

2016 Satellite meeting - *Subject Access: Unlimited Opportunities*

11 – 12 August 2016

State Library of Ohio, Columbus, Ohio, USA

Subject Access Principles in the New World: Procrustean or Procreative?

Lynne C. Howarth

Faculty of Information, University of Toronto, Toronto, Canada.

Hope A. Olson

School of Information Studies, University of Wisconsin-Milwaukee, Milwaukee, USA



Copyright © 2016 by Lynne C. Howarth and Hope A. Olson. This work is made available under the terms of the Creative Commons Attribution 4.0 International License: <http://creativecommons.org/licenses/by/4.0>

Abstract:

The “old world” of subject access emphasized compliance with standards for content and structure directed at achieving consistency. Increasingly, as epistemic perspectives have moved away from positivist to post-positivist, and constructivist worldviews, the rationale for Knowledge Organization Systems (KOS) designed from the top down has come under scrutiny. In such a shifting landscape, what can we do when we find ourselves in want of principles? Building on the conference focus of using connections between traditional and newer methods to take advantage of new opportunities, this conceptual paper re-examines theoretical underpinnings of subject access and, using design thinking approaches, looks at current applications in order to determine what is usable, what could be respectfully borrowed and adapted, and what needs to be considered and/or developed de novo?

Keywords:

design thinking, knowledge organization systems, subject access principles, social tagging, meaning-making

INTRODUCTION

In the past, the principles of knowledge organization (KO), sired by information science out of librarianship, could be applied to new media, new genres, new topics, and new user groups. We could trust our proven principles to hold as true for us as they did for Callimachis in Alexandra. However, now rapid technological change, and dynamic economic, social, and political factors construct the contexts in which we find information and those contexts construct information. Are such changes essential? Fundamental? In this new environment what can we do when we find ourselves in want of principles?

Looking particularly at longstanding principles and practices associated with subject cataloging would suggest that the “old world” of subject access emphasized compliance with standards for content and structure directed at achieving consistency. Increasingly, as epistemic perspectives have moved away from positivist to post-positivist, and constructivist worldviews, the rationale for Knowledge Organization Systems (KOS) designed from the top down has likewise come under scrutiny. In an ever-changing world with fluid boundaries, fixity has become problematic. The now commonplace activities of social tagging have likewise challenged the requirement for expertise in subject access theory and practice, as engaged, but not formally trained, individuals assign subjects and various descriptors to an array of digital objects. In a kind of compromise some repositories now accommodate surrogate records of digital objects created by subject cataloging professionals with added content from non-professional repository users.

Stand-alone social media sites, such as LibraryThing, facilitate end-user tagging, while also encouraging subscribers to create their own catalogues with “book data” from Amazon.com (ONIX standard records), from the Library of Congress, and from “1051 other libraries” (LibraryThing 2016). Whether a tagger-friendly library web catalogue, or a standard bibliographic record-hospitable social media site, it appears that the blended approach to complementing access to information through both user- and subject cataloging professional-created data has become more established and accepted within the subject access lexicon. This is one model of subject assignment and subject access in the contemporary or “new world” referred to in this paper.

This is also a world of subject access where everything old is new again, to borrow a familiar expression. The longstanding debates remain – or are perhaps intensifying – around controlled vocabularies versus natural language, between the development and application of national and international standards versus situationally-responsive non-standard approaches, between solutions designed within the profession and discipline versus those deriving from outside or from those less familiar or even unfamiliar with disciplinary practice, and between human-directed versus technology-driven applications. These debates are fueled, not only by the ubiquity of social media, but also by a more reductionist economic climate where technical solutions may be seen as more cost-effective than human expertise, by a persistent drive to understand and predict through data analytics, and by the deliberate design of solutions to longstanding problems, as evidenced by, for example, automatic categorization of large data sets, or linked data approaches to optimizing controlled vocabularies.

Underlying the debates as they relate to subject access are foundational principles which may, themselves, have been challenged and examined to varying degrees. For example, while a number of traditional classification systems have been based on some kind of warrant, i.e., literary, scientific/philosophical, educational, or cultural – as described by Beghtol (1986), what type of “warrant” might be suggested by terms assigned by a tagger in Flickr? Is there an underpinning of “warrant” in an algorithmic analysis of what topics are trending in Twitter at any given moment?

