Chapter 4
Phenomenon Versus Discipline-Based Classification

Previous chapters have developed an argument that interdisciplinary scholarship (in particular) would benefit from an approach to classification grounded in phenomena rather than disciplines. It might seem that such a recommendation represents a strong break with traditions in the field of classification research. Yet this is not the case. This chapter begins with a discussion of the historical emergence of discipline-based classifications. It then proceeds to examine a tradition within the field of knowledge organization of urging a phenomenon-based approach, and illustrates contemporary projects that are advancing in this direction. It discusses why such KOSs have not previously been widely adopted.

Disciplinary Libraries

In previous centuries, libraries were often organized by disciplines. In many contexts, this only required a very broad organization, consisting of few main classes, as the number of volumes was relatively limited, so that each one could be found within its broad group quite easily. For example, some rooms of Giacomo Leopardi's house in Recanati are covered by shelves where, at the beginning of the nineteenth century, the young poet famously used to spend many hours studying alone. On the top of each shelf is an oval sign with Roman numbers and corresponding disciplines, like sacred history, secular history, literary history, theology, and so on. The disciplines clearly reflected the culture of the time. However, this kind of system is still used in many libraries of relatively small size, both public and private, as well as in many bookshops.

From the nineteenth century on, as public and academic libraries grew to include increasingly large numbers of volumes, it became necessary to organize and catalog them in more detail. Librarians like Antonio Panizzi in London and Charles Ammi Cutter in Boston developed explicit rules for cataloging. At Amherst College, in Massachusetts, Melvil Dewey devised his Dewey Decimal System (DDC) to

R. Szostak et al., *Interdisciplinary Knowledge Organization*,
DOI 10.1007/978-3-319-30148-8_4

express the subject of any book as part of a rich hierarchy of disciplines and subdisciplines. This system allowed for potentially infinite specificity, and was also adopted with minor modifications in Europe for the Universal Decimal Classification (UDC). Other classificationists developed their own schemes on the basis of similar principles.

Classification thus meant arranging disciplines, as opposed to directly listing the objects of knowledge, expressed by controlled terms, as was done in dictionary entries and verbal subject heading systems. While by the Library of Congress Subject Headings one can express the fact that a book is about *animals*, by UDC one can say that it belongs to class *59* standing for 'zoology,' a subdivision of class *5* 'natural sciences' (a similar analysis holds for the Library of Congress Classification (LCC), DDC, and others). Indeed, these classifications are said to be based on *aspect*, the scientific perspective by which the subject matter is treated, rather than on the very phenomena discussed in the text (Broughton 2004, 18; Slavic 2007). Disciplinary classifications have remained the standard approach in twentieth-century libraries: probably this was due to an inertial process boosted by the economic advantages of sharing existing, well-known systems, especially by deriving ready-made classified records from the catalogs of such big institutions as the Library of Congress or the British Library.

This chosen path has had important consequences. Not all books dealing with animals will be grouped in class *59*. Indeed, books on animal breeding will instead be filed under *636* animal husbandry, being part of agriculture, in turn part of applied sciences; books on animal health will be filed under *614.9* veterinary hygiene, a part of medical sciences; animals as a means of transport will be filed under *656.1* road transport, part of communication industries; not to mention that paintings of animals will be filed under *75* painting, part of the arts. In other words, the phenomenon of animals will be scattered in many parts of the scheme, depending on the disciplinary perspective by which it is treated in each case. Current handbooks for classifiers explicitly state that this practice is the regular one to be adopted.

However, as we explore a disciplinary tree in more and more detail, we usually reach a point where the subclasses are illustrated by terms referring to phenomena rather than disciplines. Zoology in UDC is divided into the subdisciplines of animal physiology, ethology, systematic zoology, and so on; systematic zoology includes *599* zoology of mammals, whose technical disciplinary names 'mammalogy' or 'theriology' would be hardly known by non-experts; going on, a subclass like *599.4* will be described more often as the class of bats than as that of chiropterology, and the further subdivisions of bats would probably lack any name for a corresponding subdiscipline. We have thus gradually moved from names of disciplines to names of phenomena.

Current UDC principles are based on facet analysis, a method not yet available when the first editions of the UDC were published. It has become the UDC practice (although not published in detail yet) that the first digits of a class represent subdisciplines, while the following digits represent the object facet of that discipline. That is, in the case of zoology, the groups of animals forming its object of

study: *599.4* would thus properly mean 'systematic zoology of bats' rather than 'chiropterology.' On the other hand, no symbol makes it clear at which point in notation we move from a subdiscipline to its object facet, while the number of digits expressing subdisciplines may well vary from one class to another. Also, assuming that systematic zoology is the last disciplinary subdivision, followed by its objects, implies that there is no notational space remaining for further specifying kinds of systematic zoology (that is, to identify subdisciplines in more detail) should this be needed in the future.

The sort of implicit structures described above reflect progressive adaptations over time to systems with a very long history, and that were originally conceived with a disciplinary, non-faceted structure. Although for many practical cases DDC and UDC work well (no doubt, better than no standard classification at all), they cannot be taken as a reference when discussing the best principles needed for interdisciplinary knowledge organization.[1]

The choice of disciplines as the structural basis of the major bibliographic classifications comes from their history of following an academic approach. It is usually justified by the rationale that scholars will find it convenient to have grouped together all books sharing the same disciplinary approach, rather than all books sharing the same objects of study. Clearly, indeed, taking one of the two as the primary subdivision will make the other scattered in different points of the collection. Librarians have generally assumed that scattering disciplines in order to group phenomena would be a major disadvantage for their users. Is this true? We have doubted this conclusion in preceding chapters.

