Institution: University College London

Unit of Assessment: 36 – Communication, Cultural and Media Studies, Library and Information Management

Title of case study: Facet analysis and its influence on the major systems of library classification

1. Summary of the impact (indicative maximum 100 words)

Research on the theory and design of faceted classifications as exemplified in the Bliss Bibliographic Classification 2nd edition (BC2) has influenced the recent development of two out of the three internationally important library classification schemes, the Universal Decimal Classification, and the Dewey Decimal Classification. Collectively these are used in over 350,000 libraries worldwide, and thousands of new publications in the relevant subject classes are now classified using systems based on research undertaken in UCL Information Studies.

2. Underpinning research (indicative maximum 500 words)

Facet analysis is a powerful methodology for building knowledge organisation systems such as classification schemes, thesauri, taxonomies and ontologies [d, e]. It was originally devised by a UCL graduate, S. R. Ranganathan, in the 1930s, and taken up in the 1950s by the UK Classification Research Group (CRG), who greatly refined and elaborated the methodology, and developed a body of theoretical research which forms a significant part of the overall corpus of research literature in the fields of indexing and information retrieval. UCL research in this field is now led by Professor Vanda Broughton (UCL Information Studies, or DIS, since 1997 and a member of CRG from the early 1970s).

Facet analysis builds classification from 'the bottom up' on the basis of a detailed examination of the concepts in a given discipline or subject field [b]. Concepts are assigned to a series of categories (or facets) which are largely functional, or linguistic in nature (entities, parts, materials, processes, operations, products, agents). Relationships between concepts in a facet, and between the facets themselves, are established, and the rules for combination (the system syntax) are derived from the structure, within the conventions and needs of the particular discipline. Although there is an established body of theory, different disciplines present particular difficulties in the work of intellectual analysis. Research in DIS also involves the investigation of potential new fundamental categories, and the encoding of data to represent the structural properties of concepts and their interrelationships. This has applications not only in libraries, but in machine management, and today there are many examples of faceted applications in e-commerce, online searching, etc. Much DIS research has focused on the management and representation of faceted systems in digital formats which support such implementations [c].

DIS is the institutional base for development of the second edition of the Bliss Bibliographic Classification (BC2). The original Bibliographic Classification (BC1) was a well-regarded classification system widely used in UK academic and research libraries until the 1990s, when, for financial reasons, it was largely replaced by the Dewey Decimal and Library of Congress Classifications. The development of BC2, which started in 1977, resulted in a series of highly structured classifications for different disciplines, which provide a model of a faceted system built on theoretical principles. DIS researchers have developed new terminologies for specific subject fields including religion, the fine arts and chemistry [a].

The AHRC-funded FATKS project (2002–2003, PI V. Broughton) in DIS examined the viability of applying a faceted approach to humanities vocabulary, and developed an indexing and search tool for two humanities disciplines, religion and the arts; outputs included the creation of a relational database, and software for the automatic building of classmarks [f]. The project documentation and the demonstrator are published on the FATKS website, http://www.ucl.ac.uk/fatks/about.htm. DIS researchers working on faceted classification included Professor Ia McIlwaine (Head of DIS 1993–2001, Editor-in-Chief of the Universal Decimal Classification (UDC) from 1993-2007), who was responsible for the implementation of a faceted structure into UDC; Dr Aida Slavic (research associate on the FATKS project 2002–2003, former DIS doctoral student, and Editor-in-Chief of UDC from 2011) developed the relational database and search interface for the faceted humanities

vocabulary.

3. References to the research (indicative maximum of six references)

[a] Broughton, V. and Mills, J. and Coates, EJ. (2012) *Bliss Bibliographic Classification 2nd edition. Class C Chemistry*. deGruyterSaur: Munich. Submitted to REF2.

[b] Broughton, V. (2011) 'Facet analysis as a tool for modelling subject domains and terminologies'. In Slavic, A. & Civallero E. (eds.) *Classification and ontology: formal approaches and access to knowledge: proceedings of the International UDC Seminar, 19-20 September 2011, The Hague, The Netherlands.* Würzburg: Ergon Verlag. pp 207–228. Submitted to REF2.

[c] Broughton, V. (2010) 'Finding Bliss on the web: some problems of representing faceted terminologies in digital environments'. In: Gnoli, C. and Mazzocchi, F., (eds.) *Paradigms and conceptual systems in knowledge organization: Proceedings of the Eleventh International ISKO Conference*. Ergon: Wurtzburg. Available on request.

