The Round Table "UDC Editorial Perspectives": a report

Aida Slavic
Associate Editor, UDC Consortium

The Round Table *UDC Editorial Perspectives* was a one-day meeting for members of the UDC Editorial Team and the UDC Advisory Board organized in The Hague on 28 October 2009 in conjunction with the *UDC Seminar 2009 "Classification at a Crossroads: Multiple Directions to Usability"*. The meeting was instigated as part of the UDC Consortium efforts to build closer cooperation with UDC specialists and editors of national editions worldwide and to encourage more active involvement and communication between members of the wider UDC editorial team. This was the first face-to-face meeting of the UDC collaborators to which all members of the Advisory Board, the Editorial Team and interested observers were invited to participate.

Topics discussed

The meeting was organized around two themes: the implementation of a faceted structure in the UDC and the report on current work on common auxiliaries and other areas. The Editor in Chief of the UDC, Maria Inês Cordeiro, opened the meeting by explaining the objectives and the context of the Round Table and the current UDC management strategy. She briefly spoke about two current priorities of the UDC editorial policy: the creation of an international team of collaborators with closer connections to the UDC community worldwide and an editorial infrastructure and online collaborative environment, currently under- development, that would enable better communication, sharing and transparency of UDC work as well as wider inclusion.

Topic 1: The implementation of a faceted structure in the UDC

This session had three invited presentations followed by discussion:

- Vanda Broughton: *Looking back on the Class 2 revision: lessons learned*
  Religion poses no great difficulties in creating a faceted conceptual structure but the vocabulary is complex and terms for the same concept vary from one faith to another. How this phenomenon should be handled may be perceived differently by various user groups, or potential user groups, and a clear policy must be established for the future.
- Claudio Gnoli : *Revision of Philosophy and facet analysis*
  The first report about the ongoing work on a proposal for revision of class 1 *Philosophy*, with special reference to its facetization, and to possible implications of this for the future use of facets throughout the UDC. This report on the Philosophy revision project is available at: http://italia.udcc.org/report1.html.
- Nancy Williamson & I. C. McIlwaine: *The new UDC medicine class: progress and problems*
  A description of the classes completed, followed by identification of work still to be done. Also a discussion of problems identified in the course of the work, including notation, outdated related classes which require revision and updating, missing topics and suggested additions to auxiliary tables.

Restructuring of the UDC following the principle of facet analysis, based on broad facet categories (thing, kind, part, materials, processes etc.), has many implications for the ways schedules are presented in the UDC notation systems and in the way the classification data are controlled, managed and maintained in the UDC MRF database. Faceted schedules may be presented in different ways and the solution adopted will directly influence the size of the scheme and the way it is presented in printed or online versions. Facet analysis has been repeatedly suggested as beneficial in the restructuring of the UDC since the early 1960s. The FID/CCC Extraordinary meeting made this an official policy in 1976 stating that UDC is to be transformed into a fully faceted scheme based on a methodology put forward by A. F. Schmidt and J.-H. de Wijn (FID Internal Documentation. UDC reform work. C 75-35. The Hague, 1976).

This Round Table provided an opportunity to discuss a different approach in faceting the UDC adopted under the guidance of Prof I. C. McIlwaine, the Editor in Chief up to 2006, who chose to implement in UDC the facet analysis based on the Bliss Bibliographic Classification (BC2), i.e. for UDC to mirror the BC2 structure.

For those not familiar with BC2 it may be worth noting that, while indeed the UDC system may benefit from the better content structure featured in the BC2 scheme, the UDC scheme is far more advanced in terms of formality of data presentation, coherency of schedule organization and classification data management and distribution.

BC2 is a library classification proper which is created with the practicality of shelf arrangement and printed schedules in mind and its main priority is the logic of the linear sequence of subjects. As BC2 is not maintained in a database nor is designed to be used online, the rules imposed on its structure are created to be readable and understandable to humans. This means that BC2 notation is based on principles and not on formal rules. The way the notation is built can change slightly from class to class or new rules for building notations may be introduced in the middle of a single subject area. This may be either because the authors found a solution for creating shorter notations, or they wanted to achieve an order of concept that departs from the general principle but may be better suited for that particular case. The creativity and loose interpretation of notational rules is not a problem for BC2 as the scheme enumerates pre-combined subjects to great detail and the principal goal was not to enable easier building or understanding of the notation syntax by users. In addition, while the UDC functions as one coherent knowledge scheme, with a set of rules that apply to the entire system and in which concepts are re-used and linked across disciplines, BC2 is developed more as a series of special classifications in which each subject area is likely to list all concepts that may be required in a field whether they are single or in combination. Also differently from UDC - subjects from different disciplines cannot be freely combined in indexing. But what is most unusual to users of decimal classifications is that BC2 uses a shortened or contracted notation that does not express the hierarchical level of the class and does not reveal the components from which it was built - thus serving mainly to fix the order of classes.