What is surely true is that researchers are now *used* to finding documents grouped by discipline. So the disciplinary approach in the organization of knowledge can make them feel more comfortable with this established order of knowledge. Knowledge organization tends to represent existing, consolidated orders, rather than playing a more active role by suggesting new orders. For the purposes of interdisciplinarity, this is clearly an obstacle to the creation of innovative paths across different areas of knowledge.

Furthermore, the usual approach tends to privilege academic disciplines over domains of interest to general users. In real life, many users are less interested in Germanic literature than they are in reading a romantic novel situated in the nineteenth century, or an account of a journey across Asia; many need less to learn about motor engineering than to look at the features of a car they are considering to order through a website. Leisure is an important part of life that is poorly accounted for by traditional knowledge organization, although people increasingly use information sources regarding it (Hartel 2003). So disciplinary knowledge organization risks being useful only for a limited segment of users—

[1] Research on the updating and development of traditional schemes can happily advance in parallel with more general and speculative research on new systems: indeed, one of the authors is proud to work for the UDC Editorial Board.

maybe even the segment less in need of it, as researchers often know about the basic references in their own field already, as has often been observed.

Economic historians use the phrase 'path dependence' to refer to the importance of contingency in history. Typewriter manufacturers settled on a keyboard layout in the nineteenth century that made sense at a time when mechanical realities limited typing speeds (there is even some speculation that manufacturers may have been trying to slow typing speeds to reduce the frequency of jammed keys). It may be that a keyboard layout with the most-used letters in the middle could accelerate typing speeds significantly, but more than a century of training people on the old keyboard makes the switch difficult. It is not impossible to change paths: there is software that allows computers to utilize a different keyboard layout, and over time people might be trained on a different configuration. Likewise we may yet see the development of electric cars, but this transformation is made difficult by the century of development of technology and infrastructure to serve gasoline-powered automobiles (whereas in 1900 it was unclear which technology was superior).

The insight for information science is that once a path is chosen it develops its own momentum. Paths may be changed, but not easily. A classification system, once in place, tends to cope with new subjects in the manner that requires the least change to the system (see Chap. 7). Changes to a system that has long been in use must mean that works classified in the old way either become hard to find or must be reclassified into the new system. The conditions—disciplinary hegemony and reliance on card catalogues—that encouraged discipline-based classification in the nineteenth century have changed, just as the conditions that supported the 'QWERTY' keyboard layout have changed. If we had inherited a keyboard with the "a" and "e" in the middle, we would take this for granted. Likewise we might well have spent the last century happily fleshing out the sort of classification urged in this book.

Previous Attempts at Classifying by Phenomena

Despite the predominance of disciplines, the alternative approach of classification by phenomena was already clear in the early times of modern library classification to Paul Otlet, the pioneer of many advanced ideas in information science:

> The objects of knowledge [. . .] can be considered from two points of view. From the first point of view they can be envisaged as complete in themselves, as autonomous, as a totality, as a concrete whole. From the second point of view, they are envisaged in terms of their relations with other objects or as parts of an abstract entity. [. . .] To be complete, a classification should, therefore, enumerate both the objects and the points of view and choose as the basis of classification a sequence of one or the other as needs be. (Otlet 1896)[2]

[2] We are indebted to Thomas M. Dousa for pointing us to this source and its relevance.

Still, in practice Otlet employed the disciplinary option, which was ready-made in the DDC adopted as the main source for his UDC. A few years later, Julius Otto Kaiser (1911, para 209) recommended that users of his Systematic Indexing verbal system should focus on the phenomena studied and avoid '-ology' terms, although he did not apply this suggestion to classification.

Another early explorer of alternatives to canonical disciplines was British librarian James Duff Brown. Already a century ago, he proposed a Subject Classification, in which the main filing classes were 'concrete subjects' rather than disciplines, although the latter were still taken as main classes. This choice was aimed at placing all works dealing with one object, such as roses or coffee, in the same place independently from the 'standpoints' by which the object was considered. Subjects could then be subdivided by a Categorical Table anticipating to some extent (as was the case with Kaiser) the facets later introduced by Ranganathan. Brown's Subject Classification was also original in allowing for notational synthesis of subjects taken from different main classes, thus facilitating interdisciplinarity (Brown 1906; Beghtol 2004). Brown's system is no longer used today, but is an interesting precursor of non-disciplinary classifications.

The non-disciplinary option was then investigated in more direct and deep ways by the Classification Research Group (CRG), a sort of club of original thinkers in the domain of classification who regularly met in London from the 1950s. After developing various faceted systems focused on special domains, the CRG was granted funding by NATO in order to investigate the possibility and basic structure of a new general faceted classification system. Some members, including Foskett, Kyle, Farradane, Austin and Coates, supported the idea of building the system on main classes of entities, properties and activities, rather than on disciplines. Although unusual in classification, this was common in verbal subject heading lists and now in thesauri, a new kind of KOS inspired by linguistic tools like Roget's 1852 *Thesaurus of English words and phrases*. Indeed, terms listed in thesauri, like descriptors in alphabetical subject headings, refer to such individual concepts as (especially) objects, processes or events, rather than to disciplinary fields.

