New Course Design: Classification Schemes and Information Architecture

by Bella Hass Weinberg

Bella Hass Weinberg is a professor in the Division of Library and Information Science, St. John's University, Jamaica, NY 11439. She can be reached by phone at 718-990-1456 and by fax at 718-990-2071. More information on the program of the Division is available at www.stjohns.edu

As Andrew Dillon's columns for the Bulletin of the American Society for Information Science and Technology demonstrate, there are many definitions of information architecture. In consulting on a library automation system, I even heard a definition of information architecture that was coextensive with the field of systems analysis.

Seth Maislin points out correctly in Key Words: Bulletin of the American Society of Indexers (January/February 2002) that there are two sides to physical architecture: the functional and the aesthetic. A building may be efficient to walk through, but it may not be beautiful, or the converse may be true. The same two aspects apply to information architecture: A website may be attractive, but it may be difficult to find desired information in it.

The course described in this article emphasizes the functional aspect of information architecture, that is, the way content is structured rather than its surface presentation, in accordance with the approach of Rosenfeld and Morville's book, which serves as the textbook for the course.

The thesis of the instructor is that traditional concepts of classification theory are highly relevant to information architecture. Had it been in print, Sayers' Manual of Classification would have been the second recommended text for the course. Theoretical concepts were presented from the work of Ranganathan, but clarity was a major criterion for the selected readings, and Ranganathan's Prolegomena failed that test.

Context of the Course

Before going into further detail about the course, it is necessary to explain its relationship to other graduate courses in the program of the Division of Library and Information Science at St. John's University. Each student is required to take a core course, Organization of Information: An Introduction (formerly Introduction to Cataloging/Bibliographic Control). An elective course on indexing and abstracting has been available for decades. Several years ago, I designed a course called Language Structure and Thesaurus Development. Basic thesaurus structure is still introduced in the indexing course, but the variety of thesaurus structures is the focus of the newer course. In addition, automatic indexing and abstracting, as well as ranking algorithms, are studied in the thesaurus course.

Other elective courses on the organization of information are taught by Dr. Sherry Vellucci, author of the metadata chapter in ARIST (vol. 33, 1998). She teaches Advanced Cataloging, Technological Issues in Bibliographic Control, and Metadata for Information Professionals. During the Spring 2002 semester a Special Topics course entitled Building a Library Web Site from the Ground Up was taught by Brian Mikesell; that course included a unit on information architecture.

While many of the aforementioned courses deal with concepts of classification and hierarchy to some extent, none covers the full spectrum of systems and structures. A student in my indexing course once suggested that we offer a full course on classification. I kept this idea in the back of my head and acted on it when information architecture became a hot topic.

In a prior paper in the Bulletin (December/January 1999), I provided references documenting the failure of the paperless society to arrive. Two years later, in the medical library near my home, I observe numerous new textbooks and monographs...
being read by patrons. These printed works require a bibliographic classification scheme for arrangement on the shelves, at the same time that the categories of websites require systematic arrangement. Thus I decided to design a course that would examine these concepts in parallel. The course proposal was undergoing approval at the time that David Robins presented his survey of “Information Architecture in Library and Information Science Curricula” at the 2001 ASIST Annual Meeting; therefore, the St. John’s course is not mentioned in his Bulletin article (December/January 2002). The accompanying box provides an outline of the course units.

Assignments

Early in the course, an exercise on Library of Congress Classification (LCC) notation was given in class. The combination of integers and decimals in the notation of this scheme introduces complexity in the arrangement of call numbers. Other notational systems have alternative arrangements (for example, all single letters before all double letters), but LCC is the scheme that most librarians are likely to use, and therefore extra time was devoted to its notation.

In preparation for their term projects, the students were required to do an exercise on depth of classification by analyzing the levels of hierarchy in a website devoted to a discipline and comparing them with a bibliographic classification scheme. The students found it difficult to find sites that presented a discipline systematically; most represented the society devoted to a discipline.

Students were also required to take a field trip to a library other than that of St. John's University and analyze the classified and form sequences of the collection. Among the aspects to be considered were

- broken order;
- use of multiple classification schemes;
- form classification;
- reader interest classification;
- notation for special locators;
- reader orientation; and
- availability of an online shelflist or classified catalog.

The reports revealed a great many classified sequences, rendering the location of a given work difficult.

