David M. Pimentel
Syracuse University, Syracuse, New York, USA

Exploring classification as conversation

Abstract: Conversations are proposed as a useful lens through which to consider knowledge-organizing behaviors. Human conversations are sites of knowledge creation, where participants communicate to establish meaning that is contextual and shared. The conversations generated in collaborative online environments offer new opportunities to observe, not only how knowledge is created, but also how users participate in various knowledge-organizing activities. In a Web environment pervaded by conversational forms – social classification systems, blogs, and wikis – participatory knowledge organization is an emerging phenomenon that warrants further exploration. Other areas for research are suggested, including the potential promise to leverage participatory knowledge organization into future applications and developments of Web functionality.

1. Conversations and knowledge creation

Investigations into the theory and practice of knowledge organization must ultimately engage with questions surrounding the nature and creation of knowledge.[1] The discipline of information science has largely moved away from "information-as-brick" models, where messages can be transmitted from sender to receiver without loss of meaning. Rather, the nature of knowledge is increasingly viewed as an iterative process, with each individual attempting to make sense of the world s/he encounters (Dervin and Nilan, 1986). Such attempts result in cognitive changes for the individual, creating a contextual, personal meaning.

Various philosophies and theories have endeavored to address the nature of knowledge and knowledge creation. In an analysis of information science metatheories, Talja, Tuominen, and Savolainen (2005) draw distinctions between three broad approaches: constructivism, collectivism, and constructionism. The approaches differ on several epistemological points, but can be distinguished largely based on the role played by language. The constructionist model, in particular, characterizes knowledge as being "produced from limited viewpoints as parts of ongoing conversations" (Talja, Tuominen, and Savolainen, 2005, 90).

Outside the field of information science, and seemingly aligned with the tenets of constructionism, Conversation Theory (Pask, 1975) identifies conversational exchanges as the basis of learning and knowledge construction. While developed to model cognitive processes for machine learning, Conversation Theory also operates at a broader conceptual level: "for Pask, anything that can be sensibly said about 'conversation' is part of [Conversation Theory]" (Scott, 2001, 346). Such broad applicability is likely a result of Pask's background in cybernetics, and Conversation Theory has been described as having the "aim of unifying theories and concepts across disciplines" (Scott, 2001, 346).[2]

At its core, the framework of Conversation Theory centers on participants communicating and seeking a shared agreement, or mutual understanding. Pask's

understandings are specific to the conversation participants, as well as for the given domain and topic (Pask, 1975, 49). As a result "correctness" is relative to the participants and not measured against some external absolute (Pask, 1975, 120). Participants in the conversation can be any cognizing agents: "so, a conversation can be between two people, two organizations, two countries, or even within an individual" (Lankes, et al., 2007, 6). Conversation Theory fundamentally treats knowledge creation "as a process of knowing and coming to know" (Scott, 2001, 348).[3]

Given the nature of knowledge and knowledge creation afforded by Conversation Theory, this paper employs the notion of conversations as a lens through which to consider emerging knowledge-organization phenomena. Motivated by recent examinations of the utility of participatory networks for library environments (Lankes, et al., 2007), this paper seeks to explore how a conversation paradigm might be a valuable addition to knowledge-organization research. This paper will also suggest the potential applicability of Pask's theory in a Web environment that is rich with conversations. As technological advancements make the notion of a smarter, "semantic" Web seemingly inevitable, Pask's ideas about machines engaging in, and learning from, conversations may deserve closer scrutiny.

