

Universality is Inescapable

Claudio Gnoli
University of Pavia
Science and Technology Library
via Ferrata 1
27100 Pavia, Italy
claudio.gnoli@unipv.it

Rick Szostak
University of Alberta
Department of Economics
Tory Building 9-18
Edmonton, Alberta, T6G 2H4, Canada
rszostak@ualberta.ca

ABSTRACT

This paper argues that universal classification systems – that is, systems that are general in coverage – are necessary in an era that values interoperability. Such systems have numerous other advantages. Questions regarding the feasibility of such KOSs can only be addressed empirically. The paper outlines avenues for empirical exploration.

Keywords

Universality, general classification, interoperability, empirical investigation.

INTRODUCTION

The call for papers of this workshop references a tension between a desire for universality, or at least for “harmonization: developing the means to reconcile diversity into a coherent structure”, and various concerns that this is neither feasible nor desirable. Indeed, a satisfying universal scheme of knowledge has always been a major dream of classificationists, while being criticized by others as a utopian aspiration.

The term *universal* was especially popular during the 20th century. It was used in the names of Otlet's Universal Decimal Classification, of Wählin's Universal System of Classification, of Scheele's Universal Faceted Classification, and of Shpackov's Universal Classification. The title of Dahlberg's seminal German book was “Foundations of universal knowledge organization” (Dahlberg 1974), and “Universal Classification” was the theme chosen for the fourth International Study Conference on Classification Research and sixth annual conference of Gesellschaft für Klassifikation – from which the

This is the space reserved for copyright notices.

Advances in Classification Research, 2014, November 1, 2014, Seattle, WA, USA.

Copyright notice continues right here.

International Society for Knowledge Organization (ISKO) later emerged (Dahlberg 1982-1983). IFLA also established an international program for Universal Bibliographic Control – another classical dream of librarians and bibliographers.

Since the end of the century, however, there has been a growing awareness of the cultural biases implied in universal systems, as usually privileging the perspective of Western, Christian, white, male, and heterosexual users. Some of these concerns have also emerged from such practical tasks as producing the Korean edition of DDC, where calligraphy clearly needed to be listed among the major arts, unlike in the West (Kwasnik & Chun 2004). These needs have been subsumed in the notion of *cultural warrant*, declared to be another requirement of knowledge organization systems along with the more traditional literary warrant (Beghtol 1986).

Supporters of domain analysis claim that the only solution to these challenges is to develop a plethora of systems explicitly biased, each representing the perspective of a different community. However, the current evolution of information systems brings a further need in front of us that domain analysis alone cannot solve: that of interoperability.

INTEROPERABILITY AND UNIVERSALITY

We are now used to the coexistence in the global information network of many classifications of different origins – Western or Eastern, academic or amateurish, scientific or documentary, by librarians or by computer scientists, etc. As information is increasingly shared and published in networked forms, a new awareness is growing of the need for its organization into consistent structures. The fashionable attention recently paid to a *Semantic Web* exploiting *taxonomies* and *ontologies*, after all, is but a reincarnation of the old need for classification.

This is even more true as the information units are becoming increasingly small and combinable: after the times of monographic volumes, then of articles published in a journal or a whole Web page, they now consist of posts,

of tweets, or even of individual data in the form of an RDF triple. Such a fragmentation potentially allows for a huge amount of creative combinations generating new knowledge. However, this can only be implemented if data are complemented with networked knowledge organization systems (cfr. the series of NKOS workshops) making their meanings and relationships explicit. In other words, interoperability needs not just to be technical, as considered in the most usual discussions concerning formats, markup and syntax of data: it also has to be conceptual. And universal classifications almost certainly play an important role in this.

It is our thesis that conceptual interoperability requires universal systems. We stress that we mean by “universal” simply a system that seeks an exhaustive coverage rather than the treatment of a limited domain (Szostak 2014d discusses how “universal” is sometimes interpreted as privileging one perspective). Limiting our interest to domain-specific systems would be an illusory strategy: indeed, any special classification actually implies a broader general classification, as has been shown by Foskett (1991). Local systems classifying the same phenomena according to different perspectives – say animals in academic zoology vs. in Borges's ultimate example of a Chinese list – can interoperate only through a common universal system (Gnoli 2011).

It is not by chance that general classifications have always played an important role in the history of knowledge – and of course are never complete nor perfect, but evolve together with knowledge itself.

As we cannot do without universal systems, the question is rather which universal systems actually offer the best way towards conceptual interoperability. The “universal” classifications cited in the call for papers are universal only to a degree. They are still grounded in disciplines, and tend to use different terminology and organizational structures for different disciplines. They are more universal than individual domain-specific classifications. But they still place substantial barriers in the way of interdisciplinary exploration. A much more universal basis for classification can be provided by phenomena. Although these are not absolute, eternal units of knowledge, as their identification can be partially influenced by theories (Gnoli et al. 2013), they do allow for a wider network of reference across disciplines.

