66

Knowl. Org. 36(2009)No.1
N. J. Williamson. Classification Issues in 2007

Classification Issues in 2007

Nancy J. Williamson

Faculty of Information, University of Toronto, 140 St. George Street,
Toronto M5S 3G6 Ontario Canada <william@fis.utoronto.ca>

73rd IFLA in South Africa, 2007

At the IFLA Conference in Durban, the Classification and Indexing Section participated in two programmes—in a joint presentation with Bibliography and National Libraries on "Bibliography in the Digital Age" and in the Section's own programme. The joint programme was introduced by Ingrid Parent (Canada) in a general paper on "The importance of national bibliographies in the Digital Age," while Maja Zumer (Slovenia) addressed "The new Guidelines for national bibliographies in the digital age." These papers set the stage for more specific topics such as the paper by Françoise Bourdon (France) and Patrice Landry (Switzerland). Their interim report from the Working Group on Guidelines for Subject Access by National Bibiographic Agencies provided a brief historical introduction to the papers to follow. The work began with the collection of practices in libraries on the nature of subject access provided for various types of user groups. The paper describes the activities of the working group, its terms of reference and steps already taken to select the key criteria for defining the type of subjectaccess which should be made available for users. The goal is to be able to offer advice on the types of documents that should be indexed based on a knowledge of user groups and to offer support in the creation of indexing policies. The methodology being used is outlined and some key elements are proposed for an indexing policy. As the work continues, the group will consider the relevance and usefulness of the elements for an indexing policy. Questions that should be answered with regard to such a policy were identified as was the nature of the presentation of the tools and their application. The next steps to be taken were outlined and the Group hopes to be able to submit its recommendations for consultation in 2008. Indeed,

the programme for the 2008 IFLA Conference, in Quebec, Canada, indicates that there will be an update on the Guidelines at that time.

As one segment of the programme of the IFLA-CDNL Alliance for Bibliographic Standards (ICABS), Barbara Tillett and Corey Harper (USA) presented a paper on the "Library of Congress controlled vocabularies, the Virtual International Authority File, and their application to the semantic web." The paper, based on an earlier article written by the authors for *Cataloging & Classification Quarterly* (v.3/4, 2006) reviews the Library of Congress controlled vocabularies and describes the VIAF project (Virtual International Authority File). The goal is to use these controlled vocabularies "as building blocks for the SemanticWeb to internationally link the world's authority data from trusted sources to benefit users worldwide." The authors point out that very rich sources of authority data are available and suggest that they could be moved to Semantic Web standards using such tools as the Web Ontology Language (OWL) and the Simple Knowledge Organization System (SKOS).

In addition, the Standing Committee on Classification and Indexing presented its own programme of three papers. James Turner and Suzanne Mathieu (Canada) gave a paper on "Audio description text for indexing films." It opens with the underlying premise that "access to audiovisual materials should be as open and free as access to print-based material." and points out that this not yet a reality. The authors were investigating audio descriptive text versus written text as a basis for the indexing of moving images. Also, they considered the possibility of automatically translating keywords from audio description into other languages for purposes of indexing. The paper reports on work in the context of the E-Inclusion Research Network which has the goal of "creating powerful audio-visual tools" to permit the improvement of the "multi-media experience for the blind, the deaf, the hard of hearing and the hard of seeing." Specifically, the research focuses on kinds of informa-

Knowl. Org. 36(2009)No.1
N. J. Williamson. Classification Issues in 2007

67

tion needed by persons who are blind or have loss of vision in order to understand moving images. In the process, the authors viewed a number of films together with persons with this disability and asked them to comment on the audio description. Was it adequate? What improvements would they like to see? Another question—Could the vocabulary from the audio description be used for indexing? The study had three objectives—to validate the typology of kinds of description offered by the audio text; to compare the keywords in the audio description with keywords in users descriptions written for each shot to determine the possibility of indexing from these texts; and to compare English keywords with French in order to assess the possibilities of automatic indexing between the two languages. The methodology is detailed and the Excel sheets are explained. Results and analysis are outlined and discussion provided.

