# A Pluralistic Approach to the Philosophy of Classification

Rick Szostak

## Abstract

Any classification system should be evaluated with respect to a variety of philosophical and practical concerns. This paper explores several distinct issues: the nature of a work, the value of a statement, the contribution of information science to philosophy, the nature of hierarchy, ethical evaluation, pre- versus postcoordination, the lived experience of librarians, and formalization versus natural language. It evaluates a particular approach to classification in terms of each of these but draws general lessons for philosophical evaluation. That approach to classification emphasizes the free combination of basic concepts representing both real things in the world and the relationships among these; works are also classified in terms of theories, methods, and perspectives applied.

## Introduction

This paper has both a general argument and a particular argument. The general argument is that any classification system should be judged in terms of a range of philosophical concerns—and diverse philosophical points of view on those concerns—as well as practical considerations. That is, we should not evaluate a classification simply in terms of only one philosophical concern (say, ambiguity), and certainly not with respect to one particular philosophical point of view on that issue. The particular argument is that a certain approach to classification (to be outlined below) provides very good answers to a very wide range of philosophical concerns and thus deserves serious investigation by the information science community. The particular argument thus serves as an example of the general argument and how this can be applied.

It is not possible to deduce the ideal set of philosophical concerns against which a classification should be evaluated. This paper therefore

LIBRARY TRENDS, Vol. 63, No. 3, 2015 ("Exploring Philosophies of Information," edited by Ken Herold), pp. 591–614. © 2015 The Board of Trustees, University of Illinois

pursues an inductive approach, addressing key concerns outlined in the two works identified in the call for proposals for this present issue: Furner (2010) and the original issue on the philosophy of information science in this journal in 2004.[1] It is not pretended that this set is exhaustive—indeed, some other concerns will be referenced in the conclusion—but it is claimed that the set of issues addressed here are both important and diverse. They thus collectively comprise a very demanding set of evaluative criteria for any classification. Each issue raised is important enough to merit attention in the evaluation of any classification.

The bulk of the paper is devoted to addressing these various philosophical concerns. A couple of preliminary arguments are useful. The first involves an appreciation of a practical, or applied or perhaps pragmatic in the broad sense, approach to the philosophy of information science. The field of medical ethics provides an exemplar here, for discussions in that field are generally grounded in discussions of actual problems faced in medical practice. Likewise philosophical discussions in information science should be grounded in discussions of actual decisions that information scientists need to make. Only in this way can we hope to identify approaches to classification that are simultaneously philosophically justifiable and practical.

The second is a general plea for plurality. Philosophers appreciate that there is no one perfect philosophical argument. Yet, in practice, philosophers tend to argue for the superiority of one particular philosophical approach to an issue. As both Stock (2010) and Szostak (2011) have recently noted, it would be a mistake for information scientists to ground their work in any one philosophical approach. This preference for plurality is shared by most, but not all, of the authors in Ibekwe-SanJuan and Dousa (2014). Hjørland (2014) is in the skeptical minority in that volume and worries that "epistemological promiscuity" threatens the unity and coherence of the field of information science. This paper would urge instead a consensus around the principle of plurality and pluralistic standards of evaluation. Philosophy itself prospers through consensus regarding standards of evaluation more than consensus around particular theories. Since no philosophy is perfect, and most philosophical approaches likely have some merit, it is better for information science to try to satisfy the widest range of philosophical theorizing. Stock and Szostak were referring to diverse philosophical approaches to a particular issue: epistemology and concept theory, respectively. This paper extends the argument for plurality across issues: we should seek to address the widest range of philosophical thinking on the widest range of philosophical concerns.

That may seem like a tall order. And it might reasonably be suspected that some trade-offs between satisfying one criteria or another will be called for in any classification system. But such a possibility can only be addressed after it is seen how any particular classification survives this broad

set of tests. Although space prevents us from comparing multiple classifications in this paper, it should be noted that ideally, a classification would be evaluated comparatively. A comparative approach protects us against errors of two types: decrying a classification because it does not flawlessly address a particular concern, or alternatively celebrating a classification because it appears to address a concern reasonably well. We should always be ready to ask whether an alternative approach can be superior. While a formal comparison with an alternative classification is not carried through here, comparisons with alternatives will frequently be made.

## The Particular Approach

As noted above, this paper will subject a particular approach to classification to these various tests. The paper thus serves both as an examination of that approach and an exemplar of how other classifications might be evaluated. This approach to classification has the following key characteristics:

- *Complex concepts*, which lend themselves to differing interpretations across disciplines and groups, are broken into *basic concepts* that carry broadly similar meanings across individuals and groups.
- These basic concepts generally refer to "things" in the world or relationships among things (Szostak, 2011) but sometimes to properties of things or relationships.
- Works and ideas are then classified in terms of free combinations of any set of basic concepts.
- Most works and ideas will be classified in terms of a synthetic combination of things and relationships.
- Works will also be classified in terms of theories, methods, or perspectives applied in the work. Fairly small schedules thus allow very detailed and precise classifications of works or ideas (Szostak, 2013a).

Szostak (2011) argued that ambiguity could be substantially reduced by breaking complex concepts, which are understood differently across disciplines or groups, into their constituent basic concepts, which are understood in a broadly similar fashion across groups and disciplines. Conceptual atomism suggests that shared understanding will be most likely for things and relationships (and properties) that we regularly perceive. I thus argued in favor of a classification of works in terms of combinations of such things and relationships.

Hjørland (2014 and elsewhere) and others have suggested that classification research should *exclusively* pursue domain analysis: the careful analysis of the terminology employed in a particular field. Hjørland worries that terminology can only be understood within a community with shared theories or beliefs. An approach that argues that terminological ambiguity can be reduced enough that a comprehensive phenomenon-based

classification is feasible must therefore confront the view that this is impossible (see Fox, 2012; Hjørland, 2008; Szostak, 2008, 2011, 2013b). Szostak (2010) argued that domain analysis and the pursuit of a comprehensive classification were complementary activities: domain analysis could ensure that the terminology of any domain was translated into the basic concepts employed across a comprehensive classification.

The vast majority of scholarly works address how one or more things affect in a particular way one or more other things. The same is true for many/most general works of nonfiction: (gardener) (grows) (flowers) or (dogs) (bite) (mail carriers). The very best way to describe such works is in terms of combinations of basic things and relationships. If a work is about dogs biting mail carriers, the ideal subject heading is (mail carriers) (bitten by) (dogs). Works that merely describe the properties of one or more things or relationships can also be identified in terms of basic concepts: (steel) (is) (strong).

This approach is best facilitated by allowing "things" (dogs, mail carriers) and "relationships" (biting) and properties (strong) to be freely combined in both classification and search; this spares the classificationist from having to enumerate a vast array of combinations, and the user from having to ascertain how a particular combination was treated. The broad outlines of such a classification can be found in Szostak (2013a); the classification of things is treated in Szostak (2011), and of relationships in Szostak (2012a). This approach has been instantiated in the Basic Concepts Classification (Szostak 2013a); its key attributes are basic concepts and a fully synthetic approach. It will thus be referred to throughout the paper as the basic/synthetic approach. Note that this fully synthetic approach implies a unified comprehensive classification.

