Classification Issues

Paradigms and Conceptual Systems in Knowledge Organization, the Eleventh International ISKO Conference, Rome, 2010

Nancy J. Williamson

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

The eleventh International ISKO Conference on "Paradigms and Conceptual Systems in Knowledge Organization" was held in Rome, February 23-26, 2010. The proceedings were edited by Claudio Gnoli and Fulvio Mazzocchi and published by Ergon Verlag in 2010. This analysis follows the order of the text of the proceedings, an order prescribed from the abridged scheme for KO literature published in *Knowledge Organization,* 25, 1998, no. 4, p. 226. Some invited papers, marked with [LR], have been included and are labelled as such in the table of contents. In all, 64 papers were published.

The keynote address, "Organizing and disseminating knowledge: theoretical and instrumental innovations of Paul Otlet" was presented by Boyd Rayward (United States and Australia). In his presentation he recognizes the present day efforts to understand the relationship between knowledge organization and knowledge management. In doing so, he referred back to Paul Otlet, one of the founders of knowledge organization and "traced the developments of Otlet's ideas from bibliography through his enlarged concept of document and documentation to his attempts to envisage new organizational forms for knowledge management." There is a brief abstract of the paper in the proceedings.

The first group of two papers deals with *Order and Knowledge Organization.* Thomas Dousa (United States) deals with "The simple and the complex in E.C. Richardson's theory of classification." It focuses on an early KO model of the relationship between ontology and epistemology. He discusses the topic in general and turns to Richardson's work. There he discusses Richardson's classification as an ontological order from the simple to the complex and then as an epistemological order from the complex to the simple. He concludes with the nature and limits of the application of Richardson's epistemology. In the end he states that what may have been suitable in Richardson's day would fail "to do justice to the variety of perspectives that one encounters in the multicultural world of today." In the second paper Hope Olson (United States) tackled "Hegel's epistemograph, classification, and Spivak's postcolonial reason." In this context, she describes teleology as "a major characteristic of classification and uses Hegel's three stage of development (being, essence and idea) as set out in his *Science of Logic,* as an example of codified progression that Gayatri Spivak refers to as an epistemograph "a graduated diagram of the coming to being of knowledge" and applies it to bibliographic classification. She then looks at order in terms of DDC and UDC. Hegel's three stages are discussed followed by Spivak's approach leading into a discussion of globalization as exemplified in DDC and UDC. Olson concludes that Hegel "provides a rationale for the sequence of main classes … but it is now more of a convention than a teleological progression …. Nevertheless, bibliographic classification, at least DDC and UDC, reflects and reinforces the mainstream epistemograph even though its meaning and significance are obscured."

