

SLAVIC, A. and DAVIES, S. 2017. Facet analysis in UDC: questions of structure, functionality and data formality. *Knowledge organization* [online], 44(6): selected papers from the proceedings of the 2017 International UDC seminar: faceted classification today: theory, technology and end users, 14-15 September 2017, London, UK, pages 425-435. Available from: https://www.ergon-verlag.de/isko_ko/downloads/ko_44_2017_6_e.pdf

Facet analysis in UDC: questions of structure, functionality and data formality.

SLAVIC, A. and DAVIES, S.

2017



Facet Analysis in UDC: Questions of Structure, Functionality and Data Formality

Aida Slavic*, Sylvie Davies**

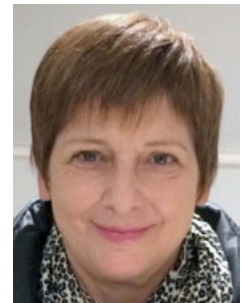
*UDC Consortium, PO Box 90407, The Hague 2509 LK,
The Netherlands, <aida.slavic@udcc.org>

**Robert Gordon University, Aberdeen,
United Kingdom, <s.davies@rgu.ac.uk>

Aida Slavic is the Editor-in-chief of the *Universal Decimal Classification (UDC)* and works on the development and the maintenance of the scheme for the UDC Consortium, based in The Hague. She is also an associate professor (visiting lecturer) in the Department of Information and Communication Sciences at the University of Zagreb (Croatia), where she is a lecturer in the master and doctoral programmes on the subject of indexing languages, knowledge organization, information retrieval and classification systems. She holds a PhD in library and information studies from University College London. Her research interest is in the use of classification in a networked environment, knowledge organization, classification, metadata and semantic technologies.



Sylvie Davies is an associate lecturer in information management at Robert Gordon University, Aberdeen and a member of the UDC Editorial Team. Her original background is in languages with degrees in German and Mandarin Chinese, followed more recently by MSc studies in information and library studies. Her teaching areas include controlled vocabularies, taxonomies and bibliographic classification. Her research interests lie in the area of cross-cultural communication (at national and corporate level) of organized knowledge with a focus of multilingual and multicultural thesauri. Locally, she has been involved in aspects of information management in the oil and gas industry via the delivery of taxonomy workshops and cross-industry forum discussion on the role of information management professionals in that industry.



Slavic, Aida and Sylvie Davies. 2017. "Facet Analysis in UDC: Questions of Structure, Functionality and Data Formality." *Knowledge Organization* 44(6): 425-435. 31 references.

Abstract: The paper will look into different patterns of facet analysis used in the UDC schedules and how these affect the scheme presentation, the underlying data structure and the management of the classification scheme. From the very beginning, UDC was designed to represent the universe of knowledge as an integral whole allowing for subjects/concepts from all fields of knowledge to be combined, linked and the nature of their relationships made explicit. In Otlet's original design, the emphasis for his new type of classification was on the coordination of classmarks at the point of searching, i.e., post-coordination, which he firmly rooted in an expressive notational system. While some UDC classes exhibit various patterns of facet analytical theory proper, others, although used in an analytico-synthetic fashion, follow less canonical structural patterns. The authors highlight the lack of connection made throughout the various stages of UDC restructuring between: a) theoretical requirements of an overarching facet analytical theory as a founding principle guiding the construction of schedules; and, b) practical requirements for an analytico-synthetic classification in terms of notational presentation and data structure that enables its use in indexing and retrieval, as well as its management online.

Received: 6 August 2017; Accepted: 7 August 2017

Keywords: UDC, facet analysis, facets, faceted classification

1.0 Introduction

Faceted classifications are often discussed in connection with information retrieval in a computerised environment and in this paper we would like to focus on this particular connection in the context of the Universal Decimal Classification (UDC) which, as a well-established analytico-

synthetic scheme, is particularly relevant to the discourse on faceted classification for effective information retrieval. By doing so, we aim to provide a better understanding of two stipulations: 1) the need for a theoretical framework that underpins an analytico-synthetic scheme based on facet analysis; and, 2) the requirements of computerised systems such as data modelling for database design,

interfacing and usability of bibliographic classification schemes online.

The longevity of the scheme and its place at the forefront of research on the mechanisation and automation of bibliographic classification provides some scope to examine the potential role of facet analysis in the online usage of analytico-synthetic classification in computerised systems. In addition, UDC schedules present examples of different approaches to facet analysis, which provide an opportunity to observe which ones appear to most influence online implementation.

1.1 Terms of reference

We plan to examine more closely the following concepts of bibliographic classification in relation to UDC: 1) notation and syntax; 2) classification termed “faceted proper” constructed on a facet analytical theory framework; and, 3) the concept of analytico-synthetic classification. In this paper, we make a distinction between the structural and the functional properties of a classification scheme, the latter referring primarily the way it functions in indexing and retrieval.

1.1.1 Notation

In the context of UDC, a classification is treated as an indexing language proper, which comprises “vocabulary” and “syntax” and in which notation plays the role of an “indexing term.” The UDC notation can be viewed as a “vocabulary label” and rules for building complex UDC expressions can be viewed as “syntax.” This, in our view, makes this notational system central to information retrieval and of most concern to classification data modelling, automation and management.

