Reviews

Edited by Melodie J. Fox

Reviews Editor

Wu, Yejun, ed. 2016 *Oil Spill Impacts: Taxonomic and Ontological Approaches*. Boca Raton: CRC Press. xv, 288 pages. References. Index.

1.0 On taxonomy and taxonomies

A very dense disambiguation page appears in Wikipedia in response to a search for "taxonomy," although one is easily led to a "Taxonomy (General)" entry, where one learns taxonomy is "the practice and science of classification of things or concepts, including the principles that underlie such classification" (https://en.wikipedia.org/wiki/Taxonomy). On that Wikipedia page WordNet is identified as a useful sort of taxonomy that can be used for Wikipedia disambiguation. In WordNet itself (I might rather have called it a lexicon than a taxonomy), there are three entries for "taxonomy," two of which are "a classification of organism into groups" and the practice of so doing (the third is specifically biological)(http://wordnetweb.princeton.edu/perl/webwn?s=taxonomy&sub=Search+WordNet&o2=&o0=1&o8=1&o1=1&o7=&o5=&o9=&o6=&o3=&o4=&h=). Further analysis of available hypernyms leads us to that for "categorization … classification, compartmentalization … assortment," which is "the act of distributing things into classes or categories of the same type," which we in the knowledge organization (KO) community might have thought was called typology. These results are similar to but not exactly as precise as what readers of *Knowledge Organization* think of as knowledge organization the science, the essence of which is the classification of things or concepts.

In KO we often turn first to Dahlberg for essential definitions, especially as they comport with concept theory. In her recent book (2014, 59), Dahlberg places taxonomies squarely in the natural sciences so as to distinguish them from classifications, which are systems of ordering larger areas of knowledge. Earlier, Beghtol (2001, 107) had described a "taxonomic relationship" as one "established by comparing and contrasting instances …" derived from a "usually bottom-up synthetic method in which low-level instances form groups … and reveals the features shared by the instances and the range of variability allowed in the group." Lambe, whose 2007 book is definitively a manual for the creation of taxonomies for knowledge management, points out some key elements with clarity; first, that "taxonomic activity exists in almost every domain of human activity" (4), second that "the term taxonomy means in general *the rules or conventions of order or arrangement*;" third, that taxonomy for knowledge management can be a form of semantic classification, equivalent to the base vocabulary of a thesaurus, which can serve as a knowledge map (5-8). Mazzocchi's (2017) recent ISKO Encyclopedia entry on KOSs places taxonomy among "classifications," as distinct from "lists," and on the lower end of the spectrum in terms of "semantic richness," with ontology occupying the high end.

In *Elements* (2014) I tried to cover similar territory with more precision, writing that "taxonomy is a framework in which elements are defined, and categories are mutually exclusive and collectively exhaustive" (4), and "at the most basic level a taxonomy is an ordered list of terms together with their definitions or other determinant characteristics. Taxonomy is a way of defining the component entities in a domain" (51), with an appropriate nod to Linnaeus, whose taxonomy is the benchmark for scientific representations. The purpose of taxonomy, as reflected in the Wikipedia entry cited above, arises from the sciences wherein the categorization, definition and structuring of observations is a core critical task. This task is one, which Foucault famously pointed out, that involves more than simple naming, but rather "by limiting and filtering the visible, structure enables it to be transcribed into language. It permits the visibility of the animal or plant to pass over in its entirety into the discourse that receives it … and ultimately, perhaps, it may manage to reconstitute itself in visible form by means of words …." (135). Limiting and filtering are critical components of scientific taxonomy because they are the means by which semantic richness is expressed ontologically, which itself is the means by which a theory-based discourse advances and evolves.

My own experience of taxonomy arose from research, specifically that by Tillett (1987, 22 ff.), who created a taxonomy of bibliographic relationships based on a close reading of existing literature, especially cataloging rules, and my own subsequent generation of a taxonomy of instantiation (1992, 28), both of which were followed by Vellucci's adaptation based on analysis of music catalog records (1997, 80). In all three cases, the taxonomies were

derived as products of research, for use as essential classification devices for empirical data analysis, and were not intended otherwise as information retrieval or content management tools.

