Report

Eighth International ISKO Conference, London UK, 13-16 July, 2004

Nancy J. Williamson

The eighth ISKO Conference held in London, UK, July 13-16, 2004, was organized by ISKO and the School of Library, Archive and Information Studies, University College London.

It was superbly arranged and a stimulating and profitable time was had by all. Vanda Broughton and Ia McIlwaine served as Conference and Programme Chairs, respectively. Fifty-five papers had been selected to be published in the proceedings and to be presented at the Conference. The keynote address was given by Clifford Lynch, Executive Director of the Coalition for Networked Information. The papers were categorized and organized into nine sessions, with concurrent presentations in each, for a total of 18 meetings covering the nine themes. A small number of papers were not presented at the Conference because the authors were unable to attend. As a result, there was some minor reorganization of the programme. In view of this, for this report, it has been decided to retain the original categories as set out in the published proceedings. Where several sessions were held in the same category these have been drawn together and all references made to page numbers refer to the published proceedings[1]. The following is a brief summary of the papers and their topics.

Under the theme "Theoretical Foundations of Knowledge Organization" three sessions were held, with a total of eleven papers on this theme. The first of these sessions focused on three aspects of the theory of knowledge organization – semantics, classification systems and concept maps. To some extent these three papers epitomized the broad coverage of the Conference in general. Hanna Albrechtsen and her colleagues addressed the "Categorical complexity in knowledge integration" in the context of an empirical evaluation of a cross-cultural film research collaboratory. Professional archivists had collaborated on the co-creation of film censorship history interacting with prototypes and addressing problems of use. Two distinct approaches were taken in the management of knowledge integration in the processes of annotation and indexing of documents. Findings suggested that deeper semantics of knowledge integration results when the annotation and indexing were linked in collaborative decision making. Clare Beghtol, in her paper, investigated "Naive classification systems and the global information society." Classification was characterized as belonging to one of two types – "professional" classifications and "naive" classifications. The latter was defined as being classifications "developed by people who have no particular interest in classificatory issues" (p. 19) but who are interested in knowledge organization for their own purposes. Specific examples of the classification of objects are included. Ultimately, the author concludes that classification serves different purposes in different contexts, but the two kinds can be seen to be dependent on each other. Communication among scholars begins at a basic level and is facilitated by naive classificatory activity. Eventually this activity accumulates and results in literary warrant leading, in some cases, to consensus among scholars, and in turn to the incorporation of scholars' original ideas into "professional information retrieval classifications." Further study could advance our knowledge as to how classification can span boundaries between cultures and between disciplines in a global information society. In the third paper of this session, Terence Smith and Marcia Zeng described "Concept maps supported by knowledge organization structures." Specifically, the focus was on the use of such maps in a particular system – the ADEPT (Alexandria Digital Earth Prototype) Learning Environment. Two approaches to the construction of such systems were described – bottom up and top down approaches. The experiment reinforced the need for heterogeneous digital materials at the conceptual level. Just having web portals and search engines with an un-integrated collection is insufficient.

In the second session on "Theoretical Foundations" the papers were somewhat more closely related. All three focused on some aspect of faceted classification. Ceri Binding and Douglas Tudhope suggested "Integrating faceted structure into the search process." Addressing the shortcomings of search interfaces, the authors propose the use of fac-

