The Hague
19-20 September 2011

UDC International Seminar

http://seminar.udcc.org/2011/

CLASSIFICATION
ontology
&

Formal
Approaches
and Access
to Knowledge
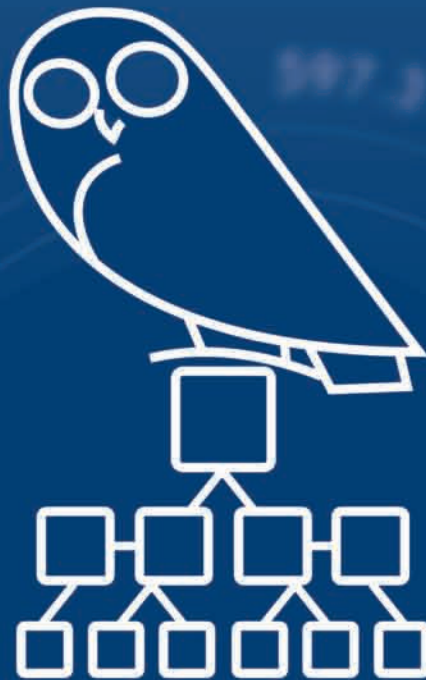
PROGRAMME

International UDC Seminar 2011
Conference programme at a glance

Day One - 19 September

09:00   Registration
09:45   Conference opening: Welcome from UDC Consortium (Alan Hopkinson, Chairman), Koninklijke Bibliotheek (representative), Programme Committee (Aida Slavic)
10:00   **On being the same: keynote address**
        Patrick Hayes
11:00   Coffee/Tea

Session 1. The role of classification and ontology on the Web
Chair: Dagobert Soergel

11:30   **Classification, collaboration and the Web of data**
        Dan Brickley
12:10   **Issues in publishing and aligning Web vocabularies**
        Guus Schreiber
12:50   **The concepts of knowledge organization systems as hubs in the Web of data**
        Thomas Baker
13:30   Lunch | **14:10 Brief Posters introduction (conference room)**

Session 2. Classifications and ontologies on their own terms
Chair: Roberto Poli

14:30   **Approaches to providing context in knowledge representation structures**
        Barbara H. Kwaśnik
15:10   **Interactions between elementary structures in universes of knowledge**
        Richard Smiraglia; Charles van den Heuvel; Thomas M. Dousa
15:40   **Demystifying ontology**
        Emad Khazraee; Xia Lin
16:10   Coffee/Tea

Session 3. Classification meets the Web
Chair: Antoine Isaac

16:40   **Interoperability of knowledge organization systems with and through ontologies**
        Daniel Kless; Jutta Lindenthal; Simon Milton; Edmund Kazmierczak
17:10   **Towards the integration of knowledge organization systems with the linked data cloud**
        Vincenzo Maltese; Feroz Farazi
17:40   **Classification and reference vocabulary in linked environment data**
        Maria Rüther; Joachim Fock; Thomas Schultz-Krutisch; Thomas Bandholtz
18:10   Reception

Day Two - 20 September

Session 4. Classification and ontology in specific subjects
Chair: Ia C. McIlwaine

09:00   **Ontologies and classification of chemicals: can they help each other?**
        Andrew Buxton

09:30   **Content analysis and classification in mathematics**
Wolfram Sperber; Patrick D. F. Ion
10:00   Coffee/Tea

Session 5. Categories and relations: key elements of ontologies
Chair: Barbara Kwaśnik

10:30   **Ontology as categorial analysis**
Roberto Poli
11:10   **Towards a relation ontology for the Semantic Web**
Dagobert Soergel
11:50   **Relations in the notational hierarchy of the Dewey Decimal Classification**
Rebecca Green; Michael Panzer
12:30   Lunch

Session 6. Modelling concepts and structures in analytico-synthetic classifications
Chair: Richard Smiraglia

13:30   **A faceted classification of general concepts**
Ingetraut Dahlberg
14:00   **Representing the structural elements of a freely faceted classification**
Claudio Gnoli; Philippe Cousson; Tom Pullman; Gabriele Merli; Rick Szostak
14:30   **Facet analysis as a tool for modelling subject domains and terminologies**
Vanda Broughton
15:00   **Analytico synthetic approach for handling knowledge diversity in media content analysis**
Devika P. Madalli; A. R. D. Prasad
15:30   Coffee/Tea