CRG members drafted main classes for their new general system, and discussed its basic principles in papers (Austin 1969; CRG 1969) as well as in their meeting bulletins:

> Mr. Farradane [. . .] said that he thought it had been accepted that the logical way to build a classification was to start from the individual concepts and build the classification through study of the relationships between them. To revert to accepting disciplines as main classes and classifying by subdivision would lead back to old confusions, since the content of 'disciplines' changes with time while the objects to which they relate, the natural entities and man, do not. [. . .] In non-scientific subjects the real problem was one of definition—to determine what the writer really meant by the words used. [. . .] Disciplines can be *derived* from various ways of observing and presenting phenomena, and some might prove to be those proposed by Mr. Langridge, but this cannot be accepted *a priori*. (CRG 1973)
>
> Coates thinks that it is dangerous to regard a structure based on division by disciplines as superior to any other; historically, people devising schemes of classification have always begun by looking at disciplines and have compromised by looking at literature. But

> disciplines are always disorderly and the general feeling of the Group was that there is not agreement on what constitutes a discipline. Langridge, however, pointed out that it is essential to distinguish between fundamental disciplines, which can clearly be recognized, and their main subdivisions, which is where the area of disagreement arises. (CRG 1978)

The new CRG approach was quite revolutionary, in that notation for a document was not just taken from a hierarchical list of classes, but could be obtained by combination of single concepts, each representing an entity or a property or an activity, by means of *operators* specifying the kind of relationship between them. 'Washing of bottles' could thus be represented as *V67(5)Z96* by connecting notations for bottles and for washing through the effect operator *5*.

Of course, combination in itself is also possible with disciplinary classifications. Ranganathan's *subject device* allowed the classifier to specify a disciplinary class by another disciplinary class written in brackets after it. Also, two disciplinary classes could be in *phase relationship*, to express meanings like 'comparison between philosophy and religion,' or 'influence of geography on history,' or 'application of mathematics to aircraft engineering.'[3] UDC especially makes extensive use of combinations between concepts from different disciplinary classes. For example, ethics as a part of philosophy can be combined with such concepts as smoking, suicide, or abortion to class books discussing ethical views on these subjects. However, these concepts have to be taken in turn from a given disciplinary class, which implies a certain disciplinary meaning: if abortion is found under medicine, the combination will mean 'ethics of abortion-as-a-medical-practice,' and if suicide is found under law, the combination will carry a legal aspect with it. In these systems, it is not possible to refer to suicide or abortion as simple objects in themselves, expressing their connotations only by the occasional combination. What would be needed for this is what Farradane called a *place of unique definition*, identifying a concept according to its nature rather than under one disciplinary perspective or another. This can only be achieved with a classification of phenomena.

Unfortunately the CRG project came to an end without producing any final version of the system. Members more supportive of the disciplinary approach, like Mills and Langridge, applied the previous research to the construction of another disciplinary system (the second edition of the Bliss Classification); although this classification also provided initial classes for phenomena treated in an interdisciplinary way, these have never been developed (Gnoli 2005). However, the idea was there. Derek Austin was hired by the British Library, where he implemented a verbal system based on the same principles to combine concepts by operators, the Preserved Context Index System (PRECIS; see Austin 1984): a very advanced tool

[3] Satija (1979) appreciates that interdisciplinary studies have made multi-phased subjects inevitable. He briefly deals with their provision in several classification schemes. Though the Colon Classification is equipped through its phase relations to deal precisely with such subjects, there are only two phases in a complex class. He suggests that through a logical extension of the phase relation rules, however, the class number for a complex class of any order may be synthesised in Colon.

based on linguistic theory and general systems theory which was applied for years to the British National Bibliography.

The CRG was not alone. De Grolier (1962) surveyed dozens of efforts at that time (including the CRG) in many countries to develop 'general categories of phenomena.' For example, Gardin in France started first in archaeology and tried to develop general categories for defining, say, the shape of a knife handle, or decorations on pottery, then moved on to pictures and texts, and developed a classification of verbs. De Grolier attributes these various efforts to the introduction of 'machines' which meant that complex terms needed to be broken into simpler terms, hierarchy should be reduced, and there should be increased emphasis on relations between terms (de Grolier 1962, 10). Yet he notes that such efforts—he refers to the CRG in particular—were hard to combine into a general classification (99). De Grolier had tried to convince UDC to adopt a couple dozen general relator terms, including one for causation, but they balked, likely because such terms infused the schedules. Perrault (1969) also urged the integration of relationship terms into UDC. Donker Duyvis of UDC noted that it is easy to subdivide but difficult to give broader meaning to an existing division in a schedule (de Grolier 1962, 42). De Grolier concluded that it was not possible to renovate existing schemes beyond a certain point (de Grolier 1962, 20–1). A common auxiliary table for *Relations, Processes and Operators* has been introduced into UDC more recently (McIlwaine 2007, 87) but involves lengthy notations and is not applied widely.

Meanwhile, other classificationists were coming to similar conclusions concerning the definition of main classes by phenomena instead of disciplines. Ingetraut Dahlberg (1974, 1978) agreed with Farradane that with disciplinary systems 'difficulties were experienced whenever the same objects were treated by different disciplines,' and so resolved 'to separate the main objects and their aspect fields (subject fields, disciplines) from each other establishing thus the first two fundamental categories.' Her Information Coding Classification first lists ten classes of general objects, and then derives disciplines by applying a set of nine perspective categories to each of them. Her work thus foreshadows the approach recommended in this book not just with respect to classifying phenomena but perspective also; it differs though in still stressing disciplines (Dahlberg 2009). Martin Scheele (1977, 1983) tested a Universal Faceted Classification of phenomena where 'all notations are freely combinable among one another.' While discussing a project of classification for community information, Robin A.B. Bonner (1982) remarked that people needing to know what to do when their spouse dies are not interested in disciplines but in practical information, ranging across political sciences, religion, psychology, economics, medicine, and so on (much like in present-day e-government websites). A.A. Shpackov (1992) developed a Universal Classification of objects and their attributes as separate from research approaches. Brian Vickery (2008) suggested that 'there could be two schedules, one listing phenomena in all their variety, the other listing "viewpoints" (or preferably, activities) in all their variety, so that each set of concepts has the same freedom. Each human activity can then, in principle, be applied to any phenomenon (e.g. we can sell anything).'All these voices encourage

classificationists to separate the dimension of phenomena from that of perspectives, whatever their preferred order. This is also recommended in the León Manifesto already mentioned in Chap. 1, which indeed emerged from a conference focused on interdisciplinarity in knowledge organization.[4]

Why Not Before?