1.1.2 Faceted classification.

The concept “faceted” deals with classification structure and is of principal interest for construction of the schedules. Facet analysis entails an analysis of concepts to be organized into mutually exclusive categories (applying a single division principle on each level of subdivision). These categories can be things (their kinds and parts), processes, properties, instruments, materials, agents, place, time, etc. Facet analysis may be intuitive, guided by common sense and basic principles of formal logic. However, when the choice and type of facets pertain to an established theoretical framework that imposes the type, number and sequence of universal facet categories, we call this framework Facet Analytical Theory (FAT) and classification systems constructed according to this principle faceted classifications proper. In some of these

schemes, such as *Bliss Bibliographic Classification (BC2)*, notation has been viewed separately from the classification structure, its only purpose being the correct mechanical ordering of classes and, therefore, the notational system did not have to be expressive in terms of syntax or hierarchy (Slavic 2008).

1.1.3 Analytico-synthetic classification.

The concept of “analytico-synthetic” is of principal interest when it comes to classification use. The function of an analytico-synthetic scheme means that the indexing terms, i.e., notation, can be composed and decomposed in the process of indexing and in the process of information retrieval. This requires full control of notational elements when managing and processing the classification automatically (Slavic and Cordeiro 2004; Slavic 2008).

The notion of facet is often linked to analytico-synthetic classification schemes, because their main purpose is to enable the combination of mutually exclusive properties that might need to be combined in describing a subject.

2.0 Fundamentals of synthesis in Otlet’s design

Paul Otlet contributed to classification theory and to the development of the analytico-synthetic principle by laying the foundation for proper understanding of bibliographic work, namely the organization and retrieval of subjects and that of documents as physical carriers (Otlet 1895-1896; 1896; 1934). He designed an indexing language with notational elements and syntax envisaged to function as a kind of universal formal vocabulary to produce an unlimited number of formalised statements about document content and a document as a physical carrier. He recognised what was later called in classification theory the “facets of a subject” (1990, 52):

Now, one quickly observes that there are certain basic ideas which are present in all part of the classification, such as geographical, historical and form categories which we already discussed. In any individual branch of the classification there are also divisions which recur regularly; thus, in Zoology, for example, each species can be envisaged from the point of view of its anatomy, its evolution, its teratology.

Otlet’s drawing in Figure 1 illustrates the function of synthesis in UDC, where facets of common auxiliaries (place, time, document form, document language) would be used to provide further specifications of main classes

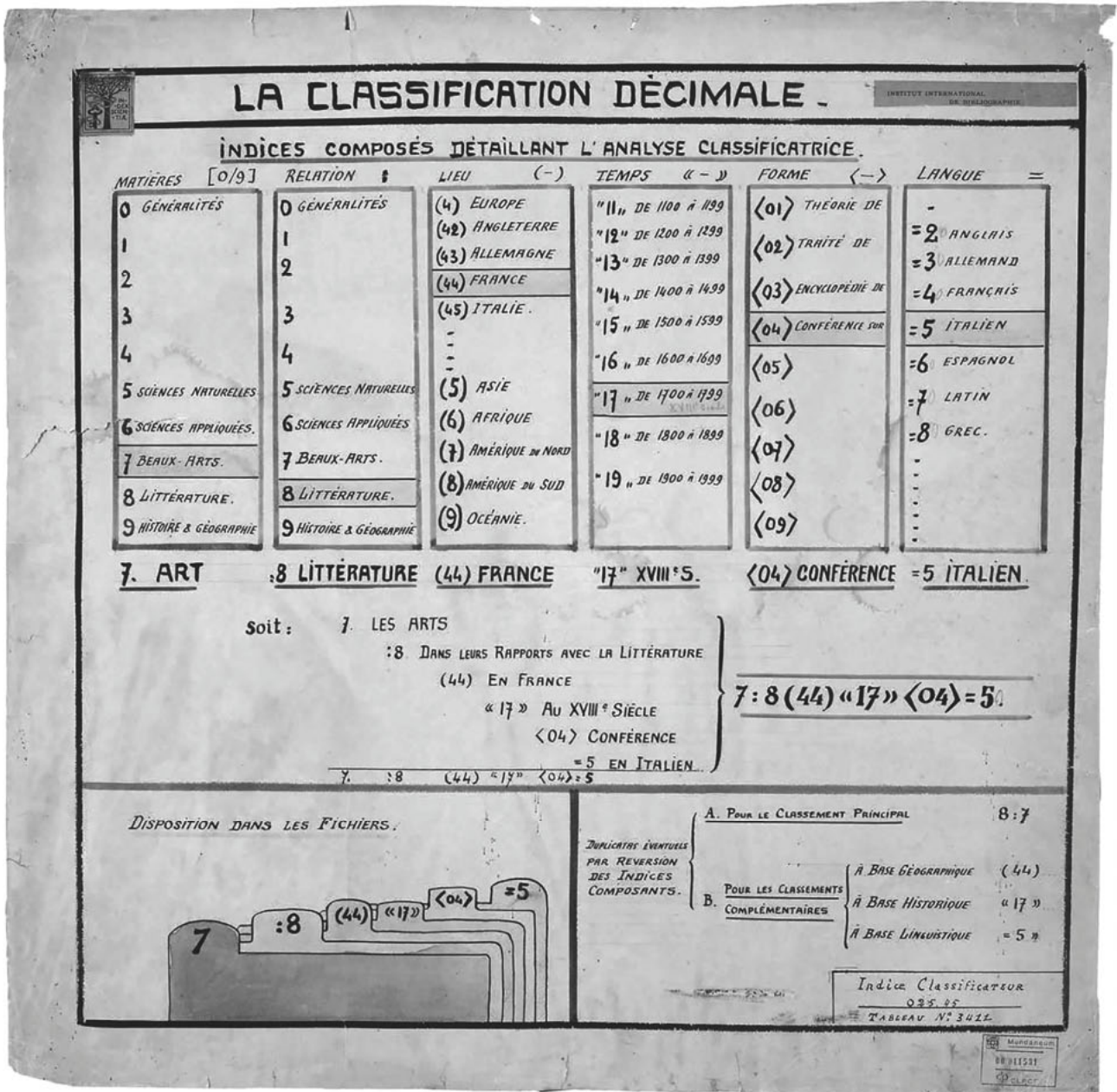


Figure 1. UDC structure as designed by Otlet (courtesy of Mundaneum).

