Traditions of facet theory, or a garden of forking paths?

Kathryn La Barre
Graduate School of Library and Information Science, University of Illinois at Urbana-Champaig, USA

Abstract

This socio-historical inquiry contrasts historical and contemporary facet theory. While examining vagaries in terminology for key concepts such as *facet*, *facet analysis*, and *facet theory* , as used by different researchers and different schools of thought, the author concludes with a call for the creation of operational definitions and functional requirements to enhance, amplify or extend current practices.

Background

The most well-known exemplars from this tradition are S.R Ranganathan's *Colon Classification* and his technique of *facet analysis* (Ranganathan, 1933). Other examples of Faceted Classification include the *Universal Decimal Classification* (Pollard, 1926) and the *Bliss Classification* (Bliss, 1929). From the early 1950s, three groups - the British Classification Research Group (CRG), the North American Classification Research Study Group (CRSG), and the Indian Library Research Circle (LRC) - worked tirelessly to promote facet theory (Brownson, 1960; La Barre, 2004; Parthasarathy, 1952; Ranganathan, 1937, Ranganathan, 1962). Broughton (2006) and others speak of a British tradition, distinguishing it from that found in the United States. Some researchers reference an Indian tradition that more closely follows Ranganathan's explications. Few would disagree that the core literary canon of facet theory is largely composed of works by Ranganathan (1933, 1937/1957/1967) and the members of the British CRG (for example Binns & Bagley, 1958/1961; CRG, 1955; Foskett, 1957; Mills, 1957; Vickery, 1960). More contemporary works from CRG members are also part of the canon (for example Broughton, 2006; Mills, 2004; Vickery, 2008). Other contributions are rarely referenced and often ignored (for example by LRC members:, Devadason, 2009; Gopinath, 1986, Neelameghan, 1992; Neelameghan & Gopinath, 1975; and by CRSG members: Anderson, 1980; Atherton, 1965; Richmond, 1981).

A Garden of Forking Paths?

> 'This web of time — the strands of which approach one another, bifurcate, intersect or ignore each other through the centuries — embraces every possibility.' (Borges, 1994, p.98).(

Borges describes a labyrinth of infinite choice, multiple possibilities and diverse simultaneous futures. We find early indications of a similar pattern of development in the trajectory of facet theory with respect to language. Although Ranganathan and the British CRG members initially used the terminology of facet theory with precision, variability increased over the intervening

years. By the 1960s de Grolier (1962) observed divergent term use among CRG members. For example, D.J. Foskett, among others, frequently interchanged the terms *category* and *facet* and described *facet analysis* as '…analysis of a subject in its entirety into a certain number of … categories of things.'

Vickery and others commonly substituted the phrase *conceptual categories* 'of high generality and application that can be used to group other concepts' for *fundamental categories*. Attributing these divergences to 'Ranganathian language', de Grolier (1962, p. 15) described use of this extremely specialized term to reference 'each facet of a subject, as well as each division of a facet' as a chief locus of confusion, an observation repeated recently by another French facet theorist (Maniez, 1999). Differences in vocabulary and understanding may also spring from geographical and cultural separations of the three research groups. Regardless of source, terminological and practical confusion is now as rampant as it is well documented.

Variant Understandings: Heritage

The term *facet* is commonly considered as analogous to category, attribute, class and concept (La Barre, 2010, 2011). Ranganathan initially used the phrase *train of characteristics* to emphasize the inherent nature of characteristics in subjects (Ranganathan, 1967). The notion of facet springs from the mathematical concept of *parameter,* a range of possible factors, aspects or elements. Parameters permit identification of sets of distinct cases. In facet theory, each parameter creates a small number of case groupings equivalent to a facet. Each facet (grouping of cases) has the potential to include multiple dimensions of a given parameter and can exist as a recursive or linguistically nested structure (Ranganathan, 1967).