If a KOS focused on phenomena is as useful as was suggested in Chaps. 2 and 3, why has it not already been developed and applied? Of course, this question can be asked of any innovation before it is successful. Yet the fact that general bibliographic classifications have not classified documents in the past in terms of theory and method applied, or in terms of a general classification of phenomena studied, suggests that these enterprises might be (too) difficult. We have seen above that many classificationists worked toward the development of phenomenon-based general classifications, and so ignorance of the very possibility of such an approach to classification cannot be the explanation. Weinberg (1988) herself, despite her concerns, made no recommendations for change, feeling that efforts to classify in terms of theory and method applied would be too complicated. However, several alternative explanations can be provided. These are listed in Table 4.1.

The last point merits further discussion. As noted in the Preface, this book is aimed at both information scientists and scholars of interdisciplinarity. It might be thought that the latter may have little interest in the details of KOSs. They can appreciate the interdisciplinary needs outlined in the first two chapters but may be quite happy to leave the details of their satisfaction to others. But KOSs are a major—in all likelihood *the* major—barrier to interdisciplinarity. And KOSs are complex creations. We will in this book outline the broad nature of desired KOSs and suggest a variety of specific strategies for their achievement. But the development of such KOSs will necessarily involve a host of detailed decisions, and these are best addressed by scholars with expertise in both knowledge organization and interdisciplinary studies. As we outline the contours of contemporary attempts to develop such classifications in this and later chapters, interdisciplinary scholars can reflect on how they might facilitate the development of these KOSs.

In sum the fact that it has not been done before is in no way an indication that it cannot be done in the future. This historical moment, characterized both by a widespread desire to facilitate interdisciplinarity and by rapid advances in digitization (and particularly the development of the Semantic Web) creates an opportunity for the development of something new. We have seen above that scholars of

[4] We do not survey here all of the contemporary authors or conferences which have voiced an attitude favorable to a phenomenon-based approach. Some of these voices of support are referenced on the website of the Integrative Levels Classification at http://www.iskoi.org/ilc/ref.php.

Table 4.1 Reasons that phenomenon-based classifications do not yet exist

As noted in the preceding section, it was simply too demanding to provide multiple entry points to the subject of a work in an age of card catalogues. Digitization makes it straightforward to classify (and search for) a work (or insight) along multiple dimensions.
Focus was mainly on the physical arrangement of volumes on shelves in a helpful linear sequence, so that a single dimension had to be chosen from the multidimensional subject of each work. With digitization the relative importance of shelving decisions has declined. Faceted classifications attempt to capture multiple dimensions of a work, while appreciating that one dimension must be privileged for shelving purposes.
While interdisciplinarity has been urged since the rise of disciplines in the eighteenth and nineteenth centuries, it has only become widely accepted within the academy in the last couple of decades (see Repko 2012; Klein 1990; Weingart 2010). That is, the general classifications in widespread use today were developed at a time of disciplinary hegemony (we borrow this phrase from Augsburg and Henry 2009).
Inertia is a powerful force in document classification (see above). We argued in Chap. 3 that the changes proposed in this book could not be accommodated by minor tweaks to existing systems. [Note in this regard that the systems proposed in this book rely heavily on linking concepts, and thus hold out the hope that novel areas of scholarly research can generally be accommodated by invoking linkages between existing concepts so that the elements of the classification will not require alteration as new subjects of study are brought within its purview.]
We will introduce in Chaps. 5, 6, and 7 several important strategies that are critical for the development of such a classification.
There may well be a need for interdisciplinarity in document classification itself. Both classificationist and classifier need little knowledge of the content of the documents they address if they will classify these only in terms of how the documents fit into a narrow disciplinary conception of subject matter. If instead these will be classified in terms of a general classification of phenomena, and in terms of theory and method applied, more careful content analysis—or collaboration between information scientists and other scholars—will be called for.[a]

[a]The inductive approach advocated by Hjørland and Nissen Pedersen (2005) and others represents one promising strategy for developing better classifications based on more careful analysis of texts. But as those authors freely admit, a multi-domain classification is unattainable with their approach. Knapp (2012) suggests that some form of crowd-sourcing might be employed to clarify the terminology in a general classification. Golub et al. (2014) argue that tagging works much better if participants utilize a controlled vocabulary. Bawden (2008) addresses the general question of whether information science should be interdisciplinary

knowledge organization have urged and explored the possibility of a comprehensive phenomenon-based classification throughout the last century (and more). Our work has built upon their efforts. We are fortunate to operate in a more propitious environment.

The Integrative Levels Classification Project

Classification by phenomena as separate from disciplines is being experimented with today in the Integrative Levels Classification (ILC) research project. This initiative, involving several researchers in various countries including two authors of this book, is mainly inspired by the work of the Classification Research Group

Table 4.2 Perspective facets in ILC

Notation	Facet
0	As for [perspective]
00	As attested in [document]
01	As known in [epoch]
02	As known in [place]
03	As studied by [method]
04	According to [theory]
05	Studied by [discipline]
06	As known in [culture]
07	Applied to [activity field]
08	Illustrated by [modality]
09	Conveying [communicative function]

briefly described above, though also informed by more recent literature in knowledge organization and original ideas.

The structure of a general classification scheme, conceived in a way similar to the CRG one but with new classes, categories and notational system, has been developed. Main classes, represented as lower case letters, are phenomena sorted by increasing level of organization, according to the integrative levels theory variously presented in twentieth century philosophy: from basic forms and physical quantities, through molecules, living beings, minds and societies, until the most complex technological and cultural products of the human spirit (the ILC main classes are reprised in Table 7.2 where the theory of integrative levels is discussed). Each class has its own subclasses (further letters) and facets (introduced by digits) (ILC 2004; Gnoli 2006; Gnoli et al. 2011).

