Additional Purposes of Knowledge Organization

Rick Szostak

University of Alberta – Economics, Tory Bldg 14103 53 Ave Edmonton Alberta T6H0S9 – Canada
rszostak@ualberta.ca

Rick Szostak is Professor Emeritus of Economics at the University of Alberta. He was President of ISKO from July 2018 to 2022. He is the author of twenty books and fifty journal articles in economics, history, interdisciplinary studies, information science, and several other fields. He has studied the theory and practice of interdisciplinarity for two decades and has emphasized in the last dozen years the ways in which knowledge organization systems might better facilitate interdisciplinary research and teaching. He has long argued that a phenomenon-based analytic-synthetic approach to classification is both feasible and desirable; he has more recently stressed the use of grammatical construction in performing that synthesis. He co-authored *Interdisciplinary Knowledge Organization* in 2016. He has created the *Basic Concepts Classification* (BCC), and is engaged in efforts to evaluate the BCC and compare it to other classification systems.

Szostak, Rick. 2024. "Additional Purposes of Knowledge Organization". *Knowledge Organization* 51, no. 5: 371-375. 14 references. DOI:10.5771/0943-7444-2024-5-371.

Abstract: The field of Knowledge Organization should recognize additional purposes beyond classifying documents for retrieval. These additional purposes can, in turn, guide us in developing Knowledge Organization Systems. A synthetic phenomenon-based approach to classification mirrors ontological reality. It thus allows writers to better comprehend how their own ideas fit within the broader structure of human understanding. It also allows students (and others) to appreciate that every idea they encounter fits within a broader whole; it should enhance their interest and ability to seek information on any topic. Such an approach serves a third purpose of enhancing social justice: Such an approach to classification is both less biased by nature and easier to navigate. There may be other purposes that we could identify which could also inform our development of KOSs.

Received: 4 June 2024; Accepted 15 June 2024.

Keywords: knowledge organization systems; synthetic phenomenon-based classification; social justice.

† This article was selected as one of the best papers at the Eighteenth International ISKO Conference, March 20-22, 2024, in Wuhan, China.

1.0 Introduction

The field of Knowledge Organization has understandably focused its attention on organizing information, primarily in the form of documents, for purposes of retrieval. It is useful, though, for the field to reflect on other purposes that Knowledge Organization might serve. These broader purposes can usefully inform our discussions of how best to structure knowledge organization systems (KOSs). That is, we might wish to structure KOSs so that they can achieve purposes far beyond document retrieval.

This article will address three additional purposes in turn: helping authors to place their ideas in context; helping students (and others) to comprehend the structure of human understanding; and to support the pursuit of social justice. The first two of these are only rarely discussed in the field. I would suggest that this likely reflects the fact that existing KOSs in widespread use do not support these goals. Thus, if we ask, "What goals can existing KOSs serve?" we will not be guided to ask about what authors or students might learn from them about the structure of human understanding. I will argue below that a synthetic phenomenon-based classification, an approach long advocated by multiple scholars in the field for facilitating document retrieval (e.g. Gnoli 2016), could far better serve these additional goals. This paper is thus simultaneously an argument in favor of a broader understanding of the goals of knowledge organization, and for the pursuit of synthetic phenomenon-based classification.

Concerns regarding social justice are often raised within knowledge organization. They are manifested most obviously in important efforts to remove biases from knowledge organization systems that treat people with particular gender identities, racial/ethnic identities, religions, or sexual orientations unfairly. The third section of this paper will argue that a synthetic phenomenon-based classification is in-

herently less biased. We may thus be better able to advance the goal of fairness in KOSs by developing new systems rather than reforming old systems. Moreover, we should worry about the access to information of disadvantaged groups. This would also be enhanced by the ease of navigation of phenomenon-based classification.

We have no pretense to being exhaustive in this paper. It may well be that other purposes for KO can be identified that would also inform our pursuit of KOSs.

2.0 Knowledge organization can help writers place their ideas in context

As noted above, Knowledge Organization has focused to date on providing access to documents. Yet if we really wish to organize knowledge – as the title of the field suggests – we should also wish that our KOSs could readily also organize "insights": the conclusions that individual pieces of research have generated over the years. Yet we rarely speak in KO of organizing insights or ideas. Why not? Since the very title of the field indicates that we should, this omission would seem to reflect the simple fact that the KOSs developed to classify documents are ill suited to classifying insights also.