2. Knowledge organization practices

Methods for organizing knowledge have a considerable history in libraries and other information environments. These methods have long been the purview of a relatively small number of individuals, with trained professionals typical in 20th-century libraries. LIS scholarship has become increasingly attuned to the limitations of these traditional knowledge-organization practices. Svenonius (1992) summarized the criticisms "leveled at the procrustean structures of our great monolithic classifications," noting "their rigidity in the face of change, the limited linearity of their relationships and their difficulty keeping pace with the dynamic and kaleidoscopic world of knowledge" (10). Similarly problematic is the bias embodied in controlled vocabularies such as the *Library of Congress Subject Headings*; the critiques against its representation of marginalized peoples span over 30 years (cf. Berman, 1971; Olson, 2002). It would seem that knowledge organization theory and praxis must find a way to evolve and overcome the shortcomings inherent in a past dominated by imposed deterministic hierarchy.

The treatment of information resources in the legacy LIS paradigm is also problematic. Typically, information resources are assigned a handful of subject representations, in the form of subject headings and classification designations. Despite such subject indexing being performed by professionals, these representations are inevitably limited to one individual's perceptions of an information object at one particular moment in time. Once such indexing has been completed, it is rarely revised or modified. Overall, the process attempts to eliminate semantic ambiguity by integrating the information object into the existing knowledge-organization structures.

Even allowing that such an approach to the organization of information objects was (somehow) optimal, it seems increasingly clear that it is not sustainable. The current explosion in digital information objects already vastly outpaces the ability for expert human catalogers to classify resources at the rate of production. Practical approaches to knowledge organization

must take into account not only the dynamic nature of knowledge, but also the alacrity required to cope with the expanding universe of resources. Knowledge organization practices in general (and perhaps, particularly those in LIS) must evolve from a univocal approach to a multi-conversational one.

Consider the myriad conversations that can be prompted by a great essay or a provocative documentary. Different individuals engage with such works from their own personal contexts; they derive meaning from the information they encounter in the resource, or they relate it to other resources or contexts. As an information object endures (or simply gains wide audience), a multiplicity of contexts centers on it through use by various people. All these contexts coexist simultaneously – and each context was relevant to one or more individuals at some point in time. This potential for an information object to relate to a multiplicity of contexts is not inherent in the document itself, but rather the result of people engaging with and using its content or ideas. In other words, the information becomes part of a conversation. Multi-conversational approaches to knowledge organization would allow for these networks of context and use to organically relate information objects to one another.

Delineating strict lines between *knowledge, context*, and *use* is beyond the scope of this paper; suffice it to say that such constructs have a complex and intricate nature. The point here is that conversation seems an apt metaphor to encompass the multifaceted, inclusive, opportunistic nature of humans engaging with ideas. These ideas may come from the person sitting by your side, or these ideas may by facilitated by an information artifact. In the latter situation, the conversation exists in the cognizing individual's mind: perhaps grappling with the thoughts of a "dead Greek," or perhaps getting inspired by a piece of Afro-Cuban music on her iPod.

3. Conversation features for knowledge organization

Consider the traditional knowledge organization process in terms of a conversation paradigm. While a document may be perceived within the context of prior conversations (the sources it cites, etc.), these are rarely made an explicit part of the knowledge organization system. By extension the object, as it is represented in the system, is also typically divorced from the multiple future conversations of which it might later become a productive part. (Namely: *a priori* enumeration of an item's use is inevitably limited to some notion of "conventional" or "expected" users – so creative and unexpected uses are difficult/impossible to predict.) Incorporation of citation networks into knowledge organization systems may help to articulate these very clear artifacts left behind as evidence of conversations. When the object's representation exposes these conversations, it helps make more transparent how knowledge is created contextually and in turn creates new knowledge.[4]

Formal, academic citations represent a linkage between documents: a statement that a relationship was made connecting the ideas in an earlier source to the current one. In the hyperlinked environment of the Web, Google's PageRank algorithm works on a similarly conversational principle, relying on millions of users creating literal links among web resources. In short order, those same users were also linking and conversing among one another: reading, writing, and commenting on personal blogs. A great deal of conversational exchange occurs on

the blogosphere,[5] and other Web 2.0 phenomena are similarly conversationally oriented. Much recent attention has been given to the Web-based phenomena of collaborative tagging and folksonomies (e.g., Golder and Huberman, 2006). The notice is certainly understandable: these social classifications offer new opportunities for researchers to unobtrusively observe real-world knowledge-organizing activities involving photos, bookmarks, and Web pages.

