuart Mill suggested the following definition:

The Linnæan arrangement answers the purpose of making us think together of all those kinds of plants, which possess the same number of stamens and pistils; but to think of them in that manner is of little use, since we seldom have anything to affirm in common of the plants which have a given number of stamens and pistils (Mill 1872, 498).

The ends of scientific classification are best answered, when the objects are formed into groups respecting which a greater number of general propositions can be made, and those propositions more important, than could be made respecting any other groups into which the same things could be distributed [...]
A classification thus formed is properly scientific or philosophical, and is commonly called a Natural, in contradistinction to a Technical or Artificial, classification or arrangement (Mill 1872, 499).

Alphabetical arrangement is also a kind of artificial classification because it is not based on relationships between objects, but on formal characteristics of the names of objects. Artificial classifications often serve important but limited practical purposes, whereas natural classifications have broader application functions. (A yet more limited kind of artificial classification is "ad-hoc classification", which is just made for a specific task, non-generalizable, and not intended to be able to be adapted to other purposes: cf. Hudon et al. 2005). Marradi found, however, that use of the term "natural classification" is often connected with a problematic positivist view. [16]

William Parry and Edward Hacker explained the concept in accordance with John Stuart Mill's understanding:

For example, one may divide rocks — or even animals — into those weighing less than ten grams, those weighing at least ten but less than twenty grams, and so on; but this is likely to be of little use, except perhaps for knowing what it would cost to mail them (Parry and Hacker 1991, 133).
And later:
[A classification] is fruitful to the extent that it suggests new hypotheses, explanations, and theories concerning its subject matter. For example, the periodic table — the classification of the elements — proved extremely fruitful, since it suggested the existence of hitherto unknown elements and even suggested what physical properties they would have. It should be noted that natural classifications, by definition, are more fruitful than artificial ones (Parry and Hacker 1991, 139).

Hjørland (2016b) considered the classification of mental disorders in the DSM system. The third edition of this system especially claimed to be atheoretical and tended to give priority to reliability in diagnosis rather than in the validity of classifications. This creates a system with doubtful functions in the understanding and treatment of mental diseases. While it is relatively easy to make a classification reliable (e.g., by classing according to weight, as in Parry and Hacker's quote), it is much more difficult, but also much more important, to make a classification useful for predicting the outcome of interventions. Therefore, the distinction between artificial and natural classification is important when natural is understood as a classification based on a substantial theory — implying that competing theories provide competing natural classifications.

There are two additional aspects to consider in relation to natural classification: the concepts of "natural kind" and "naturalism in classification".

A natural kind can be understood as a grouping that reflects the structure of the natural world rather than the interests and actions of human beings. Chemical elements — e.g., gold — are often taken as an example of a natural kind. However, the philosophical problems of natural kinds, and how it can be decided if something is or is not a natural kind, are big ones (obviously, the social constructivist denies that natural kinds exist). Here we shall not go further into this issue but leave it to a hopefully forthcoming independent article. There is a large literature on this concept: see, for example, Khalidi (2013).

Naturalism in classification may be understood as a general approach to classification theory that establishes a close connection between knowledge organization and classification in empirical science and scholarship (e.g., biological classification, classification of the chemical and physical elements, classification in arts, linguistics, psychiatry, etc.). Naturalism is

therefore opposed to the idea that the field of knowledge organization has a set of a priori classification principles or methods. Naturalism in classification is based on the corresponding concept of "naturalistic epistemology", which has been described the following way:

Broadly speaking, however, proponents of NE [naturalistic epistemology] take the attitude that there should be a close connection between philosophical investigation — here, of such things as knowledge, justification, rationality, etc. — and empirical ("natural") science (Rysiew 2016).

Naturalistic classification is therefore the attempt to learn classificatory principles by studying how the most successful classifications have been constructed in different domains, as well as the discourses and controversies about classification and its philosophy. In the field of knowledge organization, this has been relatively neglected because the field has tended to provide prescriptive principles on how to classify knowledge (e.g., facet analytic principles, standards for thesaurus construction, or user-based methodologies).

4.3b Order versus disorder of things and unity versus disunity of knowledge. It has been (and probably still is) characteristic of many researchers to believe in a fundamental order underlying the apparently confusing empirical picture of the world. Related to this view is a belief in the nature of knowledge to reflect or converge toward this underlying order. The first point is about taxonomic monism versus pluralism; the second is about descriptive monism versus pluralism (or epistemic relativism). Often taxonomic monism is associated with scientific realism, while taxonomic pluralism is associated with relativism. Chakravartty (2011) argued, however, that taxonomic monism is in opposition to contemporary science and that a form of taxonomic pluralism is consistent with realism.

Henry Bliss is a library science representative holding the belief in an underlying order of things and in the unity of and consensus of knowledge. He wrote:

The more definite the concepts, the relations, and the principles of science, philosophy, and education become, the clearer and more stable the order of the sciences and studies in relation to learning and to life; and so the scientific and educational consensus becomes more dominant and more permanent (Bliss 1933, 37).

A critique of this view was made by Satija (1992, 40-41), paraphrasing McGarry (1991, 148):

Knowledge is a cultural entity and keeps shifting its pattern like a kaleidoscope. An emergence of the new knowledge modifies the structure of the whole. Contrary to H. E. Bliss (1870-1955) there is no permanent order in knowledge. "Pattern is new every moment", said T.S. Eliot (1888-1965), with a poetic vision.

In information science and knowledge organization, this tension between the idea of order and stability and the reality of disorder and relativity is clear in the differences between traditional bibliographical classifications on the one hand, and the bibliometric maps based on, for example, co-citation patterns on the other. Whereas traditional classifications tend to provide relatively stable structures, the citation practices of researchers tend to display very unstable patterns.

While the goal of knowledge organization is to discover or construe some kind of order, the nature of the order matters. Francis Miksa, for example, wrote: "In the end, there is strong indication that Ranganathan's use of faceted structure of subjects may well have represented his need to find more order and regularity, in the realm of subjects, than actually exist" (Miksa 1998, 73). This quote may be considered a criticism of Ranganathan in accordance with Hjørland's (2014) criticism of rationalism.

Jens-Erik Mai discussed this from the perspective of post- or late modernist philosophy:

This paper traces and interrogates the shift from classification-as-ontology, in which everything is defined as it is, to a more contemporary notion of classification-as-epistemology, in which everything is interpreted as it could be — or more precisely, the paper argues for a conceptual move from modern monistic ontology to late-modern pluralistic epistemological foundation for classification theory and practice (Mai 2011, 711).

This opens many questions, and the most important claims by Mai may be considered the critique of the positivist view that the researcher and knowledge organizer are neutral agents providing an objective mirror of the universe of knowledge (see also Pando and de Almeida 2016).

The same issue was also addressed by philosopher Finn Collin, who discussed a social constructivist view of classification, which he formulated in this way:

to isolate a certain kind of thing is the same process as classifying individual things. And classification is a matter of sorting things into groups, the members of which are more similar to each other than to items outside the group. However, things are only similar or dissimilar in certain respects [...]. Classifications are not objective divisions, inherent in the nature of things, but are structures we impose upon nature. [...] *kinds* of things are indeed human creations (Collin 1993, 29: italics in original).

Then Collin wrote (29): "I believe this reasoning is mistaken. What follows from the premises is a less radical conclusion." Later, he summed up his argument:

the nominalist argument mistakes a valid anti-essentialist point for an anti-realist one. It is true that there is not, among the true descriptions of a thing, one which is privileged, in the sense that any classification of the thing has to be based upon that particular description. There is no uniquely correct classification of a thing, one that shows what the thing really is, rendering alternative classifications somehow misleading or inappropriate. But it is a mistake to infer from this that things do not in themselves belong to any classifications at all and that things only come to belong to classes when we place them there. Once we relativise similarity and dissimilarity to particular aspects of things, similarity and dissimilarity turn out to be objective, although relational, properties of things, and the predicates that are defined by the equivalency classes of things turn out to be genuine properties of those things. Reality possesses *all* the properties attributed to it in these alternative descriptions" (Collin 1993, 43; italics in original).

An issue regarding Collin's argumentation is "the true descriptions of a thing". Because descriptions are theory-dependent, different descriptions are not necessarily equally true. Also, of course, classifications are made, chosen, or used for a purpose, and therefore our classifications to a large degree will be human creations (but not therefore arbitrary or

contingent; within ornithology, for example, there seems now to be considerable optimism that a "final" classification of birds on the overall level seems within reach; cf. Fjeldså 2013). [17]

Mazzocchi wrote about the opposite of one right way to "carve nature at its joints", taxonomic pluralism:

Ontological pluralist views have also been developed. For instance, Dupré's (1993) "promiscuous realism" conceives the world (his argument refers, above all, to the biological realm) as made up of a multidimensional complexity: things are interconnected and interrelated to one another in multiple ways; there is no unique way of carving nature at its joints or one ultimately right way of classifying or hierarchizing. Rather, there are many equally legitimate ways of dividing the world into "kinds", depending on the purposes of investigations.
Returning to Greek philosophy for a moment, we find that even Aristotle, with respect to zoological classification, supported a pluralist view (see *Parts of animals* I and *History of animals*) that partially resemblances Dupré's view (Henry 2011). The biological world contains natural kinds marked by real, objective boundaries, but at the same time it is not possible to assign animals to a unique set of mutually exclusive and non-overlapping kinds. Even in Aristotle's view, many cross-cutting joints can be found in nature. Which joints are chosen to be cut along depends at least partially on the explanatory context. Since there are diverse explanatory purposes in zoology, organisms can be grouped into various (cross-cutting) kinds (Mazzocchi 2017; see also Galison and Stump 1996).

In this connection, it is worth mentioning the mental models or metaphors that govern our view of how knowledge is organized.

The tree model has been historically associated with the philosophical position of classical realism. A number of basic assumptions underlying such a position can be listed. Ontologically, there is: (1) a reality that exists independently of us, (2) a single set of categories, and (3) a single set of essential properties and therefore a unique way of dividing the world into kinds. Epistemologically, saying that also implicitly means that: (4) the ultimate order of the world is within the reach of human cognitive means, i.e. there is an absolute or neutral vantage point from which to grasp reality as it "actually" is, and this vantage point is accessible to us.
Such a model has had a strong influence on logic, philosophy, and science. In Western culture, the tree structure based on Aristotle's logic has been the dominant model of classification. Scientific taxonomic thinking (for instance, the Linnaean classification system in biology) and many contemporary semantics theories (for instance, Chomsky's sentence diagrams) also embody a similar scheme" (Mazzocchi 2017, 372).