Session 7. Transforming and extending classification systems
Chair: Gordon Dunsire

16:00   **Extending models for controlled vocabularies to classification systems: modelling DDC with FRSAD**
Joan S. Mitchell; Marcia Lei Zeng; Maja Žumer
16:40   **Transformation of a legacy UDC-based classification system: exploiting and remodelling semantic relationships**
Fran Alexander; Andy Heather
17:10   Panel discussion
18:00   Conference close

Posters
Chairs: Edgardo Civallero & Sofia Kapnisi

Posters will be displayed throughout the conference in the conference room. Following a brief introductory presentation in the conference room at 14:10 on Day 1, authors will be available to answer questions during coffee breaks on both days.

The evolution of knowledge, and its representation in classification systems
Andrea Scharnhorst; Almila Akdag Salah; Krzysztof Suchecki; Cheng Gao; Richard P. Smiraglia
Visualizing universes of knowledge: designs and visual analysis of the UDC
Charles van den Heuvel; Almila Akdag Salah; Knowledge Space Lab
UDC as a knowledge framework for building a civil engineering ontology: a practical approach to knowledge representation and visualization
Ricardo Eito-Brun; Alfredo Calosci

Special feature: Selection of posters from the exhibition "Places & Spaces: Mapping Science"

On being the same
Keynote Address

Patrick Hayes

Institute for Human & Machine Cognition (USA)

"Logic is like cricket. It is admirable so long as you are playing by the rules. But what happens to your game of cricket when somebody suddenly decides to bowl with a football or bat with a hockey-stick? Because that is what is continually happening in life".

The Manticore, by Robertson Davies

The Semantic Web and the rise of linked data is producing hundreds of millions of triples, all written in one (rather simple) logical notation by a plethora of authors. Not surprisingly, there are some oddities and internal inconsistencies in this data. One of the most common arises from assertions that two names or descriptions refer to the same thing, when in fact they are closely related but not in fact identical; and the worst of these is the so-called use/mention confusion, where a thing is said to be the same as a description of it. Use/mention confusions seem to be common largely because human thinking finds them very natural, even though formal logics find them disastrous. Why?

To deal adequately with the actual logic of human intuitions about linguistic meaning we will need semantic insights which better reflect the richly intertwined ways in which human language use weaves together concepts and descriptions. Most logical reasoning is based upon a referential style of interpretation which treats names and descriptions as ways to refer to things: the logic makes statements about these things. Equality is then a very simple matter. But human language often uses descriptions in a subtly different way, where they retain a meaning through changes in interpretation. This distinction is traditionally referred to de re – of the thing – versus de dicto – of the speech – reasoning. For example 'the number of planets' refers, in fact, to the number eight, but seems to carry more meaning than a simple numeral. The de re/de dicto distinction is most visible in modal logics formalizing statements of belief or necessity, and seems to be behind many of the use/mention confusions, though not all of them.

Classification however seems to be based upon a third mode, which we might call of the concept. We will explore this idea and its ramifications for statements of identity, using some recent ideas from formal logics designed to describe propositions.

PATRICK HAYES is Senior Research Scientist at Florida Institute for Human & Machine Cognition (IHMC). He has been a professor of computer science at the University of Essex and philosophy at the University of Illinois, and the Luce Professor of cognitive science at the University of Rochester. He has been a visiting scholar at Université de Genève and the Center for Advanced Study in the Behavioral Studies at Stanford, and has directed applied AI research at Xerox-PARC, SRI and Schlumberger, Inc. At various times, Pat has been secretary of AISB, chairman and trustee of IJCAI, associate editor of Artificial Intelligence, a governor of the Cognitive Science Society and president of AAAI. Pat's research interests include knowledge representation and automatic reasoning, especially the representation of space and time; the Semantic Web; ontology design; image description and the philosophical foundations of AI and computer science. During the past decade Pat has been active in the Semantic Web initiative, largely as an invited member of the W3C Working Groups responsible for the RDF, OWL and SPARQL standards. Pat is a member of the Web Science Trust and of OASIS, where he works on the development of ontology standards.