eted structure as a basis for mediation between searcher and indexer in guiding query formulation and reformulation by educating users about the domain. The use of structure building is demonstrated through focus on the query, on query term expansion, and on query term relationships exemplified with the use of the *Art & Architecture Thesaurus.* The ultimate conclusion was that "significant opportunities and advantages exist for the integration of faceted structure into conventional search systems, but benefits may not be realized without a complementary level of indexing" (p. 71-72). In a paper on "The Bliss Classification in action: moving from a special to a universal faceted classification via a digital platform," Vanda Broughton and Heather Lane focus on the differences in the functional requirements of a faceted system in a print-based environment (where the emphasis is on the browsing function) versus its application to digital collections (where the emphasis is on retrieval). The authors describe the nature of the scheme and the development of an online tool for indexing digital resources, using Bliss terminology and the advantages of facet analysis within the prototype tool. Findings indicated that while the conceptual structure remains similar in print and online, the two environments make different demands on content and syntax and the manipulation of the scheme. Thus the process of the content description can be markedly different. Continuing on the theme of faceted classification, in her paper entitled "Adventures in faceted classification: a brave new world or a world of confusion?" Kathryn La Barre uses a survey of definitions and current applications of facet analytical theory to develop a framework for the analysis of websites. The goal was to explore current practices, uncover common misconceptions, determine the degree of understanding and highlight developments that augment traditional practice. Findings indicated that the tools of facet analysis have much to offer organization of the web, but the author warns of the folly of incomplete understanding. "The stakes are high, confusion is rampant and much remains to be done in this brave new world!" (p. 83).

Testifying to the primary importance of "Theoretical Foundations," the third session in this category contained four presentations. Elin Jacob continued a dominant theme of the Conference in her discussion of "The structure of context: implications of structure for the creation of context in information systems," focusing on structural differences between systems of classification and categorization that lead to differences of meaning in context. The differences were analysed and the role of context discussed through the analysis of four information environments – classification, post coordination, precoordination and free-text searching. In her paper, Uta Priss argues the suitability of "A Semiotic-conceptual framework for knowledge representation," and describes such a framework and its applications. As a beginning, Priss claims the advantages of such a framework in the context of an example from an ontology language and indicates plans for further study. Giovanni Sacco proposes "Accessing multimedia infobases through dynamic taxonomies" as a solution to the failure of traditional query methods in browsing such databases. He discusses these taxonomies and uses as an example the browsing of important painters of the Renaissance in a system where the content is described by metadata. In the final paper in this session Joseph Tennis addresses "URLs and intertextuality: incumbent philosophical commitments in the development of the semantic Web."

Under the all-important theme "Linguistics and Cultural Approaches to Knowledge Organization," two sessions were held with a total of six papers. In the first session, Rebecca Green and Lydia Fraser discussed research into "Patterns in verbal polysemy." Recognizing lexical ambiguity inherent in polysemy as a major challenge in such tasks as information retrieval, information extraction, machine translation and text summarization, the authors hypothesized that "there is an inventory of recurrent semantic relationships accounting for a significant proportion of verb polysemy, much as has been proposed for nouns." (p. 29). Their belief is that awareness of these relationships will improve our understanding of language-based search systems. Previous research is outlined and methodology using approximately 600 verbs from the *Longman Dictionary of Contemporary English* is described. Results are discussed and the conclusion reached that the "most common semantic relationships underlying verbal polysemy in English are metaphor, hierarchy, semantic opposition and metonymy."(p. 34) These relationships could occur in all, or most, human languages. However, the verb senses may not parallel those in another language. Awareness of these general patterns could be used as aids to searchers and thesaurus developers. In their paper on "Terminological representation of specialized areas in conceptual structures," Maria López-Huertas *et al.* addressed the interdisciplinary subject 'gender studies.' This subject area was chosen because it is fast growing and diversified and it draws on several disciplines for its content. Four specialized

thesauri were used and two methodological approaches were applied – a study of terminology used and a structural analysis of the domain. Findings indicated that there is a lack of uniformity in the terminologies of thesauri, leading the authors to conclude that a) there is a lack of consensus on the terminology of the domain, b) common epistemological criteria are lacking, and 3) there is a need for objective methods of terminological extraction from the literature. With respect to the structure of the domain, there was severe conceptual dispersal of the models built to represent the domain. As well, there was a lack of consensus on the scope of the domain and its structural representation in the sources consulted. They conclude that there is a need for an ideal KO model for this and similar disciplines. The final paper in this session prepared by Fidelia Ibekwwe-SanJuan and Eric SanJuan was entitled "Mining knowledge chunks in a terminology network." The authors hypothesized that "syntactic variations are an interesting alternative to the clustering approach" (p. 41) and they offer meaningful ways of highlighting and organizing associated research topics in a corpus." For purposes of study, they developed a text mining and topical mapping system called "TermWatch" that clusters text prior to linguistic processing. Terms were extracted and clustered and the structure of the research topics examined. The clusters permit the domain specialist to visualize the structure of the domain topics from a very large corpus. The process is time saving and is a possible means of discovering new knowledge. Further investigation is in progress.