Four papers were presented on the topic of *Conceptology in Knowledge Organization.* Alfred Gerstenkorn (Germany) discussed "Entities and quiddities" He is concerned the binarity of concepts and their connection with two kinds of cenceptualization—the ontological and epistemological points of view. He introduces the topic of binarity, and compares some concept models. Finally, he proposes a chosen concept model. In his conclusion he states that "the binary conceptualization seems to be an operable transdisciplinary approach which facilitates the communication of KO, especially concerning paradigms and conceptual systems." Further application needs to be done before its usefulness can be proven. Birger Hjørland (Denmark) addressed "Concepts, paradigms and knowledge organization." He begins with the generally accepted agreement that concepts should be the building blocks of knowledge organization systems. This view is held by Dahlberg and many others but some researchers disagree. Hjørland sets out his understanding of concept theory with its four ideals—empiricism, rationalism, historicism and pragmatism. With this basis he proceeds with a discussion of the "criticism of concepts as units." In particular he examines the suggestion that "ontologies should not be based on concepts, but rather on universals and particulars which exist in reality and are captured in scientific laws" and on the idea that a concept is a "sign." The nature of the topic is very complex. In his conclusion Hjørland states that "it would seem very problematic not to inform users about different opinions at play. To accept 'concepts' as units in KO by implication means to accept the theory-laden nature of KO and recognize that specific KOSs are supporting specific views about the knowledge being organized." Agnes Hajdu Barat (Hungary) looked at conceptology "From paradigms of cognition and perception to phenomenon." Her paper explores the possibities of perception and cognition in the field of knowledge organization from an epistemological point of view. She is attempting to reveal some examples of new elements in the theory and practice of knowledge organization and to emphasize the necessary connection to human perception, phenomena and content dimensions. In doing so, she studies the epistemological questions, summarizes the knowledge from different sciences according to perception, phenomena and influences and makes a case for cognitivism in knowledge organization systems." She concludes that "there needs to be comprehensive rethinking of knowledge and information along new and completely different lines." In the final paper in this group, Charles van den Heuvel (Netherlands) and Richard Smiraglia (United States) discussed "Concepts as particles: metaphors for the universe of knowledge." They use the metaphor of the particle to accumulate the components of a theory of knowledge that underlies the science of knowledge organization. They outline the concepts of the universe, identify the central role of the concepts and the intertwining roles of works, instantiations and documents. This research takes a different approach in that it demonstrates a semantics that is based on structure and related forces between components rather than on content. It permits the development of mechanisms for linking related entities with so far undiscovered similarities. "The paper is an attempt to outline the central role of concepts in the knowledge universe, and the intertwining roles of works, instantiations and documents." They begin with an analysis of Paul Otlet's views of the universe of knowledge through his work in the development of UDC and follow with some later efforts on the subject. Finally, they describe their own explorations. Eventually, they hope "to demonstrate a semantics that is based on the structure of knowledge rather than on the content of documents."

The section dealing with concepts in general is followed by three sets of papers dealing with concepts in particular subject disciplines. On *Mathematics in Knowledge Organization* there was one paper by Ágota Fóris (Hungary) who examined "Change of paradigm in terminology: new method in knowledge organization." The goal of the author is "to describe the application in terminology as a scale-free network model … used in the fields of natural sciences and information technology." In the research, the author used network theory as the basis, examined and interpreted the historic development of knowledge storage and knowledge transfer. The general laws of network theory are discussed, as are the properties of a scale-free network. This leads into a description of language networks and the model of the terminological network created in the research. The author has produced a model in which "each of the three aspects of terminological approach, the cognitive, linguistic and the communication components each form a scale-free network and combining these three networks … it is possible to model the process of communication." Two papers address *Psychology and Knowledge Organization.* A paper by José Antonio Frais Montoya (Spain) entitled "Postmodernism, constructivism and knowledge organization: applications of repertory grid to knowledge construction and representation" was

not available at the time of printing. The second paper under this topic was "Perception, knowledge organization and noetic affective social tagging" by Richard Smiraglia (United States). The author points out that "studying perception and its role in the identification of concepts is critical for the advancement of KO." The purpose of the research is to advance our understanding of the role of perception in knowledge organization systems. Our present understanding of perception in KO is explained, as is its place in social tagging. Noesis, which is rooted in the *ego* is defined and noetic tagging is presented as a methodology.

At the broader disciplinary level *Science and Knowledge Organization,* four papers were presented. Rick Szostak (Canada) examined "Universal and domain specific classifications from an interdisciplinary perspective." His starting point is an exchange of views on the nature of classification between the author and Birger Hjorland which was published in the *Journal of Documentation* in 2008. Specifically there were two views—1) the view that urges the development of a superior universal classification, and 2) the view that concepts are ambiguous and it is best to classify documents only broadly in domains. These opposing views have long been known. It is possible that these views could be complementary rather than substitutes. This paper examines the possibility of their complementarity. The author first sets out the key arguments regarding both feasibility and desirability of pursuing the two types classification. The second section looks at the strengths and weakness of the two types. Then the author reviews theoretical and practical reason for the two types to be complementary. Finally, he discusses the ways and means the two methods might be set up to be used in a complementary fashion. He closes by list of five questions that might be used in setting up such a system.

