Universal Models: A History of the Organization of Knowledge

Hope A. Olson, University of Alberta, Canada

Abstract: Feminist poststructural analysis of the history of knowledge organization models and their ramifications for non-traditional knowledge reveals an assumption that universal models are necessary and desirable for information retrieval. Awareness of this assumption is raised by addressing: 1) knowledge organization as naming; 2) historical manifestations of the assumption of universality in knowledge organization; 3) examples of the limits of universality drawn from feminist literature; and 4) a theoretic for opening universal models.

1. The organization of knowledge as a process of naming

Naming nature is the special business of science. Theories, models, and descriptions are elaborated names. In these acts of naming, the scientist simultaneously constructs and contains nature. (Keller, 17)

We name knowledge with our theories, models and descriptions. Our models are structures to organize knowledge, our theories philosophical underpinnings for our models, our descriptions representations of knowledge gathered into catalogues and indexes. A poststructural perspective presupposes that realities are constructed by discourses which are the contexts of individuals' knowledges. By naming knowledge we put it into our contexts/models. The knowledge contained in these models is constructed by them just as gelatin is shaped in a mould. Thus, our organization of knowledge is an act of power. The central assumption of this paper is that naming knowledge is a discourse constructing knowledge. The methodology which I use to is an adaptation of poststructural deconstruction, an effective means of revealing assumptions. S.R. Ranganathan cited Oliver Goldsmith's "warning about the danger of one good custom corrupting humanity." (527) An unrecognized and unquestioned assumption can pose a danger to the naming/construction of knowledge.

2. A textual history of the assumption of universality

Texts by Charles Cutter, Melvil Dewey, S.R. Ranganathan, Cyril Cleverdon and standards for thesaurus construction illustrate the assumption that universal models are desirable and necessary. The texts exemplify differing approaches to knowledge organization: Cutter focuses on "the convenience of the public," Dewey on economy, Ranganathan on logic and Cleverdon on empirical evidence. The thesaurus standards are sanctioned statements regarding knowledge organization.

Detection of this assumption in the cognitive approach of B.C. Brookes and in feminist thesauri points toward its ubiquitous presence. Expression of the assumption of universality has two parts: first, a mistrust of differences in language, and, second, a need for universal models to overcome these differences. The two parts are, in deconstructive terms, a binary opposition with universality and difference being apparent opposites, each defined by not being the other. In knowledge organization, universality is the dominant concept.

Charles Cutter, in Rules for a Dictionary Catalogue, identifies the public’s linguistic differences: “As it is often the case in language usage will be found not to follow any uniform course.” (44) And the need for consistency:

The importance of deciding aright where any given subject shall be entered is in inverse proportion to the difficulty of decision. If there is no obvious principle to guide the cataloguer, it is plain there will be no reason why the public should expect to find the entry under one heading rather than another, and therefore in regard to the public it matters not which is chosen. But it is better that such decisions should be made to conform when possible to some general system, as there is then more likelihood that they will be decided alike by different cataloguers, and that a usage will grow up which the public will finally learn and profit by; as a usage has grown up in regard to the author-entry of French names containing De, Du, La, etc. (37)

Cutter proposes the universal object of collocation: “2. To show what the library has (D) by a given author, (E) on a given subject, (F) in a given kind of literature” (10). To fulfill this object, Cutter introduces hierarchy, dividing collocation into three levels: “bringing books together which treat of the same subject specifically ... bringing subjects together so as to form a class ... bringing classes together so as to form a system.” (12) Manifested in classification and in the syndetic structure of controlled vocabulary, it is the infrastructure of universal models.

Melvil Dewey, in the introduction to the first edition of his classification takes a different route to the same conclusion. Dewey suggests his universal model so that the “usefulness” of libraries “might be greatly increased without additional expenditure.” (3) “In all the work, philosophical theory and accuracy have been made to yield to practical usefulness.” (4) However, he accepts the principle of collocation as much as Cutter does:

Thus all the books on any given subject are found standing together, and no additions or changes ever separate them. Not only are all the books on the subject sought, found together, but the most nearly allied subjects precede and follow, they in turn being preceded and followed by other allied subjects as far as possible. (7)

Both Cutter (41) and Dewey (6) reinforce that the “real” subject of the book, not “the accidental wording of the title” (Dewey, 6), must be determined to put it in the
correct place. That is, authors’ words (as well as users’) need to be standardized by a universal model.

S.R. Ranganathan, in his *Prolegomena to Library Classification*, states more directly than Cutter and Dewey that natural language use differs too much for effective communication and needs to be standardized.