Two of the papers in the second session on "Linguistic and Cultural Approaches" focused on multilingual topics. Graciela Rosemblat and colleagues investigated "Adapting a monolingual consumer health system for Spanish cross-language information retrieval." The study applied a bilingual term list (BLT) approach to cross-language information retrieval in a specialized domain and compared it to the use of machine translation (MT). A list of training queries from *ClinicalTrials.gov* were translated into Spanish and submitted to a test database of English documents. The retrieval results were then compared with retrieval from the same database to which MT had been applied. The MT approach was more effective than the BLT method when measured against a monolingual English standard. The authors state that translation is only part of the problem. "English retrieval for a Spanish query is not enough and more work is needed" (p. 320). Matjaz Zalokar's paper discussed the "Preparation of a general controlled vocabulary..." in two languages, Slovene and English, for use in indexing the COBISS.SI library information system in Slovenia. This is one aspect of a long-term project in which a new subject list in Slovene is being developed from *Sears List of Subject Headings.* The first stage, the process of translating and editing *Sears,* is described. Also new terms from the *Library of Congress Subject Headings* are being added. An interesting aspect of the paper is the decision to consider faceting the *Sears* headings for use in a post-coordinate system. The final paper in this group deals with actions of users as opposed to language problems. Widad Mustafa El Hadi and her colleagues addressed "Coaching applications: a new concept for usage testing on information systems." The research uses an electronic coaching application called K-Now that observes a human/computer interface that works out the use of existing applications in a corporate environment. It has intelligence and enables the system to work out problem situations and warn users of impending problems "just in time" (p. 351). The system and the impact on participants are described. The authors suggest that the system has application in other kinds of human/computer interaction.

This Conference included numerous papers on the theme "Applications in Knowledge Representation," Four sessions were placed in this category, for a total of twelve papers. In the first session three approaches to knowledge representation were described. Jin-Cheon Na *et al.* focused on the "Effectiveness of simple linguistic processing in automatic sentiment classification of product reviews" (i.e. classification of product review documents). Focus was on text features that would enable the categorization of the reviews as to whether the products are "recommended" (positive sentiment) or "not recommended" (negative sentiment). Several methods were investigated and the use of "negative phrase" (p, 41) with simple linguistic processing gave the best results. Further study of semantic and syntactic processing and inferencing is anticipated. In his paper, Daniel O'Keefe examined "Cultural literacy in a global information society-specific language." Cultural literacy is defined and it is assumed that a discourse community must have a common vocabulary and that a taxonomy can be developed to create relationships among terms in such a vocabulary. The selection of the terms for the taxonomy and their organization into categories is described. A classification was constructed. Future study will include the extension of the taxonomy and examination of the software. In a completely different approach, Lynne Howarth investigated "Model-

ling a natural language gateway to metadata-enabled resources." The paper describes the findings of a focus group assessment of a natural language gateway derived from mapping and categorizing terminology from nine metadata schemas. Semantic ambiguities in three metadata elements were identified and discussed in relation to methods of data collection used in conjunction with the design of an *interlingua* gateway to multilingual resources. Implications for further research are addressed.