The term project was designed to bring together many aspects of the course. Students were required to select a multilevel website with a site map, site index or search engine. The following aspects of the information architecture of the site were to be evaluated:

- The number of primary categories
- The depth of the hierarchy
- The logic of the sequence of the primary categories
- The method of arrangement of the secondary categories
- The clarity of labeling of the categories
- The use of need for polyhierarchy and/or cross references between categories
- The site map or index and/or the search engine. If there is an index or search engine, how does it handle synonyms?
- Features that orient users to their location in the hierarchy
- Ease of navigation of the site
- The hierarchical structure of the site with that of a classification scheme or thesaurus in the same discipline or profession

Screen captures were essential to document the aforementioned points. In a concluding section of the term paper, students were to consider (a) whether an alphabetico-specific index would be preferable to the classified structure and (b) whether a thesaurus operating behind the scenes would facilitate the searching of specific topics on the site.

Student Reactions

At the request of the Bulletin editor, this article was written before the course was completed and before all the assignments were submitted. Without formal evaluations in hand, however, I can report on student reactions to the course.

At the beginning, I expected the students to want minimal attention focused on
traditional library classification, with greater emphasis on hierarchies in the electronic environment. This turned out not to be true, however; the students were very interested in such topics as classification theory, various types of notation, bias in classification, reader interest classification in libraries. One student observed that this was a "meat-and-potatoes" course that may have been unique in the United States. Another said that she realized how little she had formerly known about library classification schemes, and how important it was to learn them.

The most interesting reactions were elicited during analyses of the hierarchical structures of websites and of the arrangement of categories. The most common method employed is alphabetico-classed arrangement. This approach is evident in the Yahoo! directory, for example, with broad headings subarranged by specific headings, and both levels arranged alphabetically. For instance, "Economics" has the subcategories "Agricultural Economics," "Banking," "Calendars," "Commentaries" and so on (http://dir.yahoo.com/Social_Science/Economics/, accessed 3/21/2002).

Although deprecated by John Metcalfe, alphabetico-classed arrangement is commonly employed in printed abstracting services. For example, the first three "Major Concept Headings" in Biological Abstracts (1997) are "Aerospace and Underwater Biological Effects," "Agronomy," "Allergy." The first of these has the subcategory headings "Exobiology," "General Methods," "Physiology and Medicine," "Space Radiation." (In both the electronic and print examples, form subheadings are interfiled with topical ones; librarians often separate these.)

As I pointed out in a review (Journal of Academic Librarianship, May 1999), it is surprising that the 2nd edition of F.W. Lancaster's textbook on indexing and abstracting (1998) does not discuss alphabetico-classed structure, which has become central on the Web. The term is not in the index to Candy Schwartz's Sorting Out the Web either; she calls the subject categories of Web directories "ad hoc alphabetical arrangements" (p. 74).

The students were often wide-eyed as I made sardonic observations on website categories labeled Miscellaneous or Other that are filed alphabetically among more substantive headings. Who would look up such catch-all headings directly? And if such categories are necessary at all, they should come last because one cannot know their scope until one has read all the other category labels. "Gee, I never thought of that before" was a typical reaction of the class. This was certainly a fun course to teach, as the students and I laughed together at the counterintuitive arrangements of some websites. Most rewarding was the student comment that the course was enriching and had the potential to revolutionize the organization of the Web.

Conclusions

The need to study traditional bibliographic classification schemes persists, while drawing on established classification theory to enhance the organization of the Web is essential. Thus I encourage other specialists in the organization of information to apply their knowledge of classification theory to the analysis of hierarchical structures in cyberspace.

Of particular interest are the differences between implementing classification schemes in the physical and electronic environments. The need for notation in the former but not the latter is an obvious difference, but many intriguing questions remain. For example, the literature on website design includes guidelines on the acceptable number of primary categories and warns against having too many deeper levels. Bibliographic classification schemes often have a great many levels of hierarchy, however. Have the guidelines for the electronic environment been validated, or are they like the "truism" that a typeface with serifs is more readable than one without them?

