Chapter 5
The Feasibility of Developing Such Knowledge Organization Systems

Existing classification systems such as the Library of Congress (LCC) or Dewey Decimal (DDC) benefit from over a century of refinement. It is thus no simple task to develop a novel classification that might supersede (or simply complement) these. Knapp (2012) is one scholar who applauds the sort of classification being urged in this book, but worries about the feasibility of developing an entirely new classification. Yet the argument of this chapter is that it is indeed possible to do so. We will first make some general remarks regarding feasibility, and then proceed to a discussion of each of the elements of a new system that were proposed in the preceding chapters.

The Digital Revolution and the Historical Moment

The digital revolution creates a unique opportunity for the development of a new system for it highlights the limitations of classifications developed for card catalogs (Dean 2003). Moreover the proliferation of new databases with unique classifications increases the desirability of a classification that could be general in application. There is even a direct connection with interdisciplinarity: online access increases the likelihood of citation of works in other fields on a topic, articles by lesser-known authors, articles in less prestigious journals, and older articles—but only at the cost of increasing information overload (Wu et al. 2012).

While the digital revolution exposes certain weaknesses of existing systems, it at the same time creates opportunities for the new. In particular, digitization means that it is easier to classify documents along multiple dimensions. In an age of card catalogues, it was expensive to classify a work in more than one or two ways through such procedures as chain indexing (Coates 1988), for each entry point required a new card to be typed and placed in the catalogue. For example the freely faceted PRECIS system (Austin 1984) proved very demanding in terms of production work and costs. Today, the cost of adding another entry point is much smaller.

R. Szostak et al., *Interdisciplinary Knowledge Organization*,
DOI 10.1007/978-3-319-30148-8_5

Thus, even if we were not facing the rising importance of interdisciplinarity, it would make sense at this historical moment to explore the possibility that a system designed for the digital age might be preferable to systems designed for card catalogues. And in particular it makes sense to ask whether we should provide much more information about particular works than is done at present. As the trend toward the assignment of more kinds of metadata (data about documents, such as title, author, subject, and so on) attests, the digitization of KOSs encourages the indexing of documents along a wider range of dimensions. We will not reiterate this point in what follows but rather will presume that in the present age it makes sense to provide more information as long as this can easily be provided in a useful format. We can thus focus on questions of classification rather than cost.

Classifying Comprehensively with Respect to Phenomena

The main objection to the feasibility of a comprehensive classification of phenomena is that the meaning of concepts differs too much across fields. It is thus felt that a comprehensive classification would lead to confusion, for scholars from different fields would attach quite different meanings to the terms employed.

A variety of theoretical arguments can be made in support of the feasibility of a comprehensive classification. These are summarized in Table 5.1.

Questions of feasibility cannot be decided on theoretical grounds alone. It is necessary to develop at least some significant portions of such a classification in order to fully establish feasibility. There are two contemporary classifications that do precisely that:

- The Integrative Levels Classification (ILC 2004).
- The Basic Concepts Classification (BCC) (Szostak 2013a).

Each of these classifications (which differ primarily in how they employ facet analysis, and how far they proceed in the decomposition of complex concepts) builds upon the efforts of the Classification Research Group (CRG, especially Douglas Foskett and Derek Austin) to develop general classifications of phenomena. As with the earlier works of the CRG (and indeed all classifications to at least some extent), these classifications are hierarchical. A manageable set of main classes of phenomena is identified, and these are unpacked or disaggregated into multiple degrees of constituent phenomena. In both cases, phenomena can usually be well defined, either in terms of their internal essence or their function. [Gnoli and Poli (2004) argue that functional definitions only make sense for phenomena of a level of complexity equal to or beyond biological organisms; for atomic particles and chemical elements definition in terms of essence are superior.] It can thus be expected that a comprehensive classification of phenomena will not be invalidated by time: most phenomena included will retain their

Table 5.1 Theoretical justifications of comprehensive phenomenon-based classification

Ambiguity differs only by degree between general and discipline-specific classifications, though that difference of degree is likely quite significant. The degree of ambiguity lessens within groups that regularly interact (though it does not disappear). Yet it is equally clear that conversations across groups speaking different primary dialects are possible (or international tourism would be impossible).
The classification structure itself contributes to making clear the meaning of a concept and the terms expressing it, by showing for each class its hierarchical and associative relationships with related classes, as well as by providing scope notes. This is true both in a classification by disciplines and in one by phenomena: in the latter case, concepts will take more general and neutral meanings, depending on their 'unique definition' within the classification rather than their contexts of use. This will make them reusable in a variety of domains.
The degree of ambiguity generally increases with the complexity of a term. 'Dog' lends itself to a greater degree of shared understanding across fields (or indeed people) than does 'globalization.' If, as was suggested in Chap. 3, works about globalization were classified in terms of how particular economic, political, or cultural developments affected other particular economic, political or cultural elements (and also in terms of the ethical and ideological perspective pursued), the degree of ambiguity is thus much less than if they are just classified in terms of 'globalization.' We return to this line of argument in Chap. 9.
Interdisciplinary scholars recommend a strategy of 'redefinition' to confront the problems that arise when a term has different meanings across disciplines (or other groups). The term is provided with a specific meaning that will make sense to scholars from all disciplines (see Repko 2012). This strategy is employed successfully in several chapters in Repko et al. (2012). Van der Lecq (2012), for example, notes that different fields studying the emergence of language in humans employ the terms 'evolution' and 'communication' in slightly different ways; redefinition alleviates much (though not all) of the conflict with respect to identifying the causes of this important development. To be sure, the strategy is employed narrowly to allow interdisciplinary investigation of a particular problem. But its success certainly at least opens the door to a broader effort at 'redefinition.'
Psychologists suspect that there are universals in how human beings think and perhaps in how they organize their conceptual maps (especially evolutionary psychologists; Barkow et al. 1992). Neelameghan argued at length in the 1970s that it should be possible to develop subject representations grounded in these universals (Iyer 1995, 184). Evolutionary psychologists would suggest that such universals are more likely with respect to activities that humans undertook during the millennia humanity operated as hunter-gatherers. If so, general classifications will face greater difficulties in some realms than others. Yet it seems likely that any innate human classification system will be organized around phenomena rather than disciplines: 'If . . . the patron asks for a book on birds, he expects that a section in the library contains all bird books. . . . To the extent that they do not (e.g. bird books are spread out into many different places), the system is awkward and irrational from the patron's perspective, regardless of its logical consistency from the librarian's' (Donovan 1991, 26). The conceptual atomism approach advocated by some philosophers of concepts indicates that diverse people should share broadly similar understandings of concepts signifying the things (and relationships) that we regularly perceive in the world around us (Szostak 2011).