In the second paper, Jonathan Furner (USA) described research on "User tagging of library resources: Toward a framework for system evaluation." The author points out that user tagging in the searching of library resources has shown "substantial promise." However, there are still questions that need answers, such as: What are the factors that determine whether user-tagging is successful? Which systems perform the functions expected of them? The characteristics of user tagging are described; Web based sources that provide for tagging services are identified; and the characteristics of user tagging that distinguish this process from those used in conventional methods such as subject cataloguing, abstracting and indexing and bibliographic classification are listed. The question of "What specifically are the factors that determine whether user-tagging services will be successful?" is addressed. Users groups are described as belonging to one of two groups—the implementers of the service (i.e. the collection managers) and the end users who are interested in the contents of the collections to which the service provides access. Individuals vary in their motivations and goals and have different mental models of the use of such services. In this context, the author distinguishes users at three levels: by kind of primary motivation which may be individualistic or social; by kind of primary usage—tagging for creation of descriptions of resources, or as searching the location of resources; and by kind of ultimate goals of end users, be it to engage in some deeper form of interaction, or simply to complete a particular task. A final question addressed is "On what criteria should our judgment or evaluation of the quality of the tagging service's performance be based?. The author states

that "it is common to distinguish between criteria such as effectiveness, efficiency, cost-effectiviness and usability." The relationship between the quality of results and the level of effectiveness raises a whole different set of questions relative to manual vs automatic subject representation and derivation of terms from the resources vs the assignment of terms from other sources. A general indicator of effectiveness is represented in the indexer–searcher consistency (i.e. how well has the indexer been able predict those terms that will be used by searchers?). In the end, the author is not confident that focusing on retrieval effectiveness as the most important criterion is the one on which tagging service should be measured. The paper identifies a number of factors on which to focus further discussion. His final assumption is that in the future evaluation of "user-tagging devices" that there be clarification and justification of research-design choices of several kinds be provided.

In the third paper, L. Sulistyo-Basuki (Indonesia) addressed the question of "Greater subject access to Dewey Decimal Classification's notation, with special reference to Indonesia's geography, period and language notations." By 'notation" the author means the concepts and their organization into hierarchies rather than the numbering system. He cites problems for Indonesian librarians and library users in cases in which the captions and structure do not reflect "the true condition of Indonesia" with respect to three areas: Indonesian languages, geography and historical periods. It appears that the problems are bound up in the political developments within the country and the numerous events that have taken place over the years. The author noted that with respect to historical division there have been radical changes in *DDC* from its 15th to 22nd edition, not all of which he finds helpful. He cites the problem of American bias in *DDC* which he says he understands but calls for *DDC*'s editors to recognize international usage as a reason for improvement. Various solutions are discussed with respect to the three areas of concern. As an overall solution he suggested there could be compromise but concluded that the organization in *DDC* should reflect the common usage by users and librarians in Indonesia and proposes that the existing content and organization in the *DDC* 22nd edition be revised according to Indonesian views as reflected in various sources.