Likewise, many traditional subject heading systems are structured hierarchically, providing opportunities for navigating from general to specific, from narrower terms to broader terms, and across headings that are related in meaning. What happens to, or what value remains for a concept of hierarchy in an environment where multiple concepts can be chained or linked to discover, not only linguistically-related information, but other kinds of relationships –

including those that are serendipitous? The idea of a semantic web is a 21st century reality where virtually any kind of relationship that a document/object, person, place, event has to any other is discoverable, and where multiple linkages can be effected at the same time. How does this flexibility, this break with fixedness, affect how subject catalogers approach both the design and application of subject access systems? More fundamentally, how might such fluidity impact the notion of meaning-making?

Assigning subject headings or descriptors to documents (entities, objects) has traditionally been an exercise in capturing their “aboutness”. However, aboutness is not only slippery to determine, it also fails to identify the essence of a document which we traditionally seek to represent in our surrogates to enable finding specific documents and gathering documents with similarities. The essence of something is, according to the *Oxford English Dictionary*, “[t]he totality of the properties, constituent elements, etc., without which it would cease to be the same thing; the indispensable and necessary attributes of a thing.” (OED) Our efforts at subject access cannot account for that level of totality. The focus on aboutness has posed some limitations to resources that “are” rather than being “about”, as in the case of works of fiction, or a collection of objects or media, or institutional or personal archival fonds. Social tagging of museum objects or photographs has resulted in the capture of characteristics, such as color, or emotion, or association with something else, as much as with the “aboutness” of the entity or image. Are original principles relating to the requirement for representing “the subject” of a resource no longer relevant in this new world of subject access? Are we still looking for the essence of a document? Does subject contain essence or vice versa or do we need a different approach to subject representation – perhaps one that will not automatically award primacy to subject or one that focuses on how the characteristics of a document relate to each other and to a given context?

Building on the conference focus of using connections between traditional and newer methods to take advantage of new opportunities, this conceptual paper re-examines theoretical underpinnings of subject access and questions posed, above, and looks at current applications in order to determine what is usable, what could be respectfully borrowed and adapted, and what needs to be considered and/or developed *de novo*?

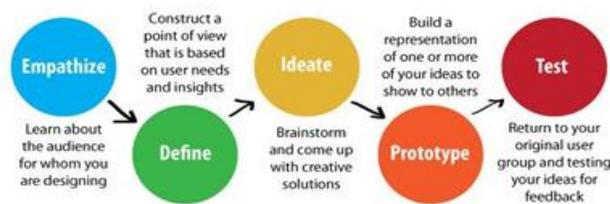
Using a design thinking approach facilitates rethinking subject access principles to assess their enduring value or usefulness – if any. Through the lens of design thinking we can explore the “unlimited opportunities” of subject access, suggesting new, different, or similar applications of core tenets for accommodating, incorporating, or even imagining new tools, technologies, and knowledge organization systems (KOS). To what extent do “old world” subject access principles serve to limit and constrain, or, conversely, to invigorate and generate “new world” subject access and KOS – i.e., principles as procrustean or procreative?

DESIGN THINKING LENS

While there is no universally adopted approach to design thinking, one, created by the design consultancy, IDEO, has been made readily available to libraries in partnership with the Chicago Public Library and Aarhus Public Libraries, and with funding from the Bill and Melinda Gates Foundation. In *Design Thinking for Libraries* (2015), the process is described as a series of steps focused on human-centered solutions to meaningful problems. A successful design thinking innovation will result when the solution is desirable, fiscally viable, and technologically feasible (p. 6). As Figure 1 illustrates, the process begins with understanding for whom one is designing (empathize), then articulating a user perspective in terms of his or

her needs, difficulties or frustrations – “pain points” – with the existing process or product (define). Putting oneself in the shoes of the user encourages focused brainstorming towards a user-centered solution (ideate). This ideation phase is normally driven by a “How might we ...?” question that aims to address a specific user need, a problem, a pain-point. Resulting initial designs can then be prototyped and introduced to the original user or user group for testing and feedback.

