

International UDC Seminar 2009

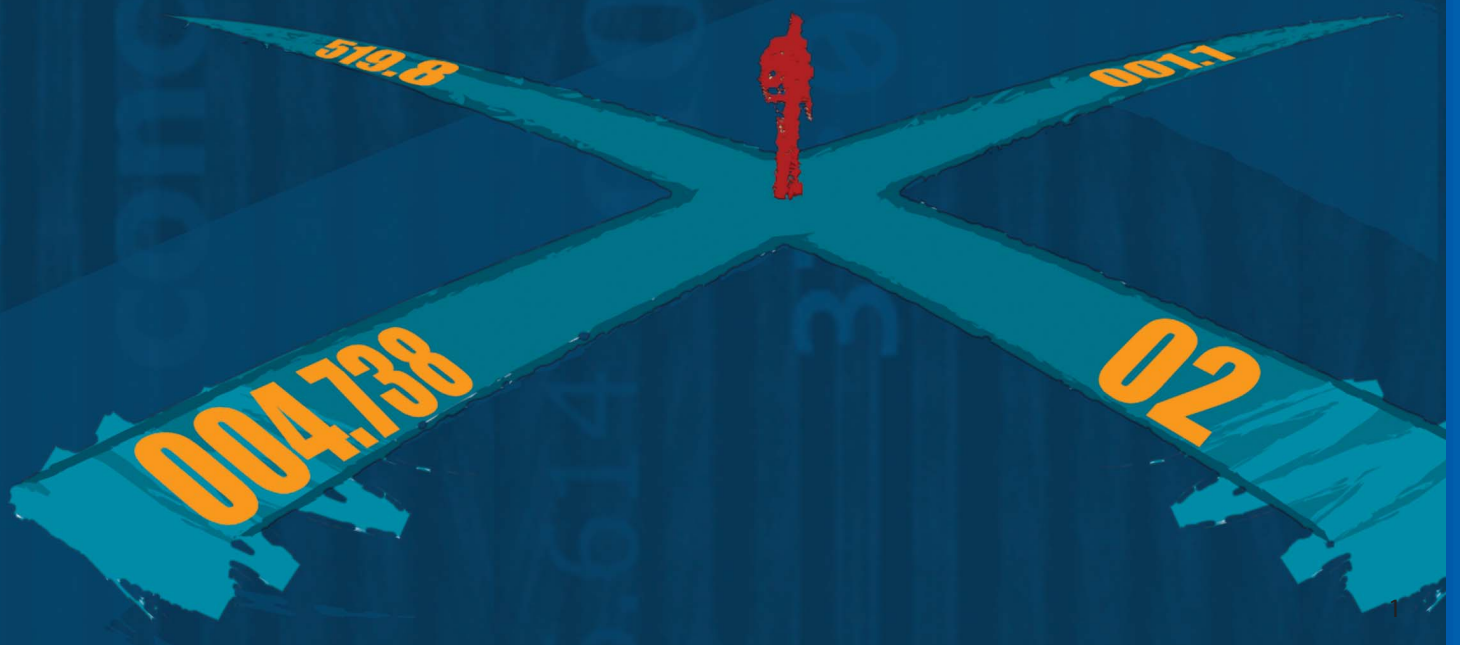
CLASSIFICATION AT A CROSSROADS

multiple directions to usability

The Hague, 29–30 October 2009

Programme

UDC
CONSORTIUM



International UDC Seminar 2009: Conference programme at a glance

Thursday 29 October

09:00	Registration. Coffee/Tea
09:30	Conference Opening and Welcome
10:00	Keynote address Dagobert Soergel , University of Maryland (USA) <i>Illuminating chaos: using classification to harness the Web</i>
11:00	Coffee/Tea
11:30	Session 1: Classifying Web resources Chair: Wouter Schallier Anders Ardö , Lund University (Sweden) <i>Automated classification: insights into benefits, costs and lessons learned</i>
12:00	Linda Kerr , Heriot-Watt University (UK) <i>Intute: from a distributed network to a unified database, lessons learned and future developments</i>
12:30	Jakob Voss , Verbundzentrale des GBV (Germany) <i>Wikipedia as knowledge organization system</i>
13:00	Lunch <i>Demo presentation (13:35–13:55) by Denis Godbout: Emerald Group Publishing and the Emerald e-Book Series Collections</i>
14:00	Session 2: Classification and thesauri Chair: Gerhard Riesthuis Marlene van Doorn and Katrien Polman , Afrika-Studiecentrum (The Netherlands) <i>From classification to thesaurus ... and back? Subject indexing tools at the library of the Afrika-Studiecentrum Leiden</i>
14:30	Victoria Frâncu , Central University Library of Bucharest (Romania) Cosmin-Nicolae Sabo , Universitatea de Nord Baia Mare (Romania) <i>Implementation of a UDC-based multilingual thesaurus in a library catalogue: the case of BiblioPhil</i>
15:00	Ágnes Hajdu-Barát , University of Szeged (Hungary) <i>Integration of thesaurus and UDC to improve subject access: the Hungarian experience</i>
15:30	Stella Dextre Clarke , independent information consultant (UK) <i>Providing for interoperability between thesauri and classification schemes in ISO 25964</i>
16:00	Coffee/Tea
16:30	Session 3: Classification frameworks, concepts, structure and relationships Chair: Peter Ohly Vanda Broughton , University College London (UK) <i>Concepts and terms in faceted classification</i>
17:00	Claudio Gnoli , University of Pavia (Italy) <i>Classification transcends library business</i>
17:30	Felix Boteram and Jessica Hubrich , Cologne University of Applied Sciences (Germany) <i>Specifying intersystem relations: requirements, strategies and issues</i>
18.00	Reception