68

Knowl. Org. 36(2009)No.1
N. J. Williamson. Classification Issues in 2007

UDC-Seminar:
Information Access for the Global Community

In 2007 a number of seminars, conferences etc. took place that were specifically related to classification. Among them was a two day seminar on the UDC organized by the UDC Consortium and held at the UDC Headquarters in The Hague on June 4th and 5th 2007. The event was entitled "Information Access for the Global Community" and it brought together publishers, editors and users of the UDC and provided a forum for information exchange and discussion of future developments. In the opening presentation, Maria Inês Cordeiro, Editor-in-Chief of UDC, stated the objectives of the seminar as follows: 1) "to create awareness and sense of community among UDC users; 2) to raise collaborative opportunities; 3) to hear about users' needs; and 4) to improve support for UDC." Presentations were made on behalf of the British Standards Institution (BSI), the British publisher, and the UDC Consortium. Inês set the stage for the remainder of the seminar in her "Introduction to the Editorial Agenda" In doing so, she briefly described the editorial work of UDC in terms of the various kinds of expertise needed—subject expertise, UDC expertise, and UDC data management. In the second segment of her presentation she referred to the work in progress. The Editorial Board has been expanded by 10 new members in order to provide help in the revision of different subjects and to reflect wider geographic representation. The new board members are primarily from Eastern Europe where UDC is used intensively. Also there are to be 7 associate editors. whose role will be to aid in the establishment of policy, to find collaborators who will take on revision tasks, to be involved in the discussion and checking of proposals for revision, and to aid in the management of the Master Reference File (MRF) and in the preparation of *Extensions & Corrections (E&C)*. Then she described the structure of the proposed new system and announced timelines for the new structure to be implemented.

The remainder of the seminar consisted of 11 papers presented in 5 sessions. Two of the presentations are available in full text—the paper by Claudio Gnoli and the "wrap up" by I.C. McIlwaine. For the others there are abstracts and power point presentations available at www.udcc.org/seminar2007.htm The power point slides are extensive and quite helpful.

In the second session of the programme, two papers provided an introduction to the remainder of the Seminar. Claudio Gnoli (Italy) presented a paper en-titled "Progress in synthetic classification: towards unique definition of concepts." This paper briefly described the evolution of bibliographic classification schemes from the end of the 19th century to the present. This shows an increasing tendency for the largely enumerative schemes to make more and more use of synthetic devices. Recent editions of UDC and the UDC-derived FATKS project were identified as examples of this trend, through their use of specific phase relationships and more common auxiliaries, like those for general properties and processes. The ideal is that each concept should have "a place of unique definition" which would stand through the system wherever it is used. He suggested that main classes "should be defined in terms of phenomena rather than disciplines." The Integrative Levels Classification (ILC) is currently addressing this possibility. In the second paper,Wouter Schallier (Belgium) addressed the question "Why organize information if you can find it? UDC and libraries in an Internet world." In his discussion, he focused on the impact of the change in knowledge organization from the print environment to the Internet. It brings us to the present situation where knowledge is "fundamentally unorganized and its content changes constantly." So, if we can find it why should we even try to organize it? "So what will be the role of UDC and libraries in this Internet environment?" The author believes that libraries "can still play a role as a major information provider, fully to the expectations of a modern end user." He emphasized "maximal accessibility, usability and active participation of the end user in the Internet environment" and the use of metadata.

Session 3 included presentations by three representatives of UDC user countries. Maria Balikova spoke about the wide use of "UDC in Czechia" in all types of libraries.

One use of the system is as a switching language to provide multilingual subject access to library catalogues. Here UDC is used to narrow the search space" in their "M-CAST system. Jiri Pika explained the use of the "Universal Decimal Classification at ETH-Bibliotek Zurich." Queries are based on "verbal three-lingual descriptors and corresponding related search terms" including synonyms, user terms and expressions from scientific journals, to provide for a dialog with OPACs. A single UDC number behind the three descriptors connects them to related document titles. regardless of language. UDC is used to sharpen retrieval from the OPAC without the user realizing it. In the third paper Darija Rozman and Boris Rifi discussed the use of the "Universal Decimal

Knowl. Org. 36(2009)No.1
N. J. Williamson. Classification Issues in 2007

69

Classification in Slovenia". The current system is described. A new translation of the Slovenian version of UDC based on the 2001 version of the MRF is being prepared and a new manual of use was published in 2006. A description of the manual was presented and proposals for future developments were outlined.

