

# Synthetic Classification and Diverse Communities

Rick Szostak  
University of Alberta  
rszostak@ualberta.ca

## ABSTRACT

This brief paper argues that a synthetic approach to classification can alleviate *all of* the major concerns that are commonly raised about how Knowledge Organization Systems (KOSs) may disserve various communities. It surveys how a synthetic approach can potentially address a variety of concerns regarding KOSs and social diversity.

## Keywords

Knowledge organization systems, synthetic, post-coordinated, social diversity.

## INTRODUCTION

There are a variety of ways in which a KOS might disserve members of particular communities. In each case, a synthetic approach to classification offers a potential solution. This paper thus explores the various ways in which a synthetic approach to classification may *liberate* the capacities of members of diverse communities. Many of the topics addressed were taken from the call for proposals.

## DEFINING A SYNTHETIC APPROACH TO CLASSIFICATION

In a synthetic (or post-coordinated) approach to classification, classifiers can formulate subject headings by combining elements from different schedules within a classification. This can be contrasted with a pre-coordinated approach in which a classifier generally needs to choose from a pre-established set of subject headings. For example, in a pre-coordinated approach, a classifier might select a subject heading of “male nurse,” whereas in a synthetic approach the classifier might combine “male” (from a schedule of genders) and “nurse” (from a schedule of occupations).

Many authors have compared pre-coordinated and post-coordinated approaches to classification. The main potential disadvantage of the latter is that a search for works on (philosophy)(of)(history) might yield many unwanted documents addressing (history)(of)(philosophy). Yet this disadvantage disappears if we employ search engines that prioritize the order in which search terms are entered.

81<sup>st</sup> Annual Meeting of the Association for Information Science & Technology | Vancouver, Canada | Nov. 10 - 14, 2018

Author(s) Retain Copyright

## BIASES IN CLASSIFICATION

If the classification system reflects the cultural biases of its designers, members of other communities may find the classification difficult to comprehend and even offensive. A classic example here is a presumption within some classifications that nurses will be female: A special subclass for “male nurse” is then created, indicating that this is thought to be the unusual case. This problem might be addressed by carefully surveying existing KOSs and addressing each case of unequal treatment of different communities individually. Alternatively, a synthetic approach to classification can address the challenge at a holistic level. In a synthetic approach, male nurse and female nurse – and indeed transgendered nurse (of various types) – are naturally treated symmetrically. We simply need to ensure that all communities are captured in our schedules.

## TERMINOLOGICAL AMBIGUITY AND TRANSLATION

Will individuals from different communities understand terms in similar ways? If not, an individual from one community may have trouble navigating a KOS designed by a member of another community. Szostak (2011) argued that the complex concepts that are understood differently across communities can generally be broken into “basic concepts”: terms for which there is enough shared understanding across individuals and communities for the purposes of classification. Globalization may have various meanings, but exports (of goods) is understood similarly by most people. A synthetic approach to classification allows subject headings to be constructed by combining basic concepts. Such subject headings should be far less ambiguous than those developed within enumerated classifications. They then allow members of all communities equivalent access to the KOS.

Basic concepts are likely also far easier to translate across languages than complex concepts. The lesser degree of ambiguity in the original language should facilitate the identification of a very similar term in other languages. Moreover, basic concepts tend to represent things that we perceive in similar ways in the world around us.

## EASE OF ACCESS

Some individuals or communities may find existing KOSs difficult to navigate, limiting their access to information. Recent decisions by some public libraries to move away from library classifications toward the BISAC employed in bookstores reflect a sense that many people find existing library classifications bewildering. One challenge here is the biases and ambiguity addressed above. But a greater challenge for many is simply not understanding how to identify

an appropriate subject heading to search. A synthetic approach to classification potentially allows users to more readily identify appropriate subject headings. Rather than needing to figure out how the KOS deals with male nurses, the user simply combines “male” and “nurse” in their search query. I have in recent research argued that we can make search even easier by following common grammatical rules in our structuring of synthetic subject headings. A user search query that employs standard grammatical construction is then readily translated into a relevant subject heading.