Friday 30 October

09.00	Starting time
09.00	Keynote address Dan Brickley , Vrije Universiteit Amsterdam (The Netherlands) <i>Open standards and classification: foundations for a hybrid approach</i>
10.00	Coffee/Tea
10:30	Session 4: Classification and the Semantic Web Chair: Dagobert Soergel Ceri Binding and Douglas Tudhope, University of Glamorgan (UK) <i>Terminology services</i>
11:00	Gordon Dunsire and Dennis Nicholson , University of Strathclyde (UK) <i>Signposting the crossroads: terminology web services and classification-based interoperability</i>
11:30	A.R.D. Prasad and Devika Madalli , Indian Statistical Institute (India) <i>Classificatory ontologies</i>
12:00	Antoine Isaac , Vrije Universiteit Amsterdam (The Netherlands) <i>Using SKOS in practice, with examples from the classification domain</i>
12:30	Lunch <i>Lottery draw – Prize: books by Ergon Verlag and Facet Publishing</i>
13:30	Session 5: New approaches to classification Chair: Vanda Broughton Veslava Osinska, Nicolaus Copernicus University (Poland) <i>Visual analysis of classification scheme</i>
14:00	Alenka Šaupel , University of Ljubljana (Slovenia) <i>UDC and folksonomies</i>
14:30	Phillippe Cousson , Lycée Camille Guérin – Poitiers (France) <i>UDC as a non-disciplinary classification system for a high-school library</i>
15.00	Coffee/Tea
15.30	Session 6: Classification in library networks Chair: I.C. McIlwaine Marie Balíková, The National Library of The Czech Republic <i>The role of UDC classification in the Czech Subject Authority File</i>
16:00	Darija Rozman , National and University Library of Slovenia (Slovenia) <i>The practical value of classification summaries in information management and integration</i>
16:30	Rosa San Segundo , Carlos III University of Madrid (Spain) <i>Using MARC classification format for UDC and mappings to other KO systems for an enriched authority file</i>
17:00	<i>Panel discussion</i>
17.45	Close of conference



DAY ONE: KEYNOTE ADDRESS

Dagobert Soergel

University of Maryland, USA

Illuminating chaos: using classification to harness the Web

The Web is a chaotic place, even more so with the appearance of Wikis, blogs, and social tagging 'in the wild'. This talk aims to provide a vision of the many ways classification can help. The first part will illustrate through a number of examples how classification – construed broadly as any approach to creating meaningful structure – can help with putting the vast amount of information on the Web to better use. People need meaningful classification to help them formulate queries; computer systems need either controlled vocabulary indexing (produced by people or automatically) or a semantic structure for mapping a semantic query into a free-text (Google) query for better retrieval. The huge amount of labour that goes into social tagging can be applied more effectively through gentle guidance, and the user-generated tags can be organized into a structure for easier search. Both people and computer systems need semantics to make sense of information once it is found. The second part will address the chaos of the many partially overlapping ontologies and other Knowledge Organization Systems that run wild on the Web. It will introduce the conceptual hub approach to KOS mapping that provides the basis for universal facet-based search of the Web.

Dagobert Soergel is Professor at the College of Information Studies, University of Maryland and in the Department of Library and Information Studies, Graduate School of Education, University at Buffalo. He has been an Information Technology and Services Consultant since 1970 and has been working in the area of IR, specifically classification (taxonomy, ontologies) and thesauri, for over 40 years. He is the author of a number of papers and textbooks in the field of indexing and information retrieval. Prof. Soergel is the recipient of the Award of Merit for 1997, the highest award of the American Society for Information Science. Read more at <http://ischool.umd.edu/people/soergel/>



Session 1

Session chair: *Wouter Schallier*

> Classifying Web resources

Anders Ardö

Lund University (Sweden)

Automated classification: insights into benefits, costs and lessons learned

While automated methods for information organization, i.e. classification, have been around for several decades, exponential growth of the World Wide Web has put them in the forefront of research in several different communities: machine learning (Artificial Intelligence), document clustering (Information Retrieval) and weighted string-matching against a controlled vocabulary (Library and Information Science). Based on the latter approach, NetLab at Lund University Library explored automated classification based on UDC in Nordic WAIS/WWW as early as 1992 and continued research during the 1990s, testing automatic classification on Engineering Index classification and DDC. Major similarities and differences between the three approaches will be discussed, and problems of automated classification recognized. There is, however, a major issue of evaluation and comparison, largely due to the challenge of identifying the *aboutness* of documents, which is related to the quality of indexing. An effort will be made, based on relevant projects and analyses, to discuss general benefits and costs, resulting quality and lessons learned.