The tree metaphor is increasingly ousted by alternatives such as net and the rhizome (see Mazzocchi 2013) and it matters which models guide our research and practice.

[top of entry]

Conclusion

The concept of classification and its associated theories is extremely wide-ranging and interdisciplinary. Many kinds of knowledge must be combined in order to make overall progress in this field. Such projects may be guided by different basic assumptions. Daniel Parrochia seems to base his suggestion for such a program on more formal and mathematical approaches. In 2016 he wrote:

In spite of these advances, most of classifications are still based on the evaluation of resemblances between objects that constitute the empirical data. This one is almost always computed by the means of some notion of distance and of some algorithms of aggregation of classes. So all these classifications remain, for technical and epistemological reasons that are detailed below, very unstable ones. A real algebra of classifications, which could explain their properties and the relations existing between them, is lacking. Though the aim of a general theory of classifications is surely a wishful thought, some recent conjecture gives the hope that the existence of a metaclassification (or classification of all classification schemes) is possible (Parrochia 2016).

Alternatively, another program may be suggested (not necessarily in conflict with Parrochia's). This program is less formal (and therefore more substantial) and views classifications as tied to (domain) theories. By implication, the study of classification involves the study of theories in different domains and the ontological claims of those theories. The justification of a good classification in this perspective is to make a justification of the theoretical premises on which it is based.

[top of entry]

Acknowledgements

I am very grateful to Fulvio Mazzocchi, who served as the editor of this article. He and the two anonymous referees provided detailed, knowledgeable, careful, and fruitful suggestions for improving the original manuscript. Also thanks to Daniel Parrochia for positive feedback on the manuscript.

Endnotes

1. Wesolek (2012, 1) stated: "He [Hjørland] thinks that concept classification should not strive to classify on the basis of the properties of objects, but rather on descriptions of objects that are loosely derived from human activity and social negotiation." In order to classify by properties, we have to know these properties and the argument is that we only know them from descriptions — our own or those of other people (or from perceptions, but such perceptions have to be transferred to descriptions) — and those descriptions or perceptions will be influenced by culture, goals, interests — in short, subjectivity. In this way, classification is always based on properties of objects. Hjørland's definition of classification is, however, correctly cited by Wesolek on the same page: "Classification, as defined by Hjørland, is the 'sorting of objects based on some criteria selected among the properties of the classified objects'."

2. Parrochia and Neuville's (2013) monograph *Towards a general theory of classifications* is written from the perspective of mathematics but demonstrates surprisingly broad knowledge of classification research, including research in the community of knowledge organization

(see also Parrochia 2016). They, too, find that their work is about the epistemology of classifications, not a mathematical textbook or monograph (vii). They write: "[T]he least we can say is that the field of a mathematical theory of classifications is not a completely stable domain, and one is led to think, finally, that a vast side of it is still to be developed. Another reason for the lack of a general theory, close to the previous one, is that scientists are faced with a very difficult problem (finding a formalism enough general to apply to any kind of classifications), for which no complete solution is known at the present. All the same, we think that the research we have carried on for more than thirty years might be of some interest for librarians, logicians, and also for scientists in the different fields of empirical science, all of whom need to devise their own classifications. But this book has a deeper stake. In fact, pure mathematics wants also a general theory of classifications to take over from classic (and too limited) versions of set theory..." (Parrochia and Neuville 2013, xv). Parrochia and Neuville (2013) seem to assume that a general theory of classification(s) must be a mathematical theory, but do not discuss if other views may also be important, or what the *relative* contribution of mathematics is. No doubt mathematics is important, but it is certainly not all there is to say about classification.

3. An anonymous reviewer wrote: "It is a bit funny to call the methods of forming classes ‘epistemology' (which is theory of knowledge). If anything, it would be 'ontology'." However, the four methodologies suggested in this article are rationalism, empiricism, historicism, and pragmatism, which are well known epistemological positions, and thus normative principles on how to obtain knowledge. It is correct, however, that these positions also rest on different ontological assumptions. The a priori of rationalism is clear ideas or logical units; for empiricism, it is sense impressions; for historicism, it is change; and, for pragmatism, the a priori is living and acting in the world.

4. An anonymous reviewer wrote: "The explanations of the Classical Theory and Prototype Theory are not strong" and provided a helpful, but lengthy improved description. The choice made here is to try to maintain the short outlines here and later to have these theories covered by independent articles.

5. A monothetic class is defined in terms of characteristics that are both necessary and sufficient in order to identify members of that class. This way of defining a class is also termed the Aristotelian definition of a class. A polythetic class is defined in terms of a broad set of criteria that are neither necessary nor sufficient. Each member of the category must possess a certain minimal number of defining characteristics, but none of the features has to be found in each member of the category. This way of defining classes is associated with Wittgenstein's concept of "family resemblances". The monothetic type is a type in which all members are identical in all characteristics; the polythetic type is a type in which all members are similar, but not identical.
The distinction between monothetic and polythetic classification is discussed by Rijsbergen (1979, 28-29): "An early statement of the distinction between monothetic and polythetic is given by Beckner (1959, 22): 'A class is ordinarily defined by reference to a set of properties which are both necessary and sufficient (by stipulation) for membership in the class. It is possible, however, to define a group K in terms of a set G of properties f1, f2, . . . , fn in a different manner. Suppose we have an aggregate of individuals (we shall not yet call them a class) such that
(1) each one possesses a large (but unspecified) number of the properties in G;
(2) each f in G is possessed by large number of these individuals; and
(3) no f in G is possessed by every individual in the aggregate.'

The first sentence of Beckner's statement refers to the classical Aristotelian definition of a class, which is now termed monothetic. The second part defines polythetic. [...]
To illustrate the basic distinction, consider the following example (Figure 3.1) of 8 individuals (1-8) and 8 properties (A-H). The possession of a property is indicated by a plus sign. The individuals 1-4 constitute a polythetic group, each individual possessing three out of four of the properties A,B,C,D. The other 4 individuals can be split into two monothetic classes, {5,6} and {7,8}. The distinction between monothetic and polythetic is a particularly easy one to make providing the properties are of a simple kind, e.g., binary state attributes. When the properties are more complex, the definitions are rather more difficult to apply, and in any case are rather arbitrary.

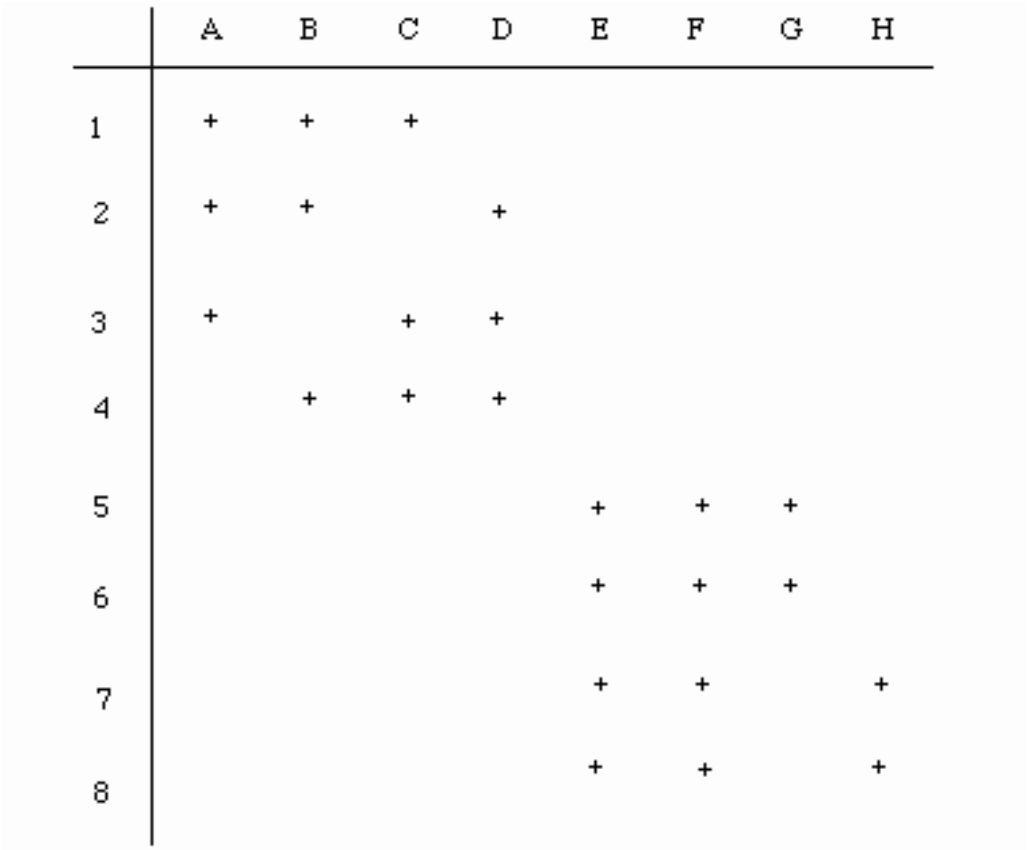

Figure 3.1. An illustration of the difference between monothetic and polythetic

Bowker (1998, 256) wrote: "Aristotelian models — monothetic or polythetic — have traditionally informed formal classification theory in a broad range of sciences, including biological systematics, geology, and physics". The first part of this quote was eliminated from the same quote in section 4.1 because it seems confusing and probably wrong that polythetic classification is part of formal classification theory and has been termed Aristotelian.

6. The idea is older, however. "The starting point is the work of the great French botanist Michel Adanson, who proposed that a member of a class of plants did not need to possess all the defining features of the class, and that a deviant specimen did not need to be assigned to a separate class (Adanson 1763, i: cliv sqq.) [note omitted]. 'The important point he made was that creatures should be grouped together on the greatest number of features in common, and there is no justification for deciding a priori on the relative importance of characters in making a natural taxonomy' (Sneath 1962: 292)" (Needham 1975, 353).

7. An anonymous reviewer wrote: "Just as a mild correction to Lakoff's verbal flourishes, the Aristotelian-Classical theory certainly had been the subject of major debate, a debate that had lasted 2000 years and involved some of the finest scholars."