meaning. While neither classification has yet been subjected to detailed user testing (a matter we will return to in the concluding chapter), both projects have proceeded far enough to establish feasibility in several respects. This empirical evidence is summarized in Table 5.2.

Table 5.2 Empirical evidence of the feasibility of a comprehensive phenomenon-based classification

Lengthy schedules of phenomena have been developed for each (with ILC having focused relatively more attention on natural science and BCC on human science). Readers are urged to peruse the schedules and contemplate whether most/all of the entries would be understood similarly across fields (or could be provided with a definition that would facilitate this).
Hundreds of records in special bibliographies as different as local culture (Gnoli 2010) and bioacoustics (Gnoli et al. 2008) have been carefully classified by ILC and can be searched by it through online interfaces.
The entries in the Dewey Decimal Classification (DDC) from 300 to 340 [and also all of ICONCLASS] have been translated into terms of the BCC (Szostak 2011; 2013b). Arguably, the BCC translation provides a less ambiguous classification than the DDC original. And this is the standard that a new classification needs to meet: not that it can provide a complete absence of ambiguity but that it can achieve at least the same degree of ambiguity as existing systems of document classification. Most DDC entries can be translated into 'very basic' concepts for which the degree of ambiguity is minimal. All DDC entries can be translated into concepts that are 'somewhat basic' in the sense that the degree of ambiguity, at least potentially, might be considered acceptable from the point of view of information science. They are certainly much 'more basic' than the complex terms translated. Recall that the translation of DDC entries often served to clarify the meaning of DDC entries themselves. Such clarification would be even greater if new DDC entries were made with translation to/from basic concepts in mind. The translation thus has the potential to enhance cross-disciplinary communication while also enhancing within-discipline communication.
Szostak (2003) established that *the arguments of* hundreds of works from across the human sciences could be classified in terms of a simple but general classification of phenomena. Indeed this was possible employing a much less sophisticated classification than the present BCC.

The Feasibility of Classifying Relationships

As we have seen, interdisciplinary research usually involves the study of relationships among phenomena studied in different disciplines. Yet it is not only interdisciplinary researchers who would benefit from increased classificatory attention to relationships. Indeed, most scholarly research studies not just one phenomenon but the influence that one or more phenomena exert on others. A scholar interested in how, say, economic productivity is influenced by the level of trust in a society will wish to consult works investigating that causal link but not every work on trust or economic productivity. This outcome can be achieved by using linked notation between phenomena to indicate the main causal link(s) pursued in a work. This focus on links would further enhance the utility of document classifications to researchers. Note in this regard that discipline-based classifications would face difficulty in capturing interdisciplinary causal links. The different terminology used for each phenomenon involved would multiply the number of ways in which the same link might be classified. But once we successfully develop a comprehensive classification of phenomena it should be straightforward to allow these to be linked to express causal or other relationships.

The information science community has displayed a renewed interest in classifying relationships in recent years (see Green 2008). Khoo and Na (2006) predict

that relationships will become increasingly important in the field. Notably, they argue that research in information retrieval has proceeded about as far as it can without addressing relationships. They appreciate that there is a great deal of experimental evidence that relationships are 'real' and that there are cross-cultural similarities in perception.[1] Moreover, relationships are of critical importance to human memory. These observations bode well for the feasibility (and desirability) of developing a comprehensive classification of relationships.[2]

The next question is how detailed should we be in classifying relationships? Much that is useful can be achieved by simply indicating that A affects B (or is associated with B, and so on). But if we hope to capture the relationships referred to across diverse fields, a much more detailed classification is surely called for. Is it feasible to develop a more detailed classification of relationships? Szostak (2012a, b) has produced such a classification, which is now incorporated into the Basic Concepts Classification (see above; Szostak 2013a). Most of the attention is devoted to causal relationships—organized into a handful of categories (three types each of physical, biological, and intentional causation, plus a general class)—but a schedule of non-causal relationships is also developed (not, or, by, of, for, of type, in, about, compared to, associated with, from the perspective of, collection of, and a handful of mathematical relators). ILC facets, by which any phenomena can be linked, are based on the set of ten fundamental categories introduced in Chap. 4 (kind, pattern, destination, origin, organ, element, process, place, time, perspective) and their subtypes; in principle, indefinitely specific subtypes can be created through categories combination (Gnoli 2008). While care must be taken in evaluating any novel classification—and we will return to questions of evaluation in Chap. 10—a few characteristics of these two relationship-oriented classifications are worthy of note, as summarized in Table 5.3.