2.0 *Oil Spill Impacts*

In this sense *Oil Spill Impacts* is a classical, research-based taxonomy, derived from empirical analysis and intended originally as an essential classification device for empirical data analysis. The extension of the attendant taxonomy into an ontological topic map, together with an extensive appendix of information resources, creates a new form of taxonomy that can stand as a model for the KO community.

The taxonomy arose from research conducted in the aftermath of the disaster precipitated by the 2010 Gulf of Mexico *Deepwater Horizon* oil spill, embracing environmental impact on the ecosystems of the Gulf of Mexico, cleanup efforts, and government and inter-agency cooperation. During this effort the need for a focused database of research relating to oil spills in general led to the compilation of a large collection of oil spill literature, and that in turn, required taxonomical and ontological work to deepen understanding about the phenomena in general as well as specific incidents. Semantically the taxonomy is particularly interesting because of the admixture of terms from the popular press with scientific terminology.

According to the contributors, the audience for the work is "students, teachers, and researchers who are interested in oil spill issues [sic] that are related to [chemistry, biological sciences, environmental science, oceanography and coastal sciences, petroleum engineering, geology and geophysics, economics, public health, law, mass communication, sociology, and anthropology]" (xiv). Also "government officials, policy makers, and the general public," as well as "students and researchers in library and information science and knowledge management." In this sense the book is designed for specific research uses and users. However, KO scholars will find in it an exemplary model for a depth taxonomy derived from a precisely-defined domain. In this sense we can consider *Oil Spill Impacts* to cross our own domain analytic taxonomic boundaries as it falls into at least two categories of domain analysis; the volume clearly presents a special classifications, but it also presents elements of a terminological study (Smiraglia 2015, 97).

The book has four distinct parts (chapters) produced by joint collaboration. Chapter 1 is the "Methodology" for the entire project by Yejun Wu who also serves as editor for the volume, Chapter 2 is a narrative of the *Deepwater Horizon* incident by Judith Sylvester, Chapter 3 is the taxonomy proper by David J. Dunaway and Amanda Lehman, and Chapter 4 is the topic map by Wu, Lehman and Dunaway. The progression from chapter to chapter serves very nicely to build the domain description ontologically from narrative to taxonomy to ontological map. The volume is rounded out with a bibliographic appendix on oil spill research, references and an index.

2.1 Defining an oil spill impacts domain

The first two chapters lay out the ontological parameters of the oil spill impacts domain and the discourse surrounding the specific event in the Gulf of Mexico that was the impetus for the underlying research. In the methodological chapter Wu describes the formalization of the event time-line, describes document collection policies and shows how concepts were extracted from source documents, normalized, ordered taxonomically, and then mapped ontologically. The methodology ends with a description of an iterative evaluation process that was employed to allow user input into what became the published text. (A detailed discussion of this process is reported in Wu Lehman and Dunaway 2015). This is followed by Sylvester's detailed description of the *Deepwater Horizon* disaster, opening with a discussion about offshore drilling, oil rig fires and oil spills that helps to situate the particular disaster at the core of what becomes this book's domain. The rest of the chapter contains a detailed time-line of specific events constituting the disaster and its impact. The resource list at the end of this chapter is itself exemplary, documenting an array of blogs and media reports that frame the event culturally as well as technically.

2.2 Oil spill taxonomy

The taxonomy itself occupies 141 pages in which concepts are organized into "12 categories and their subcategories … [which may be] developed "up to four levels" (xv)—limiting and filtering for semantic richness. A classified array serves as a table of contents. For example, section E for impacted wildlife is introduced as follows (31):

E Impacted Wildlife (Specific Species, Plants, and Animals)
- E1 General Wildlife Terms
- E2 Animals
  - E2.1 Birds
  - E2.2 Animals by Life Stage
  - E2.3 Fish and Shellfish
  - …
- E3 Microorganisms (Microbial Organisms)
- …
- E4 Plants.