that can be combined with one another, as well as the way complex classmarks would support retrieval of all aspects of a composed UDC expression. His depiction of a card catalogue (in the bottom left-hand corner) shows how multiple copies of the catalogue cards would be filed under each of the elements of the complex notation thus allowing that the document entry for the “Influence of the eighteenth century French literature on the arts” could be found by searching under either the arts, literature, france, eighteenth century, etc. Clearly, his choice of facets, their unique notational representation and their organization into tables were driven by the functions these facets have to fulfil in information retrieval.

Concepts are organized into tables according to their affinity to synthesis, i.e., their need to be combined in the process of indexing and retrieval. Each table is assigned a unique facet indicator with the exception of main UDC numbers representing main subject facet(s), which do not start with any facet code. As demonstrated below, these tables and their sequence hold the key to the syntax rules of number building, citation ordering and filing ordering. The meaning of concepts in combinations and their correct parsing are hard-coded by the presence or absence of facet indicators surrounding or preceding the numbers themselves. Indeed, UDC codes which are not preceded by facet indicators are always main numbers that repre-

sent subject areas listed in the main tables. Even at this early stage of UDC development, there is evidence that facet analysis has been applied in many parts of the schedules but this does not indicate that a single overarching theoretical framework has been applied.

2.1 Notational representation of facets in UDC

The structure of UDC is represented by a notational system as outlined below; common auxiliaries are designed to add further precision to all subjects, regardless of context. These facets have a unique notational representation (i.e., beginning and, in most cases, ending sign) and they can be directly attached to any other UDC number.

Sub-facets array in the common auxiliaries of persons (facet indicator *-05*):

- 051* Persons as agents, doers, practitioners (studying, making, serving etc.)
- 052* Persons as targets, clients, users (studied, served etc.)
- 053* Persons according to age or age-groups, sexual orientation
- 054* Persons according to ethnic characteristics, nationality, citizenship etc.
- 055* Persons according to gender and kinship
- 056* Persons according to constitution, health, disposition, hereditary traits
- 057* Persons according to occupation, work, livelihood, education
- 058* Persons according to social class, civil status

Main classes list subjects of main knowledge areas, their kinds and parts. They can be combined among themselves using connecting symbols and can be further specified by general properties using common auxiliaries or special properties using special auxiliaries (materials, properties, processes, operations).

Hierarchies are presented using simple decimal numerical notation, and they may also be structured in facets and sub-facets but without using facet indicators. Exceptions are history, geography and literature where subdivision is achieved by combining with auxiliaries of place. See an excerpt from the hierarchy of the main class of literature below.

- 821 Literatures of Individual languages [Thing, subdivided by Kinds]
- 821.1/.2 Literature in Indo-European languages
- 821.1 Literatures of Indo-European languages of Europe
- 821.11 Literature in Germanic languages
- 821.111 English literature

- 821.112 Literature in West Germanic languages (other than English)
- 821.112.2 German literature
- 821.112.28 Yiddish literature

Further specification of literature can be achieved through combination with common auxiliaries of place, time, etc. and special auxiliary facets of form, periods and literary criticism (see further in Section 3.1 below).

Special auxiliaries are concepts that are shared by subject areas but only within a specific class in the main table. Codes from these facets can be understood as “specifiers” and can only be used in their defined meaning if they are attached to a main number in their area of application. In the main tables, auxiliaries are usually processes, materials, tools, products in industries or technology, or periods, literary forms and genres within 82 “Literature” (shown in Section 3.1 below). In this sense, they can be considered as facets of a main knowledge area. Different facet indicators are used according to different levels of specificity or scope of usage. Three levels are envisaged: *-1/-9* for concepts that are normally applicable within an entire discipline; *.01/.09* introduces concepts with a narrower scope of use within a subfield of knowledge; and, *`1/`9* are reserved for special auxiliaries with a narrower or fairly specialised level of applicability. The example below shows a sub-facet array of special auxiliaries *-1/-9* that can be combined in any subdivision of class 62 “Engineering.”

- 62-1/-9 Special auxiliary subdivision for technology in general
- 62-1 General characteristics of machines etc.
- 62-2 Fixed and movable parts, components of machines
- 62-3 Fluid control parts and drives. Valves, closures etc.
- 62-4 State, condition, form of materials, surfaces, objects, products
- 62-5 Operation and control of machines and processes
- 62-6 Fuel and other heat-source characteristics of machinery and installations
- 62-7 Devices for servicing, maintenance and protection of machines
- 62-8 Machines according to motive power, propulsive force. Source of energy of machine
- 62-9 Variables, conditions and characteristics of production processes, plant and equipment

It is worth noting that some fields in UDC, e.g., medicine or technology, are using all types of special auxiliaries.

Connecting symbols comprise signs that can be used to coordinate two subjects of equal importance (+), ex-

tend the scope meaning of one class by several consecutive classes (/), relate two classes in situations when one class can be understood only in so far as it is related to the other, or when the subject of one class should always be understood as a subject of treatment by the other subject (: and ::) or subgrouping of two classes when these are related to a third class as a whole. In the process of indexing, if required, UDC numbers can also be extended by words A/Z or by codes from another system.