Anders Ardö is Associate Professor in the Department of Electrical & Information Technology at Lund University: he belongs to the Circuit and System Design research area. His main areas of interest are digital library development; vertical search engines; parallel/distributed computing; focused (topical) Web crawling; advanced methods for information discovery and retrieval; and automatic subject classification/text mining. Focus areas are knowledge discovery and development of distributed knowledge organization technologies and information systems, in particular exploring the usability of knowledge organization systems (thesauri, classifications, taxonomies, ontologies, gazetteers) in the digital library context. Dr Ardö has published more than 70 papers. (See at <http://www.eit.lth.se/staff/anders.ardo>)

Linda Kerr

Heriot-Watt University (UK)

Intute: from a distributed network to a unified database, lessons learned and future developments

Intute (<http://www.intute.ac.uk>) is a UK service, funded by the Joint Information Systems Committee (JISC), which catalogues and describes the best Internet resources for education and research. Intute catalogue of Internet resources is a unification of seven subject catalogues, previously funded by JISC as separate subject-specific services. This paper describes the process and challenges of integrating seven databases into one unified catalogue with one standard metadata schema, whilst continuing to satisfy the needs of different subject communities. Also described is the development of a 'course and theme' view onto the resources, an approach suggested by an extensive market research process. The paper outlines two projects which will evaluate the cost-effectiveness of manual and automatic metadata creation. The projects will assess what core metadata is required to provide the most effective retrieval of resources (linked to user requirements) and, by using automatic metadata generation tools, aim to improve the efficiency of metadata generation processes and improve user satisfaction in the retrieval of resources. Existing tools for automated metadata generation will be tested and evaluated to determine to what extent they can be integrated into (semi-) automated workflows.

Linda Kerr is based at Heriot-Watt University, and is a service manager for Intute, a JISC-funded UK-based service cataloguing web resources for education and research, and providing Internet Research Skills training. She has been involved in the resource discovery and delivery of web resources in Higher Education since starting as a project officer for EEVL (Edinburgh Engineering Virtual Library) in 1996. Other projects include the Institutional Repositories Search Project, the Virtual Training Suite, and EU projects creating digital resources for deaf and disabled people.

Jakob Voss

Verbundzentrale des GBV (Germany)

Wikipedia as knowledge organization system

Despite its volatile nature, Wikipedia is generally used as the major reference work on the Web. Wikipedia articles are not always perfect but they can be corrected quickly. The process by which content is created and shaped can be complex but it is always transparent because each change is tracked. Wikipedia has proven to be of practical value for the distribution and sharing of knowledge. This talk will show how Wikipedia can also be used for knowledge organization and how it is connected to other knowledge organization systems. Wikipedia can be viewed as a controlled vocabulary built of articles, languages, categories, and links. Some of its features have great potential and advantages. For instance the terms and definitions are created collaboratively with at least one article per concept and language. More importantly semantic linking can be established between articles in an unlimited way and dynamic concept hierarchies can be created from the data provided by Wikipedia. The encyclopaedia is not limited to a subject domain which makes it a top-level ontology like UDC, DDC, CyC, and WordNet. The paper outlines how Wikipedia can be used for subject indexing and how it can be linked and mapped to other controlled vocabularies, especially those available as Open Linked Data and expressed using Resource Description Framework (RDF) technology.

Jakob Voss was one of the founders and, for several years, a member of the executive board of Wikimedia Germany – a registered association for the promotion of free knowledge. Jakob has a degree in computer science, library science, and philosophy at Humboldt University, Berlin. He works at the library service centre of the Common Library Network (GBV) in Göttingen. His research interest is in Wikipedia development and social software.

> Classification and Thesaurus

**Marlene van Doorn,
Katrien Polman**

Afrika-Studiecentrum
(The Netherlands)

From classification to thesaurus... and back? Subject indexing tools at the library of the Afrika-Studiecentrum Leiden

The paper describes an approach to thesaurus creation and use. An African Studies Thesaurus was constructed for the purpose of subject indexing and retrieval in the Library of the African Studies Centre (ASC) in Leiden in 2001–2006. A word-based system was considered a more user-friendly alternative to the UDC codes which were used for subject access in the ASC catalogue at the time. In the process of thesaurus construction, UDC codes were used as a starting point. In addition, each descriptor was assigned a UDC code from the recent edition of the UDC Master Reference File, thus replacing many of the old UDC codes, some of which dated from the 1952 French edition. The presence of UDC codes in the thesaurus leaves open the possibility of linking the thesaurus to different language versions of the UDC MRF in the future. Over a period of 5 years some 8,500 UDC codes from all classes of the UDC (mostly from class 3, to a lesser extent classes 6, 8 and 9), were ‘translated’ into thesaurus descriptors and structured according to basic thesaurus relationships (BT, NT, RT, UF/USE). In a parallel but separate operation each UDC code which had been assigned to an item in the library’s catalogue was subsequently converted into one or more thesaurus descriptors.

Marlene van Doorn works as an information specialist at the Library, Documentation and Information Department of the African Studies Centre in Leiden. Her work focuses on facilitating access to African Studies resources, in particular through the library’s online catalogue, the abstracts journal *African Studies Abstracts Online* and thematic Web dossiers. She is responsible for development and maintenance of the African Studies Thesaurus, and is a member of the department’s digital library committee which serves as sounding board for digital innovations and developments.

Katrien Polman works as an information specialist at the Library, Information and Documentation Department of the African Studies Centre in Leiden. Her work focuses on facilitating access to African Studies resources and meeting the information needs of ASC library users. Her tasks include selecting, indexing, and abstracting publications, including articles from periodicals, for the library’s online catalogue and its abstracts journal, *African Studies Abstracts Online*. She evaluates and selects digital resources for the web service *Connecting Africa* and is involved in the development and maintenance of the ASC’s African Studies Thesaurus. She is also the editor of the ASC library’s Web Dossiers.