It is useful in Table 5.4 to provide some examples of the compound relators employed in BCC. These examples highlight how a synthetic approach can generate thousands of precise relationship terms from less than 100 basic relationship terms (many more examples are provided in Szostak 2013b).

As noted above, the classification developed in Szostak (2012a) devotes the bulk of its attention to causal relationships, but does include a set of non-causal relationships. This accords with the arguments made in Chap. 3 regarding the particular importance of causal relationships. Previous efforts to classify relationships as surveyed by Perrault (1994) had not placed the same emphasis on causal relationships. Moreover those classifications were generally deductive exercises not grounded in an inductive evaluation of the needs of researchers. This fact, plus the inertia associated with noun-focused classification systems that largely

[1] López-Huertas (2013) has studied cross-cultural understandings more generally, and found that there are shared understandings of some terms but not others.

[2] Khoo and Na (2006) proceed to discuss different sorts of semantic relationship (such as conjunction versus disjunction) and note that there is no scholarly consensus on a classification of these. But we have not identified that sort of distinction in our inventory of classificatory needs.

Table 5.3 Characteristics of the classification of relationships

The schedules are of very manageable length.
They are logically organized into a very manageable number of classes, each containing a handful of entries.
The distinctions among classes are transparent.
Yet these schedules and the compounds that can be generated from them capture all of the terms uncovered in a very broad inductive search for relators, as well as those implicated in previous deductive efforts to develop such a classification.
Thus all such terms can potentially be represented notationally within a mere handful of notational spaces.
Inspection of the schedules indicates that classes in the schedules would generally be interpreted in a very similar manner across disciplines and cultures (see Szostak 2011). Indeed, there may be less ambiguity associated in general with 'action' words.
This classification of relationships thus potentially allows for the first time searches by 'type of relationship' across an entire general classification.

eschewed synthetic notations, could explain why none of these previous efforts were incorporated into a larger classification. The renewed interest in relationships within the field of information science suggests that the time has come for the development and utilization of a classification of relationships that is grounded in an understanding of user needs. This will benefit all users but interdisciplinary scholars in particular.

Classifying with Respect to Theory Applied

Theories present great challenges for classification. One problem is that there is a potentially infinite number of theories, for new theories are invented every day in the scholarly enterprise. On its own this is a challenge often faced in classification, and one that can be dealt with by a hospitable (that is, expandable) classification with room for new classes.

More importantly, and as noted previously, there is considerable terminological confusion in the area of theory: the same theory name encompasses quite different types of theory, while quite similar theories go by different names (especially across disciplines). Even advocates of a particular theory quarrel about what the essence of their theory is (see, for example, Turner 2000). A classification grounded *only* in the names used by authors to describe their theories would thus generate the same sorts of confusion that the classification of subject matter in terms of disciplinary categories does. Nor is the only problem one of cross-disciplinary communication, for theories evolve through time (much more than do definitions of phenomena, though these if not carefully defined could evolve too), and thus an author using a theory today could have quite different ideas from an author using the same theory a decade before. While a classification in terms of the theoretical terms used by authors would be an important improvement over no classification by

Table 5.4 Examples of compound relators

Type of compound	Examples
Causal relator with causal relator	'Persuade' combines 'control' and 'talk'
	'Imagine' combines 'think' and 'desire'
	'Agree' combines 'decide' and 'cooperate'
	'Monitor' combines 'observe' and 'evaluate'
Causal relator with non-causal relators	'Herd' is (move) (collection of) (animals or people)
	'Attach' is (cause) (to be connected)
	'Search' is (look) (for)
Causal relator with indicators of internal changes (such as grow, decline, and maintain stability)	'Enhance' is (cause) (growth)
	'Stabilize' is (cause) (stable).
	'Include' is either (cause to be) (combined) or (cause to be) (contained in) [note that the ambiguity in the verb 'include' is thus addressed]
Causal relator with adverbial properties	'Suffer' is (experience) (bad)
	'Conspire' is (cooperate) (secretly)
	Dozens of combinations of 'cause' are possible such as (cause) (decorated) or (cause) (clean)
Causal relator with phenomena	'Possess' is (control) (by) (ownership)
	'Flying' is (moving) (by) (airplane)
	Again many combinations with 'cause' are possible such as (cause) (injury)
Complex causal chains	'Campaign' is (attempt) ((control) (by) (talking)) (associated with) (election)
	'Zoning' is (local government) (controls) (economic output) (associated with) (land)
	'Evict' is (move) (someone) (from) (home or office)

theory whatsoever, it would thus be insufficient in important respects. The researcher wondering if a particular theoretical argument has been applied to a particular set of phenomena will receive limited guidance if the literature is classified only with respect to the names of theories.