One area that may deserve closer examination is the knowledge-organizing behaviors that occur in collaborative authoring environments. Sites such as Wikipedia invite users to actively participate in the creation of content, refinement of this content, and resolution of content disputes. Wikipedia makes such conversational exchanges explicit, not only by archiving changes over time (providing access to previous edits, allowing readers to see how the content has changed), but also by maintaining a parallel discussion (or "talk") page for each article.[6] Participants exchange ideas and opinions, engage in discourse, and debate how best to change an article: quite literally carrying on conversations to produce new knowledge.

The conversations generated in such collaborative online environments offer opportunities to observe, not only how knowledge is created, but also how users participate in various knowledge-organizing activities. Some of these different types activities will be discussed shortly, yet it is important to underscore a particular distinction here: almost all the knowledge-producing and knowledge-organizing activities on wikis are subject to collective review. In terms of knowledge-organizing behaviors, this means that uniquely personal indexing terms (such as "toRead" or "me" commonly seen in tagging environments) are not part of the classificatory landscape. Since all users potentially have a voice in how content is organized, collaborative authoring environments offer a perspective on rigorously participatory knowledge organization.

These participatory systems allow users to be directly involved in organizing knowledge, shifting reliance away from institutionalized controlled vocabularies towards more democratically derived terminology. In addition to the collective agreement on terminological matters, Wikipedia has cultivated several other explicit knowledge organizing mechanisms. Some are prosaic (e.g., section headings within articles, per the manual of style[7]) or automatically generated (such as the table of contents within articles). But others knowledge-organizing features of Wikipedia display a relatively high level of sophistication with standard practices: disambiguation pages[8] resolve the problems of synonymy, while "infoboxes"[9] assist users in navigating among related articles. The most fundamental level of knowledge organization activity on the site entails the use of an extensive, collaboratively generated system of categories. [10]

Empirical study would be required to estimate the scope of knowledge-organizing activity taking place on Wikipedia. The effort involved in simply maintaining the current knowledge-organizing structures would seem to suggest that participatory knowledge organization exists on a considerable scale; whether or not Wikipedia proves an enduring feature of the Web, it seems unlikely to be the last online collaborative authoring environment.[11] Understanding the behaviors currently underpinning participatory knowledge organization may help to form theories about distributed, collaborative knowledge organization. Such a framework may benefit from leveraging a conversation paradigm, allowing for holistic

consideration of knowledge-organizing activities that range from social tagging, to participatory knowledge organization, and beyond.

4. Opportunities for knowledge organization

Others have already proposed the conversation paradigm as a useful one for information science, specifically "reorienting our knowledge organization strategies from the description of the contents of documents as relatively stable entities toward mapping and visualizing conversations, perspectives, and debates" (Tuominen, Talja, and Savolainen, 2003, 562). The knowledge organization community has a unique opportunity to focus attention onto conversations instead of solely on information objects. Such a shift requires a fundamental examination of the treatment of information objects: moving from anticipated contexts to actual contexts, from definite order to spontaneous order.

Current knowledge organization practices are based largely on *anticipating* the potential use and context of information resources. If conversations are accepted as the building blocks of knowledge, with its attendant robust context, then *actual* use and context would seem to offer a more valid approach for knowledge organization. In this regard, one can imagine navigating through a digital collection where resources are arranged based not on a single linear dimension, but instead on multiple dimensions that result from behaviors such as linking, annotation, and use. The resulting conversation-oriented environment would be rife with trails and paths.

An analogy to the physical world may prove useful here. Most educational campuses have a central outdoor area that is a focal point, e.g. the quad surrounded by notable or landmark buildings. While this area typically has a number of paved routes, it likely also has one or more shortcut footpaths worn into the grass. The path for users exploring organized information environments is provided by metadata. The LIS tradition has built a great many official (paved) routes, but we must have systems that accommodate the footpaths as well. Digital environments allow us to potentially capture a number of interesting things about those paths: who created it, when, how helpful others found the path for their particular need, etc.