In the second paper, Thomas Dousa (United States) asks "Whither pragmatism in knowledge organization?" Specifically the author is concerned with "classical pragmatism (cp) neopragmatism (np) as KO metatheories." Up to now, the two theories have appeared virtually indistinguishable in their philosophical differences. His research appears to have found that this is still virtually true. In the third paper Joliza Chagas Fernandes and Nair Yumiko Kobashi (Brazil) considered "The complexity challenge: a contribution to the epistemological reflection regarding information science." For purposes of a better understanding, this paper studies or reflects on the epistemology of information science. The authors first define epistemology and science, then identify the elements of the reflection and report on the expected results. In the final presentation of this section Maria López-Huertas and Maria José Lópes-Pérez (Spain) examined the "Epistemological dynamics in scientific domains and their influence in knowledge organization." They begin with the premise that socio-cultural context can affect the theoretical and epistemological development of a scientific domain. In doing so, it may affect not only a robust, consistent theoretical framework but also good practice. Knowledge organization systems should be concerned with this and avoid creating epistemological biases. This paper takes two domains—psychiatry and information science to demonstrate this situation. In their conclusion the authors state that "LIS and KO should look for research methodologies capable of producing integrated knowledge … KO and KOS designers should explore in a reflexive way the impact of external circumstances on the knowledge domain that will be represented and organized, trying to identify the bias that such a domain might have."

Three papers discussed *Problems in Knowledge Organization.* Sergey Zherebchevsky (United States) addresses "Formalism in knowledge organization" using a thematic analysis of the ISKO proceedings. The purpose of the study was "to improve theoretical comprehension of the domain of KO by investigating the presence of formalism and the amount of attention it receives in the field of KO as compared with the attention and efforts invested in addressing theoretical concerns." Two sets of six terms taken from an article by Elaine Svenonius in 2000 (citation not given) were used in the experiment. One set of terms were identified as Intellectual/Theoretical terms and the other as Methodological/Formalistic terms. ISKO 10 proceeding were then used to determine the number of terms of each type were used. More of the terms (61.56%) of the intellectual/theoretical type were retrieved than the methodological/formalistic type (34.44%). The finer details of the results are discussed in full. The author suggests a similar test should be tried on other KO data and comparison could be made on data in other disiciplines, e.g. cognitive psychology. Luciano de Souza Gracioso (Brazil) presented a paper on the "Pragmatic approach to virtual information action from Wittgenstein." The author is concerned with the fact that "information science is concerned with human actions" while "today, a large part of information action is configured in the virtual technological plane." He begins with a description of the repositioning of the subject and the use of language in a virtual field, followed by a discus-

sion of the philosophy of language, and the proposals of Wittgenstein. While the investigation did not lead to the use of the proposal as an application in information science, it has provided direction for further research. In the third and final, paper in this category Fidelia Ibekwe-SanJuan and Erik SanJuan (France) examined the output of "Knowledge organization research in last two decades: 1988-2008." The authors apply an automatic system to records of publication in knowledge organization over two decades. The data came from journal articles in the KO field available from the Web of Science database (WoS). Motivation for the study and data collection are described. Methodology for term extraction and term clustering is given. Previous research on this topic was examined, The trends identified in the study were located automatically without human effort and the authors plan further more intense research. Cluster maps are included with the article.