8. An anonymous reviewer wrote: "I am not sure about Kuhn and Billig in this setting. We are getting drawn off into potentially quite deep cognitive or social psychology. But isn't our interest storing and retrieving information or knowledge?" Yes, that is our interest, and the claim is that Kuhn and Billig provide important knowledge for this purpose. Kuhn provides the knowledge that we need to consider how different theories or paradigms classify knowledge and relate our decisions to a choice or a negotiation between different views. Billig help us see the problematic assumptions in the cognitive view according to which we have some built-in mechanisms on how to classify knowledge.

9. "From Cesalpino to Linnaeus, this [downward classification by logical division] was the almost universally preferred system, particularly in botany (Mayr 1982, 158-179). In this methodology, the classifier starts with the entire 'universe' — let's say, all animals — and, with the help of divisional logic, divides them into more and more homogeneous groups. A criterion like blood temperature, when applied to animals, results in two groups: warm-blooded and cold-blooded animals. By continuing dichotomy, one finally reaches the species level. Even though Linnaeus in his artificial system still employed downward classification, it had become evident by his time that a reliance on single characters — and the inevitable arbitrariness of the sequence in which these characters were chosen — could lead to rather artificial systems. At the end of the eighteenth century, downward classification was therefore replaced by upward classification. In retrospect, it eventually became evident that the downward procedure actually produces identification schemes rather than classifications. It survives today in the form of keys. Curiously, in the literature of the philosophy of science, particularly of logic, classification by logical division has been considered the method of classification up to modern times" (Mayr 1995, 420-421).

10. "This system [grouping or upwards classification] most closely conveys the basic meaning of the term classification, that is, to assemble items into classes on the basis of resemblance in observed characters" (Mayr 1995, 421) and "After earlier trials by some herbalists and by Magnol, it was particularly Adanson (1763) who promoted classification by grouping. By the first third of the nineteenth century it had become the almost universal method of classifying plants and animals (Mayr 1982, 190-208). The empirical rule guiding the taxonomist was well stated by Whewell (1840, vol. 1, 521): 'The maxim by which all systems professing to be natural must be tested is this: that the arrangement obtained from one set of characters coincides with the arrangement obtained from another set'" (Mayr 1995, 422).

11. An anonymous reviewer wrote: "There wants to be emphasis here that the classification is systematic classification [as opposed to conceptual classification]." However, the suggested principles are meant to serve conceptual classification as well as systematic classification, and a fundamental view is that conceptual classification represents the core theory on which systematic classification is based. The reviewer seems to view the two kinds of classification as too dualistic.

12. Medin and Aguilar (1999, 104; emphasis in original), for example, wrote: "Why is this notion that categories are defined by some "objective" similarity controversial? The main criticism has been that the notion of similarity is too unconstrained to be useful as an explanatory principle (Goodman 1972; Murphy and Medin 1985). Similarity is usually

defined in terms of shared properties, but Goodman argued that any two things share an unlimited number of properties (e.g. robins and elephants can move, weight more than two ounces, take up space, can be thought about, etc.). Given this apparent flexibility, it may be that we see things as similar because they belong to the same category and not vice versa. That is, maybe we can explain similarity in terms of categories."

13. About incommensurability, see Kuhn (2000), and consider the following quotes. "Incommensurability arises because it is impossible to transfer the natural categories employed within one taxonomic structure into the categorical system of another such structure. Apparently on the basis of such taxonomic incommensurability, Kuhn asserted a number of antirealist theses about truth, reference and reality. In this paper, it will be argued, however, that, far from leading to antirealist consequences about the relationship between theory and reality, the taxonomic incommensurability thesis may be incorporated unproblematically within a reasonably robust scientific realist framework" (Sankey 1998, 7). "With this theory of kinds, Kuhn redraws the picture of scientific revolutions. Since the interconnections among kind terms form a lexical taxonomy, scientific revolutions, which now are limited to the meaning change of kind terms, become taxonomic changes. A scientific revolution produces a new lexical taxonomy, in which some kind terms refer to new referents that overlap with those denoted by some old kind terms. Therefore, incommensurability does not result merely from translation failures of individual concepts. The prerequisite for full translatability between two taxonomies is not shared features of individual concepts, but a shared lexical structure (Kuhn 1990b, p. 7). Scientists from rival paradigms face incommensurability because they construct different lexical taxonomies and thereby classify the world in different ways" (Chen 1997, 260; Kuhn 1990b refers to an unpublished manuscript).

14. The following quote exemplifies the complex pattern of different interests that may be at play behind given classifications — in particular in the domain of arts: "The work of DiMaggio (1987) has provided the theoretical foundation for much research on classification systems. His concept of artistic classification systems offers a number of insights. First, the study of classification systems needs to take into account both the consumption and production of art. According to DiMaggio (1987: 441), artistic classification systems consist of 'the way that the work of artists is divided up both in the heads and habits of consumers and by the institutions that bound the production and distribution of separate genres' [italics in original]. On the one hand, classification systems arise out of processes of social distinction, whereby consumers use cultural objects to mark social boundaries. These 'ritual classifications' can thus be influenced by social structural factors at the societal level — such as stratification systems, elite cohesion, social and geographic mobility, etc. — that generate demand for cultural boundaries. On the other hand, classification systems are also influenced and mediated by ‘classification processes' at the production side. DiMaggio (1987) identifies commercial classifications (the classifications used by commercial producers to market their products), administrative classifications (created and enforced by the state), and professional classifications (classifications driven by the incentives of artists to differentiate and mark boundaries). The study of classification systems thus needs to be attentive to the extent to which the categorical demands of consumers and the categories used by producers overlap, diverge, or mutually reinforce each other" (Venrooij and Schmutz 2015, 799).

15. Nobes and Stadler (2013, 575) examine the degree to which classification is determined by who is classifying and they refer to Bloor's (1982, 268) support for the claim of Durkheim and Mauss (1903 [2010]) that the classification of things reproduces a pattern of social

arrangements more than a pattern of the things. Durkheim and Mauss found that our classifications necessarily are hierarchical because human social organization, from which our view of nature ultimately flows, is itself hierarchical. This view has been strongly criticized as 'sociologism', but is, as we saw, defended by Bloor (1982).

16. Marradi wrote: "The opposition between "natural" and "artificial" classification is a recurring theme in the last two centuries. Cohen and Nagel have cogently argued that 'any division ... according to some actual trait arbitrarily chosen is perfectly natural... [but it] may also be said to be artificial, in the sense that we select the trait' (1934, 223). Yet, many scholars have continued stressing "naturality" as a desirable property, by gradually re-interpreting it in terms of significant relationships with other classifications (Hempel 1961, version reprinted 1965, 146-7; Kaplan 1964, 50), utility 'for a wider range of inductive generalizations' (Gilmour 1940, 466), 'systematic import' (Huxley 1940; Hempel 1952; Sandri 1969), links with theory (Hempel 1952; Bunge 1967, 83), 'projectibility of discriminating concepts' (Sandri 1969, 99 ff.). In short, the concept of "natural classification" has been transferred from the ontological to the epistemological domain. However, as Tiryakian (1968, 177) has remarked, 'the reification of typologies is a frequent temptation and pitfall'. In a typical manual of the neopositivist period one can still read a statement as plain as 'A natural class is based on the fundamental character of things' (Lenzel 1938, 32). And one may suspect that, if the epistemological coat of paint were scraped off, quite substantial traces of rusty essentialism would loom through the still widespread concern for "natural" classifications" (Marradi 1990, 149).

17. Jon Fjeldså also gave a speech at the University of Copenhagen on 26 March 2014: "Får vi snart en 'endelig' fugleklassifikation?" (Do we soon get a "final" classification of birds?)

[top of entry]

References

Academia das Ciências de Lisboa. 2001. Dicionário da língua portuguesa contemporânea. vol. 1. Lisboa: Verbo.

Adams, William Y. and Ernest W. Adams. 1991. *Archaeological Typology and Practical Reality: A Dialectical Approach to Artifact Classification and Sorting*. Cambridge, UK: Cambridge University Press.

Adamson, Robert. 1901. "Classification". In James Mark Baldwin ed., *Dictionary of philosophy and psychology*. New York, NY: The MacMillan Company, vol. 1, 185. http://psychclassics.yorku.ca/Baldwin/Dictionary/defs/C2defs.htm

Adanson, Michel. 1763. *Familles des plantes*. Paris: Vincent.

Adler, Melissa and Joseph T. Tennis. 2013. "Toward a taxonomy of harm in knowledge organization systems". *Knowledge Organization* 40, no. 4: 266-272.

Andersen, Hanne, Peter Barker and Xiang Chen. 1996. "Kuhn's mature philosophy of science and cognitive psychology." *Philosophical Psychology* 9, no. 3: 347-363.

Andersen, Hanne, Peter Barker and Xiang Chen. 2006. *The cognitive structure of scientific revolutions*. Cambridge, UK: Cambridge University Press.

Aristotle. 350 BCE. *Topics*. (Translated by William Adair Pickard-Cambridge). Overland Park, KS: Digireads Publishing, 2006. http://classics.mit.edu/Aristotle/topics.1.i.html

Bailey, Kenneth D. (1994). *Typologies and taxonomies: An introduction to classification techniques*. London: Sage Publications.

Bather, Francis Arthur. 1927. "Biological classifications: Past and future". *Quarterly Journal of the Geological Society* 83 no. 1: 63- 104.

Beckner, Morton. 1959. *The biological way of thought*. New York, NY: Columbia University Press.

Beghtol, Clare. 2003. "Classification for information retrieval and classification for knowledge discovery: Relationships between 'professional' and 'naïve' classifications". *Knowledge Organization* 30, no. 2: 64-73.

Beghtol, Clare. 2004. "Exploring new approaches to the organization of knowledge: The subject classification of James Duff Brown". *Library Trends* 52, no. 4: 702-718.

Beghtol, Clare. 2010 "Classification theory". In Marcia J. Bates and Mary Niles Maack eds., *Encyclopedia of library and information sciences*. Third edition. Boca Raton, FL: CRC Press, vol. 2, 1045-1060.

Biggam, Carole Patricia. 2015. *The semantics of colour: A historical approach*. Cambridge, UK: Cambridge University Press.

Billig, Michael. 1996. *Arguing and thinking: A rhetorical approach to social psychology*. Second edition. Cambridge, UK: Cambridge University Press.

Bliss, Henry E. 1929. *The organization of knowledge and the system of the sciences*. New York, NY: Henry Holt and Company.