Victoria Frâncu

Central University Library of
Bucharest (Romania)

Cosmin-Nicolae Sabo

Universitatea de Nord Baia Mare
(Romania)

Implementation of a UDC-based multilingual thesaurus in a library catalogue: the case of BiblioPhil

The paper describes an approach to improving classification-based subject access in a library catalogue. To enhance use of UDC numbers in information retrieval, the authors have represented classification with thesaurus descriptors and implemented this solution in an automated way. In addition, descriptors in more than one language were used to interface classification. The authors illustrate a solution implemented in a BiblioPhil library system. The standard formats used are UNIMARC for bibliographic and subject authority records (the UDC-based multilingual thesaurus) with MARCXML support for data transfer. The multilingual thesaurus was built according to existing standards, the constituent parts of the classification notations being used as the basis for search terms in information retrieval. The verbal equivalents, descriptors and non-descriptors, are used to expand the number of concepts and are given in Romanian, English and French. The authors illustrate how this approach saves the time of the indexer and provides more user-friendly and easier access to the bibliographic information..

Victoria Frâncu is librarian at the Carol I Central University Library of Bucharest. She also teaches classification and indexing at the University of Bucharest. Her research interest is in the field of knowledge organization systems, in particular classifications and their role in supporting information retrieval. She is the head of the Cataloguing and Indexing Section of the Romanian Library Association and a member of the working group on the translation of the Rameau subject heading system into Romanian. She holds a PhD in information science from the University of Antwerpen.

Cosmin Sabo is a lecturer at the Nord Baia Mare University where he teaches advanced database systems, security of information systems, computer networks and object oriented programming. He holds a degree in computer science from the Faculty of Mathematics and Computer Science at the University of Sibiu and has been working as a software developer specializing in the area of computer networks, information systems security and object-oriented programming since 2000. He was head of the Automation Department at Baia Mare County Library between 2002–2006 and during that period he developed an integrated library system called BiblioPhil.

Ágnes Hajdu Barát
University of Szeged (Hungary)

Integration of thesaurus and UDC to improve subject access: the Hungarian experience

The paper explores two possible solutions for integrating a thesaurus and a classification scheme, specifically UDC, to develop a common platform for subject information retrieval through both systems. The author reports and compares experiences from two Hungarian projects aimed at creating a complex system for combining UDC and thesauri under a homogeneous theoretical framework: MÁTrIkSz (Hungarian Comprehensive Information Retrieval Language Dictionary) and the project of thesaurus construction and implementation in the Hungarian National Library (Széchényi). The role of UDC in these two projects is analyzed with respect to the features supported, classification-based retrieval functionalities, and the perceived advantages in subject access and knowledge organization. The author explains the methodology of her research based on an examination of structured and well-documented examples and literature research into the theory of UDC and its use. The paper underlines the importance of cognition as the basis for concept-building and points out some possibilities and expedients for the integration of thesauri and the UDC.

Ágnes Hajdu Barát is Professor at the University of Szeged Gyula Juhász Faculty of Education (SZTE JGYPK). She worked as an information consultant for a number of organizations and research projects. Her research interests are epistemological issues of knowledge organization, visualization of information, interdisciplinarity, and multilingual information retrieval. She is the vice president of the Hungarian Library Association, an associate editor of the UDC and the author of several books and over seventy articles in the field of knowledge organization and knowledge management.

Stella Dextre Clarke
independent information consultant
(UK)

Providing for interoperability between thesauri and classification schemes in ISO 25964

If a Knowledge Organization System (KOS) works well in every system where it is used for the indexing/classification and retrieval of information, that is no longer enough. In today's world of networked information systems, another essential requirement of KOSs is interoperability across systems. Hence ISO 25964, the international standard which will replace the existing thesaurus standards ISO 2788 and ISO 5964, covers not just the construction of thesauri but also interoperability with classification schemes and other types of controlled vocabulary. In this presentation we consider Part 2 of the standard, dealing with mapping to/from classification schemes and other KOSs. It is not about how to construct a classification scheme, but will need to describe the characteristics of existing schemes in sufficient detail to support the development of appropriate mappings between classification schemes and thesauri. Some issues to be resolved are: (i) how to handle pre-coordinated classes; (ii) how to provide for classes not enumerated in the scheme but synthesised on demand; and (iii) whether (and if so how) to include a data model for each type of KOS. As ISO 25964-2 is still in the initial drafting stage, there is every hope that ideas contributed at this conference may feed into the eventual standard.

Stella Dextre Clarke is an independent consultant specializing in the design and implementation of thesauri and other knowledge organization structures, with clients in the public and private sectors. She currently leads ISO NP 25964, the project to update and revise the international standards for thesauri. Previously she was the Convenor of the Working Group which developed BS 8723, the set of corresponding British Standards. Stella is widely known for her work on behalf of the UK Cabinet Office, as principal architect of the IPSV (Integrated Public Sector Vocabulary), a key component of the e-Government Metadata Standard. In 2006 she won the Tony Kent Strix Award for outstanding achievement in information retrieval, in recognition of this work.

> Classification frameworks, concepts, structure & relationships

Vanda Broughton
University College London (UK)

Concepts and terms in faceted classification

Recent revision of UDC classes has aimed at implementing a more faceted approach. Many compound classes have been removed from the main tables, and more radical revisions of classes (particularly those for Medicine and Religion) have introduced a rigorous analysis, a clearer sense of citation order, and building of compound classes according to a more logical system syntax. The faceted approach provides a means of formalizing the relationships in the classification and making them explicit for machine recognition. In the Bliss Bibliographic Classification (BC2) (which has been a source for both the UDC classes mentioned above), terminologies are encoded for automatic generation of hierarchical and associative relationships. Nevertheless, difficulties are encountered in vocabulary control, and a similar phenomenon is observed in UDC. Current work has revealed differences in the vocabulary of humanities and science, notably the way in which terms in the humanities should be handled when these are semantically complex. Achieving a balance between rigour in the structure of the classification and the complexity of natural language expression remains partially unresolved at present, but provides a fertile field for further research.