Facets are derived through *facet analysis*, an iterative technique that uses a list of general or *fundamental categories* to 'map a universe of subjects in a helpful sequence' (Ranganathan, 1967, p. 398). Ranganathan postulated five - personality, matter, energy, space, and time, but recognized that others were possible (Ranganathan, 1967; Vickery, 1959). Instead of defining the term *fundamental categories,* Ranganathan directed attention to the identification of isolate ideas as manifestations of the underlying fundamental category. Confusingly, Ranganathan also used the term *basic facet* when discussing *fundamental categories*.

Vickery (1959) observed that *fundamental categories* reflect an Indian worldview that differs from the Western, or Aristotelian notion of basic categories. Aristotle's *fundamental categories*: thing, kind, part, property, material, process, operation, agent, patient, product, by-product, space and time represent a 'standard set' of useful descriptive categories for science and technical subjects but are often unsuitable for humanities subjects (Vickery, 1959). For Ranganathan, each subject manifests *fundamental categories* that represent conceptual or concrete entities in any given universe. Each entity has ascertainable and stable attributes useful for dividing these entities into groups or classes. Vickery credits Ranganathan as being the first to explicitly use the idea of categories as a representation of subjects, with subject fields forming the basic classes (Vickery, 2005).

By 1960, the British CRG refined the role of *fundamental categories* by insisting they 'should not be used mechanically and imposed upon the subject, [rather] … as a provisional guide in approaching a new field… [to] give guidance in suggesting possible characteristics which should not be overlooked.' (Vickery, 1960, p. 24). As outlined in early form by Vickery (1953, p. 50), the

number of fundamental categories expanded beyond five. These are today best summarized by Aitchison (née Binns), Gilchrist and Bawden (2002): entities, things and objects subdivided by characteristics and functions; actions and activities; time; kinds or types; systems and assemblies; applications and purposes.

With a provisional set of *fundamental categories* in hand, the next step is to apply *facet analysis* to sort terms into groups or sets. The technique of *facet analysis* has been described in several places (Mills, 1962, 2004; Vickery, 1960, 2008) as a process of term collection from user queries, domain literature or item descriptions. Vickery agrees with Mills that dividing subjects into broad facets and then into arrays (or subfacets) is in accord with the principle of logical division underpinning all classification. Vickery and Mills cite critically important differences in the strictness with which the rules of logical division - isolate each distinct category, formulate each new characteristic of division, recognize every new relation - are applied in facet analysis (Vickery, 1960; Mills, 2004). Strict adherence to these principles results in a classification that allows full combinatorial freedom for facets, greater expressivity for concept relations and formation of new logical relations (Vickery, 1960).

Here too, we have a divergence among CRG members as Mills, an original member of the CRG, and editor of the 2nd edition of the Bliss Classification, notes that facet analysis results in the assignment of terms to two types of facets (Mills, 2004, p. 551):

> 'True categories: e.g. geographical location, chronological time, material of composition
> Relational categories: e.g. kind, part or property'

Vickery (2005) stated, that in his understanding, the term *fundamental category* denotes a general type of facet, a principle 'manifested' by the facet. He noted that he was not aware of this distinction (above) and found it confusing.

Variant Understandings: Contemporary

This complex tradition has paths that simultaneously bifurcate and converge within deeply variable modern understandings. Divergences emerge in the use of the term *facets* by American information architects and in some Next Generation Catalogs (La Barre, 2006, 2010). Most often bearing more resemblance to ad hoc categories, this often-superficial notion of facet leverages existing data fields as search and browse categories. Rigorous facet analysis has yet to become a standard part of OPAC system design. While the practices embedded in the design of many Next Generation Catalogs may be unique to the United States, this situation is not an indicator of unanimity. Rather, it was an expediency eagerly grasped by a few OPAC software designers. Few North American facet theorists would disagree with the statement that the true strength of facet analysis lies in the way it can 'peel the onion of an idea' (Vickery, 1966, p. 13-14). Too often, concerns about time and money work against deep subject analysis that could be most fully displayed in a faceted OPAC. In a similar vein, many facet theorists concur with Broughton's assessment that the NISO Z39.19 definition of facet as 'attributes of content objects encompassing various non-semantic aspects of a document' such as topic, author, location, format, language, and place of publication has more in common with database fields (Broughton, 2006).