FURTHER ADVANTAGES OF UNIVERSALITY

Universality is important not just for interoperability and the Semantic Web. It is also critically important for interdisciplinary research, which is of increasing importance in the academy. Scholars forced to navigate myriad domain-specific classifications would be much less likely to integrate insights from multiple fields, or to communicate any discoveries they did make to diverse audiences (Szostak 2008). Moreover, while

interdisciplinary scholars would struggle to master the terminology of diverse domain-specific classifications, a universal classification – if feasible – would serve to clarify the meaning of terminology. It was also the critical role of universality in meeting the needs of interdisciplinary research that inspired the present authors (among many others) to pursue research regarding the feasibility of universal classifications.

Szostak (2014a) argued that universality was also important for social diversity. Social understanding will be best served by a universal classification which allows users to readily access the works of other groups. Yet users can when they wish narrow their search to works pursuing a particular perspective if we will also classify works with respect to authorial perspective (a project being pursued at present). Szostak suggested that a mistaken impression that social diversity was best served by domain analysis might support (perhaps subconsciously) a belief that only domain-specific classification was feasible. It is thus particularly important that this question of feasibility be carefully evaluated.

Szostak (2014b) noted that doubts about feasibility are buttressed by various skeptical arguments in philosophy and literary theory. While it is tempting to see skepticism as the culmination of millennia of philosophical speculation, Collins (1998) has shown that various skeptical ideas – regarding ambiguity, subjectivity, the nature of reality, and a variety of other concerns – tend to arise together whenever there are “too many” competing schools of thought. If people are only aware of one way of looking at the world they tend toward epistemological optimism. If they cannot decide among many competing visions, epistemological skepticism is a natural outcome. The contemporary academy, with its congeries of disciplines, theories, and methods, seems a natural breeding ground for skepticism. Yet knowledge organization, by situating each work in a universal ‘map’ of the things studied and perspectives (including theories and methods) employed in scholarship, can substantially reduce both the reality and appearance of scholarly incoherence. If we do not assume the verity of various skeptical claims we can essay to achieve greater scholarly coherence. If we assume them to be correct we abandon the possibility of attacking this key source of skepticism. It is up to the small field of knowledge organization to decide on whether skepticism is accepted or rather investigated. [1]

Hjørland (2014) argues that KOSs should reflect accurately the way that scholars understand and categorize the world. His focus, of course, is on domains: it is not enough, he argues, for DDC to fiddle with biological classifications based on morphology and developed a century ago when biologists have turned toward cladistic classification. His argument can be extended, though, to a universal level: if the scholarly enterprise is not a congeries of isolated silos

but a reasonably coherent and interconnected enterprise, then it should be subject to universal classification.

There is, it should be noted, a dissonance between calling our field ‘knowledge organization’ and claiming that it is only possible to ‘organize’ by domain. If consensus were to be achieved that only domain analysis were possible, then ‘knowledges organization’ or ‘knowledge segmentation’ would be more appropriate titles for the field. The very naming of the field implies an interest in organizing knowledge as a whole, an interest that should not casually be set aside.

If universal classifications can reduce skepticism, enhance coherence, and facilitate interdisciplinarity, then they will enhance the rate of scholarly progress. Surveys establish that scholars – and especially interdisciplinary scholars – devote a considerable amount of time to both searching for relevant literature and identifying appropriate research outlets (e.g. Murphy 2003). If we can facilitate these tasks we can significantly enhance scholarly productivity (Szostak 2014c). The history of science literature (Root-Bernstein 1989), as well as the literatures on serendipity, undiscovered public knowledge, and literature based discovery (Makri and Blandford 2012; Davies, 1989; Swanson et al. 2001) all note that scholarly discoveries – and particularly major scholarly breakthroughs – often involve the juxtaposition of insights from diverse fields. It must be the case that a universal classification that made it easier for scholars to identify relevant information in diverse fields would facilitate this important form of scholarly discovery.

Hartel (2014) has recently urged information science to pay more attention to non-scholarly searches for answers to complex questions. She refers to the practice of people exploring issues as a ‘hobby.’ But in addition individuals investigate many issues that have an impact on their lives. Non-scholarly users are not likely to wish to master domain-specific classifications, but rather will likely prefer a universal classification that allows them to readily follow their curiosity to related topics. Better access to information will presumably make such users happier and better able to face life’s challenges. Better public appreciation of public policy issues should encourage better public policy as well.

Gnoli (2008) noted that a KOS should ideally be able to classify both texts and the objects that those texts study. A similar argument could be made with respect to ideas. Szostak (2015) has argued that texts should be classified in terms of the key ideas that they contain. But it is well known that the same objects are studied by multiple disciplines. And ideas too – think of evolutionary theory – are often pursued in multiple fields. We have suggested above that texts are best classified in a universal manner. The case is even stronger with respect to objects and ideas,

for these are far less likely to respect disciplinary boundaries.