Vanda Broughton is Senior Lecturer in Library & Information Studies, and Programme Director for the MA LIS, at University College London. She has worked on the revision of the Bliss Bibliographic Classification since 1972, and is Joint Editor of the second edition. She has been involved with the UDC since 1997, and is now Associate Editor. A member of the Classification Research Group from the 1970s, she has also been a member of the IFLA Committee on Classification and Indexing, and is the author of a number of books and articles on the theory and design of classifications and thesauri.

Claudio Gnoli
University of Pavia (Italy)

Classification transcends library business

Although bibliographic classifications usually adopt a different perspective from object classifications, the two have obvious relationships. These are becoming increasingly relevant today, as library catalogues now coexist in the global digital environment with catalogues of archives, of museums, of commercial products, etc. Therefore, a broader conception of classification is needed, that can be applied to any knowledge item. The real case of research on bagpipes in Northern Italian folklore is considered. Its most effective approach is by far a cross-media one, looking for many knowledge sources like published documents, police archives, painting details, museum specimens and ethnographic organizations. To satisfy this kind of search, the traditional disciplinary approach of classification is not enough. Tools are needed in which knowledge items can be retrieved independently from other topics with which they are combined, or the context where they occur. This is made possible if the basic units of classification are taken to be the phenomena treated, as is recommended in the Leon Manifesto, rather than disciplines or other aspect features. The concept of bagpipes should be retrieved and browsed in any combination with other phenomena, disciplines, media etc. Examples will be given using notation of the Integrative Level Classification draft system.

Claudio Gnoli has been working as an academic librarian since 1994. His main interest is classification theory. He has published papers on this subject in several international journals and conference proceedings. He is a member of the scientific advisory boards of the Universal Decimal Classification Consortium (UDCC) and of the journal *Knowledge Organization*, and vice-president of the International Society for Knowledge Organization (ISKO).

**Felix Boteram,
Jessica Hubrich**

Cologne University of Applied
Sciences (Germany)

Specifying intersystem mapping relations: requirements, strategies and issues

The paper outlines the development and improvement of intersystem relations focusing on a comprehensive international knowledge organization system and mappings between typologically different indexing languages. Ideally, intersystem relations complement highly expressive and thoroughly structured relational indexing languages. The relational structures of the participating systems contribute to the meaning of the individual terms or classes. However, when conceptualizing mapping relations, the structural and functional design of the respective systems must be fully taken into account. As intersystem relations may differ considerably from familiar interconcept relations, the creation of an adequate inventory that would be general in coverage and specific in depth demands a deep understanding of the requirements and properties of mapping relations. In the authors' experience the characteristics of specific mapping relations largely rely on the characteristics of the systems they are intended to connect. The detailed declaration of differences and peculiarities of specific mapping relations is an important prerequisite for modelling these relations. First approaches towards specifying intersystem relations are presented with special respect to linkages between universal decimal classifications and thesauri.

Felix Boteram is a librarian and research assistant working for the RESEDA project at the Institute of Information Management at the Cologne University of Applied Sciences, Germany. RESEDA aims at developing models for the representation of semantic data in systems for knowledge organisation from an interdisciplinary perspective.

Jessica Hubrich works at the Institute of Information Management at the Cologne University of Applied Sciences, Germany. She is team leader of CrissCross, a research project funded by the German Research Foundation (DFG) focused on mapping subject headings from the German Subject Authority File (SWD) to classes of the Dewey Decimal Classification (DDC).



DAY TWO: KEYNOTE ADDRESS

Dan Brickley

Vrije Universiteit Amsterdam
(The Netherlands)

Open Web standards and classification: foundations for a hybrid approach

Classification is truly at a crossroads. Library and cultural heritage institutions are increasingly making their collections accessible in machine-processable form. Movements advocating for open data sharing are helping governments publish huge public datasets, creating communally-maintained datasets via the Web, and using open Web standards to ensure these works are all cross-referenced and richly linked. New Web standards are bridging the gaps between thesauri, ontologies and databases, providing opportunities for collaboration, information sharing and new approaches to user interface design. Drawing on examples from television, subject-based information gateways and Web 2.0 trends, this talk proposes some foundational steps that will help professional subject classification remain central to resource discovery, annotation and linking.

Dan Brickley is best known for his work on Web standards in the W3C community, where he helped create the Semantic Web project and many of its defining technologies. Dan is currently working at the Vrije University Amsterdam on the NoTube EU project, developing new approaches to interactive TV that build upon SKOS, FOAF and open Web technologies. Previous work includes six years on the W3C technical staff, establishing ILRT's Semantic Web group at the University of Bristol, and more recently at Joost, an Internet TV startup. He has been involved with resource discovery metadata since 1994 when he published the first HTML Philosophy guide on the Web, and has been exploring distributed, collaborative approaches to 'finding stuff' ever since.