Classification, collaboration and the Web of data

Dan Brickley

VU University Amsterdam (The Netherlands)

This talk focuses on the relationship between subject classification and 'Web of data' trends around RDF, OWL and SKOS. In particular it sketches ways in which factual and ontological data can be used alongside subject classification and on the practical possibilities this creates — for collaboration amongst vocabulary and dataset maintainers, and in user-facing applications. Although factual ontologies and subject classification systems typically serve different purposes, they often overlap in topical coverage and are can all be expressed using shared underlying 'Web of data' technologies, such as RDF. With each passing week, new datasets—whether scientific, library, cultural heritage, governmental or social—are published as 'linked data', with RDF vocabularies, OWL ontologies and SKOS schemes as the representational 'glue' that holds the whole thing together. Factual representations of people, places and things serve as bridges between the subject classification world and the world of general Web data. Despite this, we have not yet collectively produced 'best practice' guidance that show how such linkage can be created, curated and exploited using practical, modern Web tools. A goal of this talk is to motivate such collaboration, and to suggest some priorities for the short and medium term.

Issues in publishing and aligning Web vocabularies

Guus Schreiber

VU University Amsterdam (The Netherlands)

Knowledge organization systems (KOS), such as vocabularies, thesauri and subject headings, contain a wealth of knowledge, collected by dedicated experts over long periods of time. These knowledge sources are potentially of high value to Web applications. To make this possible we need methods to publish these systems and subsequently clarify their relationships, also called "alignments". In this talk Guus discusses methodological issues in publishing and aligning classification systems on the Web. With regards to publication of Web vocabularies he explains the basic principles for building a SKOS version of a vocabulary and illustrates this with examples. In particular, he discusses how one should prevent information loss, i.e. constructing a SKOS version that contains all information contained in the original vocabulary model. The talk also examines the role of RDF and OWL in this process. Web vocabularies derive much of their added value from the links they can provide to other vocabularies. He explains the process of vocabulary alignment, including the choice of alignment technique. Particular attention is paid to an evaluation of the process: how can one assess the quality of the resulting alignment? Human evaluators often play an important role in this process. Guus concludes by showing some examples of how aligned Web vocabularies can be used to create added value to applications.

The concepts of knowledge organization systems as hubs in the Web of data

Thomas Baker

Dublin Core Metadata Initiative (USA)

The domain name system of the World-Wide Web provides a managed space of globally unique identifiers resolvable to a globally distributed set of information resources. When the concepts of a knowledge organization system (KOS) are identified using URIs, the KOS functions as a "hub" for accessing resources tagged with its concepts. Resource Description Framework (RDF) triples, consisting of a subject, a predicate, and an object, joined on the basis of matched URIs, form the spokes of these hubs. New sources of metadata can be dynamically integrated into an infinitely "expandable" description. Term-to-term alignments with other KOSs increase the conceptual reach of a KOS, while concept labels in multiple languages increase its reach linguistically. This talk illustrates the mechanics of merging linked data triples with reference to KOSs that function as hubs.

Day 1 - Session 2
Chair: Roberto Poli
Classifications and ontologies on their own terms

Approaches to providing context in knowledge representation structures

Barbara H. Kwaśnik

Syracuse University (USA)

The power of knowledge structures is to represent, to contextualize, to communicate, and to help structure knowledge in a useful way. Traditional classifications tackle the challenges of creating knowledge structures for a wide-ranging set of concepts and are set up to reflect cumulated literary and scientific warrant for many purposes, but especially the useful ordering of knowledge. Ontologies focus on modelling domains with a vigorous dedication to eliciting the most useful entities and relationships for that domain. Both leverage structure and relationships to provide a way of representing not only the entities under consideration but also the way they work in a network of meaning. At the same time the foundation of many knowledge structures is bounded by a given perspective reflecting the purposes of that structure. This paper examines two cases, the structure of knowledge as expressed in the curriculum at an American university, and the notion of "cohabitation" as a construct that shifts in meaning over time and situations. In both cases context helps define meaning.