The perspective dimension can be represented in the form of perspective facets (starting by *0*). These include communicative function, modality, activity field of application, discourse community, discipline, theory, method, place, epoch, and document (see Table 4.2). Thus the phenomenon *mqvo* 'birds' (a subclass of *m* organisms) can be combined with a particular discipline: *mqvo05tu* 'birds, studied in agronomy;' a particular method: *mqvo03et* 'birds, studied by telemetry;' a particular epoch: *mqvo01e* 'birds, as known in the Middle Ages,' and so on (Szostak and Gnoli 2008).

Of course, several combinations at a time are possible, to express very specific topics like those typical of scientific papers: ILC is freely faceted (Gnoli and Hong 2006), meaning that any concept can be combined with any other, like in *mqvt36vbh05tu03et* 'birds, affected by hunting, studied in agronomy, by telemetry.' The same phenomenon can thus occur in very different contexts, like *xs8mqvo* 'films, representing birds.' A search for birds in a digital environment will retrieve both documents, as well as any other in which birds appear in some combination or alone.

Disciplines themselves are listed in ILC at level *y* of knowledge phenomena: *ysq* 'linguistics,' *ytu* 'agronomy,' etc. Although usually expressed as perspective facets in the study of some phenomenon, they can well be taken as the main theme when this is needed: a book focusing on agronomy as a discipline cultivated in Asia can

be indexed as *ytu2k*. Thus, ILC effectively allows works to be classified either by phenomenon or by discipline, according to the collections and needs at hand. The León Manifesto claims that in the interest of interdisciplinary research phenomena should usually be given priority.

ILC still is an ongoing, experimental project. While details of its schedules are being considered and further developed, sample collections are indexed with it in order to test the system and to refine it accordingly. The most extensive tests until now have been performed with two collections.

Where the Apennine begins is a website devoted to the peculiarities of a mountainous region in northern Italy, known as the Quattro Province. Information provided in it has naturalistic as well as cultural components, including landforms, dialects, villages, local history, traditional music and dances. It is thus a good example of an interdisciplinary domain. The website includes a bibliography on the region, which is indexed by ILC. The scheme is used here in its 'free classification' version: that is, each phenomenon class is combined with others by simply listing them separated by blank spaces, much like with tags in folksonomies (Gnoli 2010). This means that, while display in the search interface is simpler, the full expressive power of facets is not utilized. However, this simple application is already enough to show that phenomena can be a good unit, not less effective than other ones, for grouping and combining concepts, browsing them and retrieving them. This application also makes use of ILC special classes for locally-preferred concepts (deictics, represented as capital letters: Gnoli 2011).

The second relevant application is for the *BioAcoustics Reference Database* (*BARD*). This is an online bibliography of several thousands of research documents in the domain of bioacoustics, with special reference to vocal communication among whales and the impact of human-produced noises on them. It mostly includes papers, but also books and technical reports, dealing with this interdisciplinary domain across physics, biology, ethology, ocean management, technology, and sometimes even military science. Many BARD records are classified with freely-faceted ILC notation, and corresponding faceted captions are automatically synthesized and displayed (Gnoli et al. 2010). The records include such complex subjects as *t8ve49t0nm(9qvtn60v25c)* 'governments, administering conservation, by civil law, in relation to: populations, of whales, tainted by technologies, in oceanic zones,' corresponding to a highly interdisciplinary report entitled *Guidelines on the applications of the environment protection and biodiversity conservation act to interactions between offshore operations and larger cetaceans*![5]

It is interesting to see how, by freely-faceted classification, any phenomenon can be given the role of the base theme of a document by expressing it at the beginning of faceted notation. In the document above, conservation law is the focus: hence it will be primarily grouped together with other documents on conservation laws (expressed as a facet of governments), although being also related to whales, oceans and so on. The document is retrievable by a search for these latter phenomena.

[5] A further example from this database was discussed in Chap. 2.

In other cases, the notation for whales, or for oceans, can be promoted to the leading position, making the item part of the documents focusing on them.

Another ILC feature that proves especially useful for BARD is the facet for methods of study: indeed, methods like remote sensing, underwater microphones, or statistical processing are often a relevant component in bioacoustics papers. Users can thus identify papers on the same phenomena studied by different methods, or on the same method applied to different phenomena, just as was recommended by the León Manifesto.

A Comparison Test

H.-Peter Ohly rightly suggested at a conference that classification by phenomena should be evaluated against a traditional disciplinary classification by some comparison test. Although no significant quantitative data are available yet, some insights can be drawn from a first set of books on nature conservation at the University of Pavia Science and Technology Library. These documents, already classified by DDC in the national online catalog SBN, were also classified by ILC, so that the resulting arrangements and indexes by the two systems could be compared. Land conservation is another domain where documents often manifest an interdisciplinary character, connecting geological, botanical and zoological components of natural areas with ecology, law, public administration, and economics including both production and tourism.

Indeed, DDC classes assigned to these documents in the national catalog belonged to various disciplinary classes such as social sciences (economy of natural resources, law, and management of environment-related social problems), natural sciences (ecology, botany), applied sciences (health, engineering, and forestry), the arts (landscape architecture), and geography. ILC classes, on the other hand, mainly belonged to phenomena at the levels of land (territories, aquifers), organisms (plants), populations and ecosystems, government institutions, and technologies (land management, industries). Classes of both systems were thus scattered, although their resulting sequence was quite different since they were based on different principles for main class order: intellectual capacities for DDC, and integrative levels for ILC. In some cases, potentially useful groupings proved possible in ILC (‘plants’) though not in DDC (‘ecology of plants’ is separated from ‘botany’).