In session two on "Applications," two of the papers were closely related. Gerhard Riesthuis and Maja Zumer's paper "FRBR and FRANAR: subject access" discusses the topic of subject access in a context that up to now has been primarily concerned with the principles of the 'other half of cataloguing' – descriptive cataloguing. In doing so, three essential documents from those earlier discussions were examined – the *Functional Requirements for the Bibliographic Record (FRBR): Final Report*, the "Functional Requirements of and Numbering of Authority Records" *(FRANAR)*, and the unpublished draft of the "Statement of International Cataloguing Principles" from the First IFLA Meeting of Experts on an International Cataloguing Code. From their analysis, they conclude that while subject cataloguing has a place in these documents, it is a somewhat minor place. This led to the conclusion that, in view of the wide use of subject access in online catalogues, subject access deserves greater attention. Five areas for further research are proposed. Closely related to this paper was Victoria Frâncu's presentation "An interpretation of the *FRBR* model." The author describes the transition from the traditional bibliographic record-to-record proposed by the *FRBR*, analyses the structure of the newer format and interprets its application to the integration of traditional and digital resources in the context of descriptive cataloguing. In the final paper in this group, Moshe Sachs and Richard Smiraglia address the global organization of knowledge "From encylopedism to domain-based ontology for knowledge management" with specific reference to the "Sachs Classification" (evolved from the *Worldmark Encyclopedia of the Nations*). The scheme provides for potentially powerful knowledge management "through the development of domain- and ecology-specific ontologies" (p. 167). The background for the study is provided, the classification described, and its use as a practical approach to knowledge management explained. The authors foresee potential for an expanded SC in the ordering of digital libraries, for use in curriculum development, in the generation of a new encyclopedia, in research for international investment and development and in the ordering of a global search engine for international trade.

In the third group of papers in the "Applications" category, there was one presentation on the "Translation of classifications." Barbara Kwasnik and You-Lee Chun addressed problems of translating *DDC* into Korean. A concept-by-concept comparison revealed differences and similarities in terms and structure that can be attributed to language and cultural differences. They observed that universal systems are now being stretched by cultural and linguistic situations quite different from those for which they were originally intended. Cultural and linguistic artefacts are shifting, making the harmonization of classifications increasingly difficult. In this same category, Hur-Li Lee and Jennifer Clyde focused on "user perspectives" in the online catalogue. In doing so they address three questions: "Is the concept of "collection" still relevant for organizing information resources in the virtual environment, where space has different connotations? Do collection structures, both physical and virtual, help users seek information and if they do, how? Do users and librarians have the same criteria for organizing information resources, both within and outside the library collection?" (p. 199). Findings suggested that a number of parameters needed by users were missing. Secondly, it was suggested that cataloguers should move away from the normal and concentrate their attention on individual needs. Thirdly, collocation by subject was found to be deficient. Users need spans of numbers they can consult to simulate browsing the library shelves. In the third paper, Jens Erik Mai's look at "The role of documents domains and decisions in indexing" reinforced the need for a catalogue that a user understands and strengthened the idea that the indexer should be paying more attention to the user. All of this suggests that the 'black box effect' has yet to be conquered.

The fourth session on "Applications" continued the focus on topics related to improving user access to information. Claudio Gnoli's paper on "Naturalism vs. pragmatism in knowledge organization" addressed the need to free knowledge from the traditional disciplines in order to achieve a more naturalistic approach to indexing and categorization in document organization.The authors refers back to approaches to language as understood by Descartes and others, and to the artificial auxiliary languages developed in the 19th century. The ultimate focus was on the possibilities of naturalistic principles of knowl-

edge originating in either of two ways – either through categories of perception (i.e. epistemology) or through structure of reality (i.e. ontology). The consideration of structure was based on the work of Dahlberg and the notion of integrative levels as endorsed by the Classification Research Group (CRG). Gnoli concluded that the dialectic between naturalism and pragmatism has always existed in knowledge organization. Further, he states that pragmatism must always be kept in mind, but that the naturalistic approach has possibilities. In "On the razor's edge," Wouter Schallier examined problems of classification systems in a different kind of dichotomy – the tension between local and over all needs in knowledge organization. The problem is seen in the context of a particular university library that uses the DOBIS/LIBIS system. In search of the solution, a new search interface for e-resources has been developed using a *UDC* authority file linked to descriptors in three languages – Dutch, English and French. In addition, *LCSH* and *MeSH* have been used to enrich the file. The OPACs using this system came into use in September 2003. This project is a small part of a much larger undertaking that will include the replacement of DOBIS/LIBIS with the Aleph 500 Library System. The system is described, illustrated and evaluated. The final goal is the provision of a retrieval system that will respond favourably to both local and overall needs. The third article in the category, by Danielle Miller, "User perception and the online catalogue," looks at what pubic library OPAC users themselves think about the catalogue. The author considers the needs of users in terms of catalogue description, display and navigation. The purpose of the study was to consider the implications of these factors for future research in online catalogue design.