The foregoing has indicated various advantages of universality. These should not be forgotten when debating the possibilities of universality. Of course, if universality cannot be achieved, these potential advantages must be foregone. But given the myriad advantages of universality, this potential should not be discarded until impossibility has been established beyond any reasonable doubt. This debate has implications far beyond the field of knowledge organization, for universality if feasible enhances the possibilities of scholarship itself, and also social understanding and cohesion, personal fulfilment, and public policy.

The foregoing also has an implication for the type of universality to be sought. Present GCSs are not truly universal: different domains are often organized like sealed compartments, so that, for example, DDC classmarks for wood in forestry, for wooden structures in engineering, for wooden buildings in construction, and for wooden buildings in architecture have no connection at all. The interdisciplinary researcher and general user still struggle to search across domains. Computers must face even greater difficulties. The case has been made, then, not just for universality but for an even more universalist approach to classification than has been achieved by extant GCSs.

FEASIBILITY

Is universality possible? And if so, how much universality is feasible? The important point to stress here is that these questions can only be answered empirically. There are theoretical arguments that can be adduced both in favor of and opposed to the idea of universality. Arguments on both sides have their merits, but none are decisive. To be sure, the authors of this paper believe that the theoretical arguments in favor of universality are stronger. But they appreciate that these arguments need to be buttressed by empirical verification. And of course the question of “How much?” begs empirical evaluation.

We can think of two internally coherent theoretical perspectives. The debate between Hjørland and Szostak over the last few years (e.g. Hjørland 2008, 2009, Szostak 2008, 2011, Fox 2013) has served to clarify what these might look like. Hjørland prefers a very pessimistic philosophical concept theory, “theory theory.” This theory argues that all concepts are grounded in theories and therefore cannot be well understood across groups with different theories and beliefs. These concepts can only be understood with reference to other concepts grounded in the same theory. The best classifications are thus domain-specific, with concepts defined internally to a domain. Translations across domains would be difficult but possible, but would have to be performed for each pair of domains. Szostak alternatively argues that information science should not be limited to a single philosophical view, given that

concept theory is a contentious field. Rather we should seek to ground KOSs in the widest range of philosophical argument possible. Most concept theories argue that complex concepts can be broken into more basic concepts. Conceptual atomism suggests that there will be much shared understanding across groups and individuals of those things and relators (and perhaps properties) that we regularly observe in the world around us. Szostak thus advocates a universal classification grounded as much as possible in these sorts of basic concepts.

The empirical nature of these theoretical conclusions deserves to be stressed. Hjørland's conclusion is effectively that there is too much ambiguity in the world for universality to be possible but not so much as to prevent domain-specific classification. Szostak's conclusion is that there is (likely) not so much ambiguity as to prevent universality. But one should be wary of leaping straight from theory to empirical conclusions. Rather, standard scientific practice is to use theory to derive empirical hypotheses and then test these. This paper essays to show how this can be done.

Faced with competing and internally consistent theories regarding what an entire scientific field should be doing – pursuing only domain analysis or pursuing also the possibility of universal classification – recourse to empirical investigation is called for. This is simply good scientific practice. Information science is often guilty of not connecting its theoretical and practical discourses (Cornelius 2014); we should do so here.

What sort of empirical evidence might we look for? We might start with evidence that the theoretical perspective articulated by Hjørland is too pessimistic. The very fact that people from different communities or disciplines regularly communicate suggests that theory theory exaggerates the difficulties in cross-domain understanding. It may well be that the explanation of this empirical reality is that sentences are far less ambiguous than individual concepts, as any ambiguous term can be explained by a periphrasis if researchers are willing to do so, but even so this would indicate the possibility of simple strategies for alleviating conceptual misunderstanding.

In the end, though, we need positive evidence of the feasibility of a universal classification. The existence and use of extant GCSs establishes that some degree of universality is associated with some degree of utility. Users find things, though not as quickly or precisely as we might like.

Is a truly universal classification feasible: one not grounded in disciplines but that employs the same structure throughout? Such a test must involve attempting to develop a truly universal classification, and evaluating it. The challenge is that a KOS grounded in the phenomena we

study rather than disciplines must be developed to a considerable extent before any sort of test is feasible. We should not accentuate this difficulty by demanding that a nascent classification match each of the capabilities of classifications that have been developed for over a century by dedicated bureaucracies. And we should ask whether any difficulties identified in such a KOS could be transcended, as have myriad difficulties been addressed over a century and a half with respect to DDC and LCC.

Happily the Integrative Levels Classification (ILC, 2004) and Basic Concepts Classification (Szostak 2013) have been developed to a point where some sorts of applications and evaluations are possible (and are beginning to be pursued). Their further development could be aided by constructive evaluation. It is thus appropriate at this point in time to reflect on the sorts of empirical inquiries that might be addressed to such classifications.