Design Thinking Process



Source: Design Thinking for Libraries

Figure 1. Design Thinking Process (IDEO)

While we are in no way suggesting that a design thinking approach is necessarily the panacea for tackling the more procrustean or limiting aspects of subject access principles, or for addressing “How might we design more effective subject access systems?” – indeed a large question! – we do view the process as somewhat consistent with traditional approaches within KO. Charles Cutter (1904) urged catalogers to keep the user uppermost in mind when providing information essential to finding a work for which an author, title, or subject is known, or for locating like materials together. Within the context of the development and application of major subject access tools within the Anglo-American worldview (notably, Dewey Decimal Classification (DDC), Library of Congress Subject Headings (LCSH), Library of Congress Classification (LCC), Cutter’s Expansive Classification), user needs were assumed by “experts” – by those who designed the KO systems, and by the subject cataloguers, indexers, bibliographers, and others who applied them.

Within our contemporary context, and from a design thinking perspective, identifying user needs and designing solutions to address them is shifting to include “the user”, per se. As noted previously, we have seen the co-mingling of “expert” and “user” contributions to records that provide for accessing physical and digital resources across a range of applications. What we have not yet witnessed is direct user engagement in the ongoing development of subject access tools, such as DDC, LCC, or LCSH. While one might argue that such tools are specific to institutional collections and applications, not to mention embedded inextricably within rich legacy catalogues and globally-shared databases, there is a kind of “social warrant” being exerted by the likes of increasingly ubiquitous social media and linked data applications. Further, the boundaries between and among the individual, the local, the national, and the international have not only been blurred, they may, in fact, be indistinguishable. Content exposed through Facebook, Twitter, Instagram, etc., is, to all extents, universal, readily

discoverable and broadly shared. At least some portion merits preservation, whether personally or institutionally archived (or both), with implications that personally or culturally-important content be made discoverable and accessible for some period of time, perhaps in perpetuity. If the two solitudes of “traditional”, standards-based, institution-maintained KO systems, and “user-generated”, non-standardized KO systems continue in parallel with occasional crossover, how might that impact a fully inclusive (historical) cultural record going forward?

To return to our design thinking framework, is there an opportunity for better understanding and defining user needs – aware of a tenet within KO to be mindful of “the user” and his or her requirements – that can, in partnership with subject access system users, lead to user-centered solutions to current KOS design problems? Understanding “the audience” for whom we are designing encompasses a vast canvas; identifying difficulties, frustrations, or specific “pain points” may be less daunting given a large body of research on information search behaviors, catalog use and user studies, and ongoing critiques of “legacy” systems. The user-centered “wisdom of the hive” promotion of social tagging, is often positioned in opposition to “traditional” subject heading/classification systems (Shirky 2005). The latter are likewise well defended by their own supporters. This oppositional stance, however, quite likely in reference to the same “user”, whose “pain points” relative to each “system” may be widely shared, does not advance a user-centered solution to the design of KO systems hospitable to deeply understood needs. There may, in fact, be more in common than different when both strengths and weaknesses of standards-based, and so-called “naïve” or non-standard KO approaches are considered as part of the same problem and also of the potential solution. With a design thinking “empathy” for the user, the focus shifts from the system per se (advantages/disadvantages) to audiences for whom, and with whom, the KO system is to be designed.

FROM INFORMATION-FINDING TO MEANING-MAKING: RETHINKING CORE TENETS

While breaking with the requirement of engaging with users who are limited by a design problem, let us, for a moment, consider who these users might be within the contemporary context, and what their needs might be relative to subject access. Users of the 19th and 20th centuries, might be characterized as needing (subject access) tools for discovering known items, or for browsing like materials. While perhaps a generalization, 21st century users continue to have a need for both known and like resources (digital or physical). In addition, however, and facilitated by increasingly sophisticated online access technologies, they have moved beyond information-finding to *meaning-making*, to chaining both information in and around a discoverable to creating new meaning, exploring links of relationships, and often viewing the same information within new or different contexts.