Bliss, Henry Evelyn. 1933. *The organization of knowledge in libraries and the subject-approach to books*. New York, NY: H. W. Wilson.

Bliss, Henry E. 1935. *A system of bibliographical classification*. New York: The H. W. Wilson Company.

Bloor, David. 1982. "Durkheim and Mauss revisited: Classification and the sociology of knowledge." *Studies in History and Philosophy of Science* 13, no. 4: 267-297.

Bowker, Geoffrey C. 1998. "The kindness of strangers: Kinds and politics in classification systems". *Library Trends* 47, no. 2: 255-292.

Bowker, Geoffrey C. and Susan Leigh Star. 1999. *Sorting things out: Classification and its consequences*. Cambridge, MA: MIT Press.

Bunge, Mario A. 1967. *Scientific research*, vol. I: *The search for system*. New York, NY: Springer.

Campbell, Joseph Keim, Michael O'Rourke and Matthew H. Slater eds. 2012. *Carving nature at its joints: Natural kinds in metaphysics and science*. Cambridge, MA: MIT Press.

Candolle, Augustin-Pyramus de. 1813. *Théorie élémentaire de la Botanique ou exposition des principes de la classification naturelle et de l'art d'écrire et d'étudier les végétaux*. Paris: Determille.

Capecchi, Vittorio and Frank Möller. 1968. "Some applications of entropy to the problems of classification". *Quality & Quantity* 2, nos. 1-2: 63-84.

Carpenter, William Benjamin. 1847. *Zoology: A systematic account of the general structure, habits, instincts, and uses of the principal families of the animal kingdom*. London: W. S. Orr.

Chakravartty, Anjan. 2011. Scientific realism and ontological relativity. *The Monist* 94, no. 2: 157-180.

Chan, Lois Mai. 1994. *Cataloging and classification: An introduction*. Second edition. New York, NY: McGraw-Hill.

Chen, Xiang. 1997. Thomas Kuhn's latest notion of incommensurability. *Journal for General Philosophy of Science/Zeitschrift für allgemeine Wissenschaftstheorie* 28, no. 2: 257-273.

Cohen, Morris Raphael and Ernst Nagel. 1934. *An introduction to logic and scientific method*. New York, NY: Harcourt.

Collin, Finn. 1993. Social constructivism without paradox. *Danish Yearbook of Philosophy* 28: 24-46.

Condorcet, Marie Jean Antoine Nicolas de Caritat, Marquis de. 1777. "Sur les familles naturelles des plantes, et en particulier sur celle des renoncules". *Histoire de l'Académie Royale des Sciences (Paris)* année 1773: 34-36.

Cranz, David. 1767. *The history of Greenland: Containing a description of the country and its inhabitants*. Vol. I-II. London: The Brethren's Society for the Furtherance of the Gospel among the Heathen.

Cutter, Charles Ammi. 1876. *Rules for a printed dictionary catalogue*. Washington: Government Printing Office.

Dahlberg, Ingetraut. 1993. "Knowledge organization: Its scope and possibilities". *Knowledge Organization* 20, no. 4: 211-222.

Dahlberg, Ingetraut. 2010. "International Society for Knowledge Organization (ISKO)". In Marcia J. Bates and Mary Niles Maack eds., *Encyclopedia of library and information sciences*. Third edition. Boca Raton, FL: CRC Press, vol. IV: 2941-2949.

Darwin, Charles. 1859. *On the origin of species by means of natural selection, or the preservation of favoured races in the struggle for life*. London: J. Murray.

DiMaggio, Paul. 1987. Classification in art. *American Sociological Review* 52, no. 4: 440-455.

Dousa, Thomas M. 2009. "Evolutionary order in the classification theories of C. A. Cutter and E. C. Richardson: Its nature and limits". In *Proceedings of the 2009 North American Symposium on Knowledge Organization, June 18-19, 2009, Syracuse, New York*: 76-90. Published online at: https://journals.lib.washington.edu/index.php/nasko/article/download/12810/11292

Drucker, Donna J. 2014. *The classification of sex: Alfred Kinsey and the organization of knowledge*. Pittsburgh, PA: University of Pittsburgh Press.

Dupré, John. 1993. *The disorder of things: Metaphysical foundations for the disunity of science*. Cambridge, MA: Harvard University Press.

Dupré, John. 2002. *Humans and other animals*. Oxford, UK: Oxford University Press.

Dupré, John. 2006. "Scientific classification". *Theory, Culture & Society* 23, nos. 2-3: 30-32.

Durkheim, Émile and Marcel Mauss. 1903. "De quelques formes primitives de classification: Contribution à l'étude des représentations collectives". *Année Sociologique* VI (1901-1902): 1-72.

Durkheim, Émile and Marcel Mauss. 2010. *Primitive classification*. (Translated from the French and with an introduction by Rodney Needham). London: Routledge.

Eiser, J. Richard. 1980. *Cognitive social psychology*. London: McGraw Hill.

Faria, Maria Isabel and Maria da Graça Pericão. 2008. *Dicionário do livro: Da escrita ao livro electrónico*. Coimbra: Almedina.

Farradane, Jason E. L. 1950. "A scientific theory of classification and indexing and its practical applications". *Journal of Documentation* 6, no. 2: 83-99.

Farradane, Jason E. L. 1952. "A scientific theory of classification and indexing: Further considerations". *Journal of Documentation* 8, no. 2: 73-92.

Feger, Hubert. 2001. "Classification: Conceptions in the social sciences". In Neil J. Smelser and Paul B. Baltes eds., *International encyclopedia of the social and behavioral sciences*. Amsterdam: Elsevier, vol. 3: 1966-1973.

Feger, Hubert. 2015. "Classification: Understandings of in the social sciences". In James D. Wright ed., *International encyclopedia of the social and behavioral sciences*. Second edition. Amsterdam: Elsevier, vol. 3, 805-810.

Fjeldså, Jon. 2013. Avian classification in flux. In Josep del Hoyo, Andrew Elliott, Jordi Sargatal and David A. Christie eds. *Handbook of the birds of the world*. Special volume 17. Barcelona: Lynx Editions, 77-146.

Fox, Melodie J. 2011. "Prototype theory: An alternative concept theory for categorizing sex and gender?" In Richard P. Smiraglia ed., *Proceedings from North American Symposium on Knowledge Organization. Toronto, Canada*, vol. 3: 151-159. http://journals.lib.washington.edu/index.php/nasko/article/view/12799

Fox, Melodie J. and Hope A. Olson. 2012. "Feminist epistemologies and knowledge organization". In Hur-li Lee and Richard Smiraglia eds., *Cultural frames of knowledge*. Würzburg: Ergon Verlag, 79-98.

Frické, Martin. 2012. *Logic and the organization of information*. New York, NY: Springer.

Frické, Martin. 2016. "Logical division". *Knowledge Organization* 43, no. 7: 539-549. A version is also available as http://www.isko.org/cyclo/logical_division

Galison, Peter and David J. Stump eds. 1996. *The disunity of science: Boundaries, contexts, and power*. Stanford, CA: Stanford University Press.

Gilmour, John Scott Lennox. 1940. "Taxonomy and philosophy", In Julian Huxley ed., *The new systematics*. Oxford, UK: Clarendon Press, 461-474.

Gitelman, Lisa ed. 2013. *'Raw data' is an oxymoron*. Cambridge, MA: MIT Press.

Gnoli, Claudio. 2006. "Phylogenetic classification". *Knowledge Organization* 33, no. 3: 138-152.

Gnoli, Claudio. 2016. "Classifying phenomena, part 1: Dimensions". *Knowledge Organization* 43, no. 6: 403-415.

Goodman, Nelson. 1972. Seven strictures on similarity. In Nelson Goodman ed., *Problems and projects*. New York, NY: Bobbs-Merrill, 437-446.

Grove, Andrew. 2010. "Taxonomy". In Marcia J. Bates and Mary Niles Maack eds., *Encyclopedia of library and information sciences*. Third edition. Boca Raton, FL: CRC Press, vol. VII: 5139-5148.

Hedden, Heather. 2016. *Accidental taxonomist*. Second edition. Medford, NJ: Information Today.

Hempel, Carl Gustav. 1952. *Fundamentals of concept formation in empirical science*. Chicago, IL: Chicago University Press.

Hempel, Carl Gustav. 1961. "Fundamentals of taxonomy". In Joseph Zubin ed., *Field studies in mental disorders*. New York, NY: Grune & Stratton. (Quotations from the version are reprinted in Carl G. Hempel, *Aspects of scientific explanation and other essays in the philosophy of science*. Glencoe, IL: Free Press, 1965, 137-154).

Hennig, Willi. 1966. *Phylogenetic systematics*. (Translated by D. Dwight Davis and Rainer Zangerl). Urbana, IL: University of Illinois Press.

Henry, Devin. 2011. "Aristotle's pluralistic realism". *The Monist* 94 no. 2: 197-220.

Hjørland, Birger. 2003. "Fundamentals of knowledge organization". *Knowledge Organization* 30 no. 2: 87-111.

Hjørland, Birger. 2009. "Concept theory". *Journal of the American Society for Information Science and Technology* 60, no. 8: 1519-1536.

Hjørland, Birger. 2011. "The importance of theories of knowledge: Indexing and information retrieval as an example". *Journal of the American Society for Information Science and Technology* 62, no. 1: 72-77.

Hjørland, Birger. 2013. "Citation analysis: A social and dynamic approach to knowledge organization". *Information Processing and Management* 49, no. 6: 1313-1325.

Hjørland, Birger. 2014. "Is facet analysis based on rationalism? A discussion of Satija (1992), Tennis (2008), Herre (2013), Mazzocchi (2013b), and Dousa and Ibekwe-SanJuan (2014)". *Knowledge Organization* 41, no. 5: 369-376.

Hjørland, Birger. 2015. "Are relations in thesauri 'context free, definitional, and true in all possible worlds'"? *Journal of the Association for Information Science and Technology* 66, no. 7: 1367-1373.

Hjørland, Birger. 2016a. "Informetrics needs a foundation in the theory of science." In Cassidy Sugimoto ed., *Theories of informetrics and scholarly communication*. Berlin: Walter de Gruyter, 20-46.

Hjørland, Birger. 2016b. "The paradox of atheoretical classification". *Knowledge Organization* 43, no. 5: 313-323.

Hjørland, Birger and Hanne Albrechtsen. 1999. "An analysis of some trends in classification research". *Knowledge Organization* 27, no. 1: 131-139.