This terminological problem might be approached inductively, in order to identify what meanings are attached by different communities to different theory names. Recall, though, that inductive methods are generally applied only within particular scholarly communities. Happily, a deductive approach is available. A variety of theory types can be identified along several key dimensions: who is the active agent(s) in the theory, what do they do, why do they do it, what sort of process through time is envisaged, and how generalizable is the theory? This five-dimensional typology of theory types was developed through recourse to one of the simplest classificatory devices: asking the 5 W questions, who, what, where, when, and why. These in the context of theory yield more precise questions, and in each

Table 5.5 Dimensions of theory types

Questions	Possible answers
Who is the agent?	There are two immediate distinctions here: non-intentional (including volcanoes or institutions) versus intentional agency (of beings that can act on purpose), each of which can take the form of individual, group, or relationship agency.
What does the agent do?	There are three broad answers: passive reaction, active action, changes in attitude.
Why does the agent do this?	With non-intentional agents, action can only be understood in terms of their inherent nature. With intentional agents, scholars can explore the five distinct types of decision-making: rational, intuitive, process oriented (which emphasizes how one behaves rather than what one achieves, and encompasses various virtues or values that one might pursue), rule-based (where one follows certain guidelines such as the Golden Rule), and tradition-based. For groups and relationships, scholars can also ask how individual preferences are aggregated.
Where does the causal process occur? (How generalizable is the theory?)	There is a continuum from nomothetic (highly generalizable), through somewhat generalizable, to idiographic (situation- or causal-link-specific) theory.
When does the causal process occur?	Though inspired by the temporal question 'when?', the possibilities refer ontologically to directions of change. There are five broad time-paths that a causal process might follow: return to the original equilibrium, cyclical oscillation, movement to a new equilibrium, change in a particular direction, or stochastic/uncertain.

case a mere handful of possible answers (see Table 5.5; Szostak 2004). Note that this set of dimensions, though logically derived, is expandable should other useful dimensions be identified.

We noted in Chap. 1 that a major driver of interdisciplinary research is the desire to tackle complex challenges. Since complex social problems involve multiple types of agency, decision-making, and so on, the typology suggests that any one theory will give incomplete guidance. Familiarity with the typology would guide analysts to recognize the limitations of a particular theory for a particular question, and to identify other theories with compensating strengths. The use of the typology is thus consistent with postmodern concerns that multiple perspectives be heard. Yet whereas postmodernists are often skeptical of the ability of different communities—either scholarly communities or societal groupings—to converse with each other, the typology indicates that it is quite feasible to understand the strengths and weaknesses of different theories and at least strive toward some consensus on how important different theoretical arguments are for particular questions.

Individual theories can be placed within the typology, and thus works can be classified simultaneously by theory and by theory type. Researchers can then search

by theory name or by a particular type of theory. The typology has been applied to the set of social 'grand' theories identified in Turner (2000) in Szostak (2004), and also to a random set of more narrowly focused theories in Szostak and Gnoli (2008). That paper surveyed a handful of research areas in social science, and showed that the theories encountered in each can be reliably classified in terms of theory types. That paper also provided ILC notation for each theory type. [The various theory types have since been incorporated into the BCC as well.] In both cases it proved straightforward to classify particular uses of a theory in terms of theory types.

The analysis of Marxian theory in Szostak and Gnoli (2008) highlights how classifying works in terms of theory type will clarify the theoretical approach taken in a work much better than simply referring to a theory name such as Marxian theory. Marxian theory at times stresses technological determinism, and other times stresses the importance of class struggle. The first is a form of non-intentional individual causation, while the latter reflects intentional group agency. The first invokes passive reaction to technological innovations. The second argument also at times assumed inevitability but at other times Marx (and especially later Marxists) stressed active action. References to class consciousness bring groups and attitudes to the fore, but assumptions of historical inevitability counteract this tendency. Inevitability implies movement to a new (socialist) equilibrium; some Marxian theories accept instead the importance of historical contingency and thus imply a stochastic process. Contingent approaches open space for intentional individual or relationship agency. As for generalizability, though Marxian theory is often perceived as universal in orientation Marx himself at least once suggested that his historical process applied only to Europe.

The approach of classifying by theory type solves immediately the problem of different theories operating under the same name. It would be hoped that users would be informed in scope notes that a particular theory name is associated with quite different theory types. The user should also be guided to quite similar theories operating under different theory names. The user might even specify an interest in alternative theories that are similar in certain respects: same type of agency or action or dynamic process.

As noted above, theories evolve through time. And theorists disagree about what a particular theory entails. As a result many theories sprawl across multiple theory types. While this fact complicates the life of the classifier—they must look beyond theory name—it has enormous advantages for the researcher who wishes to only read, say, 'Marxian theory of a particular type.'

One difficulty encountered by Szostak and Gnoli (2008) deserves mention. In some cases, authors are vague about the theory type that they are applying. But this challenge also presents an opportunity for information science to encourage an improvement in scholarly practice. Theory classification should thus act as a tool, and at the same time as a stimulus, for a clearer description and definition of theories. Authors could be encouraged to classify their own theories in terms of theory types (and perhaps comparisons with other existing theories). Classification by authors could then be overseen by classifiers to ensure

that works are placed in suitable places in classification schemes. In the meantime the classifier can simply omit theory dimensions for which the author provides no clear guidance.

Classifying with Respect to Method Applied

Classifying with respect to method applied is much easier than classifying with respect to theory applied. Broadly speaking, methods can be grouped into only a dozen general classes (see Szostak 2004, 101–2). These are summarized in Table 5.6. Some would treat 'evaluation' of programs as distinct, though this can be seen as a combination of some of these methods. Similar arguments can be made with respect to demography, case study, feminism, and perhaps also hermeneutics.

These methods generally are called by the same name in different disciplines. They can be disaggregated into a manageable set of more specific tools and techniques. These tools and techniques also are often (though not always) referred to by similar terminology in different disciplines: for example the terminology of econometrics is widely employed when statistical analysis is pursued in other disciplines. More generally, 'econometrics' involves an overlapping set of techniques from those employed in 'biometrics' or 'psychometrics.' While there is certainly value in using 'econometrics' as a subject heading (for, say, textbooks on the subject) it will also be valuable wherever possible to classify a work on a particular technique by that technique so that scholars in any field can find it readily.