Although BC2 has been widely written about and praised in research papers for the logic of its structure by its authors, the truth is that the scheme was applied in a small number of libraries for shelf arrangement and has never been implemented for online retrieval or widely used in a bibliographic information retrieval setting or international information exchange. Briefly, there are some features from BC2 that UDC may want to replicate, there are also many others, especially those related to the notation and lack of concept hierarchy that are not desirable in UDC and should be avoided.

Since 1993 two UDC classes have been selected for restructuring according to facet analysis as implemented in BC2: one is *Religion,* revised in 2000, and the other is *Medicine* the revision of which started back in 1993 and is still awaiting completion. These two subjects represent two extremes with respect to the facet analysis methodology: humanities are notoriously hard to structure according to universal categories while applied sciences (especially Medicine) are considered to be an ideal subject field to test the principle of facet analysis.

Reports at this meeting illustrate that in both cases the restructuring of schedules can be time-consuming and challenging and good results are by no means guaranteed. After pointing out some issues and solutions to facet presentation in UDC, Vanda Broughton stressed in her talk that the UDC editorial team would benefit from a clearer policy and a more elaborated and consistently applied methodology in implementing facet analysis in UDC. One of the issues touched upon in more detail in the discussion was about the way the BC2 structure, which enumerates pre-combined classes, can be reproduced in UDC - which normally prefers relational tables allowing for combination to be done at the point of indexing, not at the point of building schedules. Hence, while UDC uses a synthetic approach to reduce the size of schedule through avoidance of concept repetition, BC2 does exactly the opposite: it produces lengthy and enumerated lists of pre-combined and complex concepts.

The above and several other issues are of interest for two other projects reported at this meeting. Claudio Gnoli, regarding revision of class 1 - *Philosophy*, pointed to a variety of options in presenting a combination of concepts in UDC (with colon combinations or direct addition with special auxiliaries) and his concern was that different combinations have different effects on the logic of sequence of pre-combined numbers. Claudio's written report is published in this issue of E&C. N. Williamson and I.C. McIlwaine, regarding revision of class 61 - *Medicine*, also stressed the importance of an appropriate solution for notational representation. The size and detail of vocabulary in Medicine is not easy to manage and pre-combined classes tend to have extremely long notations and, as pointed out by N. Williamson, there are many technicalities in choosing a correct notation that have to be discussed and decided upon before the final version of Medicine is published. One of the concerns is the need for the revision of related areas such as microbiology, which is important for completing the revision of Medicine. A written report on the revision of class 61 – *Medicine* is published in the present issue of E&C.

The mention of a new Medicine class has initiated lively discussion about vacant class 4 and whether the new Medicine and Psychology schedules should be moved there. The old paradox of change versus stability split the participants into two groups. Paradoxically those in favour of keeping Medicine in class 61 hoping to preserve stability in libraries are actually proposing a solution by which the existing numbers in medicine will be assigned a new and different meaning - thus making management of reclassification extremely difficult. But most importantly this solution would completely hinder the use of UDC numbers in information exchange or cross collection searching for decades to come. Moving Medicine to class 4 would ensure that the old notation would not be reused with a new meaning. For instance, if the concept of 'liver' would be in some collections classed by either 611.36 (old number) and in another by e.g. 411.565.1, then searching across collections can be easily established by mapping two notations of the same concepts. If, however, 611.36 in some libraries is still used for 'liver' and in other libraries using the new medicine classification the same number is used in a completely different meaning - UDC numbers as a means of searching and grouping of subjects would become completely pointless. The re-use of the same notation to denote new meaning remains one of the most sensitive issues in revisions of the schedule.

It is worth noting that a proper notational solution is very important for classification management and use and it is not surprising that it is frequently discussed by those working on the schedules. UDC is a synthetic classification, designed with an expressive and hierarchical notational system that allows both great flexibility and accurate control over constituent elements; it is logical and easy to use in both the indexing process and online information retrieval. UDC notation supports semantic search expansion and semantic factoring and the complex notation can be parsed and linked to natural language terms with accuracy and ease - hence the notation is probably the system's strongest feature and is responsible for its widespread use.