> Classification and the Semantic Web

**Ceri Binding,
Douglas Tudhope**
University of Glamorgan (UK)

Terminology services

Controlled terminologies such as classification schemes, name authorities and thesauri have long been the domain of the library and information science community. Although historically there have been initiatives towards library style classification of web resources, there remain significant problems with searching and quality judgement of online content. The more recent growth of social bookmarking sites indicates a desire for the personal organisation and structuring of web resources. Social tagging produces some interesting results, but also produces ambiguous vocabularies conflating index terms with opinions. Intuitive tools incorporating established controlled terminologies in fields other than libraries remain sparse. Terminology services play a key role in opening up access to these valuable resources. By exposing controlled terminologies via a web service, organisations maintain data integrity and version control, whilst motivating external users to design innovative ways to present and utilise the data. The authors relate their experiences in creating terminology web services and associated client interface components for the archaeology domain in the STAR project (<http://hypermedia.research.glam.ac.uk/kos/STAR/>) and demonstrate how the same principles can be readily adapted to other subject areas. (http://hypermedia.research.glam.ac.uk/kos/terminology_services/)

Ceri Binding is a Research Associate in the Hypermedia Research Unit, Faculty of Advanced Technology, University of Glamorgan. Ceri graduated with a BSc in Computer Studies in 1997 whilst working as an Analyst Designer/Programmer for Hyder IT, before joining Glamorgan in 2000. He had responsibility for development work on the FACET project and implemented various standalone and web systems for the project. He is currently conducting research and development work for the STAR project, involving use of SKOS and CRM data. Related research interests include Knowledge Organisation Systems, intelligent web-based retrieval and interface design.

Douglas Tudhope is Professor in the Faculty of Advanced Technology, University of Glamorgan and leads the Hypermedia Research Unit. His area of research is Knowledge Organisation Systems and Services. He was PI on the EPSRC FACET project in collaboration with the Science Museum and Museum Documentation Association and is PI on an AHRC project (Semantic Technologies for Archaeological Resources) in collaboration with English Heritage. He was an author of the recent JISC State of the art review on Terminology Services and Technology.

**Gordon Dunsire,
Dennis Nicholson**
University of Strathclyde (UK)

Signposting the crossroads: terminology web services and classification-based interoperability

The focus of this paper is the provision by the JISC-funded HILT project of terminology and classification-based terminologies interoperability data via web services, initially using interoperability data based on the use of a Dewey Decimal Classification (DDC) spine, but with an aim to explore other possibilities in time, including the use of other spines (including, possibly, UDC), and the use of direct KOS to KOS mappings.

HILT Phase IV developed pilot web services based on SRW/U, SOAP, and SKOS to deliver machine-readable terminology and cross-terminology mappings data likely to be useful to information services wishing to enhance their subject search or browse services. It also developed an associated toolkit to help information services technical staff to embed HILT-related functionality within service interfaces. Several UK information services have created illustrative user interface enhancements using HILT functionality and these will demonstrate what is possible. HILT currently has the following subject schemes mounted and available: DDC, CAB, GCMD, HASSET, IPSV, LCSH, MeSH, NMR, SCAS, UNESCO, and AAT. It also has high level mappings between some of these schemes and DDC, and some deeper pilot mappings available.

Gordon Dunsire is Head of the Centre for Digital Library Research at Strathclyde University in Glasgow, Scotland. He is a member of the CILIP-BL Committee on AACR and the CILIP Committee on DDC, and is Chair of the Cataloguing and Indexing Group in Scotland. He is the principal developer of the SCONE collection descriptions service and other components of the Scottish Common Information Environment, and has been involved in several projects investigating the use of collection-level description and metadata aggregation in wide-area resource discovery.

Dennis Nicholson is Director of the Centre for Digital Library Research at Strathclyde University and Director of Research in Strathclyde University's Information Resources Directorate. Since 1991, he has managed a range of funded projects, either directly or in his role as CDLR manager.

**A.R.D. Prasad,
Devika Madalli**
Indian Statistical Institute (India)

Classificatory ontologies

Digital Libraries and Digital Repositories are data-intensive with large numbers of full-text resources accessible online. Activities in the area of Semantic Web development recognize the significant part played by metadata and knowledge organization systems (such as classification systems and thesauri) in capturing and communicating 'meaning'. We now have web ontology standards, such as Simple Knowledge Organization Systems (SKOS), a common data model for sharing and linking knowledge organization systems via the Semantic Web. Standards such as SKOS are also meant to be used as a vehicle for deployment of knowledge organization systems that were not born digital (or XML/RDF) such as thesauri and bibliographic classifications. This paper presents an application of the faceted classification scheme as enunciated by Ranganathan in developing ontologies. It further explores issues in modelling the faceted scheme of Ranganathan using SKOS.

A. R. D. Prasad is an Associate Professor of the Documentation Research and Training Centre. He worked on AI applications to subject indexing techniques for information retrieval and holds a PhD on the same topic. His main interests are in IT applications to information systems in general and latest research interests in Semantic Web technologies, ontologies development and deployment, digital library research and multilingual retrieval. Dr Prasad has published more than 100 papers in national and international seminars, conferences and journals.

Devika Madalli is a lecturer at the Documentation Research and Training Centre, Indian Statistical Institute, Bangalore, India and adjunct lecturer of the University of Trento, Italy Department of Information and Communication Sciences. She holds a PhD in Knowledge Representation techniques for faceted classificatory systems. Her interest is in the area of knowledge organization and application of facetization in information systems, digital libraries, Semantic Web technologies, ontologies, open access to information, content management systems and e-learning. She has published more than 30 papers in national and international seminars, conferences and journals.

Antoine Isaac
Vrije Universiteit Amsterdam
(The Netherlands)

Using SKOS in practice, with examples from the classification domain

After a brief presentation of various features of the SKOS model, and its role in relation to knowledge organization systems (KOS) and the Semantic Web, the author discusses some practical issues that have to be overcome when representing KOSs using SKOS. A particular focus is put on the classification domain, by means of examples taken from typical classification vocabularies, such as UDC. The paper illustrates how elements of the SKOS model can capture important aspects of classifications, but that some key features, such as concept coordination, are still lacking proper means of representation. Hints are given on how SKOS can be extended with elements that address these issues. The author attempts to analyse those potential elements, keeping in mind the interoperability motivation that guided the design of SKOS in the first place. In particular he will try to answer the following question: to what extent can consensual extensions be devised to use SKOS successfully with classification systems?