It should be stressed that any test of feasibility cannot demand perfection. Even domain-specific classifications do not entirely eliminate ambiguity. Importantly, though, this is not necessary. We argued above for the need to wed theory and empirics. This becomes even more important when theories are drawn from another field. Philosophers have/had for thousands of years sought absolute precision in argumentation. Their continued failure to define concepts such as 'freedom' with absolute precision (despite, it might be noted, operating within the same domain) forced a painful re-thinking within the philosophical profession. A somewhat similar story can be told in literary theory, another common source of skeptical theories: the recognition that texts were ambiguous challenged a longstanding tradition of identifying 'the theme' of a work. Neither philosophers nor literary theorists are prone in any case to asking 'how much?' With respect to conceptual ambiguity, the fact that any degree of ambiguity challenged the basic precepts of their fields ensured that it was not seen as necessary to do so. But the practical field of knowledge organization cannot simply absorb the message that 'conceptual ambiguity is a huge problem that limits what we can do.' Rather we need to ask 'how much?'

It deserves emphasis that philosophers were troubled by ambiguity even though they operated within the same domain. That is, they could not achieve the precision they desired in communicating with their closest colleagues. If knowledge organization is to be troubled by the same degree of ambiguity that has offended philosophers then it is far from clear that even domain analysis is possible. Thus, whether we suspect universality is feasible or not we cannot avoid asking 'How much?'

Indeed we need to ask two separate 'how much?' questions: How much conceptual ambiguity is left, after we have tried all ambiguity-reducing strategies?; and How much conceptual ambiguity can a KOS abide and still achieve

acceptable levels of precision and recall? We would strive for a KOS in which the level of ambiguity achieved is no greater than the level required. There is, it must be admitted, no obvious metric by which these 'how much?' questions can be evaluated quantitatively. But we can use our collective judgment in answering both questions, and especially in comparing the answers to the two questions.

The first question is the easier to address. Can the common structures employed in interdisciplinary classifications be understood in a broadly similar way across groups and individuals? Hjørland and theory theory suggest that this cannot be the case. The BCC and ILC have struggled to use wherever possible basic concepts (things and relationships) that lend themselves to broadly shared understandings. Users from different backgrounds might simply be asked to define randomly selected terms from these classifications, and content analysis employed to evaluate the similarity of these definitions. Results could be compared across classifications: domain specific; existing GCSs; interdisciplinary GCSs.. It would be useful for us to know how different these different types of classification prove to be in this respect. It will also be useful to have some idea of how ambiguity varies across fields. It could well prove that it is easier to achieve shared understandings of basic mathematical concepts than of common terminology regarding political ideology or art criticism. But we cannot know unless we ask. In the meantime, we should hardly dismiss the possibility of universality just because a few concepts are found to generate diverse understandings: we should instead essay to reduce this ambiguity.

The second question is most obviously approached through user testing. And this will likely first require that the BCC or ILC be applied to some databases of size (though we are certainly open to other research strategies). [2] It would then be possible to compare these to existing GCSs: how quickly and precisely can users find works in these databases? The comparison with domain-specific classifications is less obvious. It is certainly desirable to compare these on domain-specific inquiries. Until diverse domain-specific classifications are developed and linked by pairwise translation devices (which may never happen) a comparison of their performance on user queries across diverse fields, of the kind of the Cranfield experiments (Cleverdon 1967), is not possible. Yet the discussion above suggests that these sorts of queries are of great importance.

We might indeed imagine a host of tests, each guided by our discussion in previous sections: for the sort of juxtapositions of related insights from different fields identified in the literature-based discovery and related literatures (we could test, that is, for whether users are guided to related works in distinct fields); for the sorts of complex queries that general users often have; for queries that members of one group or discipline might make about the practices or beliefs of other groups or disciplines; or for

information on public policy questions. We might also test whether a classification is easy to understand, and if so whether it provides users with a sense of a coherent body of human understanding. We should ideally also test with respect to different types of user: academic versus general; scholars from different disciplines and those who are interdisciplinary; and novice users versus experienced users. With respect to this latter comparison it is noteworthy that some small public libraries in North America have switched from DDC to BISAC, the system used in bookstores, due in large part to a concern that their users find DDC scary.

Green (1995) notes that we tend to presume that a user seeks a particular type of work. But users with complex queries may instead need to be guided to multiple works that each address part of their query. This is obviously the case in scholarship where we inevitably cite different sources for different parts of our argument. User testing should thus not just investigate whether users find relevant works but whether they find works that address each component of a complex query. Here the ability to move along a web-of-relations from one work to a related work may prove crucial. Szostak (2014a) has argued that a synthetic approach facilitates such a web-of-relations. But such a theoretical conjecture requires empirical investigation. Users could be given multi-part queries and directed to identify a set of relevant resources.