Emblematic of this contentious situation, the OCLC research project FAST (Faceted Access to Subject Terminology) provides a 'simplified syntax' that retains LCSH vocabulary in an 'easy to understand, control, apply, and use format' that reduces the cost of indexing (OCLC FAST website) through the use of eight facets: Names: Personal, Corporate, and Geographic; Events; Titles; Time periods; Topics; and Form/Genre. Many North American facet theorists agree with Broughton that FAST 'makes some progress along the road to consistent analytico-synthesis, although it is not faceted in the sense that most UK professionals would recognise' (Broughton, 2006, p. 58). FAST lumps 'topic' (or subject) as one facet among many non-subject elements of bibliographic description. Many North American and British facet theorists recognize their genesis in MARC and Dublin Core fields, not as actual or necessarily appropriate products of rigorous facet analysis (La Barre, 2010).

Another forking path is use of the term *faceted classification* to describe website search or browsing structures for objects (La Barre, 2006). Here, some members of the British tradition draw a sharp distinction - in that faceted classification applies to subjects, not objects. Vickery (2008) dismisses concerns by stating that objects and subjects can both be represented by symbolic terms (notations). Furthermore, classifications for documents often include objects, as in the case of the subject Food (Vickery, 2008). Raghavan (2010), a member of the Indian tradition, considers a facet as one aspect of a multidimensional entity, whether subject or the carrier of knowledge. He echoes Vickery's understanding that classification deals with both objects and subjects in digital environments replete with electronic resources that have a range of attributes such as target audience, form and file format.

Perhaps codifications in current standards such as the draft of ISO/DIS 25964-1 (2009) *Thesauri and interoperability with other vocabularies* and one of its predecessors, the British Standard 8723 parts 1-5 *Structured vocabularies for information retrieval*, can provide firm ground. Each seeks to offer clear guidelines for applying facet analysis to thesauri. That standards are proprietary is not unusual - for example, the new bibliographic standard *Resource Description and Access* requires the purchase of a license to access. This state of affairs has been widely criticized by those in the Semantic Web community as erecting a barrier to cooperation and interoperability. These are important considerations in the digital environment in which these conversations are undertaken.

These two documents provide nearly identical descriptions of facet analysis as useful in generating hierarchical relationships, such as general, whole part, instance and associative relations between concepts. A facet is a grouping of concepts 'of the same inherent category' (BSI 8723-1, 2005, p. 3; ISO/DIS 25964-1, 2009, p. 4). Both describe the use of 'high-level categories: objects, materials, agents, actions, places and times' to create facet groupings. Node labels are another device designed to permit groupings by facet name, or characteristic of division (BSI 8723-1, 2005; ISO/DIS, 2009 25964-1). Facet analysis is described as a tool that permits subjects to be analyzed into concepts, and allows grouping narrower concepts by characteristics of division (BSI 8723-1, 2005; ISO/DIS, 2009 25964-1).

Another broad divergence relates to the nature of fundamental categories. Ranganathan sought 'universal' fundamental categories applicable across disciplines. For others, fundamental categories are provisional and dependent on purpose - as is evident in the special faceted classification schemes created by CRG members. Raghavan's (2010) view, that digital

environments make us increasingly aware of the need for 'universals,' is especially resonant in the context of ontologies for the Semantic Web (La Barre, 2011). This overview of the semantics of facet theory underscores the critical need for agreed-upon operational definitions and functional requirements for central concepts of facet theory if it is to contribute fully to the future of the Semantic Web and other contemporary initiatives.

Functional Requirements: Facet Analysis

In 2003, Phil Murray created the Faceted Classification listserv, in part, to expand upon a preliminary set of functional requirements for facet theory (Murray, 2003):

1. What is the most effective way to model the process of facet analysis?
2. Is there a recognized way to design and model Faceted Classification?
3. How should a human or a machine index with facets?
4. What interchange formats are best for capturing facets and facet relations?
5. What software or metadata tools are best for faceted implementations and applications?
6. What is the best approach to selecting automated categorization tools for sharing schemas, supporting facets?
7. What approaches are similar to Faceted Classification or facet analysis?