It should be stressed that the classification of a document in terms of method applied will be independent of its classification in terms of phenomena studied (which will determine its shelf placement) and theory applied. Each of these three distinct elements can be incorporated by means of notation within a single, analytico-synthetic classmark. Researchers using electronic catalogues will be able then to search by combinations of theory, method, and phenomena.

Scholars often employ multiple methods. Indeed there is an extensive literature on how and why to employ multiple methods in a particular research project. This may even occur within a discipline, as when quantitative and qualitative methods are combined in sociology. This literature overlaps in many ways with the literature on how to perform interdisciplinary research (Szostak 2015a). We should thus strive to identify multiple methods employed in a work, and perhaps also in some way to indicate the use of a multiple-method approach. As with works that engage

Table 5.6 The dozen scholarly methods

Classification itself	Experiments
Interviews	Surveys
Observation	Statistical analysis
Mathematical modeling	Textual analysis
Mapmaking (conceptual and representational)	Hermeneutics/semiotics (the study of symbols)

many relationships, it could prove necessary at times to eschew classification of a work by a method that is only superficially addressed.

Of course, many works do not involve (at least explicitly) the application of any theory or method. In such cases, these works will not be classified along those dimensions. First experiences with indexing bibliographies by ILC indicate that dimensions occur with different frequencies in different domains: in a natural science domain like bioacoustics, method is expressed far more frequently than theory (Gnoli et al. 2008), while the opposite seems to happen in the human sciences.

Classifying Authorial Perspective

In some ways, and at least along some dimensions, perspective is the easiest element from the point of view of the classificationist. We create a new dimension along which works or ideas can be classified (along with phenomenon, carrier, and so on; Gnoli 2012a). We call this dimension 'perspective.' The theories and methods addressed above might be subsumed within this broader category of perspective. Several other arrays are possible. It should be stressed that key elements of authorial perspective do not necessarily reflect disciplinary boundaries (though all elements of disciplinary perspective as outlined in Chap. 1 are subsumed). Some elements of perspective may indeed serve to facilitate cross-group communication more than cross-discipline communication. Szostak (2014) had argued that the best way to facilitate respect for social diversity is a comprehensive classification that coded works in terms of authorial perspective; it would then be easy for members of any group to find works from other groups or from their own as they wished. All of the elements of perspective might be considered by users in evaluating the relevance of a work. These elements are summarized in Table 5.7.

To be sure, it is not a trivial task to flesh out each of these arrays. Yet it should prove feasible to do so along at least some arrays. Notably, all of these elements should already be classified within the classification of phenomena. All that is needed here is an indicator of 'perspective applied' (this is *0* in ILC). Still other arrays might be imagined. Some that have been suggested in the literature include: local viewpoint (Beghtol 1998), epoch of knowledge (Tennis 2002), and application to human activity (Vickery 2008).[3]

While the task for the classificationist is straightforward, the task for the classifier may be more challenging, for authors are rarely explicit regarding perspective. This task could be alleviated if authors self-declared their perspective at the time of publication. Note in this regard that Greenberg et al. (2006) found that

[3] Condorcet had said that systems of classification that imposed a uniform view of nature were a great obstacle to science. He proposed a faceted approach that would include objects of study, methods, perspectives, uses of the knowledge, and ways of knowing (Glushko 2013, 299).

Table 5.7 The elements of authorial perspective

Element	Possible approach to classification
Discipline itself (and interdisciplinarity)	For this we need to classify all disciplines. Since disciplines are themselves phenomena (Gnoli 2005),[a] this would occur naturally if all phenomena can be freely combined. New interdisciplines could potentially be captured by combining existing disciplines, though we would want to allow the possibility of a new interdiscipline that drew upon multiple existing disciplines.
Theories applied	See above.
Methods applied	See above.
Epistemological outlook	A classification of approaches to epistemology (such as in existing domain analyses of philosophy) could be applied (see Hjørland 2005). Or we could pursue the epistemological dimensions identified in Looney et al (2014) in their efforts to facilitate interdisciplinary communication (see Szostak 2015b).
Ethical outlook	The five main ethical approaches are consequentialism, value-based, rule-based, tradition, and intuition (see Szostak 2005). These reflect in turn the five ways that humans can make any decision, which were listed above when discussing theory types.
Aesthetic attitudes	It could be that we need nothing more here than to link to various properties: (values) (beautiful).
Ideological outlook	There are several commonly recognized ideologies, which can be loosely associated with different types of ethical analysis: Conservative valuing tradition; Conservative valuing individual rights (often styled 'Classical Liberal'); Liberal stressing consequences; Socialist stressing communal values; and so on.
Rhetorical strategy	Feinberg (2011) suggests possible entries such as logical argument (manipulation of evidence), ethos (incorporation of audience beliefs and values to establish trust), and genre adaptation (adjustment of formal elements).[b] Her tripartite division parallels that often drawn by rhetoricians. Clavier and Paganelli (2012) argue that we should classify works by stance: criticism, agreement, consensus, and so on. Rhetoricians have identified over a hundred rhetorical strategies, but many of these—such as alliteration—apply to phrases rather than texts (though these may still be important for some works, such as poetry). There may be about a dozen that are commonly employed at the level of texts.

