

Faceted Subject Access: The Challenge

Sandra K. Roe

Thank you very much for inviting me back to Mexico City to attend this conference. I have fond memories of being here for the XVI Colloquium on Library Science Research in 1998. At that time I spoke of improving subject access – outlining the problems that had been identified and projects that had implemented solutions to those problems. I began with the statement that subject access functionality has yet to reach its full potential in our current online catalogs, and concluded with a call for us to continue to assign controlled vocabularies (both classification and subject headings) and to make their structures available to our catalog users.

In some respects, my topic today is an extension of that one. Some things have remained the same. Controlled vocabularies are still essential for the organization of information. It is still critical that the structures of those systems be robust and be made available to our users in ways that they can understand and utilize. If reaffirmation of this is needed, two works published last year provide it: the new ANSI (American National Standards Institute) / NISO (National Information Standards Organization) Z39.19.2005 *Guidelines for the Construction, Format, and Management of Monolingual Controlled Vocabularies*¹ and “Promoting Research and Best Practices in Subject Reference Structures: A Decade of Work by the Subject Analysis Committee.”²

Other things are different than they were in 1998. For example, we are increasingly surrounded by multiple discovery environments³ of which the library catalog is only one. Some consider

these to be in competition with the online catalog, others look to them for new ideas or new implementations of earlier ideas. The dissertation of Dr. Kathryn La Barre, now at the University of Illinois at Urbana-Champaign, explored the latter. Her research investigates the re-discovery in 2001 of a legacy form of information organization and access -- faceted analytico-synthetic theory -- by practitioners in the fields of information architecture and knowledge management who were charged with designing websites and access to corporate knowledgebases. Her methodology included interviews with individuals who explicitly claimed to use faceted analytico-synthetic theory in website design. She says, "This group has been instrumental in creating new and different ways for people to engage with the digital content of the Web; they have clearly recognized that faceted approaches have the potential to improve access to information on the web."⁴ They have recognized the value faceting can offer for searching and navigation. Some sites that were identified by this group as exemplary included Wine.com⁵, Epicurious⁶, and Barnes and Noble.⁷ I'll explain this theory more fully shortly but for now know that facet analysis permits a detailed expression of all aspects of the subject content of an item. For instance, on wine.com, one can search for a wine by these aspects or facets: price, color (red, white, pink), type of use (dinner, dessert, etc.), and location of the vineyard (California, France, Italy, etc.).

There is also new interest in facet theory and faceted displays from within the library science community. These are evidenced by the implementation of and discussion surrounding faceted browse in the North Carolina State University (NCSU) OPAC⁸ (services by Endeca⁹), the efforts of the IFLA Functional Requirements for Subject Authority Records Working Group,^{10,11} and

recent publications such as James Anderson's 2006 article, "A Fully Faceted Syntax for LCSH," and 2005 book entitled, *Information Retrieval Design*.¹²

What is faceting? In what situations is it useful? Faceting is a structured approach, a method. It is not a particular classification scheme or a particular controlled vocabulary.

Facet analysis is a procedure for indexing, and faceted analytical theory is the underlying theory for doing facet analysis correctly. These concepts were introduced to the library science community by S. R. Ranganathan as long ago as 1959.¹³

A faceted classification is considered to be one of the three main types of classifications. These are enumerative, hierarchical, and faceted.¹⁴ Enumerative classifications subdivide each subject until all possibilities are exhausted. Hierarchical classifications subdivide and order each class from general to more specific. Faceted classifications divide all concepts into mutually exclusive categories. Ranganathan likened it to a popular construction toy of the time called Meccano that consisted of a lot of pieces-- plates, rods, wheels, etc. which you could put together in different ways in order to build cranes, bridges, vehicles, and so on. A scheme in which most of the building blocks are already put together in unchangeable structures isn't a faceted scheme.¹⁵ Vickery said it like this, "the categories in a faceted classification are not locked into rigid enumerative schedules, but are left free to combine with each other in fullest freedom, so that every type of relation between terms and between subjects may be expressed."¹⁶ The Bliss Bibliographic Classification, 2nd edition (BC2), is an example of a faceted scheme. The Universal Decimal Classification (UDC) has some features of an enumerative, hierarchical, and

faceted classification with faceted main classes where punctuation in the notation is used to mark the facets. Recent revisions of the Dewey Decimal Classification (DDC) have sections based on facet analysis.