The main difference was that DDC forced the indexers to choose one class while hiding others, as it prescribes that only one theme of the document be expressed: either a document is on economics of natural resources, or it is on ecology, while the links between the two disciplines are only committed to cross references in the classification schedules. On the other hand, in ILC the base phenomenon could easily be connected with other phenomena by means of free facets, giving compound concepts like ‘ecosystems, as object of land management, by some law.’

Table 4.3 Facet categories in ILC

Notation	Facet
0	As for [perspective]
1	At [time]
2	In [place]
3	Through [process]
4	Made of [element]
5	With [organ]
6	From [origin]
7	To [destination]
8	Like [pattern]
9	Of [kind]

The nature of these relationships, expressed by the facet digits, can also prove useful for specialized searches. Table 4.3 lists the main facet categories in ILC.

These categories can combine into more specialized facets, e.g. *03* expresses process in perspective, that is method, as seen in the examples above.

Notation length can be an issue in classification, especially for shelving purposes (see Chap. 3). Although expressing a greater number of connected concepts, ILC classmarks proved to be of length comparable to DDC ones: indeed, each concept has on average a shorter notation, as being selected from arrays of 26 letters instead of 10 digits. On the other hand, DDC classmarks can be praised for their visual plainness, being formed only by a sequence of digits.

It has to be acknowledged that forcing the indexer to choose a single class is not necessarily a limitation of any disciplinary classification: indeed, another disciplinary system like UDC does allow for combinations between main classes separating them by a colon. In an online search, both UDC and ILC would then allow, in contrast to DDC, the user to retrieve particular themes connected to the one given priority. What remains different is that disciplinary schemes connect disciplines, like geography, botany or economics, rather than connecting phenomena, like territories, plants or management activities: they cannot avoid, when combining concepts, to carry disciplinary implications with them, although their meaning may represent the document content inaccurately (a plant is a plant in any context, but is not always an object of botanical study).

Furthermore, in some cases the DDC lists a very specific concept only under a given discipline, thus forcing the classificationist to choose that concept even if the actual document focus is on a combined concept that is not part of the same discipline; a phenomenon-based system, instead, can express both, and at the same time give priority to the one taken as the base theme in the document (‘governments, managing conservation’ as opposed to ‘conservation, managed by governments’ for a book on citizens participation in the management of nature parks, which was classed with DDC under ‘economics of natural resources’). Also, a guide to trade laws and legal protection of animal and plant species, going under DDC class ‘international law: protection and endorsement of natural resources,’ was collocated by ILC into ‘economies, administrated by the United Nations, of

organisms, through trade,' thus reflecting priority of the economic facet in the actual book.

Another difference is that, while disciplinary classification tends to group documents into box-like classes having a certain scope, phenomenon-based classification tends to point to precise concepts, because its notation is built as a combination of individual concepts, much as happens with a thesaurus. In many cases the result is not very different, as even disciplines can be subdivided into deep hierarchies with longer numbers, and the deepest specifications are often labeled by phenomenon terms, as was observed above while discussing the class of bats within zoology. This is what happens in such cases as DDC class *363.73* 'social welfare and security: environmental problems and services: pollution,' assigned to a book which in ILC was classified as 'ecosystems, polluted by something, as related to organisms.' In some cases, however, box-like disciplinary classes cause a partial loss of meaning: a book on spontaneous flowers in Lombardy natural reserves, going under DDC class *581.9452* 'botany, Lombardy,' could be indexed by ILC classmark *mpw5w29ed* 'angiospermae, with flowers, in Lombardy,' thus keeping track of the focus on flowers as a relevant organ of those plants. Another book on rural society and buildings in a country municipality, filed under 'history of the Como province' in DDC, could be indexed more precisely in ILC as 'villages, in the Lambro basin, in the eighteenth to nineteenth century,' thus making the specific concept of villages as artifacts free from the disciplinary context of history, and, at the same time, giving the identity of the place in question as a physical-geographical territory (i.e., 'Lambro basin'), which maintains its identity over time, and so distinguishing it from its contemporary political administration (i.e., 'Como province'), which has changed over time.

The Basic Concepts Classification

The Basic Concepts Classification (Szostak 2013a) adheres to most of the organizing principles of the ILC outlined above. It also organizes its main classes around integrative levels (though as in ILC there are several main classes at the highest level of social interaction). It allows the free combination of any terms in the classification. Disciplines are themselves classified within the classification of phenomena, as in ILC, and thus can again be designated as necessary in describing any work. As with ILC, allowing the free combination of terms allows very precise classification of works utilizing modest schedules.

Though the BCC is grounded in the literature on facet analysis, it recognizes facets structurally rather than through the use of explicit facet indicators. Its macrostructure is based on three fundamental categories. The most extensive schedule lists classes of phenomena (there are some 20 main classes at present, each designated by a capital letter, usually the first letter of the class name: *C* is Culture). These classes are then subdivided (first by another capital letter, so that *CV* denotes values, then by numbers, then by lower case letters). Subdivision

proceeds logically in terms of generic ('type of') or sometimes partitive ('part of') relationships. As noted in previous chapters this practice is facilitated by the synthetic approach: enumerative classifications often deviate from logical hierarchy in order to find a place for compound subjects. The second type of schedule classifies type of relationship. Non-causal relationships are designated with a variety of non-alphabetic and non-numeric symbols (-, <, ~, ^, and so on). Causal relationships are indicated by the use of arrows (which signal the direction of influence). Some 100 relationships are organized into four broad classes, and about a dozen subclasses in total (each relationship is designated by lower-case italicized pairs of letters). These hundred can be combined with each other, and/or with phenomena or properties (see below) or non-causal relationships, to generate hundreds of more precise relationships (persuasion—*rsrt*—combines control—*rs*—and talking: *rt*). The classification of relationships was developed in Szostak (2012a, b). The third type of schedule is the classification of adverbial/adjectival properties (these are designated by the letter *Q* for Qualities, and are subdivided in the same manner as phenomena).