Session 4 was devoted to three papers on "UDC education and training" in 3 different countries that use UDC extensively. Victoria Francu focused on the question "Does convenience trump accuracy? -The avatars of the UDC in Romania." In doing so she concentrated on major issues regarding the potential of UDC and the current controversy about its use. Specifically, she was concerned with three aspects: the importance of hierarchical structures in controlled vocabularies and their importance to the browsing function in retrieval; the lack of popularity of UDC among users of OPACs; and the situation in teaching UDC in Romania universities. Agnes Haidu Barat described "Multilevel education, training, traditions and research in Hungary" Her presentation was historical in nature going back to 1910. She concluded with brief overview of recent developments, including the situation after the publication of the new Hungarian edition, the use of UDC in Hungarian OPACS and the possibility of UDC visualization. In the third paper Rosa San Segundo Manuel described "The use of the UDC in Spain: implementation, application, teaching and research." Like the previous paper it was historical in nature. UDC was first used in Spain soon after it was introduced into Europe in 1895. but was not specifically designated for use until 1934. Its use was made officially compulsory in all Spanish libraries in 1939. In 1989 its compulsory use was repealed. "although its implementation in libraries, catalogues and bibliographies is almost complete." UDC is taught in library and information science schools from both a theoretical and a practical point of view. Research is considered to be an important factor in terms of translations, adaptations and versions, as well as in analytical works on different aspects of the UDC system.

Moving forward in time, session 5, the final paper session, looked at some "UDC prospects and innovative usages" Three papers described research projects. Antoine Isaac (Netherlands) discussed "Aligning thesauri for an integrated access to cultural heritage resources." A common interest in recent years has been the desirability of making it possible to search a variety of related resources and collections through systems with unified access. A major problem is the fact that "the objects from different collections are of-

ten described using different vocabularies—thesauri and classification schemes—and are therefore not interoperable at the semantic level." A number of methods, such as semantic links and mapping, have been tried (e.g. MACS and Renardus). As stated by this author "they have demonstrated very interesting results." However, aligning them manually is difficult and time consuming. Some work has been done making the alignment a (semi)-automatic task. Isaac described a project in which alignment techniques have been applied to build a (pilot) browser in the context of the STITCH (Semantic Interoperability To Access Cultural Heritage, Catch Program). The browser permits a unified access to two collections of illuminated manuscripts using the descriptive vocabularies of either of the two collections—Mandragore or ICONCLASS. Miguel Benito (Sweden) proposed "The Subject of medicine: the best solution today for the empty class 4." He explained the history of UDC class 4 which has been vacant since class 4 language and class 8 (literature) were combined in class 8. The author makes a strong case for using the vacant class 4 as a location for the proposed revision of Class 61 Medicine. He suggested that practically it would make all class numbers one digit shorter and stated that theoretically only DDC and UDC "have medicine together with other practical disciplines in the same division. Most systems have Medicine as a main discipline with a division on its own." Here it must be noted that length of notation is dependent not only on the breadth of space at the top level but also on the depth of analysis. As pointed out in the round up of the seminar that followed, Ia McIlwaine indicated that Medicine is not the only possible candidate for Class 4. A decision on the best use of Class 4 has yet to be made. In the final paper in this section and in the seminar in general, Erik-Jan van der Linden (Netherlands) and colleagues posed the question "Visual Universal Decimal Classification?" They cited the immense size of the UDC Master Reference File (MRF) and the difficulties of finding a class number, understanding the coding system and use by practitioners. In answer to this, they had produced an application demonstrating "interactive visualization'" The challenges were discussed and a demonstration of how those challenges might be met was presented. The results can only be appreciated by viewing the demonstration itself.