Topic 2: Current work on common auxiliaries and other areas

The second part of the meeting was intended to inform everyone about the ongoing and planned UDC work, and contained presentations on three topics:

- Edgardo Civallero: *UDC language table: overview of recent updates*
  Language tables were completely revised in 1992 after a long period of preparation, based on linguistic knowledge available from the 1970s and 1980s. In the past thirty years new knowledge emerged with respect to indigenous languages especially in South America, the richest part of the world in linguistic terms. Edgardo reported about the current work on the updating of the language tables.
- Aida Slavic (on behalf of Alan Pritchard): *Area Tables & Literary Genres - Brief summary of recent work and contributions and questions raised.*
  This is a summary of discussions over the past two years and the status of some of the work undertaken so far.
- Aida Slavic, Jiri Pika, Miguel Benito, Gerhard Riesthuis: *The new UDC Summary and scheme demonstrato*r - *Selection of classes, mapping to Dewey and translation issues* In 2008, permission was granted by the UDC Consortium Executive Committee for a new and extended UDC summary to be made available online. In the past two years this was expanded into a small project that includes the creation of the database to include translations and mappings to other systems. A report on the selection of classes, mappings to DDC summaries and translations was given.

Edgardo Civallero worked on the update of classes of South and North American indigenous languages and reported on the recent revision of UDC common auxiliaries of languages (Table 1c). These were revised in the 1980s and the complete table was introduced in 1992 (E&C 14). In the past thirty years, however, there have been many new linguistic studies that brought new understanding about language relationships and sub-groupings of the, previously less known and studied, indigenous languages. As explained by Edgardo, the classification of indigenous or autochthonous languages of South and North America that were presented in UDC had numerous flaws with respect to the current linguistic views. Due to its geography, land mass and the number of isolated areas, South America is known to be linguistically the richest part of the world. There are in existence more than ten large language families but equally a large number of isolated and unclassified languages. The outline of these languages in UDC provided a partial, skewed and occasionally wrong view, not only of the phylogenetic relationships themselves but also with respect to proportional representation of family subgroups according to the number of languages involved and the number of speakers. Because of this South and Central American languages were revised in 2008 and in 2009 the same revision continued with North American languages. The mapping between new and old numbers will be held in the UDC MRF and available to all users. The experience with South and North American indigenous languages indicates that

other language families would benefit from analysis in case they also need an alignment with the current linguistic knowledge. African languages, for instance, may be the next group to look at.

Following a brief discussion on languages, A. Slavic reported on behalf of Alan Pritchard about the ongoing work on the auxiliary numbers of place (Table 1e). Alan, who joined the UDC team in 2009, has over 15 years experience working with geographical, demographic and postcode data and has approached the tasks of consolidating, checking and developing of place tables, with great enthusiasm. In 2009 he put a great deal of effort into recording rules that he was able to derive from the existing schedule and noting inconsistencies that need to be addressed. We hope that with Alan's contribution we may be able to progress with the establishment of precise policy and guidelines with respect to this important table, which currently has 10,000 classes. One of the most crucial issues in the place table that needs resolving is the policy towards historical entities. In 2009, Alan also started collecting vocabulary for the development and revision of Literary Genres and we hope to be able to report in more detail about work done in this area soon.

Common auxiliaries of form (Table 1d) is another area of UDC that would require revision as it contains cross classifications between document form and carrier and we are going to hear soon about developments and specific plans for this table. At the end of the meeting we heard about the work done on the UDC Summary project and UDC to Dewey Summaries mapping by Jiri Pika and Miguel Benito, reported by Aida Slavic who also demonstrated the first version of the schedule available online. A lively discussion about UDC work in general filled most of the remaining time.

Concluding remarks

The number of colleagues participating in the Round table and the intensive and open discussion during the day was a clear sign that the Round Table fulfilled its objectives. The general feeling after the meeting was very positive. The Editors and Advisory Board members were more enthusiastic about collaboration. Our hope is that in the future we plan these meetings well in advance and alongside other events to enable wider participation. We think that such seminars should continue to be organized in the future, possibly in conjunction with the biannual UDC conferences and occasionally even more frequently.

Participants

There were 21 participants from 16 countries in attendance: Marie Baliková (Czech Republic), Miguel Benito (Sweden), Vanda Broughton (UK), Andrew Buxton (UK), Edgardo Civallero (Spain), Maria Inês Cordeiro (Portugal), Elita Eglite (Latvia), Victoria Frâncu (Romania), Drahomira Gavranovic (Croatia), Claudio Gnoli (Italy), Agnes Hajdu Barat (Hungary), Alan Hopkinson (UK), Devika P. Madali (India), Ia C. McIlwaine (UK), Sirje Nilbe (Estonia), Jiri Pika (Switzerland), A.R.D. Prasad (India), Gerhard Riesthuis (Netherlands), Darija Rozman (Slovenia), Aida Slavic (UK) and Nancy Williamson (Canada).

We regret that our colleagues from Ukraine, Lithuania, Slovakia, Russia, Poland and Belarus with whom we closely cooperate could not attend. Also our close collaborators whose input into the schedule is substantial - Geoffrey Robinson and Alan Pritchard, from the United Kingdom - were not able to attend and were greatly missed.