2.2 Synthesis of facets in UDC

Figure 2 illustrates the way synthesis in UDC is intended to work, taking as example class 81 “Linguistics.” The structure and its representation through the notational system impose a fair level of rigour and formality but equally allow for a high level of flexibility in implementing conceptual and semantic formality (Robinson 2003).

However, a lack of more precise structural and syntactical rules has been the object of frustration for those involved in the automation of the UDC, as far back as the late 1960s (Freeman and Atherton 1968; Caless 1969; Perreault 1969a 1969b; Rigby 1971).

3.0 Systematic implementation of facets in the restructuring of UDC

The drive to develop systematically a faceted structure for the UDC started with the Dorking conference in 1957. The FID Central Classification Committee (FID CCC) embraced the conference recommendation (UDC Reform Work 1975 6; Classification Research Group 1997) that “the most helpful form of classification scheme for information retrieval is one which groups terms into well-defined categories, which can be used independently to form compounds, and within which the terms can be

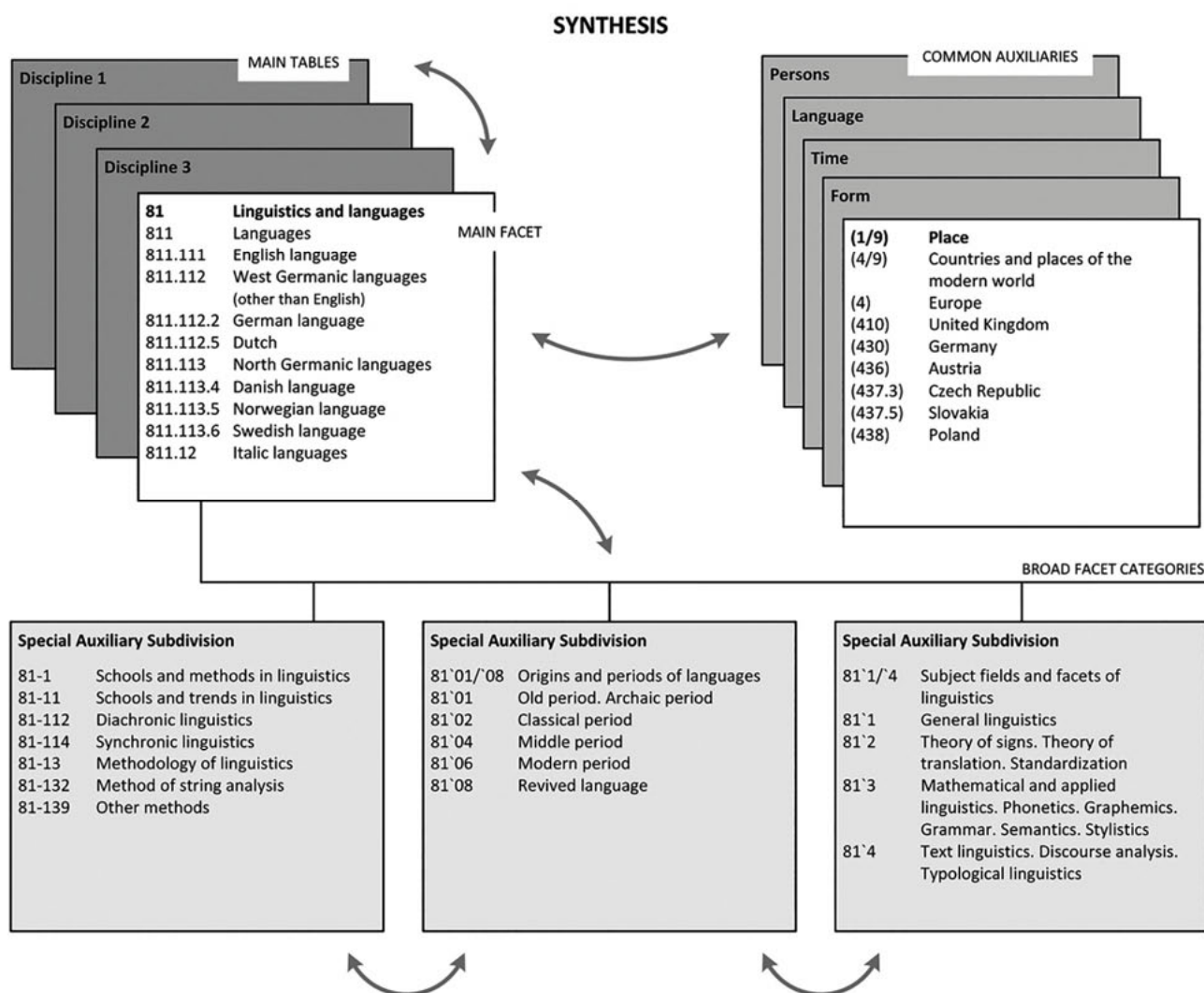


Figure 2. Facet representation and synthesis in UDC.

arranged in hierarchies where this conforms to the recognized structure of relations between them.”

The idea was pursued throughout the 1970s by the FID CCC who sought out classification experts such as members of the UK Classification Research Group (CRG) and several other scholars (Dahlberg 1971a; Neelameghan 1976). At the time, the full edition of UDC contained over 200,000 classes (in four languages) and restructuring of the scheme into a fully faceted classification was deemed by many to be too challenging. Nevertheless, changes towards a faceted structure continued to be introduced, albeit without an overarching UDC-specific theoretical framework and without procedural guidelines. Yet, faceting of the UDC is viewed as the only way forward and has become a long-term commitment of the editorial teams (McIlwaine 1990; McIlwaine and Williamson 1994; Slavic et al. 2008). This section will discuss two main approaches to the organization and presentation of faceted schedules in UDC: a) facets as “relational tables” with simple concepts; and, b) facets enumerating simple and compound concepts.