[d] Broughton, V. (2010) 'Concepts and terms in the faceted classification: the case of UDC'. *Knowledge organization*, 37, 4 pp. 270–279. Submitted to REF2.

[e] Broughton, V. (2008) 'A faceted classification as the basis of a faceted terminology' *Axiomathes* 18 (2) pp. 193–210. Springer online DOI 10.1007/s10516-007-9027-7. Submitted to REF2.

[f] Broughton, V. and Slavic, A. (2007) 'Building a faceted classification for the humanities: principles and procedures' *Journal of Documentation* 63 (5) pp. 727–754. http://hdl.handle.net/10150/105218

Output [f] emerged from an **AHRC research grant**:

'Towards a Knowledge Structure for High Performance Subject Access and Retrieval within Managed Digital Collections'. PI: Vanda Broughton. Amount: £46,226 Duration: April 2002–June 2003. AHRC B/IA/AN8003/APN13797. Graded A

4. Details of the impact (indicative maximum 750 words)

Classification research in DIS has had significant impacts on the design of new versions of two major library classification systems, the Universal Decimal Classification (UDC) and Dewey Decimal System (DDC), cumulatively used by over 350,000 libraries around the world. Most existing classification systems originate before the computer age, employ pragmatic ideas of knowledge organisation, and lack a sound theoretical basis. Facet analysis on the BC2 model offers solutions to some of these issues and enables more accurate representation of subject content, especially important for online searching and discovery. Of particular importance is the research into classification in the humanities, and the special challenges of culture-specific concepts and terminology which they present; effective ways of managing these difficulties allow for the elimination of cultural bias which is evident in many older systems, and the creation of schemes which are more acceptable to minority users.

Impact on the design of UDC library classification

The UDC is an international classification, available in 40 languages, and used in more than 150,000 bibliographic databases, documentation centres and libraries in around 130 countries, notably in central and eastern Europe, where it is the dominant scheme [1]. UDC has to exhibit a consistent and logical structure – such as that offered by facet analysis – because of its use in scientific and technical research establishments, as well as general academic collections, and its application to documentation as well as books. Collections include VINITI – All-Russian Scientific and Technical Information Institute of Russian Academy of Sciences (28 million records), NEBIS – The Network of Libraries and Information Centres in Switzerland (3m), Romanian Science and Technology Portal (3m), the national libraries and bibliographies of Slovakia (4m), Slovenia (3.5m), Hungary, Czech Republic, Croatia, Poland, Finland, Serbia, Spain and Portugal [1]. Facet analysis also provides for a data structure that is compatible with the maintenance of the classification in a database format, and for its machine management. As Ines Cordeiro (UDC Editor-in-Chief 2007-2011, Deputy Director of the National Library of Portugal and former doctoral student in DIS) reported in 2007, BC2 provided a model for UDC to achieve its organisational goal of developing a classification suitable for online environments, not just one designed for better shelf arrangement

[4].

In the 1990s UDC agreed to make BC2 the basis of all new revisions of the UDC, using its terminologies and adopting its faceted structure within the UDC format. In 1997, DIS researcher Broughton joined the Editorial Team to help achieve this goal, presenting the principles of facet analysis to editorial workshops, helping to establish editorial policy and practice, and working on the conversion of BC2 classes to UDC format. Her identification of the Genesis Problem [d] demonstrates the difficulty of representing, in a relational database, subject content and terminology that is specific to particular cultures and traditions. She also developed several new auxiliary schedules, which assign common facets across classes [2].

In 2000, Broughton created an entirely new Religion classification, based on her work on BC2 [3]. This was incorporated into the 2005 edition of UDC and applied to all accessions in religion since. Between 2008 and 2013, books in 130 countries were classified using this system.

Broughton's work on Religion [3] provided precedent for editors developing other UDC classes. In 2009, a working group led by Claudio Gnoli began updating Philosophy, explicitly 'trying to follow the pattern of facet presentation introduced in the Class 2 Religion by Vanda Broughton' [5], and using her expertise in humanities classification [f].

Revision of the important Class 61 Medicine was announced in 1995. In 2009, its editors, including Ia McIlwaine, outlined the process of adapting BC2 [6] and how auxiliary schedules created by Broughton were a major factor in rationalising its structure: 'In phase 1, a framework for the new class was established … Bliss terminology was used in the captions together with UDC notation and formatting as needed. Concepts and terms, the common auxiliaries, and classes related to medicine were used insofar as they were appropriate. There was heavy use of common auxiliary tables of general characteristics (Table 1k) -02 Properties, -04 Relations and Processes, and -05 Persons' [7].