Antoine Isaac is a post-doctoral researcher at Vrije Universiteit in Amsterdam and Koninklijke Bibliotheek where has been working for a number of years on figuring out how Semantic Web technology can be successfully used in Cultural Heritage environment. He works on the representation and interoperability of cultural heritage collections and their vocabularies (STITCH, TELplus and EuropeanaConnect projects). He is a member of the W3C Semantic Web Deployment Working Group and is involved in the design of SKOS. Read more at <http://www.few.vu.nl/~aisaac/>.

> New approaches to classification

Veslava Osinska

Nicolaus Copernicus University
(Poland)

Visual analysis of classification scheme

This paper proposes a novel methodology to visualize a classification scheme, demonstrated with the Association for Computing Machinery (ACM) Computing Classification System (CCS). The collection derived from the ACM digital library contained 37,543 documents classified by CCS. The attributes classes, subject descriptors and keywords were processed in a dataset to make a graphical representation of the documents. The general conception is based on the similarity of co-classes (themes) proportional to the number of common publications. The final number of all possible classes and subclasses in collection was 353 and therefore the similarity matrix of co-classes had the same dimension. A spherical surface was chosen as the target information space. Classes and documents node locations on the sphere were obtained by means of Multidimensional Scaling coordinates. By representing the surface on a plane like a map projection, it is possible to analyze the visualization layout. The graphical patterns were organized in some colour clusters. For evaluation of given visualization maps, graphics filtering was applied. This proposed method can be very useful in interdisciplinary research fields. It allows for a great amount of heterogeneous information to be conveyed in a compact display, including topics, relationships among topics, frequency of occurrence, importance and changes of these properties over time.

Veslava Osinska received her MSc in physics from Vilnius University, and PhD in Information Science and Bibliography at Nicolaus Copernicus University in Torun (Poland), where she teaches Information and Communication Technology and Computer Graphics. She has applied her Computer Science background and programming skills to research areas which include: effective visualization of multi-dimensional information, for example bibliographic data generated in digital libraries. She is a member of ISKO Polish Chapter as well as Polish Computer Science Society.

Alenka Šauperl

University of Ljubljana (Slovenia)

UDC and folksonomies

Social tagging systems, known as 'folksonomies', represent an important part of web resource discovery as they enable free and unrestricted browsing through information space. Folksonomies consisting of subject designators (tags) assigned by users, however, have one important drawback: they do not express semantic relationships between tags, either hierarchical or associative. As a consequence the use of tags to browse information resources means having to move from one resource to another, based on coincidence and not on the pre-established meaningful or logical connections that may exist between related resources. We suggest that the semantic structure of the UDC may be used in complementing and supporting tag-based browsing. Two specific questions were investigated: (1) Are terms used as tags in folksonomies included in the UDC? and (2) Which facets of UDC match the characteristics of documents or information objects that are tagged in folksonomies? A collection of the most popular tags from Amazon, LibraryThing, Delicious and 43Things was investigated. The universal nature of UDC was examined through the universality of topics and facets covering diverse human interests which are at the same time interconnected and form a rich and intricate semantic structure. The results suggest that UDC supported folksonomies could be implemented in resource discovery, in particular in library portals and catalogues.

Alenka Šauperl is Associate Professor in the Department of Library and Information Science and Book Studies at the Faculty of Arts, University of Ljubljana. Her teaching and research areas are in the organization of information, including descriptive and subject cataloguing, as well as abstracting. She is the author of several books and many articles in the field of cataloguing and subject indexing.

Phillippe Cousson

Lycée Camille Guérin – Poitiers
(France)

UDC as a non-disciplinary classification system for a high-school library

The paper addresses issues in establishing a user-friendly systematic collection arrangement, following the merger of two high school and college library collections classified by UDC. This scheme as it was used had some weaknesses with respect to collection usage. Due to the disciplinary nature of UDC, subjects and phenomena are dispersed in the scheme according to the disciplines in which they are the subject of study. At the same time students in a school library often seek interdisciplinary subjects and need access to clusters of documents which according to UDC may be classed in several different knowledge areas. The author illustrates how this problem was resolved by re-arranging the collection according to phenomena. This was achieved by interpreting UDC numbers as if they represented specific phenomena. Thus by superimposing some local indexing rules onto a disciplinary knowledge organization system it was possible to collocate interdisciplinary subjects under a single class number. Furthermore, by reversing subject numbers and form auxiliaries (atlases, dictionaries, textbooks etc.) which is an option envisaged in the design of UDC, documents were collocated in the way they are most frequently used by students. The author suggests that in practice one often needs to overcome the constraints of disciplinary classification and discusses the approach used in his school library collection.

Phillippe Cousson is the librarian in Lycée Camille Guérin, Poitiers. Beside librarianship (document analysis, thesaurus and classification), he has two main interests : theology and linguistics. For more than 20 years, he has been working in a collaborative indexation of periodicals of educational interest (Mémofiches). During three years he has led continuing education sessions for school librarians.