User studies have naturally tended to focus on meeting user needs. But we can start from the other end of the link between user and work: are works classified in a manner that facilitates their being located by relevant users? Soergel (1985, 233-4) discusses the challenges in doing so. [3] If interdisciplinary GCSs were applied to even a small but diverse corpus of works, users could be surveyed on whether they found the subject headings derived to be more or less evocative than those of other classifications. And we could usefully test whether users from different domains have similar opinions of particular subject headings.

We could also test whether classifiers found it easier to use one KOS or another. It might be best to employ LIS students for such a test rather than classifiers who have employed a particular KOS for years.

It was suggested above that a KOS should ideally address not just books in libraries but any phenomenon when treated as an object of knowledge (as a document), including specimens in museums or in gardens (Gnoli 2010). This is in agreement with the current trend toward grouping libraries, archives, museums and galleries in a single theoretical framework. The tests outlined in the preceding two paragraphs could then be performed not just for books but for other document types as well. Since the ILC and BCC are conceived to organize phenomena rather

than disciplines, they might be expected to perform well in these respects.

Last but not least we can test the susceptibility of each type of classification to use with linked data. Again theory can be wedded to empirics. The requirements of interoperability and the Semantic Web were reviewed at the start of the paper, and it was argued that enhanced universality would best meet these needs. We need, though, to see in practice whether this is the case.

This paper has suggested that – if certain theoretical critiques of the possibility of universality can be overcome – greater universality than is at present achieved by GCSs is both desirable and feasible. This section has outlined a set of empirical tests that could establish whether this is indeed the case.

If a new KOS were found to be superior to extant KOSs, we might even aspire to compare yet again the value of subject searching versus keyword searching. Information retrieval has gained prestige from empirical evidence that users both prefer and perform certain sorts of search better using keyword searches (Hjørland 2012). Yet scholars in the information retrieval field have grown unhappy with the limited precision accorded by keyword searches. A phenomenon-based KOS that employed a synthetic approach among basic concepts might prove attractive to users while yielding much greater precision than keyword searches. A test of this sort may be most valuable after a thesaurus is constructed in concert with these new KOSs.

TESTING STRATEGIES FOR REDUCING AMBIGUITY

The preceding section emphasized tests that could be performed at the level of entire KOSs. But a set of empirical tests of particular classificatory strategies is also both desirable and feasible. It was noted above that we should in any KOS pursue ambiguity-reducing strategies prior to evaluating the level of ambiguity achieved. But which ambiguity-reducing strategies are most important in reducing ambiguity, and how successful are they? Such strategies might potentially be utilized within different classificatory structures, even GCSs.

The practice of breaking complex concepts into basic concepts can perhaps best be evaluated by comparing KOSs that do this with KOSs that do not, as in the preceding section. But other sorts of empirical evaluation are also feasible. We can, for example, examine how users break complex concepts into basic concepts. [4] Note that one of the reasons that the BCC prefers basic concepts is that complex concepts are ambiguous: the use of basic concepts thus makes clear what was previously not. ILC allows for representation of a complex concept either by a combination of facets, or as a simple notation with links to concepts implicit in its definition recorded in the schedules,

depending on how stable and autonomous the result of the combination is (Gnoli 2013).

It should not be expected, then, that diverse users will indeed break a complex concept into the same set of basic concepts. Though both sides in the theoretical dispute outlined above might predict that different users will use different descriptors of complex concepts, they make different predictions regarding how and why this occurs. It would be particularly interesting to find out whether users within a particular domain are less likely to ascribe different sets of basic concepts to a complex concept employed in their domain. It would also be invaluable to first establish whether there are clear distinctions by domain, and if so whether users from one domain can understand the different interpretations achieved by users from other domains. If one domain translates ‘globalization’ as (American)(movies)(influence)(French)(cultural attitudes), can users nevertheless appreciate that another domain translates ‘globalization’ as (foreign trade)(affects)(wages)?

It was noted above that sentences or phrases are likely less ambiguous than individual concepts. Warner (2009) urges the abandonment of an ‘outdated’ emphasis on words in search in favor of phrases and multiword sequences (though his focus is information retrieval rather than classification). Blair (2006, 347-8) argues that the easiest way to ease the semantic ambiguity associated with the indeterminate meaning of words is to place these in sentences. [5] Frické (2012) urges a synthetic approach that would allow a clear distinction between ‘large library school’ and ‘large school library.’ The synthetic approach thus may provide much greater clarity than the common preference for very short subject headings and constrained use of synthesis. [6] Communications research has already shown empirically that phrases are less ambiguous than individual concepts (Keyton and Beck 2010). Still, further research may be desirable as to whether in the context of KOSs strings of three to five concepts provide greater clarity than one or two, and how manageable they are for the purposes of documentation. An optimal balance between analysis and synthesis of concepts should be identified.

A sentence-like structure cannot be achieved easily with exclusive reliance on noun-like terminology. GCSs tend to stress nouns and to a lesser extent adjectives. It could usefully be explored whether combinations of noun-like terms with relationship terms, (which maybe expressed in facet indicators), leads to greater clarity. [7] If this were to be the case, the likely explanation would be that such an approach mimics the sentence structure with which we regularly communicate.