The seminar ended with a panel discussion on UDC future directions involving four experts on the system—I.C. McIlwaine, Alan Hopkinson, Maria Inēs Cordiero and Gerhard Riesthius. This was fol-

70

Knowl. Org. 36(2009)No.1
N. J. Williamson. Classification Issues in 2007

lowed by a conference wrap up by I.C McIlwaine former editor of UDC. In her "Wrap up" (which is available in full text) she restated the goals of the seminar: to give the classification a higher profile, to update the users on what is going on, to suggest innovations, to provide for interchange of ideas and to share experiences and voice needs. It was also intended to encourage more participation in the work of revision by members and volunteers. She then briefly commented on the content of the various presentations. in their aspects—theoretical (Gnoli), the online environment (Balikova), the multilingual environment (Pika), and innovations, UDC as a search too,l practical implementation in several countries, etc. She challenged the point that it is difficult for a user to understand the structure of the classification and find the "right" class mark. In doing so, she stated "I do not think there is such a thing as the 'right: class mark'—one of the 'most valuable features of the UDC is its adaptability to different situations …" However, she is encouraged by new approaches that might make the system "clearer and easier" for users. In closing. McIlwaine welcomed input by all and active participation in the work of revision and outlined how participation can take place. The wrap up concluded with thank yous to all who made the seminar worth while.

North American Symposium on Knowledge Organization

The inaugural conference of the ISKO North American Chapter took place at the University of Toronto, June 14-15, 2007. The two day meeting included a symposium on "Knowledge Organization Research in North America: What have we done, what are we doing, and where do we go from here?" An executive was established and the programme consisted of 13 papers, a panel discussion and 3 poster sessions. Those present were welcomed by the Dean of the Faculty, Brian Cantwell Smith, and the symposium was chaired by Richard Smiraglia. The papers were organized into 5 sessions. While the sessions were not identified by names of sub-themes, they were appropriately grouped. Full text is available on the Web for 9 of the papers. They may be located by going to the website of ISKO-NA, and clicking on "publications." This brings up the programme of the seminar which gives access to the papers. The remaining four papers, those by Abbas, LeBarre, Menard, and Tennis, are included as extended abstracts with full text available in *Knowledge Organization*, v. 34, no. 2, 2007.

Session 1 contained two papers that provided a theoretical background for the rest of the programme. David Pimentel presented "Exploring classification as conversation" in which he proposed conversation as a "useful lens through which to consider knowledge-organizing behaviors". In it, he moves away from traditional modes of access, such as controlled vocabularies and classification schemes, into the creation of knowledge through relationships among various sources of data. As a starting place, he used "conversation theory" to look at the problems facing knowledge organization in the current information world, examining features of knowledge structures such as blogs, and social classifications (e.g. Wikipedia and knowledge from user participation). He considers opportunities for knowledge organization in this kind of environment and suggests possible areas for further research, depending on the management and analysis of data. The second paper in this session, "Ontology and the Semantic Web" by Jane Zhang discussed a new information representation system. It is "based on a more sophisticated semantic representation of information" that "aims to go well beyond the document level" and is "designed to be understood and processed by machine." Such a system would have three features: 1) turning documents into meaningful interchangeable data; 2) reflecting a rising use expectation nurtured by modern technology; and 3) presenting a unique challenge for its enabling technologies. Details include an exploration of semantic relations and their various sources (term lists, classifications and categories, and relationship lists), a discussion of the nature of granular accessibility, and of machine processibility through the use of mark-up languages. It is clear from the discussion that the Semantic Web not yet "here" in a real sense. Among the challenges is the current status of the semantic web. In reality, it is still a vision—there is a very small number pages available and the format is extremely complex and "requires enormous cost in creation and maintenance." Simplification is needed. The author also suggested that more attention should be paid to the development of "Semantic Web Services … than to annotating Web content data."

In session 2, three papers looked at new methods of approaching classification. A paper entitled "Everything old is new again: Finding a place for knowledge structures in a satisficing world" by Grant Campbell et. al described an exploratory project involving Web resources related to Alzheimer's disease in order to consider ways in which RDF metadata could be used to translate the virtues of the traditional vertical file to the Web environment form using