Embracing this type of organic approach may be particularly important when multiple, competing viewpoints focus on the same information resources. Such contested grounds are inherently biased and/or political, and efforts to describe or classify such resources rarely concede multiple knowledge claims as equally valid. By accommodating the myriad conversations that convene on an information object, conversation-based knowledge organization systems could transform an apparent ambiguity into a new kind of clarity. Exposing the nature of the classification – not a single classification, but many classifications – and contextualizing the various conversations would be at the core of such a knowledge-organizing system. Creating richer knowledge organization structures may entail discarding notions of neutrality in favor of a more perspective-explicit framework. By allowing users "to challenge existing perspectives, classifications, and vocabularies" (Tuominen, Talja, and Savolainen, 2003, 564) – in short, allowing them to participate in the practice of knowledge organization – information systems become locations where the tension of ideas is accepted as essential to the nature of knowledge.

5. Possible areas for research

Actual conversation-based knowledge organization systems for LIS require test bed implementations and evaluations. While it may be some time before a large, established research library can experiment with such a system, online environments provide a venue for more immediate investigations. Wikipedia has been offered as a community where members regularly engage in conversations about, and participate in the creation of, knowledge organization systems. Analysis of how such knowledge organization behaviors change over time, as well as how they support each other and/or cross-reference other content in the environment, could inform future implementations of conversation-based systems.

Commercial sites such as Amazon.com extensively employ user comments and rankings to build a robust conversation around books, music, movies, and other objects. The site also now involves users to "help others find" an item by making "search suggestions." By specifying relevant indexing terms that the user believes should retrieve a particular item, "Search Suggestions are intended to improve Amazon search by helping people find relevant items which otherwise wouldn't appear in search results."[12] Such involvement and trust of users stands in stark contrast to the austere nature of the surrogate records in most library OPACs. Analysis of successful commercial enterprises that employ a conversation-based paradigm may inform innovations for LIS system development.

Lastly, while Pask's Conversation Theory served as a starting point for this paper's consideration of a broader conversation paradigm, information retrieval researchers may wish to revisit the theory in a new light. Since Conversation Theory was developed by Pask as a means to facilitate machine learning, his ideas may deserve reexamination by the semantic Web community. In an examination of new technologies that will likely form the foundations of "Web 3.0," John Borland (2007) characterizes a blending of the Web's current "capacity for dynamic user-generated connections" blending together with "a dash of data mining, with computers automatically extracting patterns from the Net's hubbub of conversation" (66). Perhaps the notion of intelligent computers, learning from the conversations that we create on the Web, would be ideally suited to Pask's ideas about machines engaging in, and learning from, conversations.

6. Conclusion: Towards classification as conversation

Conversations have been suggested as a valuable way to consider the creation of knowledge, and by extension the organization of knowledge. Conversations pervade our daily lives, and artifacts of many such conversations now feature prominently in a Web environment of weblog trackbacks, wiki edits, etc. As more individuals contribute to both the implicit and explicit organization of knowledge, participatory knowledge organization behaviors will merit further investigation. Current collaborative authoring environments offer a potentially important perspective on knowledge-organizing practices, with users participating in conversations around emerging classifications. Users are becoming empowered to supplement and/or revise existing knowledge structures, adding their voices to a larger conversation – ultimately shaping how information is organized and discovered. Such collaboratively created classifications need not

necessarily replace existing or established schemes, but instead could connect numerous relevant knowledge-organizing schemes as part of a multi-vocal knowledge-organizing conversation.