Session 6

Session chair: I.C. McIlwaine

> Classification in Library Networks

Marie Balíková

The National Library of
The Czech Republic

The role of UDC classification in the Czech Subject Authority File

The paper outlines the standardization function of the Czech Subject Authority File and explores the role of the UDC as a switching language, i.e. as an intermediary between various indexing systems at institutional, national and international level. Subject indexing and classification systems used at the institutional and national level may differ from one another in their levels of specificity, syntactic features (e.g. word order of terms, subject headings versus descriptors), and in the usage of terminology. These differences raise compatibility problems and make any mapping efforts more difficult. The paper explains how such difficulties may be partially overcome by means of the UDC system. The author illustrates the potential application of UDC as a linking element between different subject organization tools used by memory institutions. In this context the author discusses subject indexing systems used in libraries, museums, galleries and archives.

Marie Balíková is the head of the national subject heading authorities and classification department of the National Library of the Czech Republic. She is responsible for the national subject authority files and the implementation of the subject analysis international standards, as well as Conspectus methods (subject access) in the Czech Republic. Marie is the head of the working group on subject access, member of the National Cataloguing Policy Committee, and member of the IFLA Standing Committee of the Indexing and Classification Section and an associate editor of the UDC. Marie is the author of many books and papers on subject indexing.

Darija Rozman

National and University Library,
Slovenia (Slovenia)

The practical value of classification summaries in information management and integration

The author discusses the value and importance of the use of short extracts from classification tables to support subject access management. While detailed classification is time consuming, complex and costly, the classification of documents into broader classes is a simpler and easier way of achieving meaningful and useful subject organization. The paper outlines the role of this type of classification use in bibliographic listings, in the organization and representation of physical documents, in the presentation of web resources, in statistical reports in collection development and use, and, last but not least, in information integration in a networked environment. This approach of subject classification is illustrated by the Slovenian union catalogue COBISS/OPAC in which a standardized set of UDC codes is used. The author emphasizes the importance of this outline for the homogeneity and continuity of the use of UDC in Slovenia and explains how this may be weakened by the changes in the top level of UDC.

Darija Rozman has been the head of the subject cataloguing department at the National and University Library in Ljubljana since 1998. She has published a number of papers on subject cataloguing and Universal Decimal Classification specifically. Darija is a member of the editorial team of the Slovenian General Subject Headings List Online (Spletni splošni slovenski geslovnik) and a member of the UDC Advisory Board.

Rosa San Segundo

Carlos III University of Madrid
(Spain)

Using MARC classification format for UDC and mappings to other KO systems for an enriched authority file

The USMARC classification format, developed in the early 1990s for the DDC and LCC systems, is also amenable for other classification systems such as UDC. This paper presents a proposal for using the MARC classification format for UDC. There are advantages in using this format for the UDC data in an authority file, e.g., for the MRF records and records for combined notations as well. There has been a trend in library catalogues for subject interoperability between traditional classification systems such as the UDC, DDC, LCC and subject headings. An example with great impact is WebDewey, which offers interlinking between classification numbers, the alphabetical index of the tables and LCSH. Another instance is the electronic version of LCC Plus, also including links to LCSH. Subject gateways built upon library authority files can support the interoperability between classification systems and subject headings. These gateways can be the backbone of a more universal access through hypertextual navigation structures supported by classification systems including UDC. To our knowledge, the MARC classification format has not yet been applied to the UDC and in this paper we are going to propose a solution supported by some examples.

Rosa San Segundo is Professor of Knowledge Organization in the Department of Librarianship and Documentation at the University Carlos III of Madrid. Her area of research is knowledge organization systems, and classification systems in particular. She is the president of the Spanish chapter of the International Society for Knowledge Organization (ISKO) and a member of the UDC Advisory Board. Prof. San Segundo is the author of many articles and books on knowledge organization systems and Universal Decimal Classification.

MagnaView UDC-viewer is OUT!

MagnaView develops innovative tools for visual analysis and powerful presentation of various types of information.

For UDC, MagnaView has developed an application which gives you interactive and visual access to the complete content of the UDC Master Reference File.

www.magnaview.com | info@magnaview.com



Included in MagnaView UDC-viewer 1.0 are:

- One named-user license of the powerful MagnaView Viewer (version 3.1).
- The MagnaView UDC-project with at least 20 different views.
- The Master Reference File in an enriched and compressed format for direct use with MagnaView and with the MagnaView UDC-project.
- Moreover, this offer includes the possibility to order additional MagnaView software with a discount of 20%.

MagnaView UDC-viewer is available for € 550,- (excl. BTW/VAT).



EMERALD GROUP PUBLISHING

Don't miss the presentation on 29th October (13:35–13:55)...

Emerald Group Publishing and the Emerald e-Book Series Collection

presented by Denis Godbout

Emerald e-Book Series Collections: Adding books to Emerald's already strong journal offering, Emerald's e-Book series collections provide a collection of over 5,000 chapters of content available through the familiar Emerald platform. The business, management and economics collection features 70 series titles and includes over 400 volumes in subjects complementing Emerald's current portfolio. The social sciences collection offers 200 volumes across 39 series titles in a diverse range of topics such as Library Science, Education, Sociology, Psychology and Healthcare.

Institutions can offer their patrons online access to the most recently published books in the field, available from multiple locations through a platform that provides the ability to search across both e-Books and journals to find relevant information quickly.

About Emerald: Established in 1967, Emerald Group Publishing Limited is the world's leading publisher of management research. In total, Emerald publishes over 710 titles, comprising 210 journals, over 300 books and more than 200 series and serials as well as an extensive range of online products and services. Emerald is COUNTER-compliant, meeting the international code of practice for reports that measure usage of online information products and services consistently. Emerald's publishing philosophy is 'Research you can use', which means that our content is both rigorous and relevant, connecting the academic world with the world of management practice.

CONFERENCE SPONSORS