Weinberg (1988) famously noted the limitations of classifying scholarly works in terms only of ‘aboutness.’ As we noted at this workshop some years ago (Szostak and

Gnoli 2009), scholarly users in particular want to know what is novel about a work. This novelty can best be captured by indicating the key arguments made (the *rheme*, cfr. Gnoli & Cheti 2013), and the perspectives applied. A synthetic approach to classifying that has a place for relationships allows works to be classified in terms of their key arguments, for these are commonly something like X affects Y in manner Z. Users could be tested as to whether such a subject heading gives them a better sense of the relevance of a work. [8] [It is somewhat beyond the focus of this paper but it would be useful to evaluate the impact of information regarding the theories, methods, and perspectives applied in a work on determining relevance; Szostak and Gnoli (2008) have developed and applied classifications of theory types and methods; a classification of perspectives is being pursued.]

Facet analysis has long been stressed in theoretical writings in the field, although it has only slowly been adopted within discipline-based GCSs. The ILC and BCC have naturally absorbed the principles of facet analysis from the outset. They differ only in the protocols by which they apply facet analysis. The ILC employs facet indicators as an option but unlike most faceted schemes allows the free combination of any terms. The BCC eschews facet indicators; while it also allows free combination of any terms it recommends a sentence-like combinatory structure such that facets are clearly indicated by a term's place in the concept string. [9] It would be useful to test the value of an explicitly faceted approach. It will also be useful to evaluate whether users should be made familiar with the nature of facets or whether these best operate behind the scenes.

SOME THOUGHTS ON USER STUDIES

Tenopir (2003) provides a useful survey of the methodology of user studies. There are two broad types: one models complex information-seeking behavior as a whole, testing and refining models through experiments and observation; the other surveys or observes large groups, often longitudinally. Both types have tended to focus on issues other than classification. The sorts of user studies recommended above would be of the second type, with an emphasis on the role of classification in search.

The earliest user studies stressed the quantitative analysis of survey results. Over time the qualitative evidence that can be obtained from open-ended interviews and observation has become increasingly appreciated. Wilson (2006 [1981]) stresses the value of the qualitative approach. Tenopir (2003) admits a personal preference for studying what users do rather than what they say. Observation, it should be noted, can be highly structured through query logs or usability testing, though it is invaluable to also debrief users after any experiment (It can also be useful sometimes for users to verbalize their experience during an experiment.) One important form of usability testing involves expert

review. While some of the empirical tests suggested above involve general users others require experts.

Studies have shown that user behavior differs by discipline, motivation, purpose, and other factors. It is thus appropriate that we have suggested above that different types of user and query be investigated separately.

Though Wilson (2006) doubts the contribution of at least the earliest user studies, it does seem that this literature has produced some robust results. The challenge is to design studies that carefully evaluate precise hypotheses. We have striven in the previous section to identify a set of carefully articulated hypotheses worthy of investigation. We have essayed to show how user studies might proceed. We fully appreciate, nevertheless, that such studies are challenging and are best pursued in the light of the lessons learned from decades of user studies.

CONCLUDING REMARKS

This paper began by making a critical argument: that the interoperability urged by contemporary developments in digital technology, and especially the Semantic Web, can only be achieved through universal classification.

Universality serves diverse other purposes as well: interdisciplinarity, scholarly cohesion and progress, social diversity, personal fulfilment, and public policy.

The question of how much universality is possible is thus of great import. Theoretical discourse on its own cannot answer this question. This paper thus outlined a set of empirical strategies for evaluating feasibility. The first set of these propose to compare nascent phenomenon-based universal classifications both with discipline-based universal classifications and with domain-specific classifications. The second set focus instead on evaluating a set of strategies for reducing ambiguity. The feasibility of universality likely depends on at least some of these strategies proving to be successful.

Information scientists have borrowed theories from philosophy and literary theory which suggest pessimism regarding universality. It is important to appreciate that in both fields the existence of any slight degree of conceptual ambiguity was highly problematic. It is highly likely that KOSs can withstand much greater conceptual ambiguity than that which offends philosophers or literary theorists. We can thus only evaluate the feasibility of any KOS, including universal classifications, empirically.

ENDNOTES

¹ When Szostak made this argument at the 2014 ISKO conference in Krakow, Hjørland noted that the skeptical ideas associated with Kuhn are now widely accepted. He suggested that Szostak would be the skeptic in rejecting these. To be sure the idea of scientific revolutions is widely accepted, and is grounded in the sort of detailed historical research that Collins is applauded for. We thus

need a more nuanced view of 'scientific progress' (and we are pleased to note that Hjørland does accept the idea of scientific progress). But Kuhn's idea of incommensurability, which might underpin an argument for only doing domain analysis, is highly contested to this day. Kuhn quite plausibly exaggerated here. As Szostak noted in Krakow, he does not argue that all skeptical arguments are misguided, but rather that some may be exaggerated.