Meaning-making chaining is facilitated through structured data, whether controlled or natural vocabulary, whether deriving from a standards-based, institutionally maintained system, or from a user-generated “naïve classification” (Beghtol 2003). We see in structured data the enabling building block of linked data. This would seem to disqualify the kinds of laissez-faire tagging associated with social media content, yet, the concept of a kind of “social warrant” might suggest otherwise. Tag clouds reveal clusters of terminology that can be seen to characterize content associated with, or perhaps dominant within a site. Is there merit to considering a de facto natural language vocabulary that is commonly shared across the site and afforded “authority” through social warrant- a communal expression of agreement on representative terminology for that content? Can traditional KO constructs of warrant offer a

potential solution for welcoming non-standard social tagging as structured data within a linked data environment, thus adding to the possibility of enhanced chaining for meaning-making? Does such an application bridge the divide between the two solitudes of subject access systems, described above, while also addressing contemporary user needs to both find existing and create new information, seamlessly, and at the same time?

Another possible bridge afforded by a traditional tenet of KO is that of facets. Core to fine tuning the subject access offered through a class number constructed with facets representing aspects essential to resource content, facets may be seen as building blocks of meaning. While it may seem glib to associate facets with tags as representative “bits” of meaningful data, such an association may, like warrant, open a door to rethinking the potential value of unstructured, user-generated, natural language tags as complementary to subject access systems. A facet comprising part of a well-formed tuple (or “triple”) within the linked data context has implicit value regardless of how it originated. Seeing tags through the lens of facets may also prompt a rethinking of how we value facets generated by “experts”, versus those by non-experts, or the creation of facets from the top down, or from the bottom up. From the perspective of a user, there may be little or no apparent difference in terminology, though there remains a requirement to learn different systems depending on where one is searching. Addressing this user “pain point” might constitute a constructive step in the rethinking and design of contemporary user-centered subject access systems.

The traditional approach of assigning subject headings or descriptors to capturing the “aboutness” of a tangible or digital object is problematic when one shifts from information discovery to creating new information or meaning-making. As observed, previously, locating or collocating an object based on its “aboutness”, while important, may fail to account for the totality of the object. If within “essence” we include the “indispensable and necessary attributes of a thing” (OED), then a subject heading or descriptor will in all instances fall short. The tangible and digital resources to which we regularly assign subjects are given meaning by those who engage with them. Essence is, so to speak, in the eyes of the beholder. Allowing for the meanings vested by “users” to be incorporated into the KOS records created as surrogates for those resources is a step in the direction from capturing aboutness to sharing meaning. Olson and Howarth (2013) explored this idea with an idea of moving the “authority” for representing an object from the KO expert to the KOS user. Associated with that was ensuring that objects and their records became only one stop in a series of meaning-making chains, linking one resource to the next, as a user determined. Shifting authority allows for the voice of the user to be heard, opening possibilities for difference, for the inclusion of those who may have been marginalized, for the promulgation of rich and diverse cultures and socio-political experiences. Whoever holds the key to naming, holds the power to name (Olson 2002). Distributing/diffusing that power broadly within subject access systems (powerful naming systems) promises greater equity and equality of voice and unfettered meaning-making.

When meaning-making is viewed from a constructivist stance, with meaning inherent not in the object or resource itself, but rather vested by those engaging with it, then the concept of hierarchically structured subject headings becomes problematic. Even irrelevant. An individual construction of meaning may, but more likely will not involve moving through well-ordered hierarchical chains. Links may be random, even serendipitous, though the rationale for moving in any direction will be meaningful to the individual. This does not negate the need or value for hierarchically structured systems, but rather suggests that, should a user choose to navigate a pre-determined series of logical links he or she would have that option. A pre-coordinated subject heading list would support this choice; using Boolean operators and post-

coordinated subject descriptors would offer a similar option. In short, a user need for flexibility may be readily accommodated.