The taxonomy then appropriately reveals domain-specific details with incremental semantic richness. For example, E2.5 Mollusks appears as follows (102):

E2.5 Mollusks
Mollusks
Polluted clams

E2.5.1 Oysters
Heavily oiled oysters
Oiled oysters
Oysters (*Crassostrea virginica*)
…
Removal of oysters from menus

The taxonomy is extensive and appears without further introduction. An article by Wu and Yang (2015) reports on some details concerning the taxonomic process that are glossed over in the book's introductory sections.

2.3 Ontological topic map

The "Oil Spill Topic Map: Concepts, Relationships, and References" occupies the next 92 pages of the volume. Once again, a classified array serves as a table of contents. In the map the classes are arranged alphabetically by symbol (e.g., A Oil Spill Incidents, B Coastal and Offshore Environments …, C Oil, Dispersants, and Other Chemicals Extent and Fate, etc.). Within each class the 128 concepts are arranged alphabetically with their associated concepts, relationships and references. The format is explained in the book's introduction (xv):

Concept 1
- Relationship 1
  - Associated concept 1
    - Reference 1
- Relationship 2
  - Associated concept 2
  - Associated concept 3
    - Reference 2
  - Associated concept 4
  - Reference 4
    - Associated concept 1
      - See associated concept 1
- Member of
  - Category label (taxonomy level x)

Thus, in the topic map oysters from E2.5 appear thus (224-5):

Oysters
- Be contaminated by
  - Spilled oil
    - (Norris, 2010b)
- Be discolored by
  - Ingested oil
    - (Norris, 2010b)
- May be impacted by
  - Sand berms
    - (Keith, 2010)
- Member of
  - Oysters (Taxonomy E2.5.1)

The topic map then points outward in two ways, serving not only to order the conceptual material ontologically, but also to connect terms directly with the nomenclature in the taxonomy and with the information resources. Another way to look at this order is to think of it as a two-dimensional display of hyperlinked functional data—taxonomy, ontology, bibliography—all interlinked as in a network.

3.0 The role of taxonomy in KO

Wu and his colleagues have compiled a very effective scientific research tool in monographic form that can serve as a standard academic reference book, a desk manual for oil spill workers ranging from scientists to journalists and beyond, the content of which supports an expansible data repository maintained by the collaborators. They also have extended the model of exhaustive domain-centered bibliography by linking their carefully curated collection of oil spill impact resources to both the taxonomy and the ontological topic map. Similarly, the taxonomy and topic map, by linking to the bibliography, provide a multi-dimensional set of tools for knowledge discovery. The work is remarkable for all of these reasons.

But the work is remarkable also for the way in which it models for the KO community an approach to both domain analysis and the construction of conceptual KOSs. KO researchers should analyze this text because of its depth of semantic richness and its example of the scientific function of what we in KO deliver—the filtering and limiting of concepts at a purely ontological level so they may be reordered in useful and diverse ways for knowledge representation and discovery.

References

Beghtol, Clare. 2001. "Relationships in Classificatory Structure and Meaning." In *Relationships in the Organization of Knowledge*, ed. Carol A. Bean and Rebecca Green. Dordrecht: Kluwer, 99-114.

Dahlberg, Ingetraut. 2014. *Wissensorganisation: Entwicklung, Aufgabe, Anwendung, Zukunft*; herausgegeben von der

Deutschen Sektion der Internationalen Gesellschaft für Wissensorganisation e.V. (ISKO). Textbooks for Knowledge Organization v. 3. Würzburg: Ergon Verlag.

Foucault, Michel. 1994. *The Order of Things: An Archaeology of the Human Sciences: A Translation of* Les Mote et les choses. New York: Vintage Books.

Lambe, Patrick. 2007. *Organising Knowledge: Taxonomies, Knowledge and Organisational Effectiveness.* Chandos Knowledge Management Series. Oxford: Chandos.