Semantic Web descriptive standards. In doing, so they endeavour to use Web technologies that lend themselves to sophisticated knowledge structures and link them to librarian's skills and practices in information handling. The authors outline the salient points of the strengths and weaknesses of the two systems. In doing so, they suggest that librarian's understanding of the traditional retrieval process, the traditional vertical file and the Resource Descriptive Framework (RDF) offer some insights into the linking process. To test this theory, they designed a project in which they carried out a pilot study using websites related to Alzheimer's disease. The details of the system—the question/answer system in RDF format and a user portal—are described in the paper. While the project is still in progress a number of implications have been identified. In conclusion, the authors state that "while the Semantic Web has not attained the wholesale adoption originally predicted for it, its commitment to rigorous information structure has made it ideal for the migration of library systems and library techniques into Web environments." In her paper on "Beyond retrieval: A proposal to expand the design space of classification" Melanie Feinberg takes a fresh approach to a traditional method of knowledge organization. She suggests that "classification researchers should investigate a wider variety of design possibilities in which the purpose of a classification is not assumed to be a retrieval tool in the traditional manner." Further she suggests that a research area might be one that facilitates the "problem-setting aspects of classification design." This might involve the description of a design language which could be operationalized as a set of product qualities. In turn, these might provide a framework within which designers could understand, evaluate and create classification systems. The paper begins by describing the conditions of classification as a retrieval tool. She gives two examples of other uses of classification—classification as an argument and classification as an element of work practice. They are very different from each other but support the author's argument that retrieval tools are just one part of the possible design space. In her conclusion, she states that "Classifications might be new interpretations of the world as shown through work in philosophy and linguistics." As such, they might structure and affect work practices, instead of reflecting them. If a design language can be articulated, it could provide a mechanism for communication and systematic criticism, while providing structure for the design process. The third paper in this group by Joseph Kasten addressed

"Knowledge strategy and its influence on knowledge organization." It takes a more analytical approach than the other two papers and looks to the business world for process rather than classification. Knowledge strategy is described as a relatively new topic derived from business literature and is further detailed as "a set of guidelines and beliefs that shape an organization's manipulation of knowledge." The author provides a literature review and poses two research questions as follows: "Does knowledge strategy influence knowledge organization methods or approaches? and "Are there organizational factors that influence knowledge organization?" A methodology was set up using "semi-structured interviews to be performed by persons involved in strategic planning of large organizations." Participants came from five organizations: 2 hospitals, 1 accounting firm, 1 bank and 1 financial services firm. Interviews of the participants were recorded and transcribed and a content analysis done. This led to the identification and classification of the knowledge strategy of each institution and enabled the identification of the knowledge processes used. The results revealed at least two important trends that were evident "Organizations that are more reactive in their knowledge acquisition tend to centralize their knowledge organization while those with a more proactive approach tend toward a distributed knowledge base. Likewise, forms that are more proactive tend toward more distributed knowledge organization methods." More research is needed to be able to understand the relationships.

The third session contained three papers, two of which targeted specific types of information—information for marginalized people, and the organization of health information. The third paper went in a somewhat different direction, addressing conceptual change in indexing languages. The title of Randall Kemp's paper—"Classifying marginalized people, focusing on natural disaster survivors" clearly describes his subject. The author has identified a class of people who are themselves underrepresented in support by society and government and as a consequence classification schemes tend to provide inadequate access to information about them. The paper explores the steps that need to be taken to devise a classification for one example of such people—natural disaster survivors. As the basis for his investigation, he posed three questions: "What classification practices, if any, marginalize people? What unintended consequences arise through well-intentioned knowledge organization practices? Within the international development arena, what classifications hinder recovery and recon-