The BCC website (Szostak 2013a) discusses how each of the facets recognized in both the Bliss Classification and the ILC are captured and clearly designated by the free combination of phenomena, relationships, and/or properties. Most obviously, the agent (cause), product (outcome), and patient (intermediate variable) in a causal relationship are clear in a classification of the form A(influences)B(influences)C.

The BCC is motivated by the recognition that the vast majority of scholarly research, and most non-scholarly works as well, focus on discussing how one or more phenomena influence one or more others. The best and easiest way to classify such works is a synthetic approach of the form (phenomenon) (type of influence) (phenomenon) (see Szostak 2012c). Note that this sort of approach was advocated by the Classification Research Group (see above). A minority of works discuss the internal composition or behavior of a single phenomenon. These are best captured through recourse to classifications of both phenomena and properties, or to the use of such relationships as 'contained in' or 'composed of.' This simple approach not only facilitates the work of both classificationist and classifier (because small schedules can in combination generate precise designations) but also the user, for they can search by combinations of simple terms.

The BCC instantiates a web-of-relations approach to classification as urged in Olson (2007). A user that starts with an interest in cats, can follow their curiosity to (cats) (compared to) (dogs), and on to (dogs) (bite) (mail carriers). Only a synthetic approach that allows the free combination of any phenomena, relationship and/or property allows users to so easily follow their curiosity from one subject to a related subject. And by facilitating this sort of movement through the web of information we greatly facilitate what is variously called 'undiscovered public knowledge,' 'literature-based discovery,' or 'serendipity:' the connection of related ideas from disparate literatures that generate a new insight (see Chap. 1).

The León Manifesto suggested that works should be classified not just in terms of phenomena and relationships, but also in terms of the theory and method applied

in a work. As noted in previous chapters it is also beneficial to classify with respect to various elements of authorial perspective. The BCC, like the ILC, captures methods and theory types within its schedules, and aspires to capture various elements of perspective as well.

The most detailed application of the BCC has been in a translation exercise in which DDC classes *300* to *345* [and later all classes in ICONCLASS] were translated into BCC. This exercise (see Szostak 2013b) was summarized in Szostak (2011). It was straightforward to translate each DDC entry. The result in each case was a notation of quite manageable length. And the translation quite often served to clarify quite vague DDC terminology (see Chap. 5). If a system such as BCC were adopted as a complement to DDC, and such translation became automatic, then concepts added in future would be clarified at the outset. Moreover, it should be stressed that such a translation exercise undervalues the BCC, relative to an exercise that would code various works directly in BCC. The DDC tries to designate works according to brief class titles whereas the BCC designates works as combinations of terms. This, quite simply, gives a much more precise description of the contents of a work.

Like the ILC, the BCC is a work in progress. Yet the broad outlines of the classification are fully developed. The DDC translation exercise indicates that the classification can cope nicely with classifying social science material. Szostak (2014a) shows how the classification can cope better than existing classifications with the demands of the humanities. The natural sciences are a challenge for any classification because of the vast numbers of species and chemical compounds that must be classified. Much clarification of the classification of species can be anticipated from biologists over the next few years, due to advances in genetic analysis, and the BCC can take advantage of these developments.

Key Points

There are historical reasons for the dominance today of discipline-based classification systems. Yet there is a long tradition in the field of knowledge organization of advocating and exploring phenomena-based classifications. It is difficult to switch from one historical path to another, but hardly impossible. The rise of both interdisciplinarity and digitization facilitate and encourage such a transformation. The Integrative Levels Classification and Basic Concepts Classification have both been developed in response to these influences. They have each been developed and tested to an extent that establishes the broad feasibility of such a classification.

References

Augsburg T, Henry S (eds) (2009) The politics of interdisciplinary studies. McFarland Press, Jefferson, NC

Austin DW (1969) Prospects for a new general classification. J Librarianship 1(3):149–169

Austin DW (1984) PRECIS, a manual of concept analysis and subject indexing, 2nd edn. British Library, London

Bawden D (2008) Smoother pebbles and the shoulders of giants; the developing foundations of information science. J Inform Sci 34(4):415–426

Beghtol C (2004) Exploring new approaches to the organization of knowledge: James Duff Brown. Libr Trends 52(4):702–718

Bonner R (1982) Community information classification research project: user oriented empirical methods of classification construction. In: Dahlberg I (ed) Universal classification: proceedings of the 4th international study conference on classification research. Indeks, Frankfurt, pp 227–234

Broughton V (2004) Essential classification. Schulman, London

Brown JD (1906) Subject classification. The London Supply, London

Classification Research Group (1969) Classification and information control. Library Association, London

Classification Research Group (1973) Bulletin n. 10. J Doc 29:51–71, 56

Classification Research Group (1978) Bulletin n. 11. J Doc 34(1):21–50, 25

Dahlberg (1974) Grundlagen universaler wissensordnung. Probleme und möglichkeiten eines universalen klassifikationssystems des wissens. Pullach bei München: Verlag Dokumentation

Dahlberg I (1978) Ontical structures and universal classification. SRELS, Bangalore, India

Dahlherg I (2009) Concepts and terms: ISKO's major challenge. Knowl Org 36(2/3):169–177

De Grolier E (1962) A study of general categories applicable to classification and coding in documentation. UNESCO, Paris