Linnaeus, Carolus. 1760. *Systema naturae* . Halle: Curt.

Mazzocchi, Fulvio. 2017. "Knowledge Organization Systems (KOS)." In *ISKO Encyclopedia of Knowledge Organization*, ed. Birger Hjørland. http://www.isko.org/cyclo/kos

Smiraglia, Richard P. 1992. Authority control and the extent of derivative bibliographic relationships. Ph.D. dissertation. University of Chicago.

Smiraglia, Richard P. 2014. *The Elements of Knowledge Organization* by Richard Smiraglia. Cham: Springer.

Smiraglia, Richard P. 2015. *Domain Analysis for Knowledge Organization: Tools for Ontology Extraction.* Chandos Information Professional Series. Oxford: Elsevier/Chandos.

Tillett, Barbara Ann Barnett. 1987. Bibliographic Relationships: Toward a Conceptual Structure of Bibliographic Information Used in Cataloging. PhD dissertation. University of California, Los Angeles.

Wu, Yejun, Lehman, Amanda, Dunaway, David J. 2015. Evaluations of a Large Topic Map as a Knowledge Organization Tool for Supporting Self-Regulated Learning. Knowledge Organization. 42(6), 386-398

Wu, Yejun and Yang, Li. 2015. Construction and Evaluation of an Oil Spill Semantic Relation Taxonomy for Supporting Knowledge Discovery. Knowledge Organization. 42(4), 222-231.

Diamond, David. 2016. *Metadata for Content Management: Designing Taxonomy, Metadata, Policy and Workflow to Make Digital Content Systems Better for Users*. CreateSpace ed v1.1. [Charleston, S.C.]: ContentManagementBook.com. 183 pages.

Stuart, David. 2016. *Practical Ontologies for Information Professionals*. Chicago: Neal-Schuman.vii, 184 pages. References. Index.

1.0 Taxonomy redux

A recent influx of interest in information architecture and its more or less casual intersection with librarianship as well as knowledge management seems to be fueling an uptick in the production of manuals for creators of taxonomies, ontologies and other formal sources of knowledge organizing metadata. Whereas the knowledge organization (KO) community has relatively settled definitions of kinds of knowledge organization and knowledge organization systems (KOSs)(Smiraglia 2017x), this new bevy of manuals takes a much more casual approach. Of course, one might take the 2007 work by Lambe as a sort of benchmark in which an effort is made to clarify key aspects of taxonomy creation in light of their importation from KO into knowledge management. The more recent texts range from deliberately casual (Hedden 2016) in which taxonomy is equivalent with KOS, requires some internal structure, and incorporates controlled vocabulary, thesauri, classifications, ontologies and overlaps metadata. Rosenfeld Morville and Arango (2015), perhaps the core textbook used in information architecture, makes no direct mention of taxonomy but does collocate as basic tools metadata, controlled vocabulary, synonym rings, authority files, classification schemes, thesauri and faceted navigation.

The two books reviewed here came on the market at about the same time in 2016. One book (Diamond) is clearly designed for content management, a term generally applied in the digital environment for activity surrounding the coordination and control of descriptive and subject content. The term is a clear parallel to what was called variously bibliographic control or cataloging and classification in the library environment, and it obviously overlaps KO to a large extent, particularly with regard to the coordination of conceptual representation. The other volume (Stuart) is designed specifically for use by information professionals. Interestingly, the two books are almost exactly the same size.