² A comparison has been performed of ILC and DDC with respect to a small set of works on ecology. It is discussed in a book under review written by the authors of this paper with Maria Lopez-Huertas.

³ Soergel notes that classifiers can only be expected to keep perhaps 1000 terms in memory. Beyond this they need access to some sort of hierarchical organization, but this needs to be easy to navigate. It is plausible that a synthetic approach utilizing basic concepts arranged in compact logical hierarchies will be easier to employ. Julien et al (2013) have found that the hierarchies of extant classifications are generally implicit rather than explicit, and often have multiple levels which differ greatly in the number of entries.

⁴ Such a project bears much similarity with the definitional exercise recommended at the start of the preceding section. Our focus here, though, is on translating complex concepts rather than basic concepts. We could here give users a list of basic concepts to choose from.

⁵ Blair notes that search engines such as google now characteristically show users the sentence in which the search word appears. But this is still not as useful as searching by sentence.

⁶ Indeed the synthetic approach arguably provides several advantages (in addition to those associated with universality itself above). It allows subject headings to capture the core arguments of a work. It is hospitable, as new combinations of existing terms can be readily generated. Combinations are often treated as subclasses within enumerative schemes (recycling as a subclass of garbage, for example); a synthetic approach facilitates logical hierarchies.

⁷ Khoo and Na (2006) address the importance of relationship terms. They stress that information retrieval has proceeded about as far as it can without some more explicit modeling of semantic relationships.

⁸ Cornelius (2013) has worried that increased access to information has not led to notably better decision-making in the world. It could be that difficulties in identifying the relevance of works is largely responsible.

⁹ Details are provided in Szostak (2013). If a work is classified as (X)(has effect Z)(on Y) it is clear that X is the agent.

REFERENCES

Beghtol, Clare (1986) Semantic validity: Concepts of warrant in bibliographic classification systems. *Library Resources & Technical Services*, 30(2), 109-125.

Blair, David, Wittgenstein, Language, and Information: 'Back to the Rough Ground!' Dordrecht: Springer 2006.

Cleverdon, Cyril W. (1967) The Cranfield tests on index language devices. *Aslib Proceedings*, 19(6), 173-194.

Collins, Randall (1998) *The Sociology of Philosophies: A Global Theory of Intellectual Change*. Cambridge, Mass: Harvard University Press.

Cornelius Ian (2013) Epistemological Challenges for Information Science: Constructing Information. In Fidelia Ibekwe-SanJuan and Thomas M Dousa, eds., *Theories of Information, Communication and Knowledge: A Multidisciplinary Approach*. Berlin: Springer, 181-203.

Dahlberg, Ingetraut, ed. (1982-1983) Universal classification. 1: Subject analysis and ordering systems: Proceedings, 4th International study conference on classification research, 6th Annual conference of Gesellschaft für Klassifikation, Augsburg, 28 June-2 July 1982. Frankfurt: Indeks.

Dahlberg, Ingetraut (1974) *Grundlagen universaler Wissensordnung: Probleme und Möglichkeiten eines universalen Klassifikationssystems des Wissens*, Dokumentation, München.

Davies, Ron. 1989. "The creation of new knowledge by information retrieval and classification." *Journal of Documentation* 45(4): 273-301.

Foskett, D. J. (1991) "Concerning General and Specific Classifications" *International Classification* 18:2, 87-91.

Fox, Melodie J. (2012) Review of two books by Szostak and the Hjørland – Szostak dialogue. *Knowledge Organization* 39(4): 300-3.

Frické Martin (2012) *Logic and the Organization of Information*. Berlin: Springer.

Gnoli, Claudio (2013) Facets, levels and semantic factoring, *SRELS Journal of Information Management*, 50(6), 751-762.

Gnoli, Claudio (2011) Animals belonging to the emperor: enabling viewpoint warrant in classification, in Patrice Landry, Leda Bultrini, Edward T O'Neill, and Sandra K Roe eds., *Subject access: preparing for the future*. Berlin: De Gruyter, p. 91-100.

Gnoli, Claudio (2010), Classification transcends library business, *Knowledge organization* 37(3), 223-229.

Gnoli, Claudio, and Alberto Cheti (2013) "Sorting documents by base theme with synthetic classification: the double query method", in *Classification & visualization: interfaces to knowledge: proceedings of the international UDC Seminar, 24-25 October 2013, The Hague, The Netherlands*, eds. Aida Slavic, Almila Akdag Salah, Sylvie Davies. Würzburg: Ergon, 225-232.

Gnoli, Claudio, Kleineberg, Michael, and Szostak, Rick (2013) The blind knowledge organizers and the elephant: Working notes on Kleineberg's levels of knowing, in *Integrative Levels Classification*, ISKO Italia, <http://www.iskoi.org/ilc/elephant.php>.