Interactions between elementary structures in universes of knowledge

Richard P. Smiraglia

University of Wisconsin (USA)

Contrasts in 20th century classification theory relate to a transition from a universe of "knowledge" system towards one of "concepts." Initiatives to develop a Simple Knowledge Organization Systems (SKOS) standard based on classification schemes and taxonomies within the framework of the

Charles van den Heuvel

Huygens ING Institute The Hague (The Netherlands)

Thomas M. Dousa

University of Illinois at Urbana-Champaign (USA)

Semantic Web (SW) are attempts to bridge the gap. Current knowledge organization systems (KOS) seem to reinforce "syntactics" at the expense of semantics. We claim that all structure is syntactic but knowledge structures need to have a semantic component as well. Therefore we consider classifications as artificial languages. The Universal Decimal Classification (UDC) constitutes a natural language-independent notation system that allows for mediating between concepts and knowledge systems. We discuss an elementary theory of knowledge organization based on the structure of knowledge rather than on the content of documents. Semantics becomes not a matter of synonymous concepts, but rather of coordinating knowledge structures. The interactions between these systems represent interactions between different universes of knowledge or concepts.

Demistifying ontology

Emad Khazraee

Drexel University (USA)

Xia Lin

Drexel University (USA)

The term "ontology" is used in different communities multifariously, in a nearly anarchic way. Ironically, the major function of ontology itself is to explicate the meaning of terms and concepts. Therefore, different conceptions of this term impede collaboration and exchange of expertise between different domains and communities. Thus, providing a clear image of the different notions of ontology is a precondition of communication. This paper studies different notions of ontology and attempts to compare these different conceptions, and to organize them into a model to facilitate collaboration in this field. The use of an ontology gamut model is proposed instead of the one-dimensional ontology spectra used in the past. This model can be used as the basis for agreement to clarify the term ontology among different communities by providing levels of formality, semantics and complexity. The coordinates of each ontology in this gamut helps with understanding the specific conception of that ontology.

Day 1 - Session 3
Chair: Antoine Isaac
Classification meets the Web

Interoperability of knowledge organization systems with and through ontologies

Daniel Kless

University of Melbourne (Australia)

Jutta Lindenthal

Information Consultant (Germany)

Ontologies are increasingly seen as a new type of knowledge organization system (KOS) besides traditional ones such as classification schemes or thesauri. Consequently, there are efforts to compare them with and map them to other KOS. This paper argues that only ontologies for reality representation are useful subjects of such comparisons and mappings. These ontologies are difficult to distinguish from other "data modelling" - types of ontology, since both can be represented through the popular Web Ontology Language (OWL). Data modelling ontologies such as Simple Knowledge Organization Systems (SKOS) are useful instruments for establishing interoperability between KOS in the sense of publishing and

Simon Milton

University of Melbourne (Australia)

Edmund Kazmierczak

University of Melbourne (Australia)

accessing data and data models in a uniform way as well as for relating them to each other. Discriminating these two understandings of ontologies particularly supports comparisons and mappings between traditional KOS and ontologies. In practice, such efforts are still impeded by the absence of standards or guidelines for vocabulary control in ontologies. Moreover, this paper emphasizes that methods for constructing and evaluating reality representation ontologies can be useful to re-engineer traditional KOS. This makes them become more interoperable in the sense of combinable, but also more useful in the sense of improving search expansion results and reusable for different purposes.

Towards the integration of knowledge organization systems with the linked data cloud

Vincenzo Maltese

DISI - University of Trento (Italy)

Feroz Farazi

DISI - University of Trento (Italy)

In representing the shared view of all the people involved, building a knowledge organization system (KOS) from scratch is extremely costly, and it is therefore fundamental to reuse existing resources. This can be done by progressively extending the KOS with knowledge coming from similar KOSs and by promoting interoperability among them. The linked data initiative is indeed encouraging people to share and integrate their datasets into a giant network of interconnected resources. This enables different applications to interoperate and share their data. The integration should take into account the purpose of the datasets, however, and make explicit the semantics. In fact, the difference in the purpose is reflected in the difference in the semantics. With this paper we (a) highlight the potential problems that may arise by not taking into account purpose and semantics; (b) make clear how the difference in the purpose is reflected in totally different semantics and (c) provide an algorithm to translate from one semantics into another as a preliminary step towards the integration of ontologies designed for different purposes. This will allow reusing the ontologies even in contexts different from those in which they were designed.