The previous section urged a practical orientation to philosophical evaluation. It should be noted, therefore, that the basic/synthetic approach has been designed to reflect and take advantage of the characteristics of the contemporary digital environment. Szostak (2014b) has suggested, for example, that it may be well-suited to the needs of the Semantic Web.

## Three Issues Raised by Furner

Furner (2010) provides a detailed survey of the philosophy of information science. Three of the key issues raised by him have important implications for the philosophy of classification and are thus addressed here.

### *The Nature of a Work*

Furner raises the vexed question of what a work is *about*. My response is that most works are about some sort of posited causal relationship. Yet, they are almost never given a subject heading that captures the essence of the causal relationship(s) that the work is about. We try and fail to classify

works about causal arguments as if they were about one complex thing. As Foskett (1996, p. 127) has urged, subject headings should be coextensive with the essence of a work.

Sentences, it should be stressed, are usually less ambiguous than single concepts, for the sentence provides context for all constituent concepts. This recognition of the importance of context lies at the heart of communication theory, which focuses on "thought units" rather than concepts. Thought units may be bigger or smaller than a sentence but will comprise multiple concepts (Keyton & Beck, 2010). The sort of causal argument suggested here would qualify as a thought unit.

While the particular causal relationship(s) addressed in a work are the key aspect of what a work is about, other important aspects include any theory or method or data that was employed, as well as the perspective or worldview of the author. These also should be captured in our classifications. These various characteristics of a work were identified as *classificatory desiderata* in the León Manifesto (2007).

Importantly, Furner (2010) appreciates that philosophers have devoted little attention to what a work is about but much to what a sentence is about. The point to stress is that the approach recommended here reduces the gap between what a work is about and what a sentence is about by describing works in terms of causal statements. (Note that works that describe a thing or an action could also be captured in terms of statements about things and/or effects.) We can then apply philosophical thinking regarding the nature of a sentence to understanding the nature of a work.

There is a further advantage of this approach: as Gnoli (2008) has stressed, a classification scheme should ideally be able to handle both works and ideas. And Börner (2006), intriguingly, envisages a not-too-distant future in which scholars no longer write stand-alone papers but rather contribute nodes or nuggets to a web of knowledge. She can be seen as operating within a long tradition in information science, from at least the work of Paul Otlet, which sought to classify both works and ideas. A classification system that classifies works in terms of ideas is clearly one way, and perhaps the best or only way, to do so.

Most philosophical inquiry into the nature of sentences occurs within the field of logic. Logicians strive to strip away stylistic elements in order to formally appreciate the semantic content of a sentence. They focus on statements that could either be true or false (or somewhere in between). Arguments involve premises and conclusions and the logical arguments connecting these. Propositions generally are mind-independent and abstract. But logicians appreciate that logic is only one form of argument; induction is another.

Happily, we need not master logical philosophy here. When statements are analyzed formally, it is usually done in forms such as "*X* is a *Y*" or "*X* has effect *Z* on *Y*." Classifying works in terms of a logical hierarchy of con-

cepts accords with the first type of statement; classifying works in terms of combinations like (thing) (effect) (thing) accords with the second. Philosophers have not generally worried about the relative importance in human conversation of different types of statement. The broad lesson for information science is that we should first identify the core message(s) of a work and then classify works in terms of the things and/or relationships that best characterize these works.

Notably, the approach to defining a work taken here accords with that advocated in Smiraglia (2001). His purpose was to distinguish the meaning of a *work* from the meanings of *text*, *document*, and other similar concepts. To this end, he surveyed not just the literature in philosophy and information science but also linguistics, sociology, and other fields. Smiraglia concludes that "the intellectual dimension of a bibliographical entry is the set of composed ideas that it conveys, which is called a work" (p. 122). Texts and documents can be defined in physical terms, but a work is necessarily abstract: it is a set of ideas. Works thus have an "inherent nature as communicative signifying objects" (p. 54); FRBR (Functional Requirements of Bibliographic Records) also treats a work as an abstract entity (p. 47).

Smiraglia does not define what he means by *ideas*. The basic/synthetic approach outlined above clarifies this vague term: *ideas* will comprise some set of descriptions of phenomena or relationships, causal arguments, theories applied, methods applied, and perspectives applied. Classifying works along these dimensions will thus capture the nature of a work. Classifications that omit any of these (and existing classifications in widespread use omit most and handle others poorly) will quite simply not capture the nature of a work.

Smiraglia's analysis nevertheless identifies certain caveats; most centrally, that the ideas that comprise a work are conveyed semantically. It is thus not possible to entirely distinguish substance from style (p. 67). A work may be appreciated as much or more for the style with which certain ideas are conveyed as for the ideas themselves. Smiraglia makes special note of music, but one could reflect also on poetry or a political speech like the Gettysburg Address. Yet, the blending of substance and style hardly obviates the value of identifying the key ideas of a work. It does, however, indicate that some attempt to capture style would be useful. This will likely prove a much harder task than classifying substance. Some descriptors might be fairly easy to apply: for example, humor, satire, sarcasm; others would present a greater challenge because rhetoricians disagree about the full set of rhetorical strategies that can be employed in a work and how each might be identified.

Smiraglia also notes that author and reader may disagree about the key ideas in a work. Deconstructionists have shown that works (especially of fiction) often contain ambiguities of which the author was not consciously

aware. Yet, it should be possible in most cases to identify the key causal arguments an author was trying to make (and new authors could be asked to supply such information) as well as any theories and methods explicitly applied.

Since works might be instantiated in multiple texts and documents, a decision must be made about when a work is transformed into a new work. Such a change might be primarily stylistic: a movie based on a novel is generally considered a new work no matter how closely it follows the original. But in general, according to Smiraglia, "the degree of change in ideational and semantic content determines the point at which a text represents a new work" (p. 50). How would we know when ideas have changed enough to declare a new work? A precise answer to such a question may never be possible. But the approach to classification pursued here would at least suggest some key questions to ask: Has the causal argument changed? Has a new theory been applied? Has a new method been applied? Positive answers to any of these would signal the creation of a new work (although we might still wonder if adding one new variable to a complex argument really generates a new work).[2]

Hjørland (2014), after reviewing some of the confusion surrounding the term *information*, suggests that information science would be well-advised to understand its focus as *documents* rather than *information*. Following Smiraglia (2001), we might suggest that a better focus would be *works* (although it is also important to identify the different documents that might instantiate a particular work). In any case, this section takes Hjørland's suggestion a step further and argues effectively that we should focus primarily upon *statements*. The best we can hope for is to organize statements and also the perspectives (including theories and methods applied) from which these emerge. We might also as a field eschew that other contested term *knowledge*. Humanity almost never "knows" anything, but we can increase our confidence in the reliability of particular statements by compiling argument and evidence. The purpose of a classification system, then, is to guide users to relevant statements and the arguments and evidence that support these. We expand on this idea in the next section.