In a single session on "Special Applications" four papers were brought together by virtue of the fact that they pertained to subjects in the medical field. Carol Bean addressed the topic "Representation of medical knowledge for automated semantic interpretation of clinical reports." Case reports from 'cardiac cauterization' were used to identify and code text for automated interpretation of semantic indicators of location and severity of disease in coronary arteries. The results of the identification will be used to develop semantic and syntactic interpretation rules for an existing automated interpretation system. In a closely related paper, Chew-Hung Lee *et al.* explored "Automatic identification of treatment relations for medical ontology learning" as part of a project to develop an automatic method for building ontologies in the medical field. A. Neelameghan and M.C. Vasudevan used a knowledge base on tumours of the central nervous system to carry out a case study on integrating image files, case records of patients and web resources. The result was a database that could be used for various purposes – patients' case registry, information retrieval, the generation of reports and statistics, and the preparation of papers for scholarly publications. While the first three papers were concerned with information storage and retrieval the final paper in this group by Nancy Williamson addressed a problem in the development of a classification scheme – where to locate the domain "Complementary and alternative medicine" in a general scheme, and in *UDC* in particular. The problematic area is the lack of a clear division between traditional and alternative medicine and the tendency for individual topics to migrate from alternative into traditional medicine. The investigation indicated that there is some evidence that a location can be provided for established schools of practice, while individual therapies can be grouped and then applied as needed under any body system, organ or disease to which they pertain. Using *BC2* as the framework for a restructured *UDC* Class 61, it appears that *UDC* is sufficiently flexible to deal with the problem.

As is typical of conferences in this field, there were a number of papers dealing with "Knowledge Organization of Universal and Special Systems," While there is an emphasis on the new and different, it was no surprise that there was one session devoted the *Dewey Decimal Classification (DDC)* and another on the *Universal Decimal Classification (UDC).* Of particular interest is the question of how these 19th century systems are adapting to the world of computers and the Internet.

In 2003, *DDC* celebrated the publication of its 22nd edition and it was appropriate that Joan Mitchell present her view of "*DDC* 22: Dewey in the world and the world in Dewey." She discusses the shaping of *DDC* by a number of social, geopolitical and technical trends in modern life. Among these are the effect of the World Wide Web and the importance of translation. A daunting question still remains – "How can we maintain the internal cohesiveness of the system and still represent the world in Dewey?" Extensions to *DDC* are many and mapping is identified as a tool to aid the broader representation of terminology and topics. The need to employ a number of methods to meet the needs of the global user community is recognized. From a different perspective, Sudatta and G.G. Chowdhury carried out a project "Using *DDC*

to create a visual knowledge map as an aid to online information retrieval." A prototype knowledge map was created. In the third paper in this session, Diane Vizine-Goetz and Julianne Beall presented a study "Using literary warrant to define a version of the *DDC* for automated classification services." The authors describe a version of *DDC* for use with three websites that use *UDC.* The sources, BUBL, Canadian Information by Subject and KidsClick, were used to create a database for the Natural Sciences and Mathematics, excluding the Life Sciences. The modified *DDC* is being tested as a source database for machine classification services.

In another session on "Knowledge Organization and Universal and Special Systems"*UDC* was the uniting theme. Using the UDC system as a starting point, Ágnes Hajdu Barát explored the challenges of the role and representation of information retrieval languages in the digital online environment and the Internet. In doing so, she responds to the problems in the use of *UDC* and considers possible new solutions and user-friendly methods. In "A question of place," Ia McIlwaine brought her experience and expertise with *UDC* to bear on the complex problem of creating and maintaining geographic tables in classification schemes. Problems identified and described are historical, grammatical, governmental administration, the legacy of colonialism, non-political regionalism and language. Hierarchical notation is posed as the most viable solution. While the paper was written in the context of *UDC,* the problems are applicable to any scheme where a geographic table is required. To conclude this session, Aida Slavic and Maria Inês Cordeiro identified the "Core requirements for automation of analytico-synthetic classifications" using the data structures from three general analytico-synthetic systems – the *Bliss Bibliographic Classification (BC2)*, the *Broad System of Ordering (BSO)* and *UDC.*