Hjørland, Birger and Karsten Nissen Pedersen. 2005. "A substantive theory of classification for information retrieval". *Journal of Documentation* 61, no. 5: 582-597.

Hlava, Marjorie M.K. 2014. *The TaxoBook: Principles and Practices of Building Taxonomies*. Vol. 1-3. San Rafael, CA: Morgan Claypool.

Hudon, Michéle, Sabine Mas and Dominique Gazo. 2005. "Structure, logic, and semantics in ad hoc classification schemes applied to web-based libraries in the field of education. *Canadian Journal of Information and Library Science/La Revue canadienne des sciences de l'information et de bibliothéconomie* 29, no. 3: 265-288.

Hull, David L. 1998. "Taxonomy". In Edward Craig ed., *Routledge encyclopedia of philosophy*. London: Routledge, vol. 9, 272-276.

Huxley, Julian ed. 1940. *The new systematics*. Oxford: Clarendon Press.

Huxley, Thomas Henry. 1869. *An introduction to the classification of animals*. London: John Churchill & Sons. https://archive.org/stream/anintroductiont00huxlgoog#page/n4/mode/2up

International Organization for Standardization. 2013. *ISO 25964-2: Information and Documentation—Thesauri and interoperability with other vocabularies, Part 2: Interoperability with other vocabularies*. Geneva: International Organization for Standardization.

ISO 5127-6. 1988. Documentation et information — Vocabulaire: Partie 6: Langages documentaires. In *Documentation et information: Recueil de normes ISO* I. Genève: ISO, 87-111.

Jacob, Elin K. 2004. "Classification and categorization: A difference that makes a difference". *Library Trends* 52, no. 3: 515-540.

Jung, Carl G. (1971). *The collected works of C. G. Jung. Volume 6: Psychological types*. Princeton: Princeton University Press.

Kaplan, Abraham. 1964. *The conduct of inquiry*. San Francisco, CA: Chandler.

Kedrow, Bonifatii Mikhailovich. 1975. *Klassifizierung der Wissenschaften* 1-2. Moskau: Verlag Progress.

Khalidi, Muhammad Ali. 2013. *Natural categories and human kinds: Classification in the natural and social sciences*. Cambridge, UK: Cambridge University Press.

Klaus, Georg. 1976. "Klassifikation". In Georg Klaus and Manfred Buhr eds., *Philosophisches Wörterbuch* Band 1-2. 12. Aufl. Leipzig: Bibliographisches Institut, vol. 1: 628-629.

Kuhn, Thomas S. 1962. The structure of scientific revolutions. Chicago, IL: University of Chicago Press.

Kuhn, Thomas S. 1974. "Second thoughts on paradigms". In Frederick Suppe ed., *The structure of scientific theories*. Urbana, IL: University of Illinois Press, 459-482. (Reprinted in Thomas S. Kuhn. 1977. The essential tension. Chicago, IL: University of Chicago Press, 293-319).

Kuhn, Thomas S. 2000. *The road since structure: Philosophical essays, 1970-1993*. Chicago, IL: University of Chicago Press.

Lakoff, George. 1987. *Women, fire and dangerous things: What categories reveal about the mind*. Chicago, IL: University of Chicago Press.

Lancaster, F. Wilfrid. 1998. *Indexing and abstracting in theory and practice*. Second edition. London: Library Association.

Landweh, Carl E., Alan R. Bull, John P. McDermott and William S. Choi. 1994. "A taxonomy of computer program security flaws". *ACM Computing Surveys* 26, no. 3: 211-254.

Lenzel, Victor F. 1938. "Procedures of empirical science". In Otto Neurath, Rudolf Carnap and Charles W. Morris eds., *International encyclopedia of unified science*. Chicago, IL: Chicago University Press, vol. I, no. 5.

Linnaeus, Carl. 1767. Fundamenta entomologiæ. Uppsala: John Edman. https://web.archive.org/web/ 20070713235756/http://huntbot.andrew.cmu.edu/HIBD-PDF/LinnaeanDiss/Liden-154.pdf

McGarry, Kevin. 1991. "Epilogue: Differing views of knowledge". In Arthur Jack Meadows ed., *Knowledge and communication: Essays on the information chain*. London: Library Association, 132-152.

Mai, Jens Erik. 2004. "Classification in context: Relativity, reality, and representation". *Knowledge Organization* 31, no. 1: 39-48.

Mai, Jens Erik. 2011. "The modernity of classification". *Journal of Documentation* 67, no. 4: 710-730.

Margolis, Eric. 1994. "A reassessment of the shift from the classical theory of concepts to prototype theory". *Cognition* 51, no. 1: 73-89.

Marradi, Alberto. 1990. "Classification, typology, taxonomy". *Quality and Quantity* 24, no. 2: 129-157. Prepublication [not identical] available at: http://web.archive.org/web/20040705070709/http://www.unibo.edu.ar/marradi/classqq.pdf

May, Ronald W. 1982. "Discriminant analysis in cluster analysis". In Herschel C. Hudson ed., *Classifying social data*. San Francisco, CA: Jossey-Bass, 39-55.

Mayr, Ernst. 1982. *The growth of biological thought*. Cambridge, MA: Harvard University Press.

Mayr, Ernst. 1995. Systems of ordering data. *Biology and Philosophy* 10, no. 4: 419-434.

Mayr, Ernst and Walter Joseph Bock. 2002. "Classifications and other ordering systems." *Journal of Zoological Systematics and Evolutionary Research* 40, no. 4: 169-194.

Mazzocchi, Fulvio. 2013. "Images of thought and their relation to classification: The tree and the net". *Knowledge Organization* 40, no. 6: 366-374.

Mazzocchi, Fulvio. 2017. "Relations in KOS: Is it possible to couple a common nature with different roles?" *Journal of Documentation* 73, no. 2: 368-383.

McKenna, Malcolm C. and Susan K. Bell. 1998. *Classification of mammals above the species level*. New York, NY: Columbia University Press.

Medin, Douglas L. and Cynthia Aguilar. 1999. "Categorization". In Robert A. Wilson and Frank C. Keil eds., *The MIT encyclopedia of the cognitive sciences*. Cambridge, MA: The MIT Press, 104-106.

Meinhardt, Helmuth et al. 1984. "Ordnung". In Joachim Ritter and Karlfried Gründer eds., *Historisches Wörterbuch der Philosophie*. Darmstadt: Wissenschaftliche Buchgesellschaft, Band 6, 1249-1310. (Meinhardt, Helmut, "Ordnung I Antike", 1251-1254; Hübener, Wolfgang, "Ordnung II Mittelalter", 1254-1279; Dierse, Ulrich, "Ordnung III Neuzeit", 1280-1303; Steiner, Hans-Georg, "Ordnung IV Mathematik Logik", 1303-1309; Sauer, Werner, "Ordnung der Begriffe", 1309-1310).

Miksa, Francis L. 1994. "Classification". In Wayne A. Wiegand and Donald G. Davis eds., *Encyclopedia of library history*. New York, NY: Garland Publishing, 144-153.

Miksa, Francis L. 1998. *The DDC, the universe of knowledge, and the post-modern library*. Albany, NY: Forest Press.

Mill, John Stuart. 1843. *A system of logic, ratiocinative and inductive: Being a connected view of the principles of evidence and the methods of scientific investigation*. London: John W. Parker, West Strand.

Mill, John Stuart. 1872. *A system of logic, ratiocinative and inductive: Being a connected view of the principles of evidence and the methods of scientific investigation* vol. 1-2. Eighth edition. London: Longmans, Green, Reader, and Dyer. http://www.archive.org/download/systemoflogicrat00milluoft/systemoflogicrat00milluoft.pdf

Moss, Wilfred Raymond. 1964. "Categories and relations: Origins of two classification theories". *American Documentation* 15, no. 4: 296-301.

Müller-Wille, Staffan. 2007. "Collection and collation: Theory and practice of Linnaean botany". *Studies in History and Philosophy of Biological and Biomedical Sciences* 38, no. 3: 541-562.

Müller-Wille, Staffan. 2013. "Systems and how Linnaeus looked at them in retrospect". *Annals of Science* 3: 305-317.

Murphy, Gregory L. and Douglas L. Medin. 1985. "The role of theories in conceptual coherence". *Psychological Review* 92, no. 3: 289-316.

Needham, Rodney. 1975. "Polythetic classification: Convergence and consequences". *Man*. New Series 10, no. 3: 349-369.

Nobes, Christopher and Christian Stadler. 2013. How arbitrary are international accounting classifications? Lessons from centuries of classifying in many disciplines, and experiments with IFRS data. *Accounting Organizations and Society* 38, no. 8: 573-595.

OED Online. 1990. *Oxford English dictionary: the definitive record of the English language*. Oxford, UK: Oxford University Press.

Padian, Kevin. 1999. "Charles Darwin's views of classification in theory and practice". *Systematic Biology* 48, no. 2: 352-364.

Pando, Daniel Abraao and Carlos Candido de Almeida. 2016. "Knowledge organization in the context of postmodernity from the theory of classification perspective". Knowledge Organization 43, no.2: 113-117.

Parrochia, Daniel 2016. "Classification". In James Fieser and Bradley Dowden eds., *The Internet encyclopedia of philosophy*. Martin, TN: University of Tennessee at Martin. http://www.iep.utm.edu/classifi/

Parrochia, Daniel and Pierre Neuville. 2013. *Towards a general theory of classifications*. Bäsel: Birkhaüser.

Parry, William T. and Edward A. Hacker. 1991. *Aristotelian logic*. New York, NY: State University of New York Press.

Peirce, Charles Sanders. 1931-58. *Collected Papers of Charles Sanders Peirce [CP]*. 8 vols. eds. C. Hartshorne and P. Weiss (vols. I–VI, 1931–35), A. Burks (vols. VII–VIII, 1958). Cambridge, MA: Belknap Press.

Pihlström, Sami. 2009. *Pragmatist metaphysics: An essay on the ethical grounds of ontology*. New York, NY: Continuum.

Plato. c.370 BC. *Phaedrus*. (Translated by Alexander Nehamas and Paul Woodruff eds.). Cambridge, MA: Hackett Publishing Co, Inc., 1995.

Popper, Karl Raimund. 1959. *The logic of scientific discovery*. London: Hutchinson. http://strangebeautiful.com/other-texts/popper-logic-scientific-discovery.pdf

Ranganathan, Shiyali Ramamrita. 1967. *Prolegomena to library classification*. Third edition. London: Asia Publishing House.