These need refinement, but help identify mission critical features and functional requirements. Slavic, who motivated the post by Murray, has also emphasized the importance of agreed-upon functional requirements to 'improve standards for the use and exchange of knowledge organization systems' (Slavic, 2008, p. 258).

Other preliminary attempts to identify functional requirements reside in unexpected places. One source is Kashyap's comparison of Ranganathan's facet postulates and principles to Chen's entity-relationship modeling. This analysis identifies key features and bridges the understanding barrier between facet theorists and those in cognate traditions (Kashyap, 2001).

The role of facet analysis in ontology development is also a critical consideration. Broughton has identified shared and complementary roles for facets in ontology building: both provide excellent vocabulary control structures, support term disambiguation, enhance browsing and searching, and frameworks for site navigation. Because Faceted Classifications rely upon the use of mathematically-based formal coding to express content and content element relations, notation and facet indicators can be leveraged by search and access systems and could be integrated or converted into a fully developed ontology (Broughton, 2006).

To take this line of reasoning a bit further, because facets represent aspects or viewpoints from which an entity may be analyzed, Sigel and others promote the use of facet analysis because it helps uncover the relationships between concepts in a domain (Foskett, 1977; Sigel, 2003; Soergel, 1985; Vickery, 1966), and supports semantic factoring, or the analysis of categories into primitives or basic level concepts at any level of an ontology (Sigel, 2003). Soergel views semantic factoring as an equivalent process to facet analysis. He envisions this process as a facet framework. If one conceives of each facet as a question, each answer thus represents one essential aspect of a given concept. This functional approach to semantic factoring is

demonstrated by the following example (Soergel, 1985, p. 258):

- 'Of which class is the concept or object a member or subclass?
- Is the object in a specific state, condition or circumstance?
- What is it capable of doing? Does it have a specific purpose?
- Does the object or concept cause, influence, produce or act upon another?
- Is X a means by which to achieve something else?
- Is it a specific aspect or viewpoint?'

This approach echoes heritage understandings such as Vickery's (1966) encouragement to ask a series of questions during the process of facet analysis: What concept does this represent? In what conceptual category should this concept be included? What are the class relations between this concept and other concepts included in the same category?

Given such a framework of analysis, it is but a small step to embed this approach in formalized knowledge representations (such as a set of IF-THEN rules) for use in a semi-automated algorithm for ontology creation. John Sowa (undated) advocates use of semantic factoring 'applied to any level of an ontology from the highest, most general concept types to the lowest, most specialized types. The methods can be automated, as in formal concept analysis, which is a systematic technique for deriving a lattice of concept types from low-level data about individual instances.' Building on this suggestion, Uta Priss (2008) also promotes Formal Concept Analysis (FCA) to generate mathematical facet models to aid in the creation of graphical representations. Might such an approach potentially serve as a formal functional model for facet analysis as well? Both Priss and Sowa reference Wille (1992) as a fundamental resource for those interested in promoting and extending the use of facet theory for ontology development.

Conclusion

Facet theorists must work together to find common ground and to regularize term use. As such, it may be useful to operationalize definitions and begin work on a set of functional requirements for facet analysis. Both deliverables are potentially useful to researchers who seek to create robust and theoretically grounded Semantic (faceted) Web applications (La Barre, 2011). This formal call to discussion was extended in June of 2011 to participants of the North American Symposium of Knowledge Organization , and to the participants at ISKO Spain. I hope that the participants of the ISKO UK biennial conference will join with representatives of all traditions at the 2012 ISKO meeting in Bangalore India, at the Documentation Research and Training Centre to create a unified vision of facet theory

References

Aitchison, J.; Gilchrist, A. & Bawden, E. (2002). *Thesaurus construction and use: A practical manual.* London: Fitzroy Dearborn.