Szostak (2022) argued that we should classify an ontological reality consisting of a large (but finite) set of phenomena that mutually exert influences on each other. Most authors – scholarly or not – will talk about how one or more phenomena influence one or more others (a minority will instead focus on describing in detail the nature of a particular phenomenon). Their works can ideally be given subject headings of the type (phenomenon X) (exerts effect N) (on phenomenon Y) (Szostak 2017). Note that we can classify both the document and its core insights in a similar fashion by synthetically combining phenomena derived from schedules of phenomena with effects derived from schedules of "relators" (that is, terms that are generally verbs or conjunctions and indicate the various ways that phenomena might interact). We then need a KOS such as the Basic Concepts Classification (Szostak n.d.) that provides schedules of both phenomena and relators (The BCC also contains a schedule of adverb/adjectival "properties" which can clarify the nature of both phenomena and relators).

A classification system structured in this way can readily alert the writer to all other posited effects of X, all other posited influences on Y, and all other examples of influence N. Likewise, it can easily both encourage and satisfy curiosity regarding the causes of X, effects of Y, or alternative causal pathways to N. Indeed, if the KOS were coupled with a search engine with an appropriate visual interface, these alternative possibilities should be just a click away (Such an interface could take the form of a flowchart with the causal link queried by the user in the middle, and related links emanating from this). This can spare the writer from some very common errors: imagining that their pet project is more important than it is, ignoring potential side effects of any behavior or policy they might recommend, and failing to draw possible lessons from similar lines of inquiry.

Pulling back a little, the writer can see that X and Y are nestled within a much larger network of influences, of which N is just one type. They can appreciate the difficulty of examining one corner of reality that is inevitably jostled by interactions with the rest of reality (natural scientists can often isolate particular influences in a laboratory; human scientists can rarely do so; see Szostak 2023). Yet they can come to see that human attempts at understanding, and the scholarly enterprise in particular, are interconnected and that we are slowly adding to human understanding by coming to understand better each link in this complex web of relationships.

Existing enumerative KOSs can hardly act to inform the pursuit of human understanding in this way. They enumerate a small fraction of the complex combinations of phenomena addressed by researchers, and generally have limited if any ability to signal what sort of influence one phenomenon might exert on another. They are thus far less precise than a KOS such as BCC in their ability to capture insights. (Smiraglia and Szostak 2017 and 2018 found that the BCC achieves precision even with respect to documents.) As a result, they tend to reinforce a mistaken idea that human understanding can be chopped into innumerable distinct little areas of exploration that can be pursued in isolation. Yet the right kind of KOS can not only better organize human understanding but also encourage a more productive approach to advancing human understanding.

We might note in closing that the Semantic Web aspires to link insights generated in one repository with insights contained in any other repository. Information is to be coded in terms of RDF Triples of format (subject)(predicate or property)(object). Note that the subjects and objects are phenomena. The predicates or properties are most often relators but may be properties (The sky is blue). A KOS such as the BCC, which has schedules of phenomena, relators, and properties, is manifestly suitable for the Semantic Web, whereas the complex subject headings of enumerative KOSs are not. The potential of the Semantic Web is severely limited at present by the lack of interoperability between the terminologies employed in different repositories. This problem might be remedied if a phenomenon-based KOS were employed widely across the Semantic Web.

3.0 A properly designed KOS can be a useful learning tool

Students in university, and even in K-12 (and indeed readers more generally), would benefit from an ability to place any piece of information they encounter within the broader

body of human understanding. They will benefit in the same ways that authors do, gaining a better sense of how confident to be in a particular piece of information, and where to look for related information. (Our understandings of human memory also suggest that we are more likely to remember any piece of information if it is connected logically to other pieces of information in our memory.)

Soergel (2013) had noted the advantages of organizing knowledge for student learning. He appreciated that most theories of learning stressed the importance of learners developing structures that tie many pieces of information together. Students were more likely to learn and remember if they could place new information within an existing structure. He worried that the instructional design literature largely ignored this important insight, providing little advice on how to provide students with organizing structures. Soergel appreciated that the KO literature had very occasionally recognized the advantages of learning KO. Yet Soergel (2013) surveyed a wide range of organizational schemes that might aid student learning. Our purpose here is instead to highlight the role that KOSs can play in enhancing learning.

A good KOS will allow users to move seamlessly from one document, idea, or object to related items. You start a search on why dogs bite mail carriers and get curious about whom else dogs bite, or whether they occasionally nuzzle mail carriers instead, and can quickly alter elements of your original search. Alternatively, you can move on down a causal chain to investigate health care options for dog bites.