72

Knowl. Org. 36(2009)No.1
N. J. Williamson. Classification Issues in 2007

struction of lives and communities devastated by natural disaster and internal conflicts?" At the beginning, he notes that it is likely that any particular classification can reflect bias for or against a given situation. To aid in framing the research he used two important values—the value placed on justice in the human situation and the emphasis on international issues. Using examples from such sources as the *Library of Congress Subjects Headings* (LCSH), the work of Hope Olson on women's studies as found in DDC, and the *International Classification of Diseases* (ICD) the author sets the stage for the remainder of his paper when he examines three types of situations which might possibly exist: marginalization of people through classification; the neutrality of classification schemes towards people; and classification of tuberculosis patients. With this background he proceeded to look at one specific kind of environmental refugee—the natural disaster survivor. He sets out four principles as requirements to guide the development of such a scheme. In a summary statement he states that "the challenge is to construct a deep (principle 1), humane (principle 2) and politically meaningful classification scheme so appropriate polices and affective immediate-need decisions emerge from a helpful scheme." Also, he mentions two systems that are in the early stages of construction—W3C and a system being worked on by OASIS. The author is passionate about the subject matter and hopes is that his discussion "contributes to the information organization practices of relief agencies, such that the agencies are better aware" and better able to deal with humanity and situations which they face. The third paper in the group, "Tagging for health information organization and retrieval" by Margaret Kipp, like the Kemp paper, dealt with a specific information domain. It examined the tagging practices evident on CiteULike, a research oriented social bookmarking site for journal articles. The articles chosen for use in the study were health information and were medically related. CiteULike is described; related studies are identified; and the methodology of the study is detailed. The text is well supported with tables and figures. Metadata was collected for a total of 1280 articles from 3 journals. There were 314 unique users and 1449 unique tags were used. A few tags were extremely popular and "represented key concepts in biology, such as proteins, evolution, and DNA or RNA." Descriptors were more heavily assigned to articles than tags and there were 2746 unique descriptors. The relationship between collaborative tagging by users of the CiteULike bookmaking system and controlled

vocabulary assigned by intermediary indexers was compared. In the second paper in the group, "The economic and aesthetic axis of information organization frameworks, " Joseph Tennis examined the how and why of decision making in indexing. Not surprisingly, two factors shape the outcome—economics and aesthetics. His extended abstract indicates that he explored a "diversity of information organization frameworks" looking at these two concerns. His investigation into the economic-aesthetic axis follows from recent developments in knowledge organization research "which is moving from prescriptive (how to design systems) to a descriptive (what systems are being built how and why) approach." The purpose was to gain great familiarity with the professional concerns of knowledge organization, expanding the scope of inquiry into knowledge organization practices and developing an understanding of the human urge to name and organize.

The two papers in Session 4 focused broadly on problems of information structures and retrieval. In "Faceted navigation and browsing features in new OPACS:" Katherine LeBarre proposed "a more robust solution to problems of information seekers?" She opens with an overview of events that have taken place over the last five years with respect to the changing scene of information handling, particularly in the bibliographic services field. Events cited include the proposal for a "World Digital Library," the rethinking of how we provide bibliographic services, the adoption of new cataloguing practices, improvements to digital access, and the whole future of the library catalogue and its possible integration with other library tools. . All of this has raised such questions as "Do we need to provide detailed cataloging information for digitized materials? Or can we think of Google as the catalog? Le Barre states "It is critical that any discussions and recommended solutions maintain a holistic view of the principles and objectives of the catalog." This paper builds on the author's previous research that studies faceted browsing and navigation in websites using wireframe analysis and will go on to further work on information seeking and seekers. In the second paper of this session Èlaine Ménard presented a "Study on the influence of vocabularies used for image indexing in a multilingual retrieval environment." Ménard cites a number of problems in accessing these materials from the web environment, In the light of these difficulties, her paper describes a research project to verify the existence of relations between two indexing approaches: traditional image indexing using controlled vocabularies and free image indexing using natural

Knowl. Org. 36(2009)No.1
N. J. Williamson. Classification Issues in 2007

73

language. The research compares image retrieval within two contexts—monolingual and multilingual and targets both known-item and category searches. Three data collection methods were used: analysis of the vocabularies used for image indexing; simulation of the retrieval process with a subset of images indexed according to each indexing approach taken; and the administration of a questionnaire on searcher satisfaction. The research is designed: 1) to identify the essential characteristics of digital image indexing required for retrieval; 2) to prescribe a suitable methodology for image indexing; and 3) to reveal the essential elements of a tailored process for collections of digital images. The overall goal is the optimization of methods used for image processing.