Anderson, J.D. (1980). Prototype designs for subject access to the Modern Language Association's bibliographic database." In J. Raben & G. Marks (Eds.) *Data bases in the humanities and social sciences: Proceedings of the IFIP Working Conference.* Amsterdam: North-Holland.

Atherton, P. (1959). Ranganathan's classification ideas: An analytico-synthetic discussion" *Library Resources & Technical Services* (9)4, 462-473.

Binns, J., & Bagley, D. (1958/1961) *A faceted classification for engineering.* 2nd / 3rd Ed. Whetstone: English Electric Company Ltd.

Bliss, H.E. (1929). *The organization of knowledge and the system of the sciences*. New York: Henry Holt.

Borges, J. L. (1994). *Ficciones*. (English Translation). Garden of Forking paths. Grove Press.

Broughton, V. (2006). The need for faceted classification as the basis of all methods of information retrieval. *Aslib Proceedings* (58)1/2, 49-72.

Broughton, V. & Slavic, A. (2007). Building a faceted classification for the humanities: Principles and procedures. *Journal of Documentation* (63)5, 727-754.

Brownson, H. (1960). Research on handling scientific information. *Science,* 132(3444), 1922-1931.

BS 8723-1:2005 (2005). *Structured vocabularies for information retrieval Part 1: Definitions, symbols and abbreviations.* London: British Standards Institution.

BS 8723-2:2005 (2005) *Structured vocabularies for information retrieval. Part 2: Thesauri*. London: British Standards Institution.

BS 8723-3:2007 (2007). *Structured vocabularies for information retrieval . Part 3: Vocabularies other than thesauri.* London: British Standards Institution.

BS 8723-4:2007 (2007). *Structured vocabularies for information retrieval . Part 4: Interoperability between vocabularies*. London: British Standards Institute.

BS 8723-5:2008 (2008). *Structured vocabularies for information retrieval . Exchange formats and protocols for Interoperability.* London: British Standards Institute.

Classification Research Group. (1955). The need for a faceted classification as the basis for all methods of information retrieval. *Library Association Record*, (57) 7, 262–268.

de Grolier, E. (1962). *A study of general categories applicable to classification and coding in documentation,* Translation by the National Science Foundation, Deventer Holland, Ysel Press.

Devadason, F.J. (2009). Online construction of alphabetic classaurus: a vocabulary control and indexing tool. *Information Processing and Management. (*17), 1-11.

Foskett, D. J. (1957) Appendix I: "Occupational safety and health documents classification scheme". In *Proceedings of the International Study Conference on Classification for Information Retrieval*. London: Aslib. pp. 115-136.

Foskett, D. J. (1977). *Subject approach to information,* 3rd ed. London: Clive Bingley.

ISO/DIS 25964-1. (2009) *Thesauri and interoperability with other vocabularies.* Geneva, Switzerland: International Organization for Standardization.

Kashyap, M. (2001). Similarity between Ranganathan's postulates and Chen's entity-relationship approach to data modeling and analysis" *DESIDOC Bulletin of Information Technology* (21) 3, 3-16.

La Barre, K. (2004). Weaving webs of significance: The Classification Research Study Group in the United States and Canada. In B. Rayward & M. E. Bowden (Eds.), *Proceedings of the Second Annual Conference on the History and Heritage of Scientific and Technical Information Systems*, 15–17 November, 2002. Medford, NJ: Information Today.

La Barre, K. (2006). The use of faceted analytico-synethic theory as revealed in the practice and construction of website design. Unpublished doctoral dissertation, Indiana University, Indiana, United States. Retrieved from https://netfiles.uiuc.edu/klabarre/www/ProposalLa_Barre.pdf

La Barre, K. (2010). Facet analysis. *Annual review of Information Science and Technology*, 44, 243-284. Medford, NJ, Information Today..

La Barre, K. (2011). A semantic (faceted) web. *Les Cahirs du Numerique (*6) 3, 103-131.