Classification and reference vocabulary in linked environment data

Maria Rüther

Federal Environment Agency - UBA (Germany)

Joachim Fock

Federal Environment Agency - UBA (Germany)

Thomas Schultz-Krutisch

Federal Environment Agency - UBA (Germany)

The Federal Environment Agency (UBA), Germany, has a long tradition in knowledge organization, using a library along with many Web-based information systems. The backbone of this information space is a classification system enhanced by a reference vocabulary which consists of a thesaurus, a gazetteer and a chronicle. Over the years, classification has increasingly been relegated to the background compared with the reference vocabulary indexing and full text search. Bibliographic items are no longer classified directly but tagged with thesaurus terms, with those terms being classified. Since 2010 we have been developing a linked data representation of this knowledge base. While we are linking bibliographic and observation data with the controlled vocabulary in a Resource Desrcription Framework (RDF) representation, the classification may be

Thomas Bandholtz

*innoQ Deutschland GmbH
(Germany)*

revisited as a powerful organization system by inference. This also raises questions about the quality and feasibility of an unambiguous classification of thesaurus terms.

Day 2 - Session 4

Chair: Ia C. McIlwaine

Classification and ontology in specific subjects

Ontologies and classification of chemicals: can they help each other?

Andrew Buxton

UDC Editorial Team, UDC Consortium (UK)

The chemistry schedule in the Universal Decimal Classification (UDC) is badly in need of revision. In many places it is enumerative rather than synthetic (giving rules for constructing numbers for any compound required). In principle, chemistry should be the ideal subject for a synthetic classification but many common compounds have complex formulae and a synthetic system becomes unwieldy. Also, all compounds belong to several hierarchies, e.g. chloroquin is a heterocycle, an aromatic compound, amine, antimalarial drug, etc. and rules need to be drawn up as to which ones take precedence and which ones should be taken into account in classifying a compound. There are obvious similarities between a classification and an ontology. This paper looks at existing ontologies for chemistry, especially ChEBI which is one of the largest, to examine how a classification and an ontology might draw on each other and what the problem areas are. An ontology might help in creating an index to a classification (for chemicals not listed or to provide access by facets not used in the classification) and a classification could provide a hierarchy to use in an ontology.

Content analysis and classification in mathematics

Wolfram Sperber

Zentralblatt MATH (Germany)

Patrick D. F. Ion

Mathematical Reviews (USA)

The number of publications in mathematics increases faster each year. Presently far more than 100,000 mathematically relevant journal articles and books are published annually. Efficient and high-quality content analysis of this material is important for mathematical bibliographic services such as ZBMath or MathSciNet. Content analysis has different facets and levels: classification, keywords, abstracts and reviews, and (in the future) formula analysis. It is the opinion of the authors that the different levels have to be enhanced and combined using the methods and technology of the Semantic Web. In the presentation, the problems and deficits of the existing methods and tools, the state of the art and current activities are discussed. As a first step, the Mathematical Subject Classification Scheme (MSC), has been encoded with Simple Knowledge Organization System (SKOS) and Resource Description Framework (RDF) at its recent revision to MSC2010. The use of SKOS principally opens new possibilities for the enrichment and wider deployment of this classification scheme and for machine-based content analysis of mathematical publications.

Categories and relations: key elements of ontologies

Ontology as categorial analysis

Roberto Poli

University of Trento (Italy)

Ontological categories are organized along a number of different dimensions. The simplest is the distinction between categories that apply to all entities, both real and ideal, and categories that apply only to some families of entities. More complicated is the analysis of the relations that connect categories one to another. Two different exemplifications of the latter case are provided, i.e., the form of duality linking some paired categories and the relations of superformation and superconstruction that connect levels of reality. Furthermore, an in-depth analysis of the category of temporality is presented. Ideas previously advanced by Nicolai Hartmann are exploited throughout the paper.