Finally, there is a lot of evidence that people frequently do not know which primary category to click on when they have a specific topic in mind, but most Web designers believe that they have to start with a "taxonomy." Perhaps this assumption should be rejected, and the alphabetico-specific index with a carefully constructed network of cross references to broader, narrower and related terms should be the preferred method of access to information on a website. The website of Consumers Union (www.consumerreports.org) made the switch from a category structure to an alphabetico-specific index with great success, as was reported by Robert Huerster at the 2001 Annual Conference of the American Society of Indexers. A book by Browne and Jermy provides guidance on creating such indexes for the Web.
If memory serves, in teaching organization of information at Columbia University’s School of Library Service in the early 1970s, Jessica Milstead observed, "We don't yet know the best way of doing anything." This statement remains true three decades later in the context of website design. There is, however, no more fascinating subject to explore than the methods of organizing information for optimal access by end-users of the Internet.

Acknowledgments

I thank Dr. Sherry Vellucci, acting director of the Division of Library and Information Science, and Catherine Hughes, a student in the information architecture course, for commenting on a draft of this article.

For Further Reading

**Classification Theory**


**Information Architecture**


**Units of Instruction**

To give a more detailed picture of the content of the course, the outline of the units of instruction follows:

**UNIT I. Introduction. Definition of terms. History of classification**

- Definition of terms: classification, taxonomy, ontology, information architecture
- Naming and classification
- History of classification
- The importance of classification in all disciplines

**UNIT II. Types of library classification schemes: philosophically based vs. literary warrant; enumerative and synthetic**

- Top-down vs. bottom-up approaches to the design of classification schemes
- The handling of compound concepts: enumeration vs. synthesis
- Facets
- Predictive power of classification

**UNIT III. Classification notation**

- "A classification is not its notation"
- Mechanizing the arrangement of hierarchies through notation
- Fixed vs. relative location
- Pure vs. mixed notation
- Notational bases: letters vs. numbers
- Hospitality of notation
- Decimal and centesimal notation
- Mnemonic devices
Use of non-alphanumeric symbols
Expressiveness of notation
Modifying the notation of standard library classification in local adaptations
Cuttering systems
Applying notation to the hierarchical structure of websites

UNIT IV. Classes vs. categories. Polyhierarchy

Strict hierarchies vs. loose categories
Concepts that belong to multiple classes: Polyhierarchy
Hypertext links and polyhierarchy
Shifting classification; revision of standard schemes
Bias in classification; local library revisions

UNIT V. Depth of classification in bibliographic schemes and websites

Coextensivity of subjects and class numbers
Number of levels in a bibliographic classification scheme
Designing the hierarchical levels of a website

UNIT VI. Thesauri and classification

Are classification schemes controlled vocabularies?
Hidden classifications in thesauri
Hierarchical displays of thesauri
Correlation of subject heading lists and library classification schemes: LCC/LCSH;
NLM/MESH; DDC & LCSH
Thesaurfacet

UNIT VII. Folk taxonomies. Reader-interest classification. User research and website design

Scholarly classification vs. folk taxonomies
Alternative placement in standard library classification schemes
Form-based classification
Reader-interest classification in libraries
Designing the hierarchy of websites in light of research on users

UNIT VIII. Concrete vs. discipline-based classification. Library classification schemes and internet organization

The Brown classification
Appropriateness of discipline-based classification schemes for organizing the Internet.

UNIT IX. Alphabetico-classed systems of organization. Analysis of the structure of internet directories

Critiques of alphabetico-classing in the indexing literature
Alphabetico-classed serial abstracting services
Yahoo! and other alphabetico-classed Internet directories

UNIT X. Field trip: Classified and form sequences in a library

UNIT XI. Print display vs. screen display of hierarchies

Resolution in print and online
Continuation headings in print and on-screen
Fish-eye views on-screen
Electronic versions of library classification schemes
Graphic displays of hierarchies

UNIT XII. Interface design issues for information architecture

Labeling of displays
Clarity of icons
The search interface
Browsing hierarchical structures
Usability studies
UNIT XIII. Alphabetico-specific indexes as alternatives to hierarchical structures.
Automatic classification.

The relationship between indexing and classification
Rearranging alphabetical lists into a classified structure
The ambiguity of the term keyword
Split indexes vs. merged indexes
Restructuring the Consumer Reports website
Algorithms for automatic classification
Visualization interfaces and classification

UNIT XIV. The profession of information architecture

Knowledge and skills required for information architecture
The future of information architecture as a profession
Keeping up with the literature on classification and information architecture
Websites on information architecture