We often hear that we live in an age of information overload. Students of all ages need to learn how to find the information they need for a wide range of questions. They will benefit enormously from access to a KOS that is easy to navigate and is used to organize multiple databases. They will benefit even more if they can learn the logical structure that guides this KOS. They will then simultaneously learn about the general structure of human understanding (which reflects the ontology described above) and how to follow their curiosity through this body. We should seek, then, to develop a KOS(s) with transparent structures so that students can easily grasp how we organize information and see beyond this to the ontological reality we are attempting to comprehend. (This idea has been around for at least a century; see Miksa 1992).

It is notable in this regard that a classification such as the BCC can be both easy to navigate and precise. This is possible because of the advantages of a synthetic approach. Enumerative classifications sprawl and defy logical order because they attempt to capture complex combinations of phenomena in a single subject heading. The schedules of phenomena in BCC are generally very flat (the exceptions being in areas like the classification of species where the ontological reality is multi-layered). They are generally also logical with the vast majority of subclasses being "types of" the superior class. The schedules of relators are short: Most relators are generated by synthesizing simpler relators (or relators and properties or phenomena). (The BCC is described in more detail in Szostak n.d. and Szostak 2019).

Imagine a class in high school (maybe earlier) where students are exposed to a logical and concise KOS that mirrors the ontological reality that we inhabit. Students simultaneously learn about the nature of their world, the nature of human attempts to comprehend that world (both natural and human worlds), and how we organize those understandings for retrieval. Students will come to appreciate that every little bit of understanding they encounter is nestled within a larger whole. By understanding its structure, they will find the body of human understanding less overwhelming. They will simultaneously enhance both their willingness and ability to seek out information (in both physical and online libraries) on any subject that interests them. Such a class is both invaluable and entirely feasible if we both broaden our understanding of the purposes of KO and develop a KOS that can achieve these broader goals.

4.0 Knowledge organization should enhance access of disadvantaged groups and eliminate biases from KOSs

Knowledge Organization has a critical role to play in the quest for social justice. Most obviously, we can strive to develop KOSs that treat various social groups fairly. Existing KOSs in widespread use have numerous social biases embedded within them, reflective of the 19th century environment in which they were designed. If we assume that nurses are female, and then create a special subclass of "male nurse" but not "female nurse," we instantiate a view of reality that discriminates. More seriously, if we view homosexuality as a mental illness, we encourage the mistreatment of a group of people. Many KO scholars have pointed out a long list of biases in existing KOSs; some of these have taken steps to alleviate or eliminate these (a classic work is Olson 2001).

Note that a synthetic approach to classification can automatically eliminate many sources of bias. If a subject heading "male nurse" combines "male" from a schedule of genders and "nurse" from a schedule of occupations, then it is classificationally equivalent to the combination "female" and "nurse." Moreover, if we have in our schedule of gender other types of gender identity, than these also can be combined with any other term in exactly the same way as "male" and "female." It is likely much easier to eliminate bias if we develop a new synthetic KOS than by trying to reform existing KOSs (Szostak 2014).

The unnecessary complexity of the schedules for existing KOSs add a further barrier to social justice: Members of disadvantaged groups find these classifications hard to navigate and are thus limited in their access to information. Many public libraries have moved away from complex KOSs

to organize their shelves like bookstores instead, with a much smaller number of very broad categories (Martínez-Ávila et al. 2014). This may enhance access to works of a general nature, but makes it harder to locate works that address a very precise topic.

My own research suggests that it is quite possible to develop a KOS that is easy for all literate users to navigate. Indeed a search interface should be able to translate most user queries into a subject heading that will guide them precisely to the works they seek. The logical structure of a phenomenon-based KOS is important here; it is entirely feasible to develop a thesaurus to guide users toward the controlled vocabulary in such a KOS (Renwick and Szostak 2020). The key here is again a synthetic approach where the interface seeks combinations of terms in the user query. Even greater precision can be achieved if our subject headings place terms in the same grammatical order in which humans speak (Szostak 2017). Note also that subject headings in a grammatical format are easier to comprehend since humans are accustomed to thinking in sentences.

A synthetic KOS can enhance social justice in two key ways: by reducing bias and by enhancing access to information. These combine for a third benefit: Members of any social group should be able to find more readily information about their own or any other social group. We can thus enhance mutual understanding.

5.0 Concluding Remarks

We can and should expand our understanding of the purposes of KO. We can, and thus should, be giving people a better understanding of how our collective understanding of the world is organized. The ontological reality is a large but finite set of phenomena that influence each other in a large but finite set of ways. A KOS that mirrors this ontological reality with logical schedules of both phenomena and relators allows us all to appreciate both reality and human understanding of reality. Authors will be better able to place their work in context. Students will learn that every piece of information is nestled within a larger structure, and will be emboldened to search out credible information on any subject.