We humans pride ourselves on our possession of articulate speech - that is, language - denied to other creatures. Yet many of our ills - social, legal, economic, political, and even domestic - are traceable to the imperfection and vagueness of language. ... Vagueness of ordinary words is one of the reasons why a Glossary of Standard Technical Terms, free from vagueness, must be established for each subject-field. (201-202)

A natural language has necessarily to grow at the pace of the common man, if not of the very last man. But the growth of ideas is in the minds of persons near the top of the intellectual scale. Therefore, when compared with the growth of ideas, the growth of natural language is sluggish. ... As a result of this sluggishness, one and the same word is often used to denote two or more ideas. Homonyms get created in this way. They often grow like weeds. The same word is also used in different senses in different disciplines. This retards the rate of thinking. It also causes aberration in communication. ... Reciprocally, there are also synonyms getting created. ... The kick-back of synonyms is sometimes even more subtle and more harmful than even the presence of homonyms. (329-330)

Fear of weeds in the garden of natural language follows from Cutter’s need to choose one name for a given subject and leads to mistrust of natural language as a vehicle for information retrieval. To resolve the problem of “vagueness,” Ranganathan follows the assumption of universality in his “Law 2: Every reader his book” and its resultant: “21 All books in one place: When a reader seeks information in a given subject, the arrangement of the books in the library will be helpful to him, only if all the books on that subject are to be found together.” (119) This consistency, like that of Cutter and Dewey, is hierarchically arranged in a genealogical tree, in “Filiatory Sequence” (51). The sequence of facets in Colon or any other faceted classification creates a hierarchy which asserts the values of some characteristics over others.

In “Part Q Classification as Transformation” Ranganathan notes the distortion of mapping a “multi-dimensional universe” of subjects onto the linear shelves of a library and the number of mappings possible (382) and that the arrangement putting related subjects near each other (Apupa; see also Dewey above) is less helpful to a minority of readers (384). His “universally applicable solutions” to this problem (540) presage the use of universal devices or structures to overcome individual differences in the cognitive approach.
Standards for thesaurus construction, for creating universal models, are formalized manifestations of the assumption of universality:

...its purpose [is] that of converting natural language, such as that used by authors of documents or posers of reference questions, into a language devised to expedite the matching process ... which is the essence of the information retrieval process. ... the users of any system are a changing group of persons whose respective interests vary widely, ... each individual may have different retrieval interests at different times, ... the content of the material and the ways in which messages are expressed also change with time ... (ANSI Z39.19, 9-10)

Four principal purposes are served by a thesaurus: (1) to control the terms used in indexing, providing a means for translating the natural language of authors, indexers and enquirers into the more constrained language used for indexing and retrieval (BS 5723,1)

These standards echo the theoretical texts' unquestioning acceptance of universality.

Cyril Cleverdon's Cranfield experiments further operationalize the assumption of universality. Cranfield I, centres on the need for consistent terminology in indexing and searching.

The crux of the problem of information retrieval is, however, the need to ensure that the description of a subject by the searcher, in terms used by the system, coincides exactly, or more or less exactly with that assigned by the indexer. If such coincidence is difficult or impossible, then the efficiency of the system will be impaired accordingly. (1960, 11)

The provision of rules to meet every case would determine unambiguously what the form of a heading should be for both indexer and searcher, and this would ensure that coincidence between the two which is the essence of successful information retrieval. (1960, 51)

It appears to be a reasonable assumption that in an ideal system the same document would have the same notation or heading whoever may do the indexing. The nearer this consistency is approached in the index, the greater the chance that the searcher will locate the required information in the expected place. (1960, 111)

Universal Decimal Classification displays inconsistency in “the many cases in which several different numbers are available for placing a given concept” (1960, 36). With faceted classification the problem is to determine the order of facets “because it is evident that no single order can cater for every requirement” (1960, 66). Subject experts demonstrate individual differences in the questions for relevance testing. “There were cases where the same document had been selected by more than one person, with the result that there were two or more questions which were
based on the same document. It was interesting to note the variety of questions which a single document could generate” (1962, 10).

Subjective relevance decisions in Cranfield II also show differences: “Individual relevance assessments, done by 182 people, and with no personal interaction with the project staff, cannot be entirely consistent.” (1966, 1:29) Differences in search questions fail to achieve the coincidence sought: “The terms which are used to express a request or a search prescription rarely coincide exactly with the terms used to describe a particular document” (1966, 1:40). Thus test results indicating that the least structured, most natural indexing language was the most successful prompted considerable surprise.

Quite the most astonishing and seemingly inexplicable conclusion that arises from the project is that the single term index languages are superior to any other type. ... This conclusion is so controversial and so unexpected that it is bound to throw considerable doubt on the methods which have been used to obtain these results, and our own first reaction was to doubt the evidence. (1966, 2:252) [emphasis added]

The „natural language, with the slight modifications of confounding synonyms and word forms, combined with simple coordination“ performs best(1966, 2:263b).