In 2011, Broughton was invited to explain the BC2 structure of physical sciences at the UDC Editorial Workshop, and the following year BC2 principles were also considered in the revision of the mathematics class, with work beginning in 2013 [8].

Impact on the design of the Dewey library classification system

The Dewey Decimal Classification (DDC) has been translated into more than 30 languages and serves library users in over 200,000 libraries in 135 countries, making it the world's most widely used library classification system.

Dewey was consistently criticised for its Class 200, Religion, in which Christianity and the Bible occupied numbers 220-289, with all other religions sharing 290-299. A 2005 user survey confirmed this general discontent with the Religion Class, and with its far greater weighting of Christianity compared with other religions [9]. In 2006, following the publication of the new UDC Class 2 Religion, the then-Editor of the DDC Joan Mitchell proposed collaboration with McIlwaine for a revision incorporating many features of Broughton's UDC schedule into the 22nd edition of Dewey [10].

When released in 2011, the top structure of the new DDC class mirrored almost exactly the UDC structure as designed by Broughton. A separate publication, 'developed for libraries with extensive religion collections' appeared in 2012 [11]. This featured 'an optional arrangement for the Bible and specific religions based on a chronological/regional view, in order to help reduce Christian bias in the standard notational sequence for the Bible and specific religions, based on a similar development introduced in the UDC in 2000' [10]. Between its release in 2012 and 31 July 2013, 178,000 books in the field of religion were added to the WorldCat database maintained by the Online Computer Library Centre (OCLC), which also publishes and maintains Dewey; the majority will have received a new DDC number using this system [12].

Between them UDC and DDC are used as organising and retrieval tools for collections in 350,000 libraries across the world. Facet analysis, the focus of DIS research, has been instrumental in addressing such issues as logical structure, currency, cultural neutrality, and machine compatibility in both of these, two of the three leading systems of library classification. The influence of faceted classification in general and BC2 in particular has ensured that its advantages are available to

millions of library users.

5. Sources to corroborate the impact (indicative maximum of 10 references)

[1] Universal Decimal Classification (UDC), (http://www.udcc.org/index.php/site/page?view=about) and list of collections (http://www.udcc.org/index.php/site/page?view=collections)

[2] Broughton, V. (2002) 'A new common auxiliary table for relations, processes and operations', *Extensions and corrections to the UDC 24* pp. 29–35. Available on request.

[3] Broughton, V. (2000) 'A new classification for the literature of religion', *International cataloguing and bibliographic control* 2000 (4) and Paper read at the 66th IFLA Council and General Conference, Jerusalem, Israel 13–18 August 2000. Available on request.

[4] Slavic, A., Cordeiro, M., & Riesthuis, G. (2008) 'Maintenance of the Universal Decimal Classification: overview of the past and preparations for the future', *International Cataloguing and Bibliographic Control* 37 (2) pp. 23–29. Available on request.

[5] Gnoli, C. (2009) *UDC Philosophy revision Report 1*. UDC Italia, 7 August 2009. (http://italia.udcc.org/report1.html)

[6] Davies, S. (2011). 'UDC Editorial Workshop, The Hague, 21 September 2011: a report' *Extensions and corrections to the UDC 33* pp.11–12. (http://arizona.openrepository.com/arizona/bitstream/10150/236492/1/Davies_E%26C33_2011.pdf)

[7] Williamson, N., McIlwaine, I. C. (2009). 'UDC Medical Sciences project: progress and problems'. *Extensions and Corrections to the UDC 31* pp. 33–36. http://hdl.handle.net/10150/199890

[8] Statement provided by Editor-in-Chief, UDC, in a personal communication dated 3 May 2013, available on request.

[9] *Dewey's Options in Religion: Survey Results* (http://www.oclc.org/nl-NL/dewey/discussion/optionsinreligion.html)

[10] McIlwaine, I., and Mitchell, J. S. (2006). 'The New Ecumenism: Exploration of a DDC/UDC View of Religion'. In *Knowledge Organization for a Global Learning Society: Proceedings of the 9th International ISKO Conference, 4–7 July 2006, Vienna, Austria*, Gerhard Budin, Christian Swertz, and Konstantin Mitgutsch (eds.), pp. 323–330. Würzberg: Ergon. Available on request.

[11] Dewey Decimal Classification web pages (http://www.oclc.org/dewey/versions/religion.en.html) Announcement of the 2012 release (http://ddc.typepad.com/025431/2012/05/200-religion-class.html)

[12] Data derived from a search of WorldCat (http://www.worldcat.org/) July 2013.