### **DIFFERENCES IN PERSPECTIVE**

Different communities or individuals may approach topics from a different perspective. If a KOS reflects one perspective, works from other perspectives may be misclassified and hard to find. One useful approach is to classify works by authorial perspective. Since authorial perspective is multidimensional, a synthetic approach to classification is best suited to classifying this: authors may differ with respect to rhetorical strategies, ideology, membership in various communities, and other ways (see Szostak 2015).

As noted above, a synthetic approach also reduces the scope for bias within the subject classification itself. An enumerated classification may assume that certain concepts naturally belong together – such as female and nurse – whereas a synthetic approach seeks to link unidimensional concepts.

Note that synthetic approaches to both subject classification and classification of authorial perspective aid users both when they seek works emanating from only one community or perspective and when they actively seek works from multiple communities or perspectives.

### **THE STRUCTURE OF CLASSIFICATIONS**

Hope Olsen (2007) famously argued that a hierarchical approach to classification may reflect a male perspective. Women may be more likely to see the world in terms of non-hierarchical relationships. A synthetic approach is grounded in a belief that authors and users should potentially be able to combine any set of concepts as they see fit. Classification systems that pursue a synthetic approach to developing subject headings still have to organize the concepts to be synthesized hierarchically, but these hierarchies can be much flatter than those within enumerated classification. The Basic Concepts Classification, for example, only rarely has more than three or four levels of hierarchy.

### **HOSPITALITY**

A KOS designed by members of one community may exclude concepts deemed important by members of other communities. The hospitality of a KOS – the ability to add new terms – is thus an important consideration here. It is not always clear where to place a new term within the multi-level hierarchies of complex terms that characterize enumerated classifications. Within a synthetic approach, new terms can

usually be created through a new synthesis of existing terms. When a new basic concept must be added to a KOS, this is easier in flat and logical hierarchies: One need not search multiple levels and wonder what the principles guiding the hierarchy are.

### **DOMAINS**

One approach to dealing with terminological ambiguity is to develop domain-specific KOSs. The obvious danger is that users then have difficulty finding information in multiple domains. This could limit both interdisciplinary and cross-community understanding. A synthetic approach to comprehensive (“universal”) classification utilizing basic concepts should facilitate any community’s access to its own literature without limiting cross-community understanding.

### **OPPORTUNITIES FOR CROWDSOURCING**

Crowdsourcing provides an opportunity for members of one community to suggest changes to a KOS designed by a member of another community. It thus provides a potentially powerful response to the imposition of a KOS with undesirable characteristics. Crowdsourcing is likely to be easier and more successful if the principles guiding the KOS are transparent. A synthetic approach to classification that employs basic concepts embedded in flat logical hierarchies should be more easy to comprehend and amend than detailed enumerated classifications. It would be interesting to explore this hypothesis empirically.

### **KOSS AS A FORM OF ADVOCACY**

The approach outlined above seeks to make a KOS as inoffensive as possible to as many communities as possible. Though the KOS itself does not seek to advocate for any one community, the precision of synthetic subject strings potentially allows works of advocacy to be more readily found by users (e.g. (fighting)(discrimination)(against)(community X)). Classification of authorial perspective would further enhance this capability.

### **CONCLUDING REMARKS**

This brief paper has purposely chosen breadth of coverage over depth of analysis. It has shown that a synthetic approach to classification can potentially alleviate a variety of challenges that KOSs may present to members of particular communities. The author thus recommends the synthetic approach as likely the best way of alleviating the adverse effects of KOSs on diverse communities.

### **REFERENCES**

- Olson, Hope A. (2007). How we construct subjects: A feminist analysis. *Library Trends*. 56 (2), 509–541.
- Szostak, Rick. (2015). Classifying authorial perspective. *Knowledge Organization* 42(7), 499-507.
- Szostak, Rick. (2011). Complex concepts into basic concepts. *Journal of the American Society for Information Science & Technology* 62(11), 2247-65.