A century after Cutter and Dewey, but only a few years after Ranganathan and Cleverdon, the new cognitive approach to information retrieval emerged, recognizing that individual people have their own individual knowledge structures. However, it does not abandon the assumption of universality. Rather, in his „The Foundations of Information Science,“ B.C. Brookes, while accepting the differences of individual knowledge structures and acknowledging that in his formula for information science $K[S] + AI = K[S + \Delta S]$ „the same $\Delta I$ may have different effects on different knowledge structures,“ (1980, 131) still calls for universal structures.

Whatever we may claim, each of us takes a subjective view of those aspects of the world and its affairs which bear on our interests. As each subjective view is liable to differ from any other, amicable social life requires us to reduce the risk of disagreement by establishing socially agreed standards, rules, regulations ... for the conduct of social affairs. Subjective assertions about any issue are critically compared and modified by those concerned until some socially agreed consensus is established. This continuing social effort to correct the errors [emphasis added] that our subjective responses lead us into is what I call objectivization. (1980, 209)

Brookes sees necessary and desirable universality as derived from consensus, rather than imposed from without or above. But the result is much the same.

All human minds too are unique. Like human bodies they grow, it seems, from almost nothing and die when the owner’s body does. But like bodies
they must share some common characteristics as otherwise communication amongst us would be impossible. This is an unchallenged assumption [emphasis added] of all human social life. (1980, 269)

This consensual universality is assumed to be essential for communication, but unlike Ranganathan, Brookes does not note the minority which will be excluded from any structure.

The assumption of universality is so entrenched that it is even found in the efforts of feminists to facilitate subject access to women-centred materials. The introduction to A Women’s Thesaurus, edited by Mary Ellen Capek, explicitly recognizes that „Language is one of the most intimate and most political of human activities. The power of naming in fact shapes and defines the institutions that structure so much of our lives. Attempts to standardize the naming process are intrinsically controversial“ (xvi). However, the attempt is made, acknowledging bias. „Our choices of terms to include and exclude, to cite as preferred or ‘do not use,’ have been shaped by our value systems and those of our organizations. ... we have few illusions about objectivity.“ (xvii) The compilers claim that „Standardization does not have to mean quashing diversity when drawn from dynamic consensus. Standards can „themselves call into question assumed norms.“ (xvii) Whether or not this approach is possible and whether or not it is achieved by this Thesaurus, its compilers at least recognize their assumption and clearly state its risks.

3. The limits of universality

Exploring the limits of an assumption involves exploring its application. From a poststructural perspective, any concept is defined by what it is not; any class or structure includes what it does not exclude. The distinctions between concepts, and the inclusions and exclusions of classes and structures are constructed by their historical and contemporary contexts. To explore the application of the assumption of universality in naming knowledge I have chosen examples from feminist literature based on the idea that difficult cases will reveal problems which occur more subtly in other instances. Feminist literature is problematic for knowledge representation being interdisciplinary, non-traditional (or anti-traditional), and dynamic in terms of language and content. The two approaches I have selected are: a historical glance at the contextual representation of women’s position in society in the Dewey Decimal Classification (DDC); and a problematization of the representation of a specific, identifiable theme.

The representation of women’s position in society in DDC illustrates diachronic differences. Most striking to a late twentieth-century sensitivity is 396 in DDC 8 (1913) in the hierarchical context of Sociology. Customs Popular life. Woman’s position and treatment; and situated between Etiquet, which trivializes, and Gipsies Nomads Outcast races, which groups women (perhaps appropriately) with blatantly marginalized groups.
By DDC 17 (1965) 302.412 Woman is in the hierarchical context of Social sciences. Sociology. Institutions and groups. The sexes. Here, Woman is situated between Man and Celibacy, followed by Courtship. The hierarchical context puts Man and Woman at the same level implying a certain equity. However, the classification of most things to do with women just before Celibacy and Courtship puts women’s activities in the context of sexual relations. Woman has a variety of subdivisions in DDC 17 offering specific categories such as women’s careers. Men’s careers are not classified in 302.411, but with the appropriate occupation.

In DDC 19 (1979) there is a general change shifting this topic to the very similar hierarchical context: 305.4 Social sciences. Sociology. Social stratification (Social structure). Women and to a place between Men and Social classes. Removing of Celibacy and Courtship from this section removes the sexual focus. In DDC 20 (1989) a change is made both to the subdivision practice and to the hierarchical context. In a feint toward androgyny, the same subdivisions are assigned to both Men and Women, implying that the same categories are equally useful for both. Since Men comes first in the schedule, the result, though not visible to the public, is that Women is divided like Men under which the subdivisions are listed. In addition, 305.3 (previously Men) is changed to a general Men and women recalling the use of the masculine as the norm and women as separate. The message is not obvious to the public, but for classifiers it ratifies prevalent biases.