Faceted analytical theory has been applied to controlled vocabularies, as a process leading to faceted *displays* of that vocabulary. The new ANSI/NISO standard Z39.19-2005 describes it like this:

Facet analysis, based on the work of Ranganathan and refinements by the Classification Research Group, is another way of organizing knowledge. It takes a bottom-up approach, forming areas of knowledge after first having pieced together their parts and determining the areas of knowledge they form, rather than the discipline-directed, top-down approach of hierarchies. Facet analysis is particularly useful for:

- new and emerging fields where there is incomplete domain knowledge or where relationships between the content objects are unknown or poorly defined;
- interdisciplinary areas where there is more than one perspective on how to look at a content object or where combinations of concepts are needed;
- vocabularies where *multiple hierarchies are required* but can be inadequate due to difficulty in defining their clear boundaries; or
- classifying electronic documents and content objects - where location and collocation of materials controlled by a notation [call number] that is used for shelf arrangement - is not an important issue.

In the case where the controlled vocabulary contains many hundreds or thousands of terms, it may be helpful to organize them according to facets. This type of facet analysis results in identifying the basic broad categories of the vocabulary and grouping the terms under them.¹⁷

Broad categories, or facets, have been proposed by a variety of individuals and groups over time. Table 1 lists fundamental facets as identified by Ranganathan, ANSI/NISO, LCSH and OCLC FAST, BC2, and Anderson.¹⁸ Ranganathan's were personality, matter, energy, space, and time (PMEST). The Art & Architecture Thesaurus uses seven facets to categorize their special fields (associated concepts, physical attributes, style and periods, agents, activities, materials, and objects). The broad categories used to subdivide topical terms in LCSH are currently topical, chronological, and geographic. At present, there are no facet indicators for main types of topical headings beyond subject, personal name, corporate name, geographic name, and title. Form, another LCSH subdivision, is an example of a non-topical facet.

BC2 uses these facets in this order: thing, kind, part, property, material, process, operation, patient, product, by-product, agent, space, and time.¹⁹ The Functional Requirements for Bibliographic Records (FRBR) Group 3 entities that are described as the subjects of works are: concepts, objects, events, and places.²⁰

Types of Facets (Topical and Non-Topical)

| Ranganathan | ANSI/NISO Generic Facets | LCSH/OCLC FAST | Bliss (BC2) | Anderson |
|-------------|-----------------------------|----------------|-------------|----------|
|-------------|-----------------------------|----------------|-------------|----------|

-----TOPICAL FACETS-----

| | | | | |
|-------------|---|---------|--|---|
| Personality | Entities Things Parts | Topical | Things Entities Kinds | Things Entities Kinds |
| Matter | Attributes Constituent materials Properties States | Topical | Materials Properties Parts | Materials Properties Parts |
| Energy | Actions Activities Operations Processes Events | Topical | Process Operations Patient Product Means | Process Operations Client Product Means |

-----NON-TOPICAL FACETS-----

| | | | | |
|-------|--------|------------|----------|----------|
| Space | Places | Geographic | Space | Space |
| Time | Times | Chronology | Time | Time |
| | | Form | Format | Format |
| | | Genre | Medium | Medium |
| | | | Audience | Audience |
| | | | Approach | Approach |

Table1. Types of Facets

Are these attempts at faceting leading somewhere? It would appear they are in commercial websites. The web designers interviewed by Dr. La Barre explicitly noted that facets provide the ability to support browsing, exploration, and discovery by helping users ask questions, and giving them choices without requiring them to know a lot about the interface or the information being searched. Faceted interfaces expose information in a simple, flexible and powerful way that is not intimidating. Faceted search exposes what is typically hidden behind search results; it gives insight into the data.²¹

On the screen shot from the NCSU catalog, note the way that one can narrow a search by facets – both topical and non-topical, while at the same time seeing the number of records that would be retrieved in each case (Figure 1). The NCSU implementation of Endeca allows us to see the potential for faceted displays in our OPACs – something Ranganathan visualized without the benefit of computers.

But please remember that facet analysis, based on faceted analytico-synthetic theory, is a procedure for indexing that teases out all aspects of the subject content of an item. A faceted browse display is only able to present the facets that have been defined and identified during subject cataloging. Unless wine.com had taken the case to select facets (price, color, type of use, location of vineyard) based on the needs of their users and then analyze each wine based on those facets, the multitude of combinations (i.e., red-inexpensive-dessert-domestic and white-high priced-dinner-Italian) would not be possible.