A third session on "Universal and Special Systems" focused on thesauri and other related tools as opposed to traditional classification systems. Stella Dexter Clarke, Alan Gilchrist and Leonard Will prepared a paper for the Conference on the "Revision and extension of thesaurus standards." Written in anticipation of the impending revision of international standards ISO 2788 and ISO 5964, and corresponding national standards, including the American ANSI/ NISO Z39.19 standard, this paper addresses some issues that are still under discussion in the various standards committees. Such issues include the treatment of facet analysis, the coverage of additional types of controlled vocabulary such as classification schemes, taxonomies and ontologies, and mapping from one vocabulary to another. Among the questions posed are: "Are thesaurus standards still needed?" (p. 213-214). and "What aspects of the standards need revision and/or extension?" (p. 214-215). Progress to date is discussed. The article is testimony to the fact that controlled vocabularies are still seen as important tools in information storage and retrieval. One of the major issues is the relationship of one controlled vocabulary with another, in both different languages and the same language. Michèle Hudon prepared a paper on the "Conceptual compatibility in controlled language tools used to index and access the content of moving image collections." Five controlled vocabularies in this field were examined to determine their degree of conceptual compatibility. Considerable overlap in the vocabularies was an indication that there is potential for a long term goal of one common basic indexing and access language in this field. In this session also, a paper by Antonio Garcia Jiménez and Félix del Valle Gastaminza entitled "From thesauri to ontologies" presented a framework for elaborating languages in the context of a digital photograph collection. Several indexing models are described, different indexing languages are proposed and a theoretical revision of ontologies in the field was carried out. Finally, in this session , Ali Asghar Shiri and Crawford Revie considered "End-user interaction with thesauri" through an evaluation of the cognitive overlap between end-users' initial queries and thesaurus structures. The investigation involves genuine search tasks in the use of the CAB Abstracts database by academic users in the field of veterinary medicine. Findings indicated a high percentage of exact and partial matches between users' terms and thesaurus descriptors. Users also found the thesaural relationships important in the narrowing down of their search topics.

Knowledge organization is always carried out in some kind of social context and the theme "Social & Sociological Concepts in Knowledge Organization" provided four papers.

Grant Campbell, in his paper "A queer eye for the faceted guy" looked at how the principles of faceted classification could be applied to a distinct subculture of society, the gay community. Similarly, Jonathan Furner and Anthony Dunbar analysed "The treatment of topics relating to people of mixed race in bibliographic classification schemes." The authors deal with the bias that is inherent in many universal schemes. Here, the particular focus of the study is

DDC. In a broader approach to the problems of social context, Peter Ohly investigated "The organization of Internet interlinks in a social science clearing house." The term 'clearinghouse' in this context is synonymous with gateways or virtual catalogues currently being developed to harness the Internet documents into manageable groups. Among the issues addressed are the selection of sources, the differences between Internet sources, the descriptive elements, the quality of source entities and the management of an Internet resource database. Finally, Chern Li Liew examined the knowledge organization issues in "Cross-cultural design and usability of a digital library supporting access to Maori heritage resources." Specifically, the paper provides background material on the culture and resources and the specific requirements for structuring, storage, organization and retrieval in a digital environment. In his conclusion, the author stresses the importance of providing a system that responds to local needs, has an appropriate classification, and uses interfaces that are amenable to user capabilities.