Reid, Thomas. 1785. "Abstraction". Reprinted in James F. Bennett ed. 1855. *Essays on the intellectual powers of man* by Thomas Reid. Sixth edition (redacted text). Boston, MA: Phillips, Sampson, and Company, 298-330. https://babel.hathitrust.org/cgi/pt?id=hvd.hwaili;view=1up;seq=17

Richardson, Ernest Cushing. 1901. *Classification, theoretical and practical: Vol. I: The order of the sciences; vol. II: The classification of books*. New York, NY: Charles Scribner's Sons. https://babel.hathitrust.org/cgi/pt?id=uiug.30112049904011;view=1up;seq=17

Rijsbergen, Cornelis Joost van. 1979. *Information retrieval*. Second edition. London: Butterworths.

Ritter, Joacim, Karlfried Gründer and Gabriel Gottfried eds. 1971-2007. *Historisches Wörterbuch der Philosophie* Bd. 1-13. Basel: Schwabe & Co.

Rosch, Eleanor. 1978. "Principles of categorization". In Eleanor Rosch and Barbara Bloom Lloyd eds., *Cognition and categorization*. Hillsdale, NJ: L. Erlbaum Associates, 27-48.

Rysiew, Patrick. 2016. "Naturalism in epistemology". In Edward N. Zalta ed., *The Stanford encyclopedia of philosophy*. Spring 2016 edition. http://plato.stanford.edu/archives/spr2016/entries/epistemology-naturalized/

Sandri, Giorgio. 1969. "On the logic of classification", *Quality & Quantity* 3, nos. 1-2: 80-124.

Sankey, Howard. 1998. Taxonomic incommensurability. *International Studies in the Philosophy of Science* 12, no. 1: 7-16.

Satija, Mohinder P. 1992. Book review of Meadows (1991): Knowledge and communication: Essays on the information chain. *International Classification* 19, no. 1: 39-41.

Satija, Mohinder P. 2000. "Library classification: An essay in terminology. *Knowledge Organization* 27, no. 4: 221-229.

Schmidt, Kjeld and Ina Wagner. 2004. "Ordering systems. Coordinative practices and artifacts in architectural design and planning". *The Journal of Collaborative Computing* 13, nos. 5-6: 349-408.

Shera, Jesse H. 1951. "Classification as the basis of bibliographic organization". In *Bibliographic Organization: Papers presented before the Fifteenth Annual Conference of the Graduate Library School July 24-29, 1950*. Ed. Jesse H. Shera and Margaret E. Egan. Chicago: University of Chicago Press, 72-93.

Shera, Jesse H. 1965. *Libraries and the organization of knowledge*. Hamden, CT: Archon Books.

Simões, Maria da Graça, M. Cristina V. de Freitas and Blanca Rodríguez-Bravo. 2016. "Theory of classification and classification in libraries and archives: Convergences and divergences". *Knowledge Organization* 43 no. 7: 530-538.

Slavic, Aida. 2000. *A definition of thesauri and classification as indexing tools: A DCMI note*. http://dublin core.org/documents/2000/11/28/thesauri-definition/

Sloan, Phillip R. 1981. "Classification". In William F. Bynum, E. Janet Browne and Roy Porter eds., *Dictionary of the history of science*. Princeton, NJ: Princeton University Press, 68-71.

Sneath, Peter Henry Andrews. 1962. "The construction of taxonomic groups". In Geoffrey Clough Ainsworth and Peter Henry Andrews Sneath eds., *Microbial classification*. Cambridge, UK: Cambridge University Press, 289-332.

Soergel, Dagobert. 2004. "Information organization". In William Sims Bainbridge ed., *Berkshire encyclopedia of human-computer interaction*. Great Barrington, MA: Berkshire Publishing Group LLC, vol. 1, 355-360.

Sokal, Robert R. and Peter H. A. Sneath. 1963. *Principles of numeric taxonomy*. San Francisco, CA: W.H. Freeman.

Spiteri, Louise F. 2008. "Concept theory and the role of conceptual coherence in assessments of similarity". *Proceedings of the American Society for Information Science and Technology* 45, no. 1: 1-12.

Stevens, Peter F. 1998. "Linnaeus, Carl von (1707-78) ". In Edward Craig ed., *Routledge encyclopedia of philosophy*, version 1.0. London: Routledge. (Electronic source, no pages). DOI: 10.4324/9780415249126-Q059-1

Stevens, Peter F. 2016. "Schools of classification". In Richard M. Kliman ed., *Encyclopedia of evolutionary biology*. San Diego, CA: California Academic Press, vol. 3, 494-498.

Suppe, Frederick. 1989. "Classification". In Erik Barnouw ed., *International encyclopedia of communications*. Oxford, UK: Oxford University Press, vol. 1, 292-296.

Sutcliffe, John Philip. 1993. "Concept, class, and category in the tradition of Aristotle". In Iven van Mechelen, James Hampton, Ryszard S. Michalski and Peter Theuns eds., *Categories and concepts*. London: Academic Press, 35-65.

Svenonius, Elaine. 2000. *The intellectual foundation of information organization*. Cambridge, MA: The MIT Press.

Svenonius, Elaine. 2004. "The epistemological foundations of knowledge representations". *Library Trends* 52, no. 3: 571-587.

Taylor, Arlene G. 1999. *The organization of information*. Englewood, CO: Libraries Unlimited, Inc.

Tiryakian, Edward A. 1968. "Typologies". In David L. Sills ed., *International encyclopedia of the social sciences*. New York, NY: Macmillan, vol. 16, 177-185.

Vandamme, Peter A. R. 2015. "Taxonomy and Classification of Bacteria". In *Manual of Clinical Microbiology*, Eleventh Edition, edited by James H. Jorgensen, Michael A. Pfaller, Karen C. Carroll, Guido Funke, Marie Louise Landry, Sandra S. Richter and David W. Warnock. Washington, DC: ASM Press (American Society for Microbiology), 255-269. doi: 10.1128/9781555817381.ch17

Veloso, Adriano and Wagner Meira Jr. 2011. *Demand-driven associative classification*. London: Springer.

Venrooij, Alex van and Vaughn Schmutz. 2015. "Classifications in popular music". In James D. Wright ed., *International encyclopedia of the social and behavioral science*. Second edition. Amsterdam: Elsevier, vol. 3, 799-804.

Venter, J. Craig, Karin Remington, John F. Heidelberg, Aaron L. Halpern, et al. 2004. "Environmental genome shotgun sequencing of the Sargasso Sea". *Science* 304 no. 5667: 66-74.

Vickery, Brian Campbell. 2008. "Faceted Classification for the Web". *Axiomathes* 18, no. 2: 145-60.

Waltinger, Ulli, Alexander Mehler, Mathias Lösch and Wolfram Horstmann. 2011. "Hierarchical classification of OAI metadata using the DDC taxonomy". In Raffaella Bernardi, Frederique Segond and Ilya Zaihrayeu eds., *Advanced language technologies for digital libraries. International Workshops on NLP4DL 2009, Viareggio, Italy, June 15, 2009 and AT4DL 2009, Trento, Italy, September 8, 2009*. Heidelberg: Springer, 29-40. (Volume 6699 of Lecture Notes in Computer Science).

Wesolek, Andrew. 2012. "Wittgensteinian support for domain analysis in classification". *Library Philosophy and Practice* (e-journal), paper 795. http://digitalcommons.unl.edu/libphilprac/795

Whewell, William. 1840. *The philosophy of the inductive sciences: Founded upon their history*. London: J.W. Parker, vol. 1-2.

Wilkins, John S. 2013. "Biological essentialism and the tidal change of natural kinds". *Science and Education* 22, no 2: 221-240. http://philpapers.org/rec/WILBEA

Wittgenstein, Ludwig. 1953. *Philosophical investigations*. Oxford, UK: Basil Blackwell.

WordNet 3.1. "Classification". http://wordnetweb.princeton.edu/perl/webwn?s=classification

[top of entry]

Appendix: A sample of definitions of classification (chronological)

The intention is to provide a comprehensive list of definitions of classification, and the idea is to update the list when new definitions are discovered in the literature. The sources for such definitions are surprisingly few and meager. The *Oxford English dictionary* provides a number of quotes back to 1767 but misses important ones such as Darwin (1859) and Huxley (1869) (see also below under 2010). *Historisches Wörterbuch der Philosophie* Bd. 1-13 (edited by Joachim Ritter) has no article "Klassifikation" or "Taxonomie" (but does have on, for example, "Kategorie, Kategorienlehre", "Ordnung" and "System, Systematik, Systematisch"). McKenna and Bell (1998, 11-33) provide an overview of the history and theory of classification and state: "The word 'classification' was not part of the scientific literature until the last decades of the eighteenth century. The first use of which we are aware occurs in a botanical paper by the Marquis de Condorcet (1777: 35)". However, earlier uses are listed below.

Plato (c.370 BC)
"Socrates: That of dividing things again by classes, where the natural joints are, and not trying to break any part, after the manner of a bad carver." (Plato c.370 BC, Phaedrus 265e)

Aristotle (350 BCE)
"A 'genus' is what is predicated in the category of essence of a number of things exhibiting differences in kind. We should treat as predicates in the category of essence all such things as it would be appropriate to mention in reply to the question, 'What is the object before you?'; as, for example, in the case of man, if asked that question, it is appropriate to say 'He is an animal'" (Aristotle 350 BCE, Book I, chap. 1, part 5).
Note: "[T]he logic of classifications, which remains, in this time, the Aristotelian logic,

receives practically no new development until the 18th century" (Parrochia 2016; electronic source, no pages; see also Sutcliffe 1993).

Michel Adanson (1763, clvi)
"...all parts and qualities, or properties and faculties of plants... barring not even one" ought to be considered before attempting a classification.

David Cranz (1767, vol. 1, ix)
"I have described what belongs to this science, not according to the classifications [Ger. *Eintheilungen*] and characteristics, which are generally adopted by, and needful for the modern naturalists, but according to a certain affinity or likeness."