- Gnoli, Claudio, and Rick Szostak (2008) "Classifying by phenomena, theories, and methods: Examples with focused social science theories." In Culture and identity in knowledge organization, Proceedings of the 10th international ISKO conference, Montréal, eds. Clement Arsenault and Joseph Tennis, 205-211. Würzburg: Ergon.
- Green, Rebecca (1995) Topical relevance relationships. 1. Why topic matching fails. *Journal of the American Society for Information Science and Technology* 46:9, 646-53.
- Hartel, Jenna (2014). An Interdisciplinary platform for information behavior research in the liberal arts hobby. *Journal of Documentation*, 70(5).
- Hjørland, Birger (2014) "Are relations in thesauri 'context-free, definitional, and true in all possible worlds?'" *Journal of the American Society for Information Science and Technology*. Published online, July, 2014.
- Hjørland, Birger. (2009) "Concept theory." *Journal of the American Society for Information Science and Technology* 60(8): 1519-36.
- Hjørland, Birger (2008) Core classification theory: A reply to Szostak. *Journal of Documentation* 64(3): 333-342.
- Integrative Levels Classification (ILC) (2004) ISKO Italia. Available at: www.iskoi.org/ilc/
- Julien, Charles-Antoine, Tirilly, Pierre, Dinneen, Jesse, and Guastavino, Catherine. 2013. "Reducing subject tree browsing complexity." *Journal of the American Society for Information Science & Technology* 64, 2201–23.
- Keyton, Joann, and Stephenson J. Beck (2010) Perspective: Examining Communication as Macrocognition. *STS Human Factors* 52(2): 335–339.
- Khoo, Christopher S. G. and fin-Cheon Na (2006) Semantic Relations in Information Science *Annual Review of Information Science and Technology*, 157-228.
- Kwasnik, Barbara H., and You-Lee Chun (2004) Translation of classifications: Issues and solutions as exemplified in the Korean Decimal Classification, in Ia C. McIlwaine ed., Knowledge organization and the global information society: Proceedings of the 8th International ISKO Conference, 13-16 July 2004, London. Würzburg: Ergon, p. 193-198.
- Makri, Stephann, and Blandford, Ann (2012) "Coming across information serendipitously – Part 1: A process model." *Journal of Documentation* 68(5): 684–705.
- Murphy, Janet (2003) "Information-Seeking Habits of Environmental Scientists" *Issues in Science and Technology Librarianship* 38 (Summer).
- Root-Bernstein, Robert (1989) *Discovery*. Cambridge Mass: Harvard University Press.
- Soergel, Dagobert (1985) *Organizing Information: Principles of Data Base and Retrieval Systems*. San Diego: Academic Press.
- Swanson, Don R., Neil R. Smalheiser, and A. Bookstein (2001) "Information Discovery from Complementary Literatures: Categorizing Viruses as Potential Weapons." *Journal of the American Society for Information Science and Technology* 52, 797-812.
- Szostak, Rick (2015) "A Pluralistic Approach to the Philosophy of Classification." *Library Trends*.
- Szostak, Rick (2014a) Classifying for Social Diversity. *Knowledge Organization* 41(2): 160-70.
- Szostak, Rick. (2014b) "Skepticism and Knowledge Organization." In Weislaw Babik, ed., *Knowledge Organization in the 21st Century: Between Historical Patterns and Future Prospects; Proceedings of the 2014 conference of the International Society for Knowledge Organization* (Krakow, May, 2014). Würzburg: Erlon Verlag.
- Szostak, Rick. (2014c) "The Importance of Knowledge Organization." ASIST Bulletin, April/May, 37-42.
- Szostak, Rick. (2014d) "How universal is universality?" [Letter to the editor.] *Knowledge Organization* 41(6): 468-70.
- Szostak, Rick (2013) *Basic Concepts Classification*. <https://sites.google.com/a/ualberta.ca/rick-szostak/research/basic-concepts-classification-web-version-2013> Retrieved Sept. 30, 2014.
- Szostak, Rick (2011) Complex Concepts into Basic Concepts. *Journal of the American Society for Information Science & Technology* 62:11, 2247-65.
- Szostak, Rick (2008) "Interdisciplinarity and Classification: A Reply to Hjørland" *Journal of Documentation* 64:4.
- Szostak, Rick and Claudio Gnoli (2009) "Beyond Aboutness: Classifying Causal Links in the Service of Interdisciplinarity" Proceedings of ASIST Special Interest Group on Classification Research 20th Workshop, Vancouver, November 7, 2009. <https://journals.lib.washington.edu/index.php/acro/article/view/12882/11378>
- Tenopir, Carol (2003) "Information metrics and user studies", *Aslib Proceedings* 55:1/2.13-17.
- Warner, Julian (2009) *Human Information Retrieval*. Cambridge: MIT Press.
- Wilson, T. (2006) "On user studies and information needs" *Journal of Documentation* 62:6, 658-70.