Maniez, J. (1999). Des classifications aux thesaurus: du bon usage des facettes. *Documentaliste Sciences de l'information,* (35) 4-5, 223-286.

Mills, J. (1957). Classification of a subject field. *Proceedings of the International Conference on Classification.* London: Aslib. 1957, pp. 29-42.

Mills, J. (1962) *A modern outline of library classification.* London: Chapman & Hall.

Mills, J. (2004). Faceted Classification and logical division in information retrieval. *Library Trends* (52) 3, 541-570.

Murray, P. (2003). Viewpoint (scope of Faceted Classification List), *Faceted Classification List.* message 97, 10 January 2003. Retrieved from http://finance.groups.yahoo.com/group/facetedclassification/message/51

Neelameghan, A. and Gopinath, M.A. (1975). Postulate-based permuted subject indexing (POPSI). *Library Science with a Slant to Documentation,* (12)3, 79-87.

Neelameghan, A. (1992). Application of Ranganathan's general theory of knowledge classification in designing specialized databases. *Libri,* 1992, (42)3, 202-226.

OCLC FAST website. (undated). *Faceted Application of Subject Terminology* http://www.oclc.org/research/activities/fast/default.htm

Parthasarathy, S. (1952). Memoirs of the Library Research Circle. *Abgila* (2), 153–159.

Pollard, A.F.C. *The Decimal Classification of the Institut International de Biliographie.* England: Cambridge University Press. 1926. [An English language translation of the UDC with introduction – the earliest comprehensive treatment in this language].

Priss, U. (2008). Facet-like structures in computer science. *Axiomathes*, (18), 243-255.

Raghavan, K. (2010). *Personal correspondence with author.*

Ranganathan, S.R. (1933). *Colon Classification.* London: Edward Goldston.

Ranganathan, S. R. (1937/1957/1967) *Prolegomena to library classification*, 2nd ed. London: Library Association.

Ranganathan, S.R. (1933), (1952, 4th ed.). *Colon Classification.* Madras, India: Madras Library Association.

Ranganathan, S. R. (1962) Classification research no longer a toddler. *Abgila* (2), 88–89.

Richmond, P. (1981). *Introduction to PRECIS for North American usage.* Littleton, CO: Libraries Unlimited.

Sigel, A. (2002). KO as a use case for TMs. In *XML topic maps: Creating and using topic maps for the Web,* Addison-Wesley Professional.

Slavic A. (2008). Faceted Classification: Management and Use. *Axiomathes*, (18)2, 257-271.

Soergel, D. (1985). *Organizing Information: Principles of database and retrieval systems,* San Diego: CA, Academic Press.

Sowa, J. (undated). Information interchange and interpretation. *NCITS T2 Com listserv* Retrieved from http://www.jfsowa.comontology/gloss.com

Tudhope, D. (2004). *Personal correspondence,* public-esw-thes, message 0051. Retrieved from http://lists.w3.org/Archives/Public/public-esw-thes/2004Feb/0051.html

Vickery, B. C. (1953). Systematic Subject Indexing. *Journal of Documentation 9* (1), 48-57.

Vickery, B. C. (1959). *Classification and indexing in science.* London: Butterworths.

Vickery, B. C. (1960). *Faceted classification: A guide to construction and use of special schemes.* London: Aslib.

Vickery, B. C. (1966). *Faceted Classification Schemes.* Rutgers series on systems for the intellectual organization of information, v. 5. New Brunswick NJ: Graduate School of Library and Information Science at Rutgers University.

Vickery, B. C. (2005). *Personal correspondence with the author*.

Vickery, B. C. (2008). Faceted classification for the web. *Axiomathes (*18) 2, 145-160.

Wille, R. (1992). Concept lattices and conceptual knowledge systems. *Computers and mathematics with applications.* (23) 6-9, 493-515.

Wordnet 2011 *definition: tradition*. Retrieved from http://wordnetweb.princeton.edu/perl/webwn?s=tradition

Corresponding author:
Kathryn La Barre can be contacted at kathryn.labarre@gmail.com