2.0 *Metadata for Content Management*

Diamond's book is 183 pages of focused instruction and advice about tools for content management. Diamond is an experienced expert from the digital asset management industry. His approach is deliberately pragmatic and his advice is concise. Fourteen unnumbered chapters form the core of the book, beginning with a discussion of what is wrong with content management systems and ending with some future gazing. In between are chapters about "Taxonomy Design from Scratch," "Metadata Field Considerations," "Define Some [Content Management] Policy," "The Value of Synonyms," "Localization Considerations," "Workflow Metadata," and user-based applied research "Metadata and Taxonomy Reviews." The specifics are consistent with KO practice. For example. a first step is "determine your terms," which can be done ("derived") by analyzing transcripts of or listening to conversations about manageable content. Scope should be defined, hierarchy should be coordinated and vocabu-

lary controlled. Metadata fields should be reverse engineered from user scenarios, and once defined, fields should be coordinated in schema that distinguish between controlled vocabulary, formatted data and natural text. Localization and workflow discussions cover actual problems of managing specific content, always with regard for user expectations. Examples are helpful and instructions are clear. Exercises for practice are good ways for readers to become engaged in the work of content management.

3.0 *Practical Ontologies for Information Professionals*

Stuart's book is 184 pages of ontology-based description and instruction for creating and managing knowledge organization systems in information professional contexts, with the specific aim of reinforcing the importance of ontologies for knowledge discovery (24). Stuart is identified on Amazon.com as an independent information professional and frequent author. This book has more of the apparatus associated with a standard academic text; chapters are numbered, references are gathered at the rear in a bibliography, and an index is present. Chapters are "What is an ontology," "Ontologies and the semantic web," "Existing ontologies," "Adopting ontologies," "Building ontologies," "Interrogating ontologies," and "The future of ontologies and information professionals." Illustrations range from the British National Bibliography to FRBR entities and relationships to a word cloud from BARTOC. The bibliography includes citations to works familiar to the KO community as well as to the information architecture community thus forming a bridge of sorts. Linked data for bringing libraries, archives, museums and other cultural heritage institutions into the semantic web community are surveyed, as are existing tools (RDF, XML, etc.), ontologies (OWL, OWL2, SKOS, Dublin Core, etc.), and upper ontologies and data models, particularly for cultural heritage implementation (BFO, Europeana, CIDOC-CRM, DBpedia, FOAF, etc.). The writing is clear and concise. The emphasis is introductory rather than instructional.

4.0 Conclusion

Both of the books reviewed here are concise and informative introductions to core problems in the organization of knowledge for knowledge discovery and content management. The Diamond book is clearly oriented as a how-to manual and could be used for class exercises in taxonomy and metadata development at a very basic level. The Stuart book is a good introduction to ontology and ontologies in the cultural heritage information community and could be a good survey text for a segment of an introductory course in knowledge organization. Terminology in Stuart's book is more consistent with standard KO terminology. Diamond's book is self-consciously non-academic in structure and approach.

One reason to review works like these in this journal is to keep the KO community abreast of the slow mulching of core concepts and tools into the consciousness of the semantic web and digital content community. Many of these works take deliberate pot-shots at academics in our discipline with off-hand references to "library science." It would behoove them to understand that most KO scholarship takes place in information schools and finds application in academic knowledge discover and cultural heritage information management and data curation. Another reason to review works like these is to understand the softening of definition of concepts such as taxonomy, metadata and ontology—the three that form the majority of the content of these two books. One hopes our new ISKO Encyclopedia (http://www.isko.org/cyclo/) will contribute sufficient clarity that our core definitions are harvested by search engines and eventually written into Wikipedia.

References

Hedden, Heather. 2016.*The Accidental Taxonomist*, 2d ed. Medford, N.J.: Information Today.

Lambe, Patrick. 2007. *Organising Knowledge: Taxonomies, Knowledge and Organisational Effectiveness.* Chandos Knowledge Management Series. Oxford: Chandos.

Rosenfeld, Louis, Peter Morville and Jorge Arango. 2015. *Information Architecture: For the Web and Beyond*, 4th ed. Sebastapol, CA: O'Reilly.

Smiraglia, Richard P. 2017x. "Review of *Oil Spill Impacts: Taxonomic and Ontological Approaches*, ed. Yejun Wu. *Knowledge Organization* 44: 677-80.

Richard P. Smiraglia
University of Wisconsin-Milwaukee, Wisconsin, USA
smiragli@uwm.edu