

### Thinking about Classification

On 28 June 2017, a workshop was held at CILIP's London HQ on the topic of classification. The day was split equally between a series of inspiring and informative presentations, and group discussions.

The day formed a collaborative workshop, with lots of opportunities for librarians to learn from each other's practice and gather advice to apply to their own collections. There was as much communal discussion as formal presentation, and group activities on localisations, ethics, and classification advocacy. Speakers presented and conducted surveys on the consumption of the various schemes to find out which were being used in different libraries and why, and where schemes had been tailored to suit specific needs, or replaced by an entirely local scheme.

### What is classification?

Deborah Lee of the Courtauld Institute of Art opened the workshop with an introductory discussion distinguishing classification from categorisation and cataloguing. She defined the classification scheme as a source for location and retrieval.

Lee identified a classification scheme as a system of knowledge organisation, which could be defined as a document in its own right, or an aesthetic object, whose anatomy comprises schedules, instructions and citation orders, an introduction of philosophy and context, and an index. A scheme notation's desired attributes were identified as hospitality, brevity, expressiveness, and mnemonics.

She also covered the concept of facet analysis – the breaking down of complex subjects into elemental concepts – and identified this as a key attribute of sophisticated classification schemes, whose ordering can either collocate or scatter a subject. The process of classification moves from subject analysis, to translating the subjects into the scheme's language and concepts, to assigning notation, then adding a suffix or 'Cutter' to create the classmark.

This workshop also asked *why* we classify. The answer is for retrieval, of either a known item or to make browsing more relevant. Lee noted that classification often operates so that like stays with like.

### UDC

Aida Slavic, editor-in-chief of the UDC, spoke about new members of the consortium, including the national libraries of Czech Republic, Slovakia, Slovenia and Croatia.

UDC is markedly international in its focus, simultaneously developed in 4 languages and updated nearly every year. There are 3000 UDC classes in 57 languages available as Linked Data under a Creative Commons licence, 13 languages in an abridged subscriber and 9 languages in the Complete 72000 classes, as of September 2017.

She noted that the process of classification is moving away from libraries and towards publishers with pre-purchased metadata. Emerging stakeholders are looking for tools for managing multilingual access, semantic expansion and subject access via a switching language that can facilitate mapping for a cross-collection search.

## Revisions

Slavic observed that the primary motives for revision include adding new terminology for new areas or geopolitical subdivisions (when the administrative division of modern countries change) and correcting biases (which usually involves changing the terminology but not the class). Despite the best intentions, she also noted the difficulties of revision – when a scheme is used in at least 150 000 libraries in over 135 countries (as is the case for UDC), then many will continue to use cancelled numbers 20-50 years after the classes have become obsolete, because reclassifying requires extra work.

## Reclassification

The taxonomy of reclassification includes merging, updating, localisation, standardisation, or a change of scheme. Lee discussed different types of reclassification including 'guerrilla reclassification'.

Siobhan Cottom from Nottingham College provided a case study of a classmark for fashion growing too full at her library. She needed to introduce more specific bibliographic divisions to enhance efficient retrieval. The group recommended soliciting student warrant and working from the top down, rather than book up, to aim for collocation rather than dispersal, and to aim for a maximum of ten books per number.

## Ethical issues

While working within historical constraints, classifiers can effect changes in schemes and knowledge communities dictate the organisation of their knowledge. The 'Thinking about classification' workshop discussed the nuances of neutrality and objectivity, pejoration, error, and the importance of factual resources. Unethical classification can lead to a disengaged and disenfranchised user base, and an attentive attitude towards voices and ownership should propagate enhanced representation and retrieval.

Historical errors and pejoration include out-of-date borders and colonial geographic hierarchies, such as the imposition of colonial US subdivisions on indigenous nations in earlier editions of Dewey Decimal Classification. Alongside the ethical issues of descriptive terminology, there is the potential for offence in hierarchical position. Melissa Adler was cited as a recent contributor to the study of problematic parent classes in her study of the classification of sexuality at the Library of Congress in *Cruising the Library* (2017). Finally there is the comparative order and size of classes, which often favour a US-centric view of history, and an overwhelmingly Christian representation of religion.

## The Future of Classification

Vanda Broughton questioned the assumptions that classification is an obsolete, artificial, 19<sup>th</sup> century practice and whether hierarchy should be abandoned in favour of links. She argued that systems must proceed to enhance discoverability and that pattern and predictability of a map-domain model allows inference into the unknown. Running through the history of classification, she covered H.E. Bliss' theory and Otlet's advocacy of analytico-synthetic tools designed for retrieval rather than physical organisation and the representation of complex content, returning again to Ranganathan's concept of idea, verbal and notational planes (the final one might form the control device of URLs).

Broughton conducted her own survey of 50 of the top UK University libraries to find that 48/50 use LCSH for subject retrieval and, 20 use DDC, 18 LCC, 6 local, 2 Garside, 2 UDC, 1 Bliss, 1 NLM, and 15 use more than one major scheme.

She noted that the role of classification is more limited with online discovery and without browsing, and that booksellers' categorisation systems are gaining momentum through libraries' outsourced processing. The future of classification faces its own technical problems: automatic classification from natural language processing, extraction relationships and hierarchies, and the automation of images. Despite these challenges, Broughton maintained that classification underpins most tools that work in digital faceted information architecture: classification for reasoning and the semantic web relies on access to structured data.