Towards a relation ontology for the Semantic Web

Dagobert Soergel

University at Buffalo (USA)

The Semantic Web consists of data structured for use by computer programs, such as data sets made available under the Linked Open Data initiative. Much of this data is structured following the entity-relationship model encoded in RDF for syntactic interoperability. For semantic interoperability, the semantics of the relationships used in any given dataset needs to be made explicit. Ultimately this requires an inventory of these relationships structured around a relation ontology. This talk will outline a blueprint for such an inventory, including a format for the description/definition of binary and n-ary relations, drawing on ideas put forth in the classification and thesaurus community over the last 60 years, upper level ontologies, systems like FrameNet, the Buffalo Relation Ontology, and an analysis of linked data sets.

Relations in the notational hierarchy of the Dewey Decimal Classification

Rebecca Green

OCLC Online Computer Library Center, Inc. (USA)

Michael Panzer

OCLC Online Computer Library Center, Inc. (USA)

As part of a larger assessment of relationships in the Dewey Decimal Classification (DDC) system, this study investigates the semantic nature of relationships in the DDC notational hierarchy. The semantic relationship between each of a set of randomly selected classes and its parent class in the notational hierarchy is examined against a set of relationship types (*specialization*, *class-instance*, several flavours of *whole-part*). The analysis addresses the prevalence of specific relationship types, their lexical expression, difficulties encountered in assigning relationship types, compatibility of relationships found in the DDC with those found in other knowledge organization systems (KOS), and compatibility of relationships found in the DDC with those in a shared formalism like the Web Ontology

Language (OWL). Since notational hierarchy is an organizational mechanism shared across most classification schemes and is often considered to provide an easy solution for ontological transformation of a classification system, the findings of the study are likely to generalize across classification schemes with respect to difficulties that might be encountered in such a transformation process.

Modelling concepts and structures in analytico-synthetic classifications

A faceted classification of general concepts

Ingetraut Dahlberg

ISKO Honorary Member (Germany)

General concepts are all those form-categorial concepts which – attached to a specific concept of a classification system or thesaurus – can help to widen, sometimes even in a syntactical sense, the understanding of a case. In some existing universal classification systems such concepts have been named "auxiliaries" or "common isolates" as in the Colon Classification (CC). However, by such auxiliaries, different kinds of such concepts are listed, e.g. concepts of space and time, concepts of races and languages and concepts of kinds of documents, next to them also concepts of kinds of general activities, properties, persons, and institutions. Such latter kinds form part of the nine aspects ruling the facets in the Information Coding Classification (ICC) through the principle of using a Systematiser for the subdivision of subject groups and fields. Based on this principle and using and extending existing systems of such concepts, e.g. which A. Diemer had presented to the German Thesaurus Committee as well as those found in the UDC, in CC and attached to the Subject Heading System of the German National Library, a faceted classification is proposed for critical assessment, necessary improvement and possible later use in classification systems and thesauri.

Representing the structural elements of a freely faceted classification

Claudio Gnoli

University of Pavia (Italy)

Tom Pullman

University of Cambridge (UK)

Philippe Cousson

Lycée Camille Guérin (France)

Freely faceted classifications allow for free combination of concepts across all knowledge domains, and for sorting of the resulting compound classmarks. Starting from work by the Classification Research Group, the Integrative Levels Classification (ILC) project has produced a first edition of a general freely faceted scheme. The system is managed as a MySQL database, and can be browsed through a Web interface. The ILC database structure provides a case for identifying and representing the structural elements of any freely faceted classification. These belong to both the notational and the verbal planes. Notational elements include: arrays, chains, deictics, facets, foci, place of definition of foci, examples of combinations, subclasses of a faceted class, groupings, related classes; verbal elements

Gabriele Merli

University of Pavia (Italy)

Rick Szostak

University of Alberta (Canada)

include: main caption, synonyms, descriptions, included terms, related terms, notes. Encoding of some of these elements in an international mark-up format like SKOS can be problematic, especially as this does not provide for faceted structures, although approximate SKOS equivalents are identified for most of them.