[a]We should strive, though, for a logical categorization of disciplines. Vickery (2008) sagely observed that general documentary classifications are organised into main classes that are often called "disciplines", perhaps because we think of them as the knowledge domains studied in an academic setting. But such classes as Mining and Education are, in the first instance, not domains studied but activities pursued, by the mining industry and by teachers. Only secondarily do they become objects of academic study. Vickery found it helpful to replace the idea of disciplines with that of "fields of human activity", activities in which people engage practically and are then led to study and write about (Gnoli 2012b)

[b]Her purpose is to argue for an explicit authorial voice in classifications, but her argument can perhaps be used for classifying works by authorial voice

authors could do a good job of reporting Dublin Core metadata, often better than metadata professionals. Even without self-declaration, discipline will usually be easy to identify (by, say, institutional affiliation). And the other elements will be as well for those works that self-consciously pursue a particular ethical, epistemological, or ideological outlook. The challenge will be greatest for works that are not self-conscious in their perspective. Some arrays will be of greater importance for some works than others, and thus author or classifier can focus only on those arrays that seem of particular importance to a particular work (see Szostak 2015b).

A Hybrid KOS?

In Chap. 3, we suggested that a KOS suited to interdisciplinarity could likely not be achieved with minor tweaks to existing discipline-based KOSs. We did hold out some hope that some sort of hybrid classification might prove feasible. Our analysis in this chapter has supported the feasibility of a phenomenon-based KOS. It is appropriate to revisit here also the possibility of a hybrid classification.

As noted above, Szostak (2011, 2013b) showed how class numbers *300* through *339.9* in DDC could be translated into the terminology of the BCC. A similar exercise could undoubtedly be performed for ILC. It is thus quite possible that a new system containing the elements described above could emerge as a complement to rather than a substitute for existing classification systems. In a hybrid system, users could be allowed to search for documents (or ideas) using either system. If class numbers in one system translate automatically into another, we gain increased retrieval without incurring additional costs in classification. While subject analysis by phenomena as performed by human classifiers would likely produce different, more refined classmarks, such an automatic procedure would provide at least a first, cheap approximation.

But of course not all class numbers translate flawlessly. Happily, Szostak did find it straightforward to translate the vast majority of DDC entries. In many cases, the DDC heading itself was a combination of two or three basic concepts [Yet within DDC it is not generally straightforward to search by the concepts within such compounds.]. As an example, DDC *302.35* 'Social interaction in complex groups' can be translated as (Interpersonal relationships) (in) (groups) [It is noteworthy here that the subdivisions in DDC refer to (formal) organizations rather than (informal) groups. The schedules of BCC list types of both groups and organizations (Szostak 2013a). In this case as many others, translating into basic concepts serves to clarify DDC terminology.] In all cases, a manageable handful (often one or two) of basic concepts was all that was required, and thus these could potentially be represented notationally in a quite limited notational space.

When it was not obvious how to translate a DDC entry, this was usually the result of ambiguity in DDC terminology. Take *332.41* 'Value of money.' Such an entry appears at first glance to be an obvious compound of simpler terms. Yet in fact this vague term seems in the context of the *332* hierarchy to mean 'inflation/

deflation;' and thus is best captured by a more precise (increase/decrease) (price level). Philosophical considerations of economic value are captured elsewhere in DDC and would be translated by linking ethical analysis to economic elements. This example provides further evidence that translation into basic concepts can serve to clarify the meaning of terms in existing classifications. And if such translation became standard practice then terms added in future would be clarified at the outset.

It should be stressed that the purpose of this section was merely to establish the feasibility of a hybrid system. It is not at all the authors' intent to suggest that this is the preferred option. It may well be best to focus future classificatory efforts on a stand-alone phenomenon-based classification that is truly comprehensive. This is an empirical question that can only be answered after user testing of completed systems. But it should be stressed that in a hybrid system users would necessarily have better access than at present, for they could always search as at present but also in terms of the synthetic approach we have recommended. Once users became familiar with the comprehensive, phenomenon-based classification they might rarely, if ever, wish to search in terms of the traditional, discipline-based ones.

Comprehensive Thesaurus

A comprehensive phenomenon-based classification would be most useful if accompanied by a comprehensive thesaurus providing preferred captions, synonyms and related terms for each class. Such an integration of a classification and a thesaurus has been described as a 'classaurus' by G. Bhattacharyya (Gnoli et al. 2011). Indeed, the existence of such a thesaurus would not only be useful to the interdisciplinarian but might banish any fears that the disciplinarian would have trouble navigating the new system.

It is useful to reflect here on the nature of the relationship between thesauri and classification systems. 'There are obvious advantages of a conceptually well-structured classification when generating a thesaurus, since the clear identification of relationships allows some degree of mechanical handling of the process…' (Broughton 2010, 275). Though Broughton's emphasis was on hierarchical relationships, her insight has broader application: development of the sort of classification urged above would facilitate the development of a comprehensive thesaurus.

Chan and O'Neill (2010, 18–9) recognize that thesauri and subject heading lists overlap in purpose. Thesauri are usually more strictly hierarchical because each entry is a single concept. Notably, thesauri are often multilingual because it is easier to translate simple terms. It would seem then that a classification that allowed simple terms to be combined would bridge the gap between subject headings and thesauri, and by extension disciplines and languages.

López-Huertas and Torres Ramírez (2007) note that many existing domain-specific thesauri are organized around general rather than domain-specific classes. This can be problematic in that scholars in a particular domain might be far more

interested in some general classes than others (see Chap. 6). But just as we have advocated a combination of general and domain-specific approaches to classification, we could urge the same for thesauri. A domain-specific thesaurus could employ concepts from a general classification or thesaurus while organizing these in a way that reflected the dominant interests of the domain. It would contribute to and be consistent with a comprehensive thesaurus.