3.1 Facets as “relational tables” with simple concepts

The first of the UDC classes to be developed in a fully faceted fashion were linguistics and literature, both derived from the classification of languages in the Common auxiliaries of languages (Table 1c). Revision planning started at the end of the 1970s and was completed at the end of the 1980s resulting in the publication of the fully faceted classes 80 “General questions relating to both linguistics and literature. Philology,” 81 “Linguistics” and 82 “Literature” in *Extensions and Corrections to the UDC 14* (1992) and in the first version of the UDC Master Reference File in 1993. This revision is significant as it has revealed concerns associated with matters of structure and its presentation.

Languages represent the main facet (thing) in both classes linguistics and literature, and they are assumed under 811 “Individual languages” and 821 “Individual literatures.” However, the schedules do not list languages in these two classes. These are enumerated in UDC only once, in “Table 1c - Common auxiliaries of language,” which contains hierarchies of languages with the facet indicator = (e.g., =111 ‘English’). These common auxiliaries are used to indicate that a document is written in a certain language (e.g., 51(07)=111 “Mathematics - textbook - in English language”). More importantly, they are also used to build hierarchies of “Table 1f - Common auxiliaries of human ancestry, ethnic grouping and nationality” and main classes 811 “Linguistics” and 821 “Literature.”

The UDC notation for main subjects in linguistics and literature is derived (when required) by adding notation from the table of language to the basis notation for linguistics and literature as follows:

811.111 English language [as a subject, i.e., Linguistics]
821.111 English literature

In terms of class subdivision presented in the schedules, the two main classes 81 “Linguistics” and 82 “Literature” contain only facets of special auxiliaries listing kinds and parts in the case of literary forms and genres for literature as follows:

MAIN FACET

82 Literature

SPECIAL AUXILIARY FACETS

82-1/-9 Special auxiliary subdivision for literary forms, genres

82-1 Poetry. Poems. Verse

82-2 Drama. Plays

82-3 Fiction. Prose narrative

82-4 Essays

82-5 Oratory. Speeches

82-6 Letters. Art of letter-writing. Correspondence

82-7 Prose satire. Humour, epigram, parody etc.

82-8 Miscellanea. Polygraphies. Selections

82-9 Various other literary forms

82.02/.09 Special auxiliary subdivision for theory, study and technique of literature

82.02 Literary schools, trends and movements

82.09 Literary criticism. Literary studies

821 Literatures of individual languages

[to be derived from common auxiliaries of languages Table 1c]

This approach is similar to the way tables in relational databases are designed; it is clean conceptually, logical and easy to understand. Schedules are very concise and structurally simple to automate and use. All combinations of concepts are happening at the users’ end, i.e., in the process of indexing. Consequently, the schedules show these classes as being virtually empty, consisting only of rules for derivation, examples of derivation and examples of combinations. Obviously, this has a huge impact on the usability of the schedules as a source of terminology which is why many have proposed that, should UDC be completely faceted, the schedules should have an addition, i.e., a “separate part,” which would list compounds and complex numbers. This was voiced particularly clearly by Dahlberg (1971a) and by a FID CCC report *UDC Reform Work* (1975).

In 1990, however, when the UDC Master Reference File (UDC MRF) database was designed, it was decided that it would be sufficient if all these compounds and

complex UDC number combinations containing important search terms were presented as “examples of combinations” within the record of a class to which combination can be hierarchically subsumed. The provision was made for a fairly limited field structure to allow for “examples of combinations” parsing and sorting. The main UDC classmark field is designed to hold “a simple UDC number” only. The data model and data element schema in the UDC MRF database reflect the UDC schedule organization that was decidedly stretched by the development described in the following section.

3.2 Facets enumerating simple and compound concepts

This is a new approach of presenting UDC schedules related to the most recent phase in the faceting of UDC, which was based on the FAT framework applied in *BC2*. The experiment started on UDC classes medicine and religion, the latter being completed and introduced in UDC in 2000 and throughout subsequent years. The main facet (thing) contains a hierarchy of faiths represented by simple UDC notation (main numbers) and all other facets (property, processes, operations, etc.) are presented as an array of facets within special auxiliary tables -1/-9. Given the long-term revision plan and the resolution of applying FAT as exemplified in the *BC2* structure, it is certainly appropriate to compare the treatment of the “new” class religion according to the *BC2* model with its treatment according to UDC principles.

3.2.1 Presentation of compound and/or complex classes

BC2 schedules present broad facets containing simple concepts but then expand by enumerating synthesised notations. This has affected UDC in so far as the new religion class enumerates simple, compound classmarks as if they belong to the main hierarchy of simple numbers. An extract from the faceted schedules 27 “Christianity” shows the list of classes in orthodox church displayed in the form of a main UDC hierarchy:

271.2 Orthodox Church
 271.2-1 The Orthodox ‘Tradition’
 271.2-284 Doctrinal statements. Symbolical Books
 271.2-284.7-247 The Gospel Book
 Example(s) of combinations:
 271.2-282.7-247-536.36 Prostration before the Gospel Book
 271.2-472-022.43 The Longer Catechism
 271.2-523.46 Side rooms, chambers: diaconicon
 ...