Could the sum of knowledge-organizing conversations approximate the conceptual coherence of traditional classification schemes? Could experts and novices both successfully navigate conversation-based knowledge systems? Actual implementations and evaluations will be needed to answer such questions. But if noted technology commentator David Weinberger (2006) is correct that "*better* knowledge is a property of conversations" (17), then perhaps conversations can also be the source of better knowledge organization.

Acknowledgements

Notes

1. For the purposes of this paper, any hard distinctions between *information* and *knowledge* are largely ignored. Knowledge is generally construed as information in context.
2. Pask (1976) applied Conversation Theory in the realm of educational pedagogy.
3. Law provides an illustrative example: the field of law is constructed around a community conversation, decisions and precedents (as well as dissents) form the basis of development and progress.
4. Clearly the idea of citation networks and citation indexing originates with Garfield, but is raised here to be considered as an aspect of the conversation paradigm.
5. As of May 2007, Technorati claims to track 80.3 million blogs – http://technorati.com/about/
6. Parallel discussion/talk pages also exist for users, as well as for non-article content (such as categories).
7. See Wikipedia Manual of Style,
http://en.wikipedia.org/wiki/Wikipedia:Manual_of_Style_(headings)
8. See Wikipedia guideline on disambiguation,
http://en.wikipedia.org/wiki/Wikipedia:Disambiguation
9. See Wikipedia Manual of Style,
http://en.wikipedia.org/wiki/Wikipedia:Manual_of_Style_(infoboxes) ; see also Wikipedia's navigation templates, http://en.wikipedia.org/wiki/Wikipedia:Navigational_templates
10. Wikipedia guideline on categorization, http://en.wikipedia.org/wiki/Wikipedia:Categorization
11. In fact, the new expert-oriented Citizendium wiki launched in March of 2007.
12. Amazon.com. *Search Suggestion FAQ*. Retrieved May 10, 2007, from
http://www.amazon.com/gp/associations/help/faq.html

References

Berman, S. (1971). *Prejudices and antipathies: A tract on the LC subject heads concerning people*. Metuchen, N.J.: Scarecrow.

Borland, J. (2007). A smarter Web. *MIT Technology Review*, 110 (2): 64-71.

Dervin, B. (1977). Useful theory for librarianship: Communication, not information. *Drexel library quarterly, 13* (3): 16-32.

Dervin, B. and Nilan, M. S. (1986). Information needs and uses. *Annual Review of Information Science and Technology, 21*: 3-33.

Golder, S. and Huberman, B. (2006). Usage patterns of collaborative tagging systems. *Journal of Information Science, 32* (2): 198-208.

Lankes, R. D., et al. (2007). *Participatory networks: The library as conversation*. Information Institute of Syracuse. Retrieved April 20, 2007 from http://iis.syr.edu/projects/PNOpen/ParticiaptoryNetworks.pdf

Olson, H. A. (2002). *The power to name: Locating the limits of subject representation in libraries*. Boston: Kluwer.

Pask, G. (1975). *Conversation, cognition and learning: A cybernetic theory and methodology*. New York: Elsevier.

Pask, G. (1976). *Conversation theory: Applications in education and epistemology*. New York: Elsevier.

Scott, B. (2001). Gordon Pask's Conversation Theory: A domain independent constructivist model of human knowing. *Foundations of Science, 6* (4): 343-360.

Svenonius, E. (1992). Classification: Prospects, problems and possibilities. In: *Classification research for knowledge representation and organization*, edited by N.J. Williamson and M. Hudon. Amsterdam: Elsevier Science Publishers.

Talja, S., Tuominen, K., and Savolainen, R. (2005). "Isms" in information science: Constructivism, collectivism and constructionism. *Journal of Documentation, 61* (1): 79-101.

Tuominen, K., Talja, S., and Savolainen, R. (2003). Multiperspective digital libraries: The implications of constructionism for the development of digital libraries. *Journal of the American Society for Information Science and Technology, 54* (6): 561-569.

Weinberger, D. (2006). Conversation and the cult of expertise. *KMWorld,* November/December: 16-17.