Another category of special interest at this Conference was "Knowledge Organization in Corporate Information Systems." Three papers were presented on this theme. Anita Colman investigated "Knowledge structures and the vocabulary of engineering novices" in which the language of undergraduate students in civil engineering was used. Concepts and relationships in the area of 'soil consolidation' were used to understand the knowledge of novices and compare them with the knowledge of human experts and a thesaurus. "Results show that there is little similarity between the knowledge structures of the novice, the expert and the tool" (p. 281). In their paper, Evelyn Mounier and Céline Paganelli dealt with "The representation of knowledge contained in technical documents." FAQs (frequently asked questions) were used in the study. The model proposes a method for automatic recognition of information unit types contained in documents. In the third paper, Martinus van der Walt addressed the possibility of designing "A classification scheme for the organization of electronic documents in small, medium and micro enterprises (SMEEs)." Facet analysis would be used. However, it was recognized that such a scheme would need to be tailored to each individual situation. Thus, the specific question is whether a "standard" scheme could be developed that would provide "a backbone classification" usable in any type of enterprise, so that each enterprise would not have to start from scratch.

There is a growing interest in the organization of images and three papers were selected for presentation under the theme "Knowledge Organization of Non-Print Information: Sound, Image, Multimedia." Laura Bartolo *et al.* investigated "Information management of microstructures" with particular reference to non-print multidisciplinary information in a materials science digital library. The study considers the management of a range of document specific materials used by novice and expert materials science users. More specifically, these are microstructures of soft materials such as polymers, liquid crystals, colloids, DNA, proteins and connective tissues. The investigation is relative to the development of a Materials Digital Library as part of the National Science Foundation's National Science Digital Library Program (NSDL). In the same general domain, but in more general terms, Pauline Rafferty and Rob Hidderley carried out "A survey of image retrieval tools." Recognizing the fact that document retrieval tools do not work for all documents, this paper discusses the different "flavours"of three image retrieval tools – *Art & Architecture Thesaurus*, *Iconclass*, and the *Library of Congress Thesaurus for Graphic Materials.* In the third paper, Richard Smiraglia addresses "Knowledge sharing and content genealogy: extending the "works" model as a metaphor for non-documentary artefacts with case studies of Etruscan artefacts." It aims to extend the concept of "works" as paper documents to artefacts. The paper reviews the characteristics of documentary works and extends the metaphor from the documentary environment to the artefactual environment. Then, by altering the terms slightly, it moves from the documentary domain into the artefactual domain, harmonizing case studies of Etruscan artefacts from the University of Pennsylvania Museum of Archaeology and Anthropology are used to demonstrate the inheritance of "works" in non-documentary artefacts.

Theoretical foundations provided a background in many of the pragmatic discussions during the Conference. However one session focused specifically on"Theories of Knowledge and Knowledge Organization" per se. In the first of three papers, Keiichi Kawamura examined the work of well known modern theorists in a paper on "Ranganathan and after: Coates practice and theory." The author looks at Eric Coates' contributions through an examination of the *British National Bibliography,* the *British Technology Index* and the *Broad System of Ordering.* In doing so, he focused on how Coates' works relate to each other, why his achievements are important globally

and which of Coates' problems throw light on the still unsolved problems of knowledge organization. Of particular interest is the fact that Eric Coates was present in the audience when this paper was presented. In this same group, Shiyan Ou *et al.* presented a paper on "Automatic discourse parsing of sociology dissertation abstracts as sentence categorization." Decision induction was used for the automatic categorization and three models were developed and are described. Findings suggest that sentence position information increases categorization accuracy. Further research is planned. Finally, Iolo Jones *et al.* researched "Natural language processing and knowledge organization systems as an aid to retrieval." Here, the authors address several aspects of this broad topic including disambiguation of homographs and nominal compounds in free text. The use of *Roget's Thesaurus* as an intermediary in the process is discussed. A brief review of relevant literature is included and design considerations are presented. Some success was achieved but further refinement is needed.

This was indeed a very full conference that brought forth many interesting findings and perspectives. The organization of the sessions into themes and topics presents one approach to the content, However, close examination of the content and the index to the proceedings indicates that there is a good deal of cross fertilization among the various component parts. Many theories, tools and applications are present in numerous papers across the whole Conference. The Conference was very rich in content.

Note

1. McIlwaine, Ia. C., ed (2004), *Knowledge Organization and the Global Information Society: Proceeding of the Eighth International ISKO Conference, 13-16 July 2004. London, UK.* Würzburg, Germany: ERGON Verlag.