271.3 Separated churches
 271.4 Uniate churches

Normally UDC would treat all these compounds as examples of combination:

272.2 Orthodox Church
 Example(s) of combinations:
 271.2-1 The Orthodox ‘Tradition’
 271.2-284.7-247 The Gospel Book
 271.2-284.7-247-536.36 Prostration before the Gospel Book
 271.2-284 Doctrinal statements. Symbolical Books
 271.2-472-022.43 The Longer Catechism
 271.2-523.46 Side rooms, chambers: diaconicon
 271.3 Separated churches
 271.4 Uniate churches

3.2.2 Development of new hierarchies using compound notations as a basis

In *BC2*, a class introduced within a broad facet can not only be repeated in combination with other classes and be listed as a part of the main hierarchy, but it can be further subdivided at the place where it was used as a compound, and alphabetical extension is freely used to form hierarchy. For example:

2-15 Named gods
 [special auxiliary, from facet 2-1 ‘Nature of religion. Phenomenon of religion’]
 233 Hinduism [Main simple number]
 233-13 The Holy. Brahma Absolute being
 233-14 God(s) and goddess(es)
 233-158D Devi
 233-158G Ganesh
 233-158K Kali

Unlike *BC2*, UDC does not allow that pre-combined notations to be used as bases for further subdivision. Only simple classes can have subdivisions in UDC. Furthermore, alphabetical extension is treated by UDC as a “non-UDC” code and cannot be used for the development of a hierarchy. Only simple classes can have subdivisions in UDC. Named Gods in this case Devi, Ganesh or Kali can be presented only as examples of alphabetical extension. As follows:

233 Hinduism
 Example(s) of combination:
 233-158 Hinduism - Named gods

234-158KAL Hinduism - Named gods - Kali
 234-158SHI Hinduism - Named gods - Shiva

- [skipped hierarchical level]
 2-265 Other literary forms
 [the end of subdivision]

3.2.3 Caption formulation

In *BC2*, it is normal to describe with a single verbal expression a compound subject, relying on the fact that the terms describing the class in full are already listed in the hierarchy above or in facets where the terms were introduced for the first time, for example:

2-23 Sacred books. Scriptures
 [special auxiliary, from facet 2-2 'Evidences of religion']
 27 Christianity [Main number]
 27-23 Bible

This approach leads to a "loss" of terms in the captions of compound classes. Synthesis in UDC requires that class descriptions be expressive of concept combinations so that online users can search the presence of component terms in all verbal descriptions, a crucial point since the UDC MRF does not have a subject alphabetical index. The way UDC would normally deal with this situation is to display fully descriptive captions as follows:

27 Christianity
 Example(s) of combination:
 27-23 Christianity - Sacred books. Bible

3.2.4 Increased use of differential facets (two-level facets)

The usage of differential facets occurs when facets built within special auxiliaries can be further developed to accommodate a more specific subject. This is a source of many unnecessary duplications especially in class 6 "Applied Sciences," whereby one can find a facet of special auxiliaries of processes, operations or tools being repeated under each technology. One priority in UDC maintenance is to reduce these situations and make provision for each individual concept to be presented in one place only, and controlled by a single rule of synthesis to allow automation of synthesis and semantic linking. However, this commitment was not followed in the revision of the class religion as we find special auxiliaries that are originally introduced at the beginning of class 2, being further developed within individual religions.

2-1/-9 Special auxiliary subdivision for religion
 [presented at the beginning of class 2 and applicable to all faiths]
 2-1 Evidences of religion

Later under 233 "Hinduism," the special auxiliary 2-265 is developed as follows:

233-265 [skipped]
 233-265.3 Itihasa. Epics and sagas
 233-265.32 Ramayana
 233-265.33 Mahabharata
 233-265.34 Bhagavadgita
 233-265.35 Puranas

The "differential facets" as applied in religion are rather difficult to manage. When these pre-combined classes appear in a class subdivision one cannot determine whether the notation created, e.g., -265.32 is a completely new UDC notation that represents a new special concept only to be used in hinduism or whether this notation (i.e., the concept it presents) already exists in its general sense in the main special auxiliary table at the beginning of class 2, i.e., on the level above. In other words, it is not clear whether the compound concept is an example of combination of two concepts that already exist in UDC, or a concept that is introduced for the first time (cf. Broughton 2010 and Gnoli 2011).

4.0 Recommendations

If we observe the literature dealing with issues of UDC restructuring, we can see that concerns regarding the connections between faceted structure, notation and data presentation for automation purposes were voiced from the very early stages of the restructuring of the UDC. There are seven areas of concern, and they are all interconnected.

4.1 Facet analysis based on a FAT framework

Principles of structuring UDC according to universal facet categories should be properly described in terms of the theoretical framework applied and in terms of how it fits into the UDC system as a whole (e.g., relationship of the existing classes, common auxiliary facets and schedule presentation). This also entails an implementation of more rigorous principles for hierarchical subdivisions (Kyle and Vickery 1961; Vickery 1961; Dahlberg 1971a 1971b; Neelameghan 1976; McIlwaine 1990; McIlwaine and Williamson 1994; Slavic 2008; Broughton 2010; Vukadin and Slavic 2014).

4.2 Formalization of notational system

The UDC notational system can be further improved through expressiveness, formalization and strict rules for notation use and re-use (Perreault 1969a; Caless 1969; Freeman and Atherton 1968 1969; Lloyd 1972 Dahlberg 1971a; Strachan 2000; Slavic and Cordeiro 2004; Gnoli 2011).

4.3 Rules for UDC number formation

A high level of synthesis requires an equally high level of formalization of syntax based on rules for phase relationships, extensions, grouping and subordination. These should be applied in conjunction with guidelines for content analysis and indexing (Perreault 1969a 1969b; Caless 1969) and in relation to automation of rules for the purpose of notation building, parsing and association of notation with verbal expressions.