More generally, the arguments in favor of only domain-specific thesauri are broadly similar to arguments for domain-specific classifications, and rest on the assumption that conceptual ambiguity is too great for a general thesaurus to be feasible. If ambiguity can be overcome in classification, then this should prove also be the case with regard to thesauri. And recall that the act of classification reduces ambiguity. This will inevitably reduce the barriers to a general thesaurus. And success there would further reduce ambiguity.

The success of WordNet deserves mention. This website aspires to identify a variety of thesaurus-like relationships (especially but not exclusively hierarchical and equivalence relationships) across all words in the English language. It is widely consulted, and efforts are now underway to duplicate this effort in and across other languages.

We could also make note of the scale and scope of thesauri in the medical field. The NCI Metathesaurus, sponsored by the National Cancer Institute in the United States, has over 2 million terms in its controlled vocabulary (and over 4 million entry terms) and describes over 22 million relationships between different cancers, therapies, and other phenomena. Its goal is to standardize terminology across cancer research in particular but biomedical research more generally.

In Chap. 2 we also discussed the value of being more specific regarding the different types of Related Term (RT) identified in a thesaurus. Since guidelines for thesauri already identify the sorts of relationships that are to be captured in a RT reference, and indicate that these specific types of RT can be indicated when applied, it is presumably issues of cost that prevent these from being indicated in practice. The developer of the thesaurus will presumably have a good idea of the type of RT in question when they identify an RT. A greater appreciation of the benefits—especially in a digital age—of specifying particular RT relationships, should encourage this practice. Several distinctions could usefully be drawn:

- Causal relationships of various types. ISO2788 distinguishes operation/process, agent/instrument, cause/effect, cause/affected, and thing/counter-agent.
- Other types of relationship. The UMLS thesaurus, used in the medical field distinguishes physically related to, spatially, temporally, functionally [a type of causation], and conceptually.
- Properties. ISO2788 distinguishes concept/property, concept/origin, action/property, and unit of measurement. Discipline/object could be considered a very specific type of property.

Other more specific types of related terms could likely be added. And improvements are also possible in other thesaural practices. With respect to hierarchical relationships, it is important to distinguish ‘type of’ (some would also distinguish

'instance of') from 'part of.' With respect to equivalence relationships, it would be helpful to distinguish the degree of equivalence.

Ontologies

It is clearly possible to develop upper-level (that is, comprehensive in coverage) ontologies. Several of these already exist (see Masolo et al. n.d.). The main challenge at present is to achieve any degree of consensus regarding these. It is also not clear that these ontologies actually serve the needs of the Semantic Web. Indeed Hart and Dolbear (2013) suggest that research on the Semantic Web has turned away from the exploration of ontologies.

The classifications of both phenomena and relationships provided for above could be useful in the development of upper level ontologies. These characteristically are each comprised both of things and relationships (and properties; Masolo et al. n.d., 43; Almeida et al. 2010). While precise definitions have not been attempted of each entry in the schedules of either BCC or ILC, this could be done. And then precise definitions of the much larger set of compound terms would follow. Relationships between classes in terms of integrative levels—that is, existential dependence between concepts—are a feature of phenomenon-based classifications still hardly considered in ontologies, where they could bring significant improvements.[4]

The classification urged in this book might thus serve as a stepping stone to an ontology. That ontology would inherit the various advantages of the classifications itself, while enhancing the ability of computers to navigate the classification. In later chapters it will be worthwhile to look at exactly what is required of an ontology for the Semantic Web and see to what extent this sort of classification is appropriate.

Key Points

The coincidental rise of digitization and interdisciplinarity in recent years makes it both desirable and feasible to develop a comprehensive and multi-dimensional classification. Works (and ideas) can be classified in terms of phenomena and relationships, theories and methods applied, and various other elements of authorial (including disciplinary) perspective. Numerous detailed practices for achieving

[4] The General Formal Ontology developed by Heinrich Herre and others considers levels of reality as one structuring principle, under influence of continental philosophy. Dependence between levels as a promising additional feature in ontologies has been discussed in a formal meeting between one of the authors (Gnoli) and ontologists at the University of Trento (Fumagalli, Maltese, Farazi and others).

each of these types of classification were outlined in this chapter. Such a classification system could either complement or substitute for existing systems. It in turn could be complemented by a comprehensive thesaurus and perhaps ontology.

References

Almeida MB, Souza RR, Fonseca F (2010) Semantics in the Semantic Web: a critical evaluation. Knowl Org 38(3):187–203

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

Barkow J, Cosmides L, Tooby J (eds) (1992) The adapted mind: evolutionary psychology and the generation of culture. Oxford University Press, New York

Beghtol C (1998) Knowledge domains: multidisciplinarity and bibliographic classification systems. Knowl Org 25(1/2):1–12

Broughton V (2010) Concepts and terms in the faceted classification: the case of UDC. Knowl Org 37(4):270–279

Chan L, O'Neill E (2010) FAST: faceted application of subject terminology: principles and application. Libraries Unlimited, Englewood, CO

Clavier V, Paganelli C (2012) Including authorial stance in the indexing of scientific documents. Knowl Org 39(4):292–299

Coates EJ (1988) Subject catalogues: headings and structure, 2nd edn. Library Association, London

Dean RJ (2003) FAST: development of simplified headings for metadata. http://www.oclc.org/research/projects/fast/international_auth200302.doc