Optimal faceted browse displays will not result unless facet analysis is done at the subject indexing stage. It is the desire of some web designers to create these displays, and this is what has them seeking out assistance from the writings of S. R. Ranganathan, Brian C. Vickery, and Jack Mills,

This year a new article by Anderson and Hofmann proposes some working rules to apply LCSH terms using the broad categories of BC2 in order to insure that we capture all aspects of subject content of an item when we catalog it.²² They suggest a bottom-up approach using LCSH terminology so that fully faceted headings can be integrated in catalogs with older LCSH headings.

These are promising developments that challenge us to revisit a theory that our community previously abandoned related to subject cataloging in order to make our OPACs better discover tools for our users.

¹ ANSI/NISO Z39.19-2005 Guidelines for the Construction, Format, and Management of Monolingual Controlled Vocabularies. Bethesda, MD: NISO Press, 2005, 11.

² David Miller, Tony Olson, and Sara Shatford Layne. "Promoting Research and Best Practices in Subject Reference Structures: A Decade of Work by the Subject Analysis Committee." *Library Resources & Technical Services* 49(3): 154-166.

³ Lorcan Dempsey. "The Library Catalogue in the New Discovery Environment: Some Thoughts." *Ariadne*, no. 48 (July 2006).

⁴ Kathrin [sic] La Barre. "A multi-faceted view: Use of facet analysis in the practice of website organization and access." *Knowledge Organization for a Global Learning Society: Proceedings of the Ninth International ISKO Conference 4-7 July 2006 Vienna, Austria*, 359.

⁵ <http://www.wine.com/>

⁶ <http://www.epicurious.com/>

⁷ <http://www.barnesandnoble.com/>

⁸ <http://www2.lib.ncsu.edu/catalog/?Nty=1&N=0&Ntk=Keyword&Ntt=baseball>

⁹ <http://endeca.com/>

¹⁰ <http://www.ifla.org/VII/s29/wgfrsar.htm>

¹¹ Tom Delsey. "Modeling Subject Access: Extending the FRBR and FRANAR Conceptual Models." *Cataloging & Classification Quarterly* 39(3/4) 2005: 49-61.

-
- ¹² James D. Anderson and Melissa A. Hofmann. "A Fully Faceted Syntax for Library of Congress Subject Headings." *Cataloging & Classification Quarterly* 43(1) 2006: 7-38. Electronic prepublication is available online at: http://www.haworthpress.com/store/E-Text/View_EText.asp?sid=EUH5ECK43SH09H0U8HBTFFXWX4HH3N60&a=3&s=J104&v=43&i=1&fn=J104v43n01%5F03 (accessed Sept. 7, 2006). James D. Anderson and José Pérez-Carballo. *Information Retrieval Design: Principles and Options for Information Description, Organization, Display and Access in Information Retrieval Databases, Digital Libraries, Catalogs, and Indexes*. St. Petersburg, FL: Ometeca Institute, 2005.
- ¹³ "FAST Transitions in OPACs – From subject heading lists to faceted strings: a talk by Pauline Atherton Cochrane to the UIUC Library Colloquium Group, May 12, 2006."
- ¹⁴ Kathryn La Barre. "The Use of Faceted Analytico-Synthetic Theory as Revealed in the Practice of Website Construction and Design" (Ph.D. dissertation, Indiana University, 2006), viii. Available online at https://netfiles.uiuc.edu/klabarre/www/LaBarre_FAST.pdf.
- ¹⁵ Vanda Broughton. *Essential Classification* (New York: Neal-Schuman Publishers, Inc., 2004), 257, 258.
- ¹⁶ Brian C. Vickery. *Faceted classification schemes*. In S. Artandi (ed.). *Rutgers Series on Systems for the Intellectual Organization of Information* (v.5). New Brunswick, NJ: Graduate School of Library Science at Rutgers University.
- ¹⁷ NISO, 14, 15.
- ¹⁸ Pauline A. Cochrane. FAST transitions in OPACs: From subject headings to faceted strings, a talk by Pauline Atherton Cochrane to the UIUC Library Colloquium Group on May 12, 2006. Unpublished document available from author.
- ¹⁹ <http://www.sid.cam.ac.uk/bca/bcclass.htm>
- ²⁰ Barbara Tillett. *What is FRBR?* (Washington, DC: Library of Congress, 2003), p. 3. Available online at: <http://www.loc.gov/cds/downloads/FRBR.PDF>
- ²¹ La Barre, 155, 156.
- ²² Anderson and Hofmann.