Floridi's well-known efforts to define *information* are broadly consonant with an emphasis on statements and perspectives. For him, information must be about something, it must be meaningful, and it must be well-formed. He also stresses what he terms the *truthfulness* of information. (Furner [2014] summarizes Floridi's thinking on these issues.) Statements of the sort that we stressed above would meet the first three criteria: they would be about something and be well-formed, and their meaningfulness would reflect whether they were statements about something that some user valued and whether it came from a perspective the user valued (see below). Furner argues that information science should

devote much more attention than it has to truthfulness. Our attitude toward truthfulness can be the same as our attitude toward knowledge: that the best that humans can do is to evaluate the arguments and evidence that support any statement. Again, we are guided to reflect on how we can both guide users to relevant statements and facilitate their efforts to evaluate these.

We should note in closing that Hjørland (2014) appreciates that documents need not be text. Knowledge-organization systems must cope with pictures, videos, audio, and various sorts of online databases. Can one classification scheme cope with all of these? It is notable in this regard that a basic/synthetic approach potentially could. Works of art can be identified as (woman)(riding)(horse), museum artifacts as (axe)(for)(fighting), and audio or video also classified in terms of the main statements being made.[3]

*Evaluating the Value of a Statement*

Furner (2014) also discusses what makes some statements more important than others. Information scientists should wish to direct users to statements that they will consider important. Having identified works in terms of the statements above, Furner's query takes on added significance: what does a user need to know about a work in order to assess its likelihood of being valuable? The particular causal argument being made will be of primary importance to the user. As noted above, identifying works in terms of such characteristics as theory applied, method applied, and philosophical and disciplinary perspective of author is also of critical importance here: users are likely to judge a work important if it makes causal arguments of interest and applies a theory, method, and perspective that the user appreciates.[4]

Notably, the sort of classification urged here guides users to important works whether the user has strong methodological preferences or whether the user is seeking (as interdisciplinary researchers should; see Repko [2012]) to integrate across multiple theories, methods, and perspectives; that is, they can search for only one theory, method, and perspective or across many, as they wish. Likewise, they may wish to focus on only one causal link or on a system of related links.

At present, we rarely classify works in terms of the theory, method, or perspective applied. Users will thus retrieve works that take an approach they disdain, and may have to investigate these in some detail before ascertaining their nature.[5] This may, from time to time, encourage the user to broaden their mind; so too will a classification that alerts them to the availability of alternative approaches but does not frustrate them by failing to distinguish these. As for causal arguments themselves, these are also rarely signaled in our classifications. Even if cause and effect receive distinct subject headings, a Boolean search for these will also retrieve any

works that discuss the two subjects regardless of whether a causal relation in a particular direction is posited. In sum, existing classifications do a very poor job of directing users to important works. The basic/synthetic system would do a much better job.[6]

Yet, even the basic/synthetic system will be imperfect. Particular authors may apply a particular theory or method badly; a user may then mistakenly think that they have found what they are looking for. They may then become disillusioned after consulting the work. Or, the user may lack the expertise to appreciate the shortcomings of the work. Ideally, we should be able to signal to users the quality of a particular work (Budd, 2011). This is a task that will challenge the classificationist: we can aspire to fairly objective definitions of theories, methods, causal arguments, and even perspectives (and can fairly comfortably ask authors to self-declare these), but work "quality" is inherently subjective. Of particular note, the perceived quality of a work is not static: a work may rise or fall in the esteem of others as theories and methods evolve through time. We may need to be open here to strategies that fall outside the task of classification (at least as usually defined). We could aspire to providing ever-easier access to citation trails in both directions—and perhaps someday it will be possible to distinguish favorable citations from criticisms. We could, as many OPACs (online public access catalogs) already do, provide ready access to other users' reviews and ratings of a work. We might even try to organize these in terms of key attributes, such as reliability, validity, completeness, and so on. But in the end, since quality is a moving target, we will likely need to accept and even celebrate the fact that users will be guided to works that are not widely embraced. One thing that classificationists and classifiers might do is recognize when works have been subjected to a peer-review process.[7]

Huang and Soergel (2013) have recently addressed the issue of *document relevance* to user needs. They note that information science as a field still has a limited view of what *relevance* involves. There are hundreds of characteristics of relevance noted in the literature; Huang and Soergel suggest that these can be appreciated under several headings. Among functional characteristics, they list several that have been noted above: the cause/effect relationship, conditions (that influence causal relationships), topicality (they appreciate that previous approaches have been limited in expressing only nouns and being word-based; these can be improved by embracing verb-like terms and relationships),[8] context (which captures the elements of perspective connected to an author's place in the world), purpose (which captures other elements of perspective associated with authorial intent), and method. The two other functional headings are comparison (which means drawing lessons from other times and places) and evaluation (which captures the significance of a work). Comparisons can only meaningfully be made if causal relationships are first

specified. And evaluation, as we have suggested, will depend primarily on the causal relationship investigated and the theory and method applied. In addition to functional characteristics, Huang and Soergel mention reasoning-based and semantic characteristics. The former captures how a work affects a user's thinking. The point to stress here is that users think in terms of causal arguments and evaluate in terms of theories, methods, and perspectives. Therefore, a document is most likely to change their thinking if it addresses questions they care about in a manner they appreciate. Note that alerting users to works that address closely related links may be particularly valuable in changing their thinking: finding exactly what they looked for may not change their thinking, and completely unrelated material will simply be ignored.[9] Semantic characteristics—how a document relates to the user's interest semantically—involves such things as hierarchy and adverb/adjective descriptors. Issues of hierarchy are addressed below. While we have stressed the importance of things and relators in this paper, we have noted that an approach stressing linked notation can also readily provide for adjectives and adverbs (see Szostak, 2014b). In sum, the approach to classification examined in this paper fares well under each of the headings identified by Huang and Soergel (2013).

*Philosophy through Information Science*

In his conclusion, Furner (2014) notes that information science has contributed little to philosophy. He extends this argument by saying that theories of information within information science are little noted anywhere outside the field. I would argue, however, that the field can provide an important empirical input to philosophy and beyond. Philosophical inquiry has established that some degree of ambiguity is inevitable in human communication, but by its nature cannot answer the question, how much? Information scientists should not (but often do, generally implicitly) assume some particular degree of ambiguity. Information science may be better placed than any other field to establish how much ambiguity is inevitable. It was noted above that breaking complex concepts into basic concepts reduces ambiguity, so also does then classifying basic concepts hierarchically, which establishes what sort of thing something is and what sort of thing it is not. But hierarchy can only serve this function if a strictly logical classification is pursued. Hierarchy is abused in extant classifications in order to capture elements of causation; for example, *recycling* is treated as a subclass of *garbage* because there is no other convenient place to put it (Mazzocchi, Tiberi, De Santis, & Plini, 2007). This would not be necessary in a classification of basic things and relationships intended for use in a synthetic classification of works.