4.4 Semantic linking

Further improvement and formalization is required for the managing UDC's syndetic structure ("see" and "see also" relationships). In addition to the formalization of hierarchical relationships (mentioned in point 1 above), the supporting UDC MRF database should be able to manage concept identification, terminology control and semantic expansion (Rigby 1971).

4.5 Composition and organization of the schedule (system as a whole).

Although with UDC MRF, we do not speak of different "parts" of UDC (as it would be in the printed schedule), the database must support access to different views of the system (Dahlberg 1971a, 1971b; UDC Reform Work 1975).

4.6 Data management system

Since the UDC MRF was migrated to a relational database in 2009, the database and associated tools are continuously being improved especially in relation to the UDC online service and open linked data exposure. As these developments are nowadays considered to be technical, they are reported in internal documentation rather than in published papers (cf. Freeman 1964).

4.7 Documentation on revision policy, guidelines and procedures

This is possibly the weakest point of the UDC system and very much the priority for the editorial team. Al-

though the requirements have been put forward many times in the past, the progress here is very slow (Lloyd 1972; Wellisch 1975; Slavic et al. 2008). Examples of issues shown in Section 6 and put into historical perspective, here in Section 7, have been dealt with in the UDC MRF database by developing additional data elements and control fields in order to be able to validate the rules and combinations but this has mostly been a damage-control exercise. To prevent further problems, from 2010 onwards, it was decided that major restructuring of the UDC would be postponed until a facet analytical framework is analysed properly and aligned with the existing UDC system, its notation, schedule presentation and the supporting data model and data element schema.

5.0 Conclusion

We are still lacking a standard ontology that would allow the computerised representation of analytico-synthetic classifications. This may be due to the lack of understanding of the requirements for classifications to function in computerised environments. As we have shown, there is a long history of reporting problems with respect to management and use of analytico-synthetic classifications in an online environment. Via examples taken from the latest revision in UDC, we have illustrated the lack of connection between the theoretical framework proposed for building faceted classifications and the required level of formality in the presentation of structure and syntax in schedules. In the case of UDC, the theoretical model adopted from *BC2* was not aligned properly with the underlying data, and lacked the level of notational and presentation formality needed to fit in the existing database system that had been designed to manage synthesised UDC notations.

Increased synthesis means greater reliance on syntax rules. There is a greater need to understand and define the type of synthesis with respect to facet analysis and synthesis. Equally one needs to understand in which way synthesised numbers will be entered in the UDC MRF database or presented in printed schedules. UDC has a well-defined and well-understood notational system and, if this is to support a specific FAT framework, one would need to understand which kind of notation and notation building are used for which kind of facet categories (Gnoli 2011).

Increased synthesis also affects the formulation of captions and the changes needed to support the management of class description and concept management in the UDC MRF database. Our experience with UDC shows that true progress with classification can only be made if there is appropriate data schema and database structure to support the required functionality. Any sub-

stantial structural changes come with further requirements in the careful management of these changes, and the need to create additional tools to manage structural alignments and concordances between the old and new structures.