Facet analysis as a tool for modelling subject domains and terminologies

Vanda Broughton

University College London (UK)

Facet analysis is proposed as a general theory of knowledge organization, with an associated methodology that may be applied to the development of terminology tools in a variety of contexts and formats. Faceted classifications originated as a means of representing complexity in semantic content that facilitates logical organization and effective retrieval in a physical environment. This is achieved through meticulous analysis of concepts, their structural and functional status (based on fundamental categories), and their inter-relationships. These features provide an excellent basis for the general conceptual modelling of domains, and for the generation of KOS other than systematic classifications. This is demonstrated by the adoption of a faceted approach to many web search and visualization tools, and by the emergence of a facet based methodology for the construction of thesauri. Current work on the Bliss Bibliographic Classification (Second Edition) is investigating the ways in which the full complexity of faceted structures may be represented through encoded data, capable of generating intellectually and mechanically compatible forms of indexing tools from a single source. It is suggested that a number of research questions relating to the Semantic Web could be tackled through the medium of facet analysis.

Analytico-synthetic approach for handling knowledge diversity in media content analysis

Devika P. Madalli

Indian Statistical Institute (India)

A. R. D. Prasad

Indian Statistical Institute (India)

Knowledge space is diverse and thus extremely complex. With increased means for online publishing and communication world communities are actively contributing content. This augments the need to find and access resources in different contexts and for different purposes. Owing to different socio-cultural backgrounds, purposes and applications, knowledge generated by people is marked by diversity. Hence, knowledge representation for building diversity-aware tools presents interesting research challenges. In this paper, we provide an analytico-synthetic approach for dealing with topical diversity following a faceted subject indexing method. Illustrations are used to demonstrate facet analysis and synthesis for use in annotations for Media Content Analysis within the European Commission (EC) funded 'Living Knowledge' project.

Extending models for controlled vocabularies to classification systems: modelling DDC with FRSAD

Joan S. Mitchell

OCLC, Inc. (USA)

Marcia Lei Zeng

Kent State University (USA)

Maja Žumer

University of Ljubljana (Slovenia)

The Functional Requirements for Subject Authority Data (FRSAD) conceptual model identifies entities, attributes and relationships as they relate to subject authority data. FRSAD includes two main entities, *thema* (any entity used as a subject of a work) and *nomen* (any sign or sequence of signs that a *thema* is known by, referred to, or addressed as). In a given controlled vocabulary and within a domain, a *nomen* is the appellation of only one *thema*. The authors consider the question, can the FRSAD conceptual model be extended beyond controlled vocabularies (its original focus) to model classification data? Models that are developed based on the structures and functions of controlled vocabularies (such as thesauri and subject heading systems) often need to be adjusted or extended to accommodate classification systems that have been developed with different focused functions, structures and fundamental theories. The Dewey Decimal Classification (DDC) system is used as a case study to test applicability of the FRSAD model for classification data, and as a springboard for a general discussion of issues related to the use of FRSAD for the representation of classification data.

Transformation of a legacy UDC-based classification system: exploiting and remodelling semantic relationships

Fran Alexander

BBC Information and Archives (UK)

Andy Heather

Dods political information (UK)

This paper reviews a project to remodel and unify diverse BBC Archive classification schemes, including the large Universal Decimal Classification (UDC) - based classification, Lonclass, as part of the BBC's Digital Media Initiative (DMI). The aims of the remodelling included migrating classification data from legacy systems and using the faceted structure of the classifications as a basis for proto-ontological relationship building. The processes of analysis and development of a methodology to decompose and reassemble the classifications raised such challenges as how to adapt bibliographic classifications for use as digital asset management tools and how to preserve the legacy intellectual property to enable continuing use of taxonomic classification as an access route to multimedia content. These objectives required the sophisticated semantics of the UDC-based classification to be retained during migration to an off-the-shelf taxonomy management product that could be integrated with diverse systems to form the basis of an enterprise-wide framework. The decompositions and reclassification process informed ways of preserving the high precision semantics of bibliographic classifications for use as a foundation for natural language-based retrieval and for translation into ontologically expressive formats, such as Resource Description Framework (RDF).