The changing placement of Women in DDC points to the changing contextual construction of knowledge. Both Dewey and Ranganathan explicitly indicate that one of the purposes of their models is to place closely related topics next to each other. Placing Women between Etiquet and Gipsies Nomads Outcast races makes a clear statement. Consistently placing Women after Men in 301.412 and 305.4 and in the tables for subdivision reinforces a hierarchical division of the sexes in the name of consistency.

Application of a universal model is also illuminated by the delineable feminist literature prompted by Carol Gilligan’s research suggesting that women generally solve moral problems with an ethic of care, instead of the ethic of justice suggested by research based on male subjects. The ethic of care/ethic of justice dichotomy generated a literature linked to a dichotomy of women’s relational orientation and men’s goal orientation. It is also linked to standpoint feminism which posits that women, as a marginalized group, must understand both our own marginal context and the mainstream culture. We cannot afford to be linear and our supporting role fosters an ethic of care. Originating in developmental psychology, this literature also includes social welfare, philosophy, literary criticism, home economics, education, and nursing. Titles, tables of contents and back-of-the-book indexes in this area include the concepts: ethics, moral judgment, values, worldview, caring, justice, feminine qualities, masculine qualities, feminist perspectives, interpersonal relations, goal orientation, relational orientation, separateness, inde-
pendence, individuality, connectedness, and interdependence. Library catalogues on the Internet yield subject headings such as Ethics, Women—Psychology, Women—Social conditions, Feminism, Caring, Self (Philosophy), Femininity (Psychology), Sex differences (Psychology), Moral development, Moral education, Developmental psychology, etc. There are considerable differences among of subject headings in any given catalogue (recalling indexer consistency research). The Women’s Thesaurus and The Canadian Feminist Thesaurus both include the descriptor ethic of care, but not the ethic of justice. Other descriptors in these thesauri are as diverse as the subject headings. None of these controlled vocabularies has succeeded in imposing a universal model able to collocate the materials in this literature. The subject headings demonstrate interdisciplinarity, referring to philosophy, psychology, education and social aspects. Records from the Library of Congress reveal DDC and Library of Congress Classification (LCC) numbers in sociology, education and ethics. Subject headings and classification following disciplinary boundaries fail to collocate.

Representing a binary opposition is difficult because the two classes are defined by not being each other. The binary opposition is inseparable; two sides of the same coin. This topic looks at culturally feminine and masculine views of the world, theoretically divided between two classes; but because a masculine perspective is the norm the feminine is the different part of the topic. Thus, a discussion of differences between the sexes is classed with women because it will be perceived as what makes women different from the norm.

4. A theoretic for opening the universal model

Defining and questioning the assumption of universality does not condemn existing models in knowledge organization. It raises the dilemma of representing difference without relinquishing the successes of existing models. Deconstruction suggests that as every concept is defined by what it is not and every structure is defined by what it includes and excludes, there can be no universal model. Any model excludes something/someone. Drucilla Cornell in her Philosophy of the Limit advocates an ethical relation between a system and what it excludes, the Other, by breaching the limits of the system to allow space for the voice of the Other to speak. The limit between the inside and outside must be made permeable by employing specific strategies (not grand designs). This endeavour to create an ethical relationship with the Other is not teleological. Rather it is a perpetual effort to ameliorate existing exclusivities.

The Other in our universal models is most likely to be users (or potential users) or authors. Strategies for breaching the limit may develop from our existing literature. The cognitive approach offers possible strategies for implementing user warrant taking advantage of technologies such as hypertext, dynamic interfaces and transaction logs. Voices of authors may be heard through citation analysis em-
ployed for self-classification, possibly collocating by citation clusters with authors creating their own contexts. Computerized manipulation of terms offers ways of enhancing the voices of marginalized authors. Finally, the large literature on the advantages of free text and controlled vocabulary retrieval merits close scrutiny in studying the limits of our models.

Ranganathan describes classification as a mould which is filled up with recorded ideas. "It may happen that the mould is unsuited and cracks or even bursts. Then we have to pick out a more suitable mould." (569) The philosophy of the limit suggests that instead of replacing the mould it be made permeable. This image of oozing gelatin might be replaced by a house, with additions for new family members, a garden for beauty and food, and windows for fresh air and distant views. A variety of strategies for accommodating difference while maintaining and enhancing the habitability of the home.

References