As to the question of whether subject access principles or core tenets in the new world are proactive or procrustean, the answer is, “both”, or perhaps more cautiously, “it depends”. We have suggested that a design thinking framework may be helpful for rethinking core tenets of KOS, generally, and subject access systems, in particular, insofar as it establishes a deep understanding of – and empathy with – users and user needs as necessary. This mindfulness of users is not anathema to KO professionals. To the contrary, user needs underpin approaches to major subject access systems developed during the late 19th, and continuing into the 20th and 21st centuries. How such needs are assessed, and the extent to which “the user” is actually engaged in the design of subject access systems has changed across time, such that the ultimate authority of the KO professional to assume and address user needs specifically has been called into question. More empowered contemporary users who can devise their own systems to satisfy their own needs are the norm.

SUBJECT ACCESS SYSTEMS IN THE NEW WORLD

With a shift away from finding known or like resources to creating new information and progressive meaning-making, we have suggested the utility of rethinking core tenets, such as warrant, facets, assignment of “aboutness”, and structured hierarchies to foster user-centered subject access system design – that is, principles as proactive. It may also benefit us to consider epistemic shifts in our approaches to assigning meaning, looking to other cultures or long-standing traditions to destabilize and rethink procrustean boundaries. How, for example, might indigenous privileging of relationships over entities, of “family” over “things”, of ancestral roots over tangible possessions alter the design of subject access systems? So-called traditional knowledge may offer a new way; information retrieval that relies on a culture of verbal story-telling and human memory for transition of essential “facts” may offer insights into how to better preserve and communicate the cultural record across time.

In this “new world” of subject access, we might consider moving from a model where the KO professional is “expert” to one where users define and KO professionals facilitate. For example, where a user tags a resource for meaning, we could assist with integrating that natural language, perhaps unstructured term to align with standardized, structured vocabularies where required to facilitate “talking” with other systems, or to conform with requirements for well-formed data within the linked data/semantic web environments. Such alignment mechanisms might take the form of recommender systems, alerting a user that his or her term “equates to” or is “related to” a descriptor favored in a (specified) subject access system. Suggested alternatives would open doors to legacy and other KOS, enhancing the potential for linking to yet more resources for creating new information and meaning-making. While users would benefit, two-way links and the potential for moving beyond “aboutness” by incorporating user-generated tags that help to better capture the “essence” of a resource would likewise enrich “legacy” KOS. In the new world, this represents a win-win scenario!

CONCLUSION

A design thinking framework defines the success of an innovation in terms of its desirability, fiscal viability, and technological feasibility. The current landscape of subject access and KO systems would seem to suggest that engaging in user-centered, innovative design can be

fiscally viable and technologically feasible. KO professionals have a history of user-mindfulness, and core principles that have guided the development of subject access tools and content-rich resource repositories. We have the way, if we determine we have the will.

References

Beghtol, Clare. 2003. Classification for Information Retrieval and Classification for Knowledge Discovery: Relationships between 'Professional' and 'Naïve' Classifications. *Knowledge Organization* 30: 64-73.

Beghtol, Clare. 1986. Semantic Validity: Concepts of Warrant in Bibliographic Classification Systems. *Library Resources & Technical Services*, 30(2): 109-125.

Cutter, Charles A. 1904. *Rules for a Dictionary Catalog*. 4th ed. Washington, D.C.: Government Printing Office.

Design Thinking for Libraries. <http://designthinkingforlibraries.com/>

LibraryThing 2016. <https://www.librarything.com/> Accessed 15 May, 2016.

Olson, Hope A. 2002. *The Power to Name: Locating the Limits of Subject Representation in Libraries*. Dordrecht: Kluwer.

Olson, Hope A., and Howarth, Lynne C. 2013. A Space of Transition: Rethinking Surrogates. In *Transition Cultures, Transition KO: Evolving Exploration, Critical Reflection, and Practical Work: Proceedings of the 4th North American Symposium on Knowledge Organization (NASKO 2013), June 13-14, 2013, Milwaukee, WI, USA*.

Shirky, Clay. 2005. *Ontology is Overrated: Categories, Links, and Tags*. http://www.shirky.com/writings/ontology_outrated.html Accessed 15 May, 2016.