Gnoli C (2005) BC2 classes for phenomena: an application of the theory of integrative levels. Bliss Classif Bull 47:17–21

Gnoli C (2006) The meaning of facets in non-disciplinary classification. In: Budin G, Swertz C, Mitgutsch K (eds) Knowledge Organization for a global learning society: proceedings of the 9th ISKO conference. Ergon, Würzburg, pp 11–18

Gnoli C (2010) Themes and citation order in free classification. IASLIC Bull 55(1):13–19

Gnoli C (2011) Animals belonging to the emperor: enabling viewpoint warrant in classification. In: Landry P, Bultrini L, O'Neill ET, Roe SK (eds) Subject access: preparing for the future. De Gruyter, Berlin, pp 91–100

Gnoli C, Hong M (2006) Freely faceted classification for Web-based information retrieval. New Rev Hyperm Multimedia 12(1):63–81

Gnoli C, Merli G, Pavan G, Bernuzzi E, Priano M (2010) Freely faceted classification for a Web-based bibliographic archive: the BioAcoustic Reference Database. In: Sieglerschmidt J, Ohly H-P (eds) Wissensspeicher in digitalen Räumen: Nachhaltigkeit, Verfügbarkeit, semantische Interoperabilität: Proceedings der 11. Tagung der Deutschen Sektion der Internationalen Gesellschaft für Wissenorganisation, Konstan [Knowledge storage in digital space: persistence, availability, semantic interoperability]. Ergon, Würzburg, pp 124–134

Gnoli C, Pullmann T, Cousson P, Merli G, Szostak R (2011) Representing the structural elements of a freely faceted classification. In: Slavic A, Civallero E (eds) Classification and ontology: formal approaches and access to knowledge: proceedings of the International UDC Seminar, The Hague. Ergon Verlag, Würzburg, pp 193–206

Golub K, Lykke M, Tudhope D (2014) Enhancing social tagging with automated keywords from the Dewey Decimal classification. J Doc 70(5):801–828

Hartel J (2003) The serious leisure frontier in library and information science: Hobby domains. Knowl Org 30(3/4):228–238

Hjørland B, Nissen Pedersen K (2005) A substantive theory of classification for information retrieval. J Doc 61(5):582–595
Integrative Levels Classification (ILC) (2004) ISKO Italia. www.iskoi.org/ilc/
Kaiser JO (1911) Systematic indexing. Pitman, London
Klein JT (1990) Interdisciplinarity: history, theory and practice. The Wayne State University Press, Detroit
Knapp JA (2012) Plugging the 'whole': librarians as interdisciplinary facilitators. Libr Rev 61 (3):199–214
McIlwaine IC (2007) The universal decimal classification: a guide to its use, Revised ed. UDC Consortium, The Hague
Olson H (2007) How we construct subjects: a feminist analysis. Library Trends 56(2):509–541
Otlet P (1896) Rules for developing the decimal classification. Office international de bibliographie, Bruxelles, Translated in W. Boyd Rayward (1990) International organization and dissemination of knowledge: selected essays of Paul Otlet. Elsevier, Amsterdam. pp 63–70
Perrault JM (1969) Towards a theory for UDC. Bingley, London
Repko AF (2012) Interdisciplinary research: process and theory, 2nd edn. Sage, Thousand Oaks
Satija MP (1979) Extension of two-phased subjects in the Colon Classification. Herald Libr Sci 18 (4):344–348
Scheele M (1977) Ordnung und Wortschatz des Wissens, v. 1: Das Ordnungssystem, Universal Faceted Classification. Schlitz, Guntrum
Scheele M (1983) Automatic indexing of titles and keywords on the basis of a model for an overall thesaurus of knowledge. Int Classif 10:135–137
Shpackov AA (1992) The nature and boundaries of information science(s). J Am Soc Inform Sci Technol 43(10):679
Slavic A (2007) On the nature and typology of documentary classifications and their use in a networked environment. El profesional de la información 16(6):580–589
Szostak R (2011) Complex concepts into basic concepts. J Am Soc Inform Soc Technol 62 (11):2247–2265
Szostak R (2012a) Classifying relationships. Knowl Org 39(3):165–178
Szostak R (2012b) Toward a classification of relationships. Knowl Org 39(2):83–94
Szostak R (2012c). Classification in terms of basic concepts. Advances in Classification Research 2012, proceedings of the ASIST SIG/CR Workshop, October 26, 2012, Baltimore, MD, USA. http://journals.lib.washington.edu/index.php/acro/article/view/14234
Szostak R (2013a) Basic concepts classification. https://sites.google.com/a/ualberta.ca/rick-szostak/research/basic-concepts-classification-web-version-2013
Szostak R (2013b) Translation table: DDC [Dewey Decimal Classification] to basic concepts classification. http://www.economics.ualberta.ca/en/FacultyandStaff/~/media/economics/FacultyAndStaff/Szostak/Szostak-Dewey-Conversion-Table.pdf
Szostak R (2014) Classifying the humanities. Knowl Org 41(4):263–275
Szostak R, Gnoli C (2008) Classifying by phenomena, theories, and methods: examples with focused social science theories. In: Arsenault C, Tennis J (eds) Culture and identity in knowledge organization, proceedings of the 10th international ISKO conference, Montréal. Ergon, Würzburg, pp 205–211
Vickery BC (2008) The structure of subject classifications for document retrieval. http://www.iskoi.org/ilc/vickery.php
Weinberg BH (1988) Why indexing fails the researcher. Indexer 16(1):3–6, http://people.unt.edu/~skh0001/wein1.htm
Weingart P (2010) A short history of knowledge formations. In: Frodeman R, Klein JT, Mitcham C (eds) The Oxford handbook of interdisciplinarity. Oxford University Press, New York, pp 3–14