The synthetic approach to classification that achieves these purposes simultaneously enhances social justice. It increases the capacity of disadvantaged groups to access information. Moreover, a synthetic approach to classification is inherently less biased since it treats all groups in exactly the same way.

There may be yet other purposes of KO. We should seek to identify any additional purposes that KO might serve, and then design KOSs that serve the widest range of purposes.

References

Gnoli, Claudio. 2016. "Classifying Phenomena Part 1: Dimensions." *Knowledge Organization* 43, no.6: 403-15. doi.org/10.5771/0943-7444-2016-6-403

Martínez-Ávila, Daniel, Rosa San Segundo, and Hope A. Olson. 2014. "The Use of BISAC in Libraries as New Cases of Reader-Interest Classifications." *Cataloging & Classification Quarterly* 52:2, 137-55.

Miksa, F. L. 1992. "The Concept of the Universe of Knowledge and the Purpose of LIS Classification." In *Classification Research for Knowledge Representation and Organization: Proceedings of The 5th International Study Conference On Classification Research, Toronto, Canada, June 24–28, 1991*, edited by N. J. Williamson & M. Hudon. Elsevier: Amsterdam, Netherlands,101-126.

Olson, Hope A. 2001. "The Power to Name: Representation in Library Catalogs." *Signs* 26, no.3: 639-668.

Renwick, Tobias, and Rick Szostak. 2020. "A Thesaural Interface for the Basic Concepts Classification". *Knowledge Organization at the Interface*: *Proceedings of the Sixteenth International ISKO Conference, 2020 Aalborg, Denmark, edited by* Marianne Lykke, Tanja Svarre, and Daniel Martínez-Ávila. Advances in knowledge Organization 17. Ergon Verlag, *527-31*

Smiraglia, Richard, and Rick Szostak. 2017. "Comparative Approaches to Facets in Interdisciplinary KOSs: UDC and Basic Concepts Classification." In *Faceted classification today: theory, technology and end users*: *Proceedings of the International UDC Seminar 14-15 July 2017, London, UK*, edited by Aida Slavic and Claudio Gnoli. Würzburg: Ergon Verlag, 279-84.

Smiraglia, Richard and Rick Szostak. 2018. "Converting UDC to BCC: Comparative Approaches to Interdisciplinarity" In *Challenges and Opportunities for Knowledge Organization in the Digital Age: Proceedings of the Fifteenth International ISKO Conference 9-11 July 2018 Porto, Portugal*, edited by Fernanda Ribeiro and Maria Elisa Cerveira. Advances in knowledge organization 16. Baden-Baden: Ergon. 530-38.

Soergel, Dagobert. 2013. "Knowledge Organization for Learning. Conjectures and Methods of Study." *Ciencia da Informação* 42, no.2: 232-54. https://doi.org/10.18225/ci.inf.v42i2.1384

Szostak, Rick. n.d. *Basic Concepts Classification*. (2010, revised 2013, regularly updated since). https://sites.google.com/a/ualberta.ca/rick-szostak/research/basic-concepts-classification-web-version-2013

Szostak, Rick. 2014. "Classifying for Social Diversity." *Knowledge Organization* 41, no2*:* 160-70. doi.org/10.5771/0943-7444-2014-2-160

Szostak, Rick. 2017. "A Grammatical Approach to Subject Classification in Museums." *Knowledge Organization*

44, no.7: 494-505. doi.org/10.5771/0943-7444-2017-7-494

Szostak, Rick. 2019. "The Basic Concepts Classification (BCC)". In *ISKO Encyclopedia of Knowledge Organization*, edited by Birger Hjørland and Claudio Gnoli. Also available in *Knowledge Organization* 47, no. 3: 231-43.

Szostak, Rick. 2022. "Toward a General Theory of Knowledge Organization." In *Knowledge Organization across Disciplines, Domains, Services and Technologies. Proceedings of the Seventeenth International ISKO Conference, July 2022 Aalborg, Denmark*, edited by Marianne Lykke, Tanja Svarre, David Haynes, Mette Skov, Martin Thellefsen, and Daniel Martínez-Ávila. Advances in knowledge Organization 19. Ergon Verlag, 261-72.

Szostak, Rick. 2023. *Integrating the Human Sciences: Enhancing Cohesion and Progress across the Social Sciences and Humanities.* London: Routledge.