The topic *Knowledge Organization Systems: General Questions* includes seven papers focusing on a number of subtopics, including: faceting, categories, ontologies, integration and genres. Renato Sousa (Brazil), Douglas Tudhope (Wales) and Mauricio Barcellos Almeida (Brazil) gave a presentation entitled "The KOS spectra: a tentative typology of knowledge organization systems." The purpose of the paper is "to discuss why and how the KOS's should be tentatively classified on a new basis, aiming to shed some light to the discussion." Representation and knowledge organization systems are discussed, a taxonomy is drawn up and illustrated in a diagram of "KOS's by type." The question, what to evaluate? is addressed and the evaluation dimensions or spectra is illustrated. As a result of the research, the authors found that "we are far from having a consensus on KOS taxonomies and the related terminology." This paper is seen as a step in the evolving discussion, "presenting some of the most important aspects to be taken into account when evaluation and choosing a specific KOS." Further research will publish a full model as a high level ontology. The next four papers by Gnoli, Scognamiglio, Poli and Kameas were invited and originally given at a workshop on levels of reality as a KO paradigm The paper presented by Claudio Gnoli (Italy) was entitled "Levels, types, facets: three structural principles for KO." The three major principles involved in the structure of knowledge organization are identified and discussed. Then they are considered as to how "they interact with each other, working as substructures that determine the macro-structure of a KOS." In what order should they be combined? (e.g. types, then levels? Facets, then type; Levels, then facets; etc.). Six possible orders for interaction are examined. It was found that types is the most classical principle while facets, studied and applied later, can be seen as another classical principle. Levels were acknowledged even later, by studies of the CRG. Levels interact as an implicit principle and their interaction with the other principles "is believed to be a useful contribution to the development of a more complete theory of KO." Carlo Scognamiglio (Italy) prepared a paper entitled "Strata and top categories for an ontologically oriented classification." It addressed communication between ontological research and knowledge organization. This paper has as its theme "the something" (the matter) of classification. It begins with discussion of the debate on the alternative between the classification of documents and the classification of entities. The author shows how every kind of rigorous classification must be supported by an ontology. An example of the ontological approach is a theory of levels of reality. In turn, the author states that a theory of levels needs a set of general categories common for all levels. In this paper the exploration of top categories uses a combination of the critical ontology elaborated by Nicolai Hartmann and some notions from General Systems Theory. The paper by Robert Poli offered "Domain theory: a preliminary proposal." The author states that domain theory can be built both on a theory of levels of reality and the theory of wholes. To test this approach he uses biology as a test domain. The domain is analyzed and core entities and facets are identified. Four types of domains were located. Two of the types (domain in general and sub-domain or facet frames) are analyzed. The other two types will be described in later papers. In conclusion, he advances the hypothesis "that the structure of the top levels is different for each of the four types of domain." This reinforces the importance of distinguishing the various domain types. A fourth invited paper, entitled "Ontologies in adaptive systems supporting every human activity" by Achilles Kameas (Greece) was not available at the time of printing. The last paper under this topic is "An integrative approach to the design of knowledge organization" systems presented by Melanie Feinberg (United States). "It presents a design process that negotiates between the communicative goals of an author, the information needs of an audience, and the structure of existing subject literature." Her design process is a theoretical model of the six primary activities that take place in the development of a research project as follows: 1) Envisioning: persona and scenario development or the identifica-

tion of the target audience for the product of the research being involved in research; 2) Strategizing: making a plan to achieve the nacent vision, or a scenario for identifying the research problem presented in a brief; 3) A learning process in which the designer surveys the subject literature and compiles a source book of concepts to use as raw material to be used in the strategy and in which changes in strategy may take place; 4) Sketching, a process in which categories are sketched out, definitions developed and potential hierarchical and associative relationships created; 5) Revisiting, refining and reflecting until the interconnection of the documents which represent the user experience occurs; and, 6) Finally, analysis and critique in preparation for further implementation. She concludes with a diagram showing the interrelationships of the six steps showing how " the design process is neither linear nor circular … while all activities are independent they may be occurring simultaneously, and knowledge gained in one activity may necessitate revisiting. In the final paper of this section Amelia Abreu (United States) presented a papers entitled "Medium cool: Genres, attitudes and affect." "Cool" is used as a metaphor to examine the role that affect and social context play in the design and use of knowledge organization systems. The author provides a theoretical basis for her study, considers how knowledge organization might produce affect and how it may be used for further study. Finally, she considers the value "cool" may have as a commodity in the larger economy of information.

Another popular area of interest was *Knowledge Organization Systems Structure and Elements*, *Facet Analysis.* Five papers were presented in this category. Uta Priss and John Old (Scotland) gave a paper on "Concept neighbourhoods in knowledge organization systems." They began by considering previous research on the topic. A brief overview describes the FCA technologies for concept neighbourhoods and neighbourhood lattices and its application using the electronic version of *Roget's Thesaurus* It was logical that the application might also be used with other lexical databases such as WordNet and Wikipedia. The paper describes further research using the application and the word "sleep" from Roget to create