The evolution of knowledge, and its representation in classification systems

Andrea Scharnhorst

e-Humanities Group & DANS, KNAW (The Netherlands)

Almila Akdag Salah, Cheng Gao, Krzyzstof Suchecki

e-Humanities Group, KNAW (The Netherlands)

Richard P. Smiraglia

University of Wisconsin (USA)

Classification systems are often described as stable reference systems. Sometimes they are accused of being inflexible concerning the coverage of new ideas and scientific fields. Classification as an activity is the basis of all theory-generating research, and also plays a powerful role in social ordering. It is obvious that the ways in which we seek information and in which information is provided has changed dramatically since the emergence of digital information processing and even more with the internet, and web-based technologies. The purpose of this paper is to illustrate the notion of a stable knowledge organization classification as a temporary stationary manifestation of an open and evolving system of classification. We compare the structure of the main classes in the Universal Decimal Classification (UDC) according to their usage of special auxiliaries to demonstrate the dynamic evolution of the UDC over time, as a stable reference system representing published organized knowledge. We view the ecology of the UDC, and discover that most changes are to the ecology itself as numbers are re-interpreted. This subtle type of change is a key to monitoring the evolution of knowledge as it is represented in the UDC's stable reference system.

Visualizing universes of knowledge: design and visual analysis of the UDC

Charles van den Heuvel

Huygens ING Institute and Royal Netherlands Academy of Arts and Sciences (The Netherlands)

Almila Akdag Salah

Royal Netherlands Academy of Arts and Sciences (The Netherlands)

Knowledge Space Lab

Royal Netherlands Academy of Arts and Sciences (The Netherlands)

 In the 1950s, the "universe of knowledge" metaphor returned in discussions around the "first theory of faceted classification", the Colon Classification (CC) of S.R. Ranganathan, to stress the differences within an "universe of concepts" system. Here we claim that the Universal Decimal Classification (UDC) has been either ignored or incorrectly represented in studies that focused on the pivotal role of Ranganathan in a transition from "top-down universe of concepts systems" to "bottom-up universe of concepts systems." Early 20th century designs from Paul Otlet reveal a two directional interaction between "elements" and "ensembles" that can be compared to the relations between the universe of knowledge and universe of concepts systems. Moreover, an unpublished manuscript with the title "Théorie schématique de la Classification" of 1908 includes sketches that demonstrate an exploration by Paul Otlet of the multidimensional characteristics of the UDC. The interactions between these one- and multidimensional representations of the UDC support Donker Duyvis' critical comments to Ranganathan who had dismissed it as a rigid hierarchical system in comparison to his own Colon Classification. A visualization of the experiments of the Knowledge Space Lab in which main categories of Wikipedia were mapped on the UDC provides empirical evidence of its faceted structure's flexibility.

UDC as a knowledge framework for building a civil engineering ontology: a practical approach to knowledge representation and visualization

Ricardo Eito-Brun

Universidad Carlos III de Madrid (Spain)

Alfredo Calosci

Universidad Europea de Madrid (Spain)

This short paper analyzes the use of the Universal Decimal Classification (UDC) as a knowledge framework for building Web-enabled ontologies based on the Web Ontology Language (OWL), and an approach for the visualization of the relationships between the different concepts that make up the target ontology. Traditional use and applications of Universal Decimal Classification have been restricted to the physical arrangement of books within libraries, and although different research projects have been executed to adapt UDC for web-searching in OPACs (Online Public Access Catalogue) and other information services, current professional practice shows that UDC in the context of online retrieval has not been widely implemented. As the Web evolves to a knowledge-based, data-driven repository of repositories, it raises the following question: what is the role that UDC and other classification schemas play in the information services we expect to use and deliver in the future? The authors describe the use of UDC to generate a basic ontology for the representation of civil engineering knowledge. The need of this ontology was raised during the development of a web-based portal for historical documents on civil engineering developed for the Spanish Centre for Historical Studies of Public Works and Town Planning (CEHOPU).

UDC CONSORTIUM

International UDC Seminar 2011 is sponsored by

KNAW

Koninklijke Nederlandse Akademie van Wetenschappen

Ergon

EBSCO

poolparty

protégé

DANS

Data Archiving and Networked Services

LIBER · LIBER

Association of European
Research Libraries

KB

Koninklijke Bibliotheek

IFLA

International Federation
of Library Associations