In the final session there were 3 presentations. "In the margins" by Jane Abbas provided "Reflections on scribbles, knowledge organization and access." She describes scribbles in the margin as "notes to self, keywords, subject headings? tags?—data to remind us as to the relevance of a passage?" Does this change the meaning for the original reader? For the next user? What is said or not said can change the data. Knowledge is not static and the systems continually change. In the past we have used formal methods of organization—subject headings, classification systems, etc. The author notes that this has not solved all the access problems. Moreover times are changing. "On the surface, it seems the Web has taken much of knowledge organization out of our hands." Users can search the web with a few keywords and they are not concerned with Boolean strings and the more formal tools of access. The author points out that it is now within our power to find out more about the "scribbles" of users through sources such users' search terms gathered from OPACS, and users' tagging. Answers could be found to such questions as: 'What does tagging mean to users? Is it a way to describe a text, a scribble in the margins, or a search term? Are these potential uses different to users? What are users' motivations for tagging (personal findability or organization; communal or familial sharing; meaning making)?" Indeed, what can we learn that will inform knowledge organization? And how can we incorporate our findings into our formal systems? These are some of the questions asked by the author. Finally she concludes that what we need to do is to consider how we can use these sources "to adapt, augment, revitalize our knowledge organization structures." There is plenty of data available to do so. Richard Smiraglia presented a paper entitled "Performance works: continuing to comprehend instantiation." At the outset he points out that re-

search in knowledge organization has been primarily conceptual but empirical methods have also been used. In this paper the author's intention is to "demonstrate the efficacy of the empirical model for category generation by taking one category of instantiation—the performance work—and submitting it to analytical scrutiny." He begins by explaining what constitutes a "performance" As set out in the *Oxford English Dictionary* performance is always active—an action taking place in space and in time It can be seen in 2 senses—as the action itself and as the instance of an action. As such a performance work is "a work created for the purpose of generating a performance." As a "work" its can be more than a single text.. It may embody a complex grouping of documents by and about the text in a variety of physical formats such as the materials about the performance.—the expressions, manifestations, and related and adjunct materials and recorded performances. For purposes of the analysis he compares the data from two other studies supported by tables. The result is a basic naïve model which could be used to move forward in a more critical empirical analysis of the phenomenon of performance works. The final paper presented by Rebecca Green and Nancy Falligen was entitled "Anticipating new media: a faceted classification of material types." It addresses the problems of, and challenges to, the existing methods and schemes of knowledge organization to accommodate new media types now emerging and proposes a faceted classification to solve the problems. The paper begins with an analysis of the current situation as to bibliographic description as defined in AACR2 an its successor CDA. Specifically the concern relates to the "differentiation between the content and its carrier." Recent views suggest that rather than casting the content and carrier as distinct and orthogonal entity types, they should be treated as interdependent. It is this change of views that this paper addresses. A literature review was cared out on content and carrier issues and the significance of the findings discussed. Based on the results it was felt that "faceted classification of material types would allow for and encompass emerging and future media types without needing major rule revision to accommodate them, thus allowing bibliographic control to keep pace with changing technologies more efficiently." A facet analysis was undertaken in which seven facets were recognized: content, generations of content, recording of content, publication/distribution, physical characteristics, perception/use and relationships. The result was applied to two new material types –wikis and blogs. The final analysis "has shown … that content and carrier are inseparably

74

Knowl. Org. 36(2009)No.1
N. J. Williamson. Classification Issues in 2007

connected." The authors have concluded that this method of approach will better accommodate new media that emerge in the future.

In all of these papers, the authors have recognized the importance of the past and have endeavoured to investigate and develop innovative methods for projecting knowledge organization into the future. In doing so, they hope to modify, develop, and replace current methods while avoiding the dangers of throwing out the baby with the bathwater.