Carl Linnaeus (1767, 152)
"Natura Insectorum per plures eorum ætates jam perspecta, superest ut systematice eadem contemplemur. Recta autem eorum CLASSIFICATIO vitam huic scientiæ & facultatem conciliat, ubi singula insecta suum quasi nomen ipsa produnt"
= "The nature of insects through their several ages [stages of life] having already been examined, it remains to contemplate them systematically. Now the correct CLASSIFICATION of them [sci., insects] furnishes life and means to this science, where individual insects themselves produce, as it were, their own name." (majuscule in original)

Thomas Reid (1785, 191)
"Our ability to distinguish and give names to the different attributes belonging to a single thing goes along with an ability to observe that many things have certain attributes in common while they differ in others. This enables us to put the countless hordes of individuals into a limited number of classes, which are called 'kinds' and 'sorts' — and in the scholastic language called 'genera' and 'species'" (here quoted from Frické 2012, 25).

William Whewell (1840, vol. 1, xxxiii, XCV)
"The attempts at Natural Classification are of three sorts; according as they are made by the process of *blind trial*, of *general comparison*, or of *subordination of characters*. The process of Blind Trial professes to make its classes by attention to all the characters, but without proceeding methodically. The process of General Comparison professes to enumerate all the characters, and forms its classes by the *majority*. Neither of these methods can really be carried into effect. The method of Subordination of Characters considers some characters as *more important* than others; and this method gives more consistent results than the others. This method, however, does not depend upon the Idea of Likeness only, but introduces the Idea of Organization or Function" (italics in original).

John Stuart Mill (1843, vol. 2, Bk. IV, Ch. 7, 299-300)
"Classification, thus regarded, is a contrivance for the best possible ordering of the ideas of objects in our minds; for causing the ideas to accompany or succeed one another in such a way as shall give us the greatest command over our knowledge already acquired, and lead more directly to the acquisition of more. The general problem of Classification, in reference to these purposes, may be stated as follows: To provide that things shall be thought of in such groups, and those groups in such an order, as will best conduce to the remembrance and to the ascertainment of their laws."

William Benjamin Carpenter (1847, I. §2)
"The object of all Classification ... [is] to bring together those beings which most resemble each other and to separate those that differ."

Charles Darwin (1859, 420)
"... all true classification is genealogical ..."
(See also Mayr and Bock 2002; Padian 1999.)

Thomas Henry Huxley (1869, 1)
"By the classification of any series of objects is meant the actual, or ideal, arrangement together of those which are like and the separation of those which are unlike, the purpose of this arrangement being to facilitate the operation of the mind in clearly conceiving and retaining in the memory, the characters of the objects in question."

Charles Ammi Cutter (1876, 10)
"Class, a collection of objects having characteristics in common".

Robert Adamson (1901, vol. 1, 185)
"The process of arranging the objects of some province of experience into kinds or groups, characterized by the possession of common marks.
As ordinarily defined, it involves more than logical DIVISION (q.v.), the rules of which furnish the minimal conditions of the process. In addition, classification takes into account (1) either the specific purpose of the arrangement, or (2) the natural conjunctions of marks which are of most importance. In either case, the aim of classification is to render possible the greatest number of general propositions regarding the objects, and so to facilitate the complete and systematic survey of them. The ideal of a classification that is not determined by special, human ends, as e.g. in classification of occupations in a census return, is to copy in its systematic arrangement the real order of interdependence in the things themselves. What is called 'artificial,' as opposed to natural classification, differs in degree only, not in kind.
Literature: MILL, Logic, Bk. IV. chaps. vii, viii; VENN, Empirical Logic, chap. xxx; JEVONS, Princ. of Sci., chap. xxx. (R.A.)."

Ernest Cushing Richardson (1901, 1)
Classification is the "putting together of like things, or more fully described, it is the arranging of things according to likeness and unlikeness. It may also be expressed as the sorting and grouping of things. It is convenient sometimes, to speak of ‘likeness and unlikeness' but really in classification it is 'likeness' which rules while 'unlikeness' is merely what is left over when likeness has been defined. The 'putting together of like things' is therefore the fullest and most exact form of the definition."

Charles Sanders Peirce (1902/1931, Vol. 1, § 231, p. 103)
"All classification, whether artificial or natural, is the arrangement of objects according to ideas. A natural classification is the arrangement of them according to those ideas from which their existence results. No greater merit can a taxonomist have than that of having his eyes open to the ideas in nature; no more deplorable blindness can affect him than that of not seeing that there are ideas in nature which determine the existence of objects."

Henry E. Bliss (1935, 3)
"In dealing with the multiplicity of particular things, actualities, and specific kinds, we find

that some are alike, in general characters and in specific characteristics; and we may consequently relate them in a class, or classes, that is classify them."

Jason E. L. Farradane (1950, 83).
"Classification is a theory of the structure of knowledge, i.e. of the relations between different parts of knowledge. No arbitrary method of grouping, however carefully applied, is true classification. The problem is primarily epistemological. What is true knowledge, and what are true relations between the parts of knowledge? It is essential to define these if the classification is to be true and logically sound."

Jason E. L. Farradane (1950, 87).
"A classification indicates the relations between items of knowledge."

Jason E. L. Farradane (1952, 73-74).
"A classification represents a theory of the structure of knowledge, i.e. of the relations between the different parts of knowledge. The basic problem is to determine what these relations are and how they link the different concepts from our knowledge into a coherent structure. The arbitrary or 'deductive' subdivision of an assumed total of knowledge cannot give a true representation of these relations, which do not consist only of groupings of a class and its members, or division of a whole into its parts. It was shown that a classification must be constructed 'inductively', or upwards, piecing together known fragments of relations."

Jesse H. Shera (1965, 120)
"Classification is the crystallization or formalizing of inferential thinking, born of sensory perception, conditioned by the operation of the human brain, and shaped by human experience. It lies at the foundation of all thought, but it is pragmatic and it is instrumental."

Jesse H. Shera (1965, 127)
"He [the librarian] must appreciate classification, not as a tool, but as a discipline in which is to be studied the reaction and response of a living mind to the record left by a distant and usually unknown mind; a discipline that seeks to achieve a better understanding of the changing patterns of thought and the points of contact at which they can be related to specific units of recorded information."
(Drucker 2014 is, however, a work demonstrating that this demand is not specific to the librarian but describes, for example, the biologist and sexologist Alfred Kinsey equally well.)

Shiyali Ramamrita Ranganathan (1967, Chapter CP, 77-78)
Ranganathan distinguished five senses of 'classification'.
1. "Division. (See Chap CC).
This is the primitive meaning of the term 'classification'. Even a child practises classification in Sense 1 with its playthings. Even early man had practised it.
2. Assortment [grouping of things of the same sort]. (See Chap CD).
Classification in Sense 2 is inherent in Man. Perhaps it is a concomitant of the finiteness of the speed of neural impulses in the human body. When the speed is finite, structure emerges. Wherever there is a structure, sequences emerge. When sequence is helpful to the purpose at hand, it is Classification. The sequence inevitable inside of the skin, so to speak, gets expressed extraneurally also. To classify in Sense 2 is thus a neural necessity. Sharpness in thinking, clarity in expression, unerringness in communication, expedition in response, and exactness in service depend ultimately on helpful sequence or Classification in Sense 2. The work of philosophers and of taxonomists in the field of classification is generally restricted to

Classification in Sense 2. (See Chap CL).
3. Classification in sense 2 *plus* Representing each entity by an ordinal number taken out of a system of ordinal numbers, designed to mechanise the maintenance of the sequence,

1. Either when an entity has to be replaced after having been taken out of its position;
2. Or when a new entity has to be interpolated or extrapolated in the correct place in the sequence.

This ordinal number is the Class Number. (*See* Chap CG and CM).
Classification in Sense 3 is usually practised by large business concerns having to handle a large number of commodities. The Customs Authorities too use Classification in Sense 3 in their published list of commodities liable to customs duty.
4. Classification in Sense 3 when complete assortment is made of an amplified universe — that is, when the entities and the pseudo-entities arising in the process of successive assortment stand arranged in one filiatory sequence, each with its Class Number. (*See* Chap CH and CK).
Classification in Sense 4 is not used very much. It is only classification in Sense 3 and Sense 5 that are frequently in demand.
5. Classification in Sense 4 with all the entities removed but only the pseudo-entities or classes retained — each class having the number representing it. (*See* Chap CM).
It is classification in Sense 5 that is used,

1. Either when the universe classified is infinite;
2. Or when some of the entities are unknown and unknowable at any moment, even though the universe classified is finite.

In particular, it is Classification in Sense 5 that is practised by the library profession.
It should be recalled that that in classification in Sense 5

1. The individual entities do not figure in the complete assortment;
2. Classes take the place of entities ; and thus,
3. Each class including the Original Universe is a Class of Classes.

Classification in Sense 2 has only a Scheme for Classes associated with it. But Classification in Sense 5 has a Scheme for Classification associated with it. We shall restrict the meaning of the term 'Classification' to Classification in Sense 5" (italics in original).

Bonifatii Mikhailovich Kedrow (1975, vol. 1, 3)
"Der Klassifizierung der Wissenschaften bedeutet den *Zusammenhang der Wissenschaften*, der in ihrer Stellung in einer bestimmten Reihenfolge oder in einem System entsprechend einigen allgemeinen Grundsätzen zum Ausdruck kommt" (italics in original).
This can perhaps be generalized in this way: Classification of objects means the display of connections between the objects in a certain order or in a system reflecting certain basic principles. Kedrow found that the principle of historicism must govern all natural classifications (see section 4.2c γ Genetic/historicist approaches to classification).

Georg Klaus (1976, 628-629)
"Klassifikation. Verfahren zur Unterteilung eine Klasse K von Dingen usw. in Teilklassen" (628). "Die dialektisch-materialistische Einstellung zur Klassifikation besteht also nicht etwa darin, dass sie im Gegensatz zur antidialektischen die Berechtigung und den Wert von

Klassifikationen bestritte, sondern darin, dass sie die These von der zeitlichen und strukturellen Relativität der Klassifikationen vertritt, während die antidialektische Auffassung die Klassifikationen in jeder Hinsicht als absolut betrachtet."

Phillip R. Sloan (1981, 68)
"The arrangement of objects or entities into groups or classes, usually on the basis of perceived similarity and difference."

ISO 5127-6 (1988, 93)
A classification system is an "indexing language intended for a structured representation of documents or data, through the use of indexes and corresponding terms, in order to allow systematic access, resorting to an alphabetical index, if necessary" (here cited from Simões et al. 2016, 531; note that this standard has been revised by ISO 5127:2001).