Donovan JM (1991) Patron expectations about collocation: measuring the difference between the psychologically real and the really real. Catalog Classif Q 13(2S):23–43

Feinberg M (2011) How information systems communicate as documents: the concept of authorial voice. J Doc 67(6):1015–1037

Glushko RJ (ed) (2013) The discipline of organizing. MIT Press, Cambridge, MA

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

Gnoli C (2008) Categories and facets in integrative levels. Axiomathes 18(2):177–192

Gnoli C (2010) Themes and citation order in free classification. IASLIC Bull 55(1):13–19, http://arizona.openrepository.com/arizona/handle/10150/111813

Gnoli C (2012a) Metadata about what? Distinguishing between ontic, epistemic, and documental dimensions in Knowledge Organization. Knowl Org 39(4):268–275

Gnoli C (2012b) Vickery's late ideas on classification by phenomena and activities. In: Gilchrist A, Vernau J (eds) Facets of Knowledge Organization: proceedings of the ISKO UK second biennial conference. Emerald-Aslib, Bingley, pp 11–24

Gnoli C, Poli R (2004) Levels of reality and levels of representation. Knowl Org 31(3):151–160

Gnoli C, Merli G, Pavan G, Bernuzzi E, Priano M (2008) Freely faceted classification for a Web-based bibliographic archive: the BioAcoustic Reference Database. Repositories of knowledge in digital spaces: proceedings of the Eleventh German ISKO conference, Konstanz. Ergon, Würzburg

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

Green R (2008) Relationships in knowledge organization. Knowl Org 35(2/3):150–159

Greenberg J, Pattuelli MC, Parsia B, Robertson WD (2006) Author-generated Dublin Core metadata for web resources: a baseline study in an organization. J Digit Inform 2(20)
Hart G, Dolbear C (2013) Linked data: a geographic perspective. CRC, Boca Raton, FL
Hjørland B (2005) Empiricism, rationalism and positivism in library and information science. J Doc 61(1):130–155
Integrative Levels Classification (ILC) (2004) ISKO Italia. www.iskoi.org/ilc/
Iyer H (1995) Classificatory structures: concepts, relations and representation. Indeks Verlag, Frankfurt/Main
Khoo C, Na J-C (2006) Semantic relations in Information Science. Annu Rev Inform Sci Technol 40:157–228
Knapp JA (2012) Plugging the 'whole': librarians as interdisciplinary facilitators. Libr Rev 61 (3):199–214
Looney C, Donovan S, O'Rourke M, Crowley S, Eigenbrode SD, Rotschy L, Bosque-Perez NA, Wulfhorst JD (2014) Using Toolbox workshops to enhance cross-disciplinary communication. In: O'Rourke M, Crowley S, Eigenbrode SD, Wulfhorst JD (eds) Enhancing communication and collaboration in interdisciplinary research. Sage, Thousand Oaks, pp 220–243
López-Huertas MJ (2013) Transcultural categorization in contextualized domains. Inform Res 18 (3), http://InformationR.net/ir/18-3/colis/paperC16.html
López-Huertas MJ, Torres Ramírez I (2007) Gender terminology and indexing systems: the case of woman's body. Libri 57:34–44
Masolo C, Borgo S, Gangemi A, Guarino N, Oltramari A (n.d.) Ontology Library. Laboratory for Applied Ontology - ISTC-CNR. http://wonderweb.semanticweb.org/deliverables/documents/D18.pdf
Perrault JM (1994) Categories and relators: a new schema. Knowl Org 21(4):189–198
Repko AF (2012) Interdisciplinary research: process and theory, 2nd edn. Sage, Thousand Oaks
Repko AF, Newell WH, Szostak R (eds) (2012) Case studies in interdisciplinary research. Sage, Thousand Oaks
Szostak R (2003) A schema for unifying human science: interdisciplinary perspectives on culture. Susquehanna University Press, Selinsgrove, PA
Szostak R (2004) Classifying science: phenomena, data, theory, method, practice. Springer, Dordrecht
Szostak R (2005) Unifying ethics. University Press of America, Lanham, MD
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 (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 for social diversity. Knowl Org 41(2):160–170
Szostak R (2015a) Interdisciplinary and transdisciplinary multi-method and mixed methods research. In: Hesse-Biber S, Johnson RB (eds) The Oxford handbook of mixed and multi-method research. Oxford University Press, Oxford, pp 128–143
Szostak R (2015b) Classifying authorial perspective. Knowl Org 42(7):499–507
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
Tennis J (2002) Subject ontogeny: subject access through time and the dimensionality of classification. In: López-Huertas MJ (ed) Challenges in knowledge representation and organization

for the 21st century. Integration of knowledge across boundaries. Proceedings of 7th International ISKO Conference, Granada. Ergon, Würzburg, pp 54–59

Turner B (2000) Introduction: a new agenda for social theory? In: Turner B (ed) The new Blackwell companion to social theory. Blackwell, Oxford, pp 1–16

Van der Lecq R (2012) Why we talk: an interdisciplinary approach to the evolutionary origin of language. In: Repko A, Newell WH, Szostak R (eds) Case studies in interdisciplinary research. Sage, Thousand Oaks, pp 191–224

Vickery BC (2008) The structure of subject classifications for document retrieval. http://www.iskoi.org/ilc/vickery.php

Wu L-L, Huang M-H, Chen C-Y (2012) Citation patterns of the pre-web and web-prevalent environments: the moderating effects of domain knowledge. J Am Soc Inform Soc Technol 63(11):2182–2194