References

- Broughton, Vanda. 2010. "Concepts and Terms in the Faceted Classification: The Case of UDC." *Knowledge Organization* 37: 270-79.
- Caless, Thomas W. 1969. "Subject Analysis Matrices for Classification with UDC." In *Proceedings of the First Seminar on UDC in a Mechanized Retrieval System conducted by R. R. Freeman and P. Atherton, Copenhagen, 2[nd]-6th September, 1968*, ed. R. Mølgaard-Hansen and Malcolm Rigby. FID/CR Report 9. Copenhagen: Danish Centre for Documentation, 114-20.
- Classification Research Group 1997. "The Need for a Faceted Classification as the Basis of All Methods of Information Retrieval: Memorandum of the Classification Research Group." In *From Classification to "Knowledge Organization": Dorking Revisited or "Past is Prelude"*, ed. Alan Gilchrist. FID occasional paper 14. The Hague: International Federation for Information and Documentation, 1-9. First published in *Library Association Record* 57, no. 7: 262-8.
- Dahlberg, Ingetraut. 1971a. "Possibilities for a Reorganization of the UDC." In *Proceedings of the Second Seminar on UDC and Mechanized Information Systems, conducted by R. R. Freeman, Frankfurt, 1-5 June, 1970*, ed. R. Mølgaard Hansen and Margit Westring-Nielsen. FID/CR Report 11. Copenhagen: Danish Centre for Documentation, 193-210.
- Dahlberg, Ingetraut. 1971b. "Possibilities for a New Universal Decimal Classification." *Journal of Documentation* 27: 18-36.
- Freeman, Robert R. 1964. "Computers and Classification Systems." *Journal of Documentation* 20: 137-43.
- Freeman, Robert R., and Pauline Atherton. 1968. "File Organization and Search Strategy Using UDC in Mechanized Reference Retrieval Systems." In *Mechanized Information Storage, Retrieval and Dissemination: Proceedings of the FID/IFIP Joint Conference, Rome, June 14-17, 1967*, ed. K. Samuelson. Amsterdam: North-Holland Publishing, 122-59.
- Freeman, Robert R., and Pauline Atherton. 1969. "Final Report of the Research Project for the Evaluation of the UDC as the Indexing Language for a Mechanized Retrieval System." In *Proceedings of the First Seminar on UDC in a Mechanized Retrieval System conducted by R. R. Freeman and P. Atherton, Copenhagen, 2[nd]-6th September, 1968*, ed. R. Mølgaard-Hansen and Malcolm Rigby. FID/CR Report 9. Copenhagen: Danish Centre for Documentation, 39-56.
- Gnoli, Claudio. 2011. "Facets in UDC: A Review of Current Situation." *Extensions and Corrections to the UDC* 33: 19-36.
- Kyle, Barbara. 1961. "The UDC: A Study of the Present Position and of its Future Developments, with Particular Attention to those Schedules which Deal with the Humanities, Arts, and Social Sciences." *UNESCO Bulletin for Libraries* 15: 53-69.
- Lloyd, G. A. 1972. "UDC: Revise or Relegate? FID's Standard Reference Code Project and UDC Improvement Programme." *Aslib Proceedings* 24: 580-7.
- McIlwaine, Ia C. 1990. "The Work of the System Development Task Force." In *The UDC: Essays for a New Decade*, ed. Alan Gilchrist and David Strachan. London: Aslib, 19-28.
- McIlwaine, Ia C., and Nancy J. Williamson. 1994. "A Feasibility Study on the Restructuring of the Universal Decimal Classification into a Fully-Faceted Classification System." In *Knowledge Organization and Quality Management: Proceedings of the Third International ISKO Conference 20-24 June, 1994, Copenhagen, Denmark*, ed. Hanne Albrechtsen, S. Oernager. Frankfurt: Indeks Verlag, 406-13.
- Neelameghan, Arashanipalai. 1976. "A Theoretical Foundation for UDC: Its Need and Formulation." In *Proceedings of the international symposium 'UDC in Relation to Other Indexing Languages' held in Herceg Novi, Yugoslavia, June 28-July 1, 1971*. Beograd: Jugoslovenski centar za tehničku i naučnu dokumentaciju, 137-73.
- Otlet, Paul. 1895-6. "Sur la structure des nombres classificateurs." *Bulletin de l'Institut international de bibliographie* 1: 230-43.
- Otlet, Paul. 1896. *Règles pour les développements à apporter à la Classification Décimale*. Bruxelles: Office international de Bibliographie.
- Otlet, Paul. 1934. *Traité de documentation: le livre sur le livre; théorie et pratique*. Bruxelles: Editions Mundaneum. http://lib.ugent.be/fulltxt/handle/1854/5612/Traite_de_documentation_ocr.pdf
- Otlet, Paul. 1990. "On the Structure of Classification Numbers." In *International Organisation and Dissemination of Knowledge: Selected Essays of Paul Otlet*, trans. W. Boyd Rayward. Amsterdam: Elsevier, 51-63. <http://hdl.handle.net/2142/8645>
- Perreault, Jean. M. 1969a. *Towards a Theory for UD: Essays Aimed at Structural Understanding and Operational Improvement*. London: Clive Bingley.
- Perreault, Jean M. 1969b. "On the Colon in UDC and on the Boolean Operators." In *Proceedings of the First Seminar on UDC in a Mechanized Retrieval System conducted by R. R. Freeman and P. Atherton, Copenhagen, 2[nd]-6th Sep-*

- tember, 1968, ed. R. Mølgaard-Hansen and Malcolm Rigby. FID/CR Report 9. Copenhagen: Danish Centre for Documentation, 182-92.
- Rigby, Malcolm. 1971. "The role of the UDC in Automated Information and Data Systems." In *Proceedings of the Second Seminar on UDC and Mechanized Information Systems, conducted by R. R. Freeman, Frankfurt, 1-5 June, 1970*, ed. R. Mølgaard Hansen and Margit Westring-Nielsen. FID/CR Report 11. Copenhagen: Danish Centre for Documentation, 182-92.
- Robinson, G. (2003). "Citation Order in UDC," *Extensions and Corrections to the UDC* 25:19-27.
- Slavic, Aida. 2008. "Faceted Classification: Management and Use." *Axiomathes* 18: 257-71. <http://arxiv.org/abs/1705.07047>.
- Slavic, Aida, and Ines. Cordeiro. 2004. "Core Requirements for Automation of Analytico-Synthetic Classifications." In *Knowledge Organization and the Global Information Society: Proceedings of the Eighth International ISKO Conference, 13-16 July 2004 London, UK*, ed. Ia C. McIlwaine. Advances in Knowledge Organization 9. Würzburg: Ergon Verlag, 187-92. <http://hdl.handle.net/10150/105814>
- Slavic, Aida, Maria Inês Cordeiro, and Gerhard Riesthuis. 2008. "Maintenance of the Universal Decimal Classification: Overview of the Past and Preparations for the Future." *International Cataloguing and Bibliographic Control* 37: 23-9.
- Strachan, David. 2000. "A Chance to Make UDC Notation More Computer-Friendly." *Extensions and Corrections to the UDC*. 22: 43-6.
- UDC MRF11. 2012. Dataset. The Hague: UDC Consortium.
- UDC Reform Work. 1975. Internal FID document C 75-35 (30 December 1975) to all CCC members for discussion in CCC-meeting on 2-4 February 1976.
- Vickery, B. C. 1961. "The Universal Decimal Classification and Technical Information Indexing." *UNESCO Bulletin for Libraries* 15: 126-38.
- Vukadin, Ana, and Aida Slavic. 2014. "Challenges of Facet Analysis and Concept Placement in Universal Classifications: The Example of Architecture in UDC." In *Knowledge Organization in the 21st Century: Between Historical Patterns and Future Prospects; Proceedings of the Thirteenth International ISKO Conference 19-22 May 2014, Kraków, Poland*, ed. Wielaw Babik. Würzburg: Ergon Verlag, 236-43.
- Wellisch, Hans H. 1975. "Dewey Decimal Classification, Universal Decimal Classification, and the Broad System of Ordering: The Evolution of Universal Ordering Systems." In *Classification Systems: The Dewey Centennial*, ed. Kathryn Luther Henderson. Urbana-Champaign, Illinois: Graduate School of Library Science, 113-23.