Lois Mai Chan (1994, 259)
Classification is "the multistage process of deciding on a property or characteristic of interest, distinguishing things or objects that possesses that property from those which lack it, and grouping things or objects that have the property or characteristic in common into a class. Other essential aspects of classification are establishing relationships among classes and making distinctions within classes to arrive at subclasses and finer divisions."

Francis L. Miksa (1994, 144)
"[Bibliographic] Classification is the activity of *creating categories* into which bibliographic items of all kinds may be placed (i.e., the work of the classificationist) and also the activity of *identifying* bibliographical items in terms of the categories already extant in a given system (i.e., the work of the classifier). It encompasses systems for arranging items on the shelves of libraries (sometimes called 'bibliothecal' classification), as well as systems for arranging the surrogates of items in catalogs (sometimes called 'bibliographical' classification). It includes classificatory systems based on all kinds of item characteristics (subject, form, author, citation, size, etc.), in all forms of order (logical and systematic, alphabetical, faceted, etc.), with all kinds of operating methods (pre- and post-coordinated, statistically based clustering and identification, etc.), and differing in scope from the universal to the very narrow. Finally, library classification embraces a wide range of purposes, although most often its chief purpose has been to facilitate document retrieval."

Frederick Wilfrid Lancaster (1998, 17)
Classification is "sorting items into 'conceptual classes'" and "forming classes of objects on the basis of their subject matter".

Geoffrey Bowker and Susan Leigh Star (1999, 10)
"A classification is a spatial, temporal, or spatio-temporal segmentation of the world. A 'classification system' is a set of boxes (metaphorical or literal) into which things can be placed to perform some kind of work — bureaucratic or knowledge production."

Arlene G. Taylor (1999, 237)
"The placing of subjects into categories; in organizing of information, classification is the process of determining where an information package fits into a given hierarchy and then assigning the notation associated with the appropriate level of the hierarchy to the information package and its surrogate."

Satija (2000, 222)
"Classification means to divide objects/entities (both abstract and concrete) on the basis of their differences or, conversely, the grouping of entities on the basis of their similarities. Classification is any process of dividing, sorting, grouping, arranging, ordering, ranking, mapping and correlating."

Elaine Svenonius (2000, 10)
"Organization can take many forms. Its prototypical form is classification. Classification brings like things together. In traditional classifications, like things are brought together with respect to one or more specified attributes. Any number of attributes can be used to form classes of documents embodying information, such as same size or color, same subject, or same author. However, the most important attribute for a system whose objectives is to organize information is the attribute of 'embodying the same work'. "

Hubert Feger (2001, 1966)
"Classification is the assignment of objects to classes".
Later on, this was expanded: "The fundamental goal of classification is to find structures common to a group of objects, using properties to classify the objects into subgroups based on the similarity of their properties" (Feger 2015, 805).

The Portuguese Language Dictionary of the Academy of Sciences (Academia das Ciências de Lisboa 2001, 837).
Classification is the "action of distributing in classes, by categories ... according to precise criteria" (cited from Simões et al. 2016, 531).

Ernst Mayr and Walter Joseph Bock (2002)
"The logical consequence of the definition of class is that classification must be defined as the ordering of diversity into classes of similar entities. And this has been traditionally the almost universally accepted concept of classification. [...] A classification of organisms is based on the shared possession of their diverse attributes. The units of similarity in a Darwinian classification are called taxonomic characters that have the property of being homologous to one another in the several entities or groups. The claim of a few modern authors that there is no agreement on the definition of the word 'classification' is quite misleading. Actually, prior to 1950 there was virtually total unanimity on the usage (in classification) of the words classification and class, as referring to the grouping of similar items.
A classification is defined as "The arrangement of entities in a hierarchical series of nested classes, in which similar or related classes at one hierarchical level are combined comprehensively into more inclusive classes at the next higher level" (176).
"Class — (in classification) A group or collection of entities (individuals), possessing attributes or traits in common ('being similar'), a kind or sort, grouped together under a general or class name.
Classification — The arrangement of similar entities (objects) in a hierarchical series of nested classes, in which each more inclusive higher-level class is subdivided comprehensively into less inclusive classes at the next lower level.
Darwinian classification — The ordered grouping of organisms into classes, according to their similarities and consistent with their inferred evolutionary history.
Downward classification — Establishing groups by logical division.
Evolutionary classification — A classification that duly considers both evolutionary processes, the ecological adaptiveness of evolutionary divergence (degree of difference) and the genealogy (phylogeny) of the taxa. Basically equal to a Darwinian classification.

Hierarchical classification — The system of ranks that indicates the categorical level (level of difference) of each taxon" (191).

Kjeld Schmidt and Ina Wagner (2004, 392)
"Classification, in turn, is a special practice of categorization, involving pre-established and systematic systems of signs. That is, classification is a linguistic operation of applying a classification scheme, i.e., an ordered set of signs that is pre-established according to (a) some general principles and criteria of ordering and (b) some procedures of identification and naming. In short, an act of classification is an application of a classification scheme. Classification systems (such as thesauri) can thus be seen as instantiations of classification schemes.
[...] Classifications and categorizations are both convention-based practices and equally so. But classifications are convention-based in a quite specific sense. In the case of categorization there are no pre-established principles and criteria for determining the correctness of an act of categorization. With acts of classification, however, such pre-established principles and criteria exist, in that they specify relationships between items in terms of, for example, class/membership, part/whole, composition, cause/effect, origin/fate, function, ownership, value/risk, location, or state. Accordingly, an actor applying a classification scheme in a particular case can be held accountable in terms of the principles, criteria, and procedures of the classification scheme."

Dagobert Soergel (2004, 358)
"A classification is a structure that organizes concepts into a meaningful hierarchy, possibly in a scheme of facets. The classification of living things is a taxonomy. (The term taxonomy is increasingly used for any type of classification.) A classification is now often called an ontology, particularly if it gives richer concept relationships" (bold in original omitted here).

Faria and Pericão (2008, 258)
Classification is a "group of ordered concepts, distributed systematically in classes, forming a structure" and a "structuring of concepts into classes and subdivisions to express the existing semantic relationships between them" (here cited from Simões et al. 2016, 531).

Brian Campbell Vickery (2008, 145-6)
"A classification is a hierarchical structure of symbolic terms. The symbols point to, or represent, entities in the real or an imagined world. The structure may be used as a ‘‘finding aid’’, or information retrieval tool—to pinpoint a particular sought entity."

Clare Beghtol (2010, 1045)
"To classify means to put things into meaningful groups. Things can be physical objects, ideas, events, or anything else that human beings can perceive or imagine, and a meaningful group can be formed using any characteristic or combination of characteristics of the things. Groups can be considered to be permanent or they can be considered temporary responses to a need of the moment."

Ingetraut Dahlberg (2010, 2941)
"With this journal [*International classification*, 1974-1992, thereafter *Knowledge organization*], 'classification' was understood as a multi-meaning word that includes the following concepts:

1. classification in the sense of 'classification system', i.e., a system of classes arranged in hierarchical or faceted order;
2. classification in the sense of classifying, i.e., establishing a system of classes;
3. classification in the sense of classing, i.e., relating the classes of a classification system to objects or subjects of reality; and
4. classification in the sense of classification science, i.e., relating to this field of study and its activities" (list typography added).

Oxford English Dictionary (update from 2010)
"classification, n.
Origin: A borrowing from Latin. Etymon: Latin classificatio.
Etymology: < post-classical Latin classificatio (1673 in a German source; 1767 in Linnaeus) > classical Latin classis class n. + -ficātiōn- , -ficātiō suffix. Compare German Klassifikation (1760 or earlier as †Classification), Swedish klassifikation (1740 as †classification), Danish klassifikation (1748), French classification (1780), Italian classificazione (1796). Compare slightly later classify v.

1. The result of classifying; a systematic distribution, allocation, or arrangement of things in a number of distinct classes, according to shared characteristics or perceived or deduced affinities. Also: a system or method for classifying.
2. The action of classifying or arranging in classes, according to shared characteristics or perceived affinities; assignment to an appropriate class or classes.
3. A category to which something is assigned; a class."

Adriano Veloso and Wagner Meira (2011, 9)
"In a classification problem, there is a set of input-output pairs (also referred to as instances or examples) of the form $z_i = (x_i; y_i)$: Each input x_i is a fixed-length record of the form $\langle a_1, \ldots, a_l \rangle$; where each ai is an attribute-value. Each output y_i draws its value from a discrete and finite set of possibilities $y = \{c_1, \ldots, c_p\}$, and indicates the class to which z_i belongs. Cases where $y_i = ?$ indicate that the correct class of z_i is unknown. There is a fixed but unknown conditional probability distribution $P(y|x)$, that is, the relationship between inputs and outputs is fixed but unknown"
"This formulation implies that the classification problem corresponds to the problem of function approximation." (Veloso and Meira 2011, 11).

Daniel Parrochia and Pierre Neuville (2013, 21)
"Definition 1.9.1 We call 'classification' the operation consisting of sharing, distributing or allocating objects in classes or groups which are, in general, less numerous than them. [10] This is also the result of this operation. We want, as much as it is possible, this result to be constant, i.e. the classification must remain stable for a little transformation [11] of data".
Note 10: "In the case of infinite classifications, this requirement, of course, must be weakened: we may only want the (infinite) cardinal of the classification to be less than or equal to the (infinite) cardinal of the set of objects to be classified."
Note 11: "The sense of it will have to become clearer."

WordNet (3.1) [downloaded 2016-05-21] defines four senses of the noun 'classification' of which three are relevant for this entry:

- "S: (n) categorization, categorisation, classification, compartmentalization, compartmentalisation, assortment (the act of distributing things into classes or categories of the same type)
- S: (n) classification, categorization, categorisation (a group of people or things arranged by class or category)
- S: (n) classification, categorization, categorisation, sorting (the basic cognitive process of arranging into classes or categories)".

[top of entry]

Version 1.8 (= Version 1.3 plus section 3.6 and quotations from Adams and Adams (1991), Peirce 1902, Shera (1951), Vandamme (2015) and Vickery (2008); version 1.0 published 2017-02-09, this version 2019-10-30.
Article category: Core concepts in KO

This article (version 1.0) is also published in *Knowledge Organization*. How to cite it: Hjørland, Birger. 2017. "Classification". *Knowledge Organization* 44, no. 2: 97-128. Also available in *ISKO Encyclopedia of Knowledge Organization*, eds. Birger Hjørland and Claudio Gnoli, http://www.isko.org/cyclo/classification