If we can achieve a classification of basic concepts, and then clarify the meaning of complex concepts in terms of these, we significantly decrease

the degree of ambiguity in cross-group communication. Lambe (2007) is one information scientist who has argued that the first task of classification is to clarify the meaning of concepts in order to facilitate conversation. Scholars of interdisciplinarity, in turn, have long worried about ambiguity and sought strategies to reduce it (O'Rourke, Crowley, Eigenbrode, & Wulfhorst, 2014). Philosophers have striven for perfect clarity and failed to achieve it. Although they have rarely asked, how much?, they should nevertheless be very interested in the efforts of information scientists both to measure and to reduce ambiguity. This will especially be the case if we find that it is possible to decrease the level of ambiguity, such that both a comprehensive classification and cross-group understanding are facilitated.

## Revisiting the 2004 *Library Trends* Issue on the Philosophy of Information Science

I now turn to several concerns raised in the original *Library Trends* issue on the philosophy of information science in 2004.

### *The Nature of Hierarchy*

Olson (2004) raises concerns about the nature of hierarchy. She expands on these concerns in a 2007 article in *Library Trends*. In the latter article, she draws on feminist philosophy in order to urge an approach to classification that relies less upon hierarchy and more on a web of relationships. I have recently (Szostak 2014a) shown that the sort of approach discussed in this paper fits Olson's criteria and responds to many particular concerns she raises.

Of particular note, Olson appreciates that existing classifications handle paradigmatic relationships best. Yet, since such relationships are enduring, we rarely need state the obvious. It is syntagmatic relationships (where the connection is not essential, as in embroidery of Christmas ornaments) that we will often wish to express (search for), but these are handled poorly; Boolean searches will yield many hits that do not capture the desired relationship. Again, the solution involves allowing us to freely connect any set of concepts in classifying a work.

It is often noted that pharmacologists may wish to classify chemical compounds in terms of their physiological effects, whereas chemists will wish to classify them in terms of their chemical nature. The point to stress here is that pharmacologists are interested in classifying a *causal relationship*; they wish to be able to find, say, (chemical)(reduce)(blood pressure). More generally, they wish to link types of chemical, types of effect, and various human organs. This is best done by allowing these two types of things and one type of relator to be freely combined.

Notably, this approach also reduces and perhaps even eliminates a concern often voiced by Olson (and others): that there are multiple ways of

classifying things. The simple fact is that many of these are actually attempts to classify causal relationships rather than things. As the example above illustrates, we need not develop a separate classification of chemicals in order to guide pharmacologists to the works or ideas they need. Once we instantiate a web-of-relations approach, the pressure (at least much of it) to (mis-)classify things in multiple ways simply goes away.

And note that free combinations do indeed instantiate a web of relationships. Someone interested in how a certain human capability can be amplified can link that ability to relevant organs and thence to chemicals that improve their functioning; or, they can link readily to exercises that enhance performance. All things get linked indirectly to all other things through a web of (all types of possible) influences. A classification that employs different terminology and strategies for different disciplines will not facilitate user movement from one node to a related though distant node in the web of understanding.

Users can readily follow their curiosity along this web, a task not greatly facilitated by existing classifications. Nor does full text searching identify the nodes and links in the web. But the approach pursued here does so, and the relatively small schedules involved (Szostak, 2013a) will allow the user to quickly apprehend the contours of the web. Users will thus be better able to find what they are looking for. Of crucial importance, they will also be much better able to make the sort of discovery justifiably celebrated in the literatures of undiscovered public knowledge, literature-based discovery, and serendipity. Quite simply, for any thing or relator that the user is interested in, they will have ready access to any other work that discusses these in combination with any other thing or relator. The rate at which new discoveries emerge from drawing new connections across diverse literatures must advance.

Olson (2007) appreciates that hierarchy is valuable, but merely wishes to lessen its importance. By excising inappropriate uses of hierarchy, we better allow it to serve its functions. Soergel (1985) lists many of these: hierarchy facilitates choice of the appropriate level of generality; it facilitates inclusive searches (for example, search by meat, get documents on pork); it facilitates aggregate analysis (of meat consumption, say); it allows for different degrees of specificity for different purposes; and it facilitates sharing across systems that may employ different levels of generality (pp. 246–247).[10] Each of these functions will be unnecessarily complicated if hierarchies are cluttered with relationships. For example, Soergel discusses "packaging" as part of a hierarchy of foods though this is clearly not a kind of food, but rather something done to food.[11] An approach that would handle relationships like "packaging" through the use of linked notation will allow for logical hierarchies.

## Ethical Evaluation

Olson, like Furner, discusses ethical issues. Information scientists clearly have ethical responsibilities. But what are these? The thrust of this paper is that we should be open to a variety of philosophical perspectives.[12] It might then seem that ethical evaluation would be extremely complicated. Happily, this need not be so. I argued (Szostak, 2005a) that there are only a handful of valid and complementary approaches to ethical analysis. A similar approach was taken in a recent article by Fox and Reece (2012), who establish a handful of criteria for evaluating the ethics of a classification.[13] I will apply their standards to the classification recommended here.

Fox and Reece first stress that we should help others in need, which implies that we should pay special attention to users that have difficulty accessing relevant literature. Such users may be unfamiliar with the disciplines around which all major classifications are structured, hence they will appreciate a classification that allows them to combine simple terms in searches (see the discussion of postcoordination below). Fox and Reece also urge us to keep borders porous. I have showed (Szostak, 2014a) that the approach taken here could encourage both cross-group and within-group understanding. More generally, as we have seen, this approach encourages users to make connections with other literatures. Fox and Reece (2012) suggest combining control with tagging: the middle path to be advocated below between pre- and postcoordination seeks also to balance structure and freedom. They urge us to ensure that users are getting what they need through user testing. This is not a criterion that can be addressed in a theoretical paper, but I would stress here that we need to evaluate not just whether users find what they knew to look for but whether they find relevant literatures of which they were previously unaware.

We should appreciate user rights. This, I would suggest, is a powerful argument for some degree of postcoordination. Users have a right to information and should not have to struggle with an unnecessarily complex classification, and especially an unnecessarily ad hoc classification, in order to find it. Moreover, they should have access to a comprehensive classification that gives them equal access to all relevant literatures. A classification grounded in disciplines can never provide equal access to all literatures. Appreciating rights segues into the final criterion: do no wrong. Here, I would stress the importance of nondiscriminatory subject headings. Olson (2007) rightly complains that existing subject headings treat certain groups (such as male nurses) as anomalies. One advantage of relying upon synthetic links across hierarchies is that such discrimination disappears; male nurses are treated in exactly the same way as female nurses.

## Philosophy of Pre- versus Postcoordination

Jacob (2004) appreciates that there are strengths and weaknesses of both pre- and postcoordination. It is thus not possible to make a strong philosophical argument in favor of either. The basic/synthetic approach to classification recommended above combines elements of each: schedules of things and relationships would be developed that could be freely combined by classifier and user. I will argue here that this approach captures key strengths of both pre- and postcoordination while avoiding most weaknesses. Sauperl (2009) investigates whether technological advances have caused one type of system to become preferable over others. She finds the same result as Jacob and as Svenonius (1995): that each of the two systems has advantages.

The literature on pre- versus postcoordination rarely even allows for the possibility that there is something in-between. But Sauperl (2009), following Svenonius, notes that there are three characteristics that distinguish the two:

- *who* performs the coordination of terms
- *when* the coordination is performed
- *how* the coordination is performed (p. 818)

With these three distinguishing criteria, it must be possible for a classification system to possess elements of each. Sauperl proceeds to appreciate that a postcoordinated system has a shorter list of headings and no grammatical syntax (that is, only Boolean logic is possible). Although the basic/synthetic system is indeed characterized by short schedules, it does imply a sort of sentence structure on user queries—albeit a sentence structure that reflects the very nature of user queries.

Sauperl, again following Svenonius, appreciates that expressing the type of relationship would improve retrieval over a postcoordinated system; it is simply not possible to express sophisticated relationships within postcoordinated systems. For example, it is not possible to distinguish the "philosophy of history" from the "history of philosophy" with a simple Boolean search for "philosophy and history." Since precoordinated systems include more specific terms, they are richer in information than postcoordinated ones. The basic/synthetic system allows the "philosophy of history" to be clearly distinguished from the "history of philosophy" without requiring the user to first ascertain what search terms are allowed within the classification. It thus combines the advantage of logical syntax with the advantage of the user's construction of search terms.

The speed of searching, both authors argue, is better in precoordinated systems because specific terms allow for more focused retrieval and hence for better precision. *Precision* is indeed critical for search speed; otherwise, the user must sort through a large number of false drops. But we must not neglect the requirement within precoordinated systems that

the user must first identify allowable search terms. Postcoordinated systems have an advantage at the front end of searches, for the user can combine their own search terms. The basic/synthetic system mimics the front-end advantages of postcoordinated systems while delivering the same back-end degree of precision as precoordinated systems.

Sauperl (2009) notes that users generally appreciate neither the long subject strings of precoordinated systems nor the contours of Boolean searches required by postcoordinated systems. (She speculates that artificial intelligence might be the answer: a computer familiar with a user may be able to surmise whether they are more interested in the philosophy of history than the history of philosophy.) Although user testing is clearly necessary, it is at least plausible that searching in terms of a syntax that mimics common sentence structure would be superior. A user interested in why dogs bite mail carriers can enter the search (mail carriers) (bitten by) (dogs). No tutorial on Boolean searches is required, nor does the user need to investigate how a classificationist might have attempted to render this relationship within a precoordinated system. All that is required is a thesaurus within the search interface that can "translate" uses of "postmen," "bite," and "canine" in queries.

Postcoordinated systems have various advantages: they have shorter schedules, and they are more hospitable as new combinations can be facilitated automatically.[14] Such systems thus grow in size much more slowly. For similar reasons, postcoordinated systems age more gracefully, for the complex strings within precoordinated systems gradually fail to capture how complex ideas are expressed. Postcoordinated systems are easier to learn and employ by classifiers; they are thus less expensive to create and maintain. In all of these ways, the basic/synthetic system shares the advantages of postcoordinated systems. Precoordinated systems also have various advantages: they clarify terminology (such as homonyms) by placing these in context; and more generally, they allow for greater precision.[15] As noted above, the basic/synthetic system also provides clarity and precision.

There are also several criteria for which pre- and postcoordinated systems fare equally well. Svenonius (1995) claimed that they were equivalent with respect to recall—the percentage of relevant documents obtained. Arguably, the proposed system can outdo both by allowing users to specify a search of their choice, but yet using precise syntax. Both pre- and postcoordination can be handled by computers; so also can the basic/synthetic system.[16] One of the challenges in the computerization of precoordinated systems is that these are all characterized by ad hoc decisions difficult for a computer to master; the proposed system is resolutely logical. Yet, as already noted, syntax makes user queries more readily understood than in a postcoordinated system. The basic/synthetic system is well-suited to universality and is eminently browsable; it excels in suggestibility, familiarity, simplicity, and size.

Jacob (2004), Svenonius (1995), and Sauperl (2009) each concluded that there was no clear winner between pre- and postcoordinated systems because each approach might prove best-suited to different circumstances. Interdisciplinary scholarship teaches us to always look past dichotomies and ask if there is some continuum or middle ground between alleged opposites (Repko, 2012). It is indeed possible to develop a system that combines the strengths of pre- and postcoordination. One might call such a system "poly-coordinated." Since we have shown that poly-coordination is feasible, it should become a characteristic to be sought generally in classification.

## Relevance Redux

Budd (2004) investigates the concept of *relevance*. He notes that the relevance of a document will depend not only on a user's needs but on their preferences regarding types of information (are they empiricists or rationalists?). The sort of information advocated above regarding the theories and methods applied should enable a user to better identify relevant works. Furthermore, relevance depends on *context*: the user will judge whether the document is both appropriate and reliable on the basis of their understanding of how and why it was constructed. (Kleineberg [2013] also advocates capturing the *how* and *why* along with the *what*.) In particular, relevance will depend on whether the author and reader share key assumptions about the world; again, information regarding the theories, methods, and perspectives applied are critical.

Budd also stresses that user needs are not static but dynamic. A user may start with general curiosity about a topic and develop more precise queries as they read. The desideratum here, then, is a classification system that allows the user to follow their curiosity. At times, curiosity may simply guide the user into more detailed explorations of a broad topic—logical hierarchies will be critical here. But often in our complex interrelated world, we can imagine that a user will be guided to investigate how their original topic of interest is related to others (and Budd appreciates that any piece of information is necessarily part of a larger network). And this sort of search is best facilitated by the sort of synthetic notation urged elsewhere in this paper of the sort (thing)(relationship)(thing). A user that began with curiosity about one thing and becomes curious while reading about how that thing relates to another thing should be able to easily search for that particular thing/relationship/thing combination.[17]

## Grounding Information Science in the Lived Experience of Librarians

Librarians try to guide users to relevant works, both when the user knows what they want and when they are unsure. Both the user who knows what they want—for example, (mail carriers)(bitten by)(dogs)—and the

user curious about what dogs do more generally will find a classification grounded in combinations of basic concepts easier to navigate.

About thirty years ago, Batty and Bearman (1983) suggested a possible conflict between librarians and information scientists: the former were primarily interested in shelving similar items together, while the latter were more interested in signaling the uniqueness of each item. This tension is undoubtedly exaggerated for librarians who also wish to guide users to the works best-suited to their interests. In any case, the purpose of subject catalogs (and other search engines) is to guide users to relevant information. Users will wish to know the particular arguments made in a work (and perhaps its theory, method, and perspective). A classification that captures the true nature of a work (see above) will serve the diverse needs of librarians best.

Several authors in the 2004 *Library Trends* issue, but especially Cornelius, urged us to look past theoretical arguments to the actual experience of librarians. Unfortunately, the avenues of communication between practicing librarians and scholars of information science are weaker than they might be, and arguably weaker than they once were. This journal is one important venue in which the needs of practicing librarians are voiced. A survey of the journal since this last issue on the philosophy of information science uncovers several items that focus on the needs of all or some librarians in a manner that is, at least implicitly, related to matters of classification. While this inductive approach cannot by its nature claim to be exhaustive, we do uncover an intriguing range of concerns that should be appreciated when evaluating any classification system.

The journal's spring 2005 issue was devoted to search engines. Part 1 complained about how search results may be biased by pressure from commercial advertisers. The second part argued that the solution is to bypass search engines by allowing subject-centered searches. Concerns about possible biases in the unknown algorithms guiding search engines continue to this day. The search engine Blekko strives to remove spam from search results; notably, it encourages searches within a set of subject headings. While librarians may have limited advice to give to users perusing commercial search engines, they will, of course, have much advice to provide if users turn toward subject searches (see Fabos, 2005). And we have argued above that the basic/synthetic classification will render subject searches both easier and more successful.

*Library Trends*'s summer 2009 issue was devoted to school libraries. In her introduction, Mardis (2009) worried that school librarianship was viewed as low status by both librarians and teachers. The purpose of the issue was "to shift staid conceptions of school librarianship in the LIS academy to the idea of dynamic educational informatics in schools" (p. 1). Mardis noted that even young students are likely to search the internet outside of school. There is thus a tremendous opportunity for school

librarians to take on an important educational role in guiding students on how to find relevant and reliable information on any topic, whether for school assignments or not.[18]

This task will be much easier if school librarians have access to a classification system that is easy to master and that allows students to readily follow their curiosity from one topic to another. But the benefits hardly stop there. The classification system discussed in this paper succeeds precisely because it actually captures the essence of scholarship. The vast majority of scholars investigate how one or more things affect (in particular ways) one or more others. As noted above, the classification thus captures the nature of works. And the short schedules capture the things that are studied (whereas existing classifications confuse the user by jumbling together things and relationships, basic and complex). We can thus teach students to find not just information but the broad contours of the body of human understanding (see Szostak, 2003).[19]

The journal's fall 2009 issue (and later, the summer/fall 2010 issue) addressed workforce issues. While much of the discussion emphasized career paths, wages, and working conditions, there was one important point made for our purposes: "In addition to offering collections in support of pleasure and other reading, public librarians play an important role in providing access to technology and meeting the general, employment, health, legal, financial, leisure, and educational information needs of their patrons" (Marshall, Solomon, & Rathbun-Grubb, 2009, p. 125). This simple quote highlights the critical fact that librarians need to guide users to information on a host of very complex subjects. The diversity of user needs means that the librarian cannot aspire to provide detailed subject-area understanding for every user request but must instead be able to show them where to look. The complexity of individual user needs means that the librarian needs to be able to show users how to connect diverse bits of information; existing classifications do not make it easy for users and librarians to find answers to multiple complex queries. The basic/synthetic system facilitates complex queries. This may, at times, allow users to find what they need without librarian assistance, but it will allow librarians to provide more focused assistance when asked. That is, while the proposed classification can be self-taught, it might also reinvigorate the perceived value of reference librarians by making subject searches seem a more valid and valuable option.

The journal's spring 2010 issue explored parliamentary libraries. Since modern governments legislate on a wide range of issues, these libraries need to provide information on a wide range of subjects. Politicians and their advisors often face very tight timelines: they need to find relevant information very quickly (Sandgrind, 2010). Although Sandgrind did not stress it, the issues addressed are generally complex; indeed, one of the key challenges in contemporary public policy is to avoid undesirable side

effects (Szostak, 2005b). The users of parliamentary libraries and the librarians they consult thus must be able to very quickly find relevant information on a connected web of causal relationships. They need a web-of-relations approach to classification.

Sandgrind made an interesting observation: "The explosion of information made available by the Internet makes the role of parliamentary libraries and research services more, not less important, as busy parliamentarians need people to filter information for them and to do so in a timely, accurate, and politically neutral way" (p. 416). This point resonates far beyond parliamentary libraries. As we have seen above, users need to be guided to good information and be facilitated in evaluating this. The internet creates challenges, as well as opportunities, for users of all types. And the best antidote to information overload is information organization.

*Library Trends*'s fall 2011 issue addressed information literacy advocacy. In his introduction, Crawford noted that it is difficult to either advocate or obtain research funding for information literacy. Research libraries had become the focus of information advocacy, but these are encouraged to emphasize their research role. While the focus of the issue was on advocacy, it is worth stressing that a classification that is easier to comprehend, improves searching, and shows how things are related will make information literacy easier to achieve. As noted above, users will become much more interested in mastering subject-search techniques if these are more rewarding. And public policy will likely follow user interests.

The journal's spring 2012 issue looks at small-town libraries. Wiegand (2012) opens with a cautionary note. The sorts of information-seeking behaviors we have stressed above may often be secondary goals for such libraries, whose primary purpose is supporting social harmony by providing a shared meeting space and shared stories. But social harmony is not static in a world of dynamic change. And so a more transparent classification system need not just aid what Wiegand sees as secondary goals but may well aid communities as they seek to navigate societal change and update their shared stories.

There are undoubtedly many other concerns that would emerge from a broader engagement with the lived experiences of librarians. This section has nevertheless hopefully established both the value and feasibility of doing precisely that.

### Natural Language versus Formalization

Svenonius (2004) worried that there is a trade-off between employing natural language in a classification system versus applying very precise meanings to controlled vocabulary that differ from the common understandings of the terms used. As with our above discussion of pre- versus postcoordination, we should ask whether this trade-off can be minimized

in practice through a novel approach to classification. And here again, the basic/synthetic approach does, at least, decrease the tension identified by Svenonius. After all, *basic concepts* are defined as those that carry very similar understandings across groups. It is thus possible to employ natural-language basic concepts in a very precise manner. The challenge identified by Svenonius really comes into play with *complex concepts* because then, different users will understand the terminology in quite different ways and the classification will have to carry extensive scope notes indicating how terms are defined. In the classification examined in this paper, complex concepts are treated as combinations of basic concepts. Since the latter evoke shared understandings, so also will explicit combinations of these.

For example, the word *globalization* means many things to many people; however, phrases like (movies) (transmit) (cultural values) and (international trade) (increases) (job turnover) are much less ambiguous. A classification that employs *globalization* must either accept a high level of ambiguity or provide a detailed definition. A classification that involves only terms like *movie* or *international trade* can achieve a low degree of ambiguity without the necessity of detailed definitions.

It may be that some complex concepts prove hard to disambiguate into basic concepts. As always, we should be careful of drawing empirical conclusions from theoretical arguments. But it must be the case that the tension identified here by Svenonius is lessened for all complex concepts that, in practice, can be broken into basic concepts.

## Conclusion

This paper has performed an inductive survey of several distinct though important issues that should be addressed in evaluating any classification system. The general conclusion to be drawn is that there is no single philosophical issue or perspective that should dominate such an evaluation; rather, we must be thoroughly eclectic in approach, evaluating any existing or proposed classification in terms of a variety of issues and perspectives.

The paper illustrates this general argument by evaluating a particular approach to classification described in detail above involving universality and the free combination of basic concepts. This approach fared very well in our evaluation because

- it captures the nature of a work much better than existing classifications;
- it thus also guides users to works and arguments they consider to be important;
- it reduces the ambiguity of the concepts employed;
- it employs hierarchy logically, and it requires hierarchy less than other classifications;

- it accords with a handful of critical ethical considerations;
- it captures the key strengths of both pre- and postcoordination;
- it reflects (at least theoretically) the needs of librarians;
- it serves an important educational purpose by capturing the contours of human understanding; and
- it allows natural language to be employed precisely.

I have not claimed that the criteria addressed in this paper are exhaustive; they are, however, reflective of the issues raised in Furner (2010) and the original survey of philosophical issues in *Library Trends* in 2004. The classification evaluated here has been evaluated (favorably) with respect to other criteria as well (see Szostak, 2014a).[20] I speculated at the start of this paper that there might be trade-offs in satisfying these various criteria, but this does not seem to have been the case in practice. The results here suggest that it is possible to classify works and ideas much better than we do at present along a wide variety of dimensions. Given the incredible importance of knowledge organization to not only scholarship but democratic citizenship and societal innovation of all types, the possibility of a broadly superior approach should be pursued.

## Notes

1. This paper also engages along the way with several arguments from Ibekwe-SanJuan and Dousa (2014).
2. I will, of course, also classify works by author and title and generally assume that a change in authorship represents a new work. But the key insight of Collins (1998) should be recalled: that scholarship is "a conversation," and that the key figures in a field become emblems of the thinking of that field. For example, Plato and Aristotle are associated with a set of ideas that were much discussed in ancient Greece. A new author may thus change the nature of a work less than we might assume.
3. Hjørland worries elsewhere in that paper that epistemological pluralism threatens the coherence of the field of information science. Yet, he opens by citing Buckland to the effect that the challenge to identifying the nature of information science lies in the fact that different approaches are employed for each type of document. Pursuit of a truly comprehensive classification might thus provide the coherence that Hjørland seeks.
4. Kleineberg (2013) urges us to classify the *what*, *how*, and *why* of a work (and indeed of the things discussed in a work). Arguably, the approach recommended here captures each of these elements: what an author studied, how, and why.
5. "Information . . . is without meaning outside the social practice that produced it" (Cornelius, 2014, p. 188). Cornelius may exaggerate somewhat, but the quote signals the importance of capturing how a statement is generated.
6. Cornelius (2014) worries that increased access to information has not obviously resulted in better decision making in the world. Yet, as Hjørland (2014) appreciates, we guide users to information with the goal of encouraging human progress. Part of the solution lies within the field of information science: users need to be guided to good information and the tools to evaluate it. Other parts of the solution involve encouraging better political processes and becoming acquainted with interdisciplinary techniques for addressing conflicts in the literature (see Repko, 2012; Szostak, 2012a, 2012b).
7. They can also provide access to measures like journal impact factors and authorial H-indexes. But they should tread carefully, for impact factors vary by field, and H-indexes vary by field and age (and recall that some authors become emblems for collective ideas).
8. Friedman and Smiraglia (2013) find that most concept maps employed in knowledge organization have nouns as nodes and verbs as arcs. But our classifications do not reflect this synergy.

9. The paper discusses below how a web-of-relations approach to classification can achieve this end.
10. According to Soergel (1985): "To sum up: hierarchy must never be a strait jacket in which the universe of knowledge has to fit somehow or other. On the contrary a properly designed hierarchy shows the manifold relationships between concepts and thus assists in indexing and searching. Whenever a hierarchy sets constraints it is faulty; whenever it helps the indexer or searcher, it serves its purpose" (p. 256).
11. Soergel (1985) notes that the DDC (Dewey Decimal Classification) has no class for packaging but has for "economic aspects of" and "technological aspects of." This sort of illogical practice severely limits the ability of users to pursue their curiosity.
12. The aesthetic standards proposed by Ojennus and Tennis (2013) could also be applied. There is significant overlap between these standards and some of the issues discussed in this paper; for example, the accessibility of a classification is aesthetically valuable.
13. Their list of ethical approaches ignores the arguments from intuition and tradition that are discussed in Szostak (2005a). And they treat some ethical approaches, such as the philosophy of caring, separately, whereas I view these as part of broader approaches.
14. Various information scientists, including Jens-Erik Mai and Grant Campbell, have argued that flexibility is of increasing importance as new subjects multiply.
15. But can you search by all elements of a precoordinated string? If not, this is a major drawback. Sauperl (2009) suggests that this may becoming possible.
16. This possibility is addressed further in Szostak (2014b).
17. As noted above, Szostak (2014a) discusses how the approach to classification investigated here instantiates the web-of-relations approach urged by Olson (2007).
18. One paper in the issue, by Ritzo, Nam, and Bruce, urges students to act as information gatherers for community organizations. This would require students to know how to find information on very complex issues.
19. As noted above, Börner (2006) predicts a future in which scholars, rather than writing stand-alone papers, add nuggets or nodes to the web of knowledge. The classification pursued here instantiates such a web.
20. Szostak (2014a) showed that the classification facilitated both cross-group and within-group understanding, responded very well to a feminist critique of classificatory practice by instantiating a web-of-relations approach, and transcended concerns regarding the arbitrary nature of existing hierarchies (by classifying relationships as *relationships* rather than trying to treat them as *things*). Along the way, several other advantages were noted: shorter schedules, expressive notation, ability to search by type of relationship, amenability to translation from other classifications, and placing works (but not concepts) in multiple hierarchies.

## References

Batty, D., & Bearman, T. C. (1983). Knowledge and practice in library and information services. In F. Machlup & U. Mansfield (Eds.), *The study of information: Interdisciplinary messages* (pp. 365–369). New York: Wiley.

Börner, K. (2006). Semantic association networks: Using semantic web technology to improve scholarly knowledge and expertise management. In V. Geroimenko & C. Chen (Eds.), *Visualizing the semantic web: XML-based internet and information visualization* (pp. 183–198). London: Springer.

Budd, J. M. (2004). Relevance: Language, semantics, philosophy. *Library Trends, 52*(3), 447–462.

Budd, J. M. (2011). Meaning, truth, and information: Prolegomena to a theory. *Journal of Documentation, 67*(1), 56–74.

Collins, R. (1998). *The sociology of philosophies: A global theory of intellectual change.* Cambridge, MA: Harvard University Press.

Cornelius, I. (2004). Information and its philosophy. *Library Trends,* 52(3), 377–386.

Cornelius, I. (2014). Epistemological challenges for information science: Constructing information. In F. Ibekwe-SanJuan & T. M. Dousa (Eds.), *Theories of information, communication and knowledge: A multidisciplinary approach* (pp. 181–203). New York: Springer.

Crawford, J. (2011). Introduction. *Library Trends, 60*(2), 257–261.

Fabos, B. (2005). Introduction. *Library Trends, 53*(4), 519–523.

Foskett, A. C. (1996). *The subject approach to information.* London: Library Association.

Fox, M. J. (2012). Book Review [of Szostak 2003 and 2004]. *Knowledge Organization, 39*(4), 300–303.
Fox, M. J., & Reece, A. (2012). Which ethics? Whose morality? An analysis of ethical standards for information organization. *Knowledge Organization, 39*(5), 377–383.
Friedman, A., & Smiraglia, R. P. (2013). Nodes and arcs: Concept map, semiotics, and knowledge organization. *Journal of Documentation, 69*(1), 27–48.
Furner, J. (2010). Philosophy and information studies. *Annual Review of Information Science and Technology, 44*, 161–200.
Furner, J. (2014). Information without information studies. In F. Ibekwe-SanJuan & T. M. Dousa (Eds.), *Theories of information, communication and knowledge: A multidisciplinary approach* (pp. 143–179). New York: Springer.
Gnoli, C. (2008). Ten long-term research questions in knowledge organization. *Knowledge Organization, 35*(2–3), 137–149.
Hjørland, B. (2008). Core classification theory: A reply to Szostak. *Journal of Documentation, 64*(3), 333–342.
Hjørland, B. (2014). Information science and its core concepts: Levels of disagreement. In F. Ibekwe-SanJuan & T. M. Dousa (Eds.), *Theories of information, communication and knowledge: A multidisciplinary approach* (pp. 205–235). New York: Springer.
Huang, X., & Soergel, D. (2013). Relevance: An improved notion for explicating the notion. *Journal of the American Society for Information Science and Technology, 64*(1), 18–35.
Ibekwe-SanJuan, F., & Dousa, T. M. (Eds.). (2014). *Theories of information, communication and knowledge: A multidisciplinary approach.* New York: Springer.
Jacob, E. K. (2004). Classification and categorization: A difference that makes a difference. *Library Trends, 52*(3), 515–540.
Keyton, J., & Beck, S. J. (2010). Perspective: Examining communication as macrocognition. *STS Human Factors, 52*(2), 335–339.
Kleineberg, M. (2013). The blind men and the elephant: Towards an organization of epistemic contexts. *Knowledge Organization, 40*(5), 340–362.
Lambe, P. (2007). *Organising knowledge: Taxonomies, knowledge and organisational effectiveness.* Witney, UK: Chandos.
León Manifesto, The. (2007). Retrieved September 2014, from http://www.iskoi.org/ilc/leon.php
Mardis, M. (2009). Introduction: A gentle manifesto on the relevance and obscurity of school libraries in LIS research. *Library Trends, 58*(1), 1–8.
Marshall, J. G., Solomon, P., & Rathbun-Grubb, S. (2009). Introduction: Workforce issues in library and information science. *Library Trends, 58*(2), 121–125.
Mazzocchi, F., Tiberi, M., De Santis, B., & Plini, P. (2007). Relational semantics in thesauri: Some remarks at theoretical and practical levels. *Knowledge Organization, 34*(4), 197–214.
Ojennus, P., & Tennis, J. T. (2013). Modelling the aesthetic axis of information organization frameworks, part 1: Theoretical basis. *Journal of Documentation, 69*(6), 807–826.
Olson, H. (2004). The ubiquitous hierarchy: An army to overcome the threat of a mob. *Library Trends, 52*(3), 604–616.
Olson, H. (2007). How we construct subjects: A feminist analysis. *Library Trends, 56*(2), 509–541.
O'Rourke, M., Crowley, S., Eigenbrode, S. D., & Wulfhorst, J. D. (2014). *Enhancing communication and collaboration in interdisciplinary research.* Thousand Oaks, CA: Sage.
Repko, A. (2012). *Interdisciplinary research: Process and theory.* Thousand Oaks, CA: Sage.
Ritzo, C., Nam, C., & Bruce, B. (2009). Building a strong Web: Connecting information spaces in schools and communities. *Library Trends, 58*(1), 82–94.
Sandgrind, G. (2010). Introduction: The purpose, present situation, and future of the parliamentary library. *Library Trends, 58*(4), 413–417.
Sauperl, A. (2009). Precoordination or not? A new view of the old question. *Journal of Documentation, 65*(5), 817–833.
Smiraglia, R. P. (2001). *The nature of «a work»: Implications for the organization of knowledge.* Lanham, MD: Scarecrow Press.
Soergel, D. (1985). *Organizing information: Principles of database and retrieval systems.* San Diego, CA: Academic Press.
Stock, W. G. (2010). Concepts and semantic relations in information science. *Journal of the American Society for Information Science and Technology, 61*(10), 1951–1969.

Svenonius, E. (1995). Precoordination or not? In *Subject indexing: Principles and practices in the 90's. Proceedings of the IFLA Satellite Meeting Held in Lisbon, Portugal, 17–18 August 1993* (pp. 231–255). Munich: Saur.

Svenonius, E. (2004). The epistemological foundations of knowledge representations. *Library Trends, 52*(3), 571–587.

Szostak, R. (2003). "Comprehensive" curricular reform: Providing students with an overview of the scholarly enterprise. *Journal of General Education, 52*(1), 27–49.

Szostak, R. (2005a). *Unifying ethics.* Lanham, MD: University Press of America.

Szostak, R. (2005b). Interdisciplinarity and the teaching of public policy. *Journal of Policy Analysis and Management, 24*(4), 853–863.

Szostak, R. (2008). Interdisciplinarity and classification: A reply to Hjørland. *Journal of Documentation, 64*(4), n.p. (letter to the editor).

Szostak, R. (2010). Universal and domain-specific classifications from an interdisciplinary perspective. In C. Gnoli & F. Mazzocchi (Eds.), *Paradigms and conceptual systems in knowledge organization: Proceedings of the 2010 Conference of the International Society for Knowledge Organization* (pp. 71–77). Wurzburg, Germany: Ergon.

Szostak, R. (2011). Complex concepts into basic concepts. *Journal of the American Society for Information Science and Technology, 62*(11), 2247–2265.

Szostak, R. (2012a). Classifying relationships. *Knowledge Organization, 39*(3), 165–178.

Szostak, R. (2012b). *Restoring human progress.* Reading, UK: Cranmore Publications.

Szostak, R. (2013a). *Basic concepts classification.* Retrieved September 2014, from http://sites.google.com/a/ualberta.ca/rick-szostak/research/basic-concepts-classification-web-version-2013

Szostak, R. (2013b). Speaking truth to power in classification. *Knowledge Organization, 40*(1), 76–77.

Szostak, R. (2014a). Classifying for social diversity. *Knowledge Organization, 41*(2), 160–170.

Szostak, R. (2014b, November). *Classification, ontology, and the semantic web.* Paper presented at the ASIST Conference: Special Interest Group on Classification Research, Montreal, Canada.

Wiegand, W. A. (2012). Introduction. *Library Trends, 60*(4), 651–654.

---

Rick Szostak is professor of economics at the University of Alberta. His research, interdisciplinary in orientation, has focused over the last decade on the feasibility and desirability of developing a new approach to classifying both documents and ideas. He has published five books and numerous articles that develop and/or apply such a classification. He is developing the Basic Concepts Classification, which could serve as a substitute for or complement existing systems of document classification. He is also working on a coauthored book titled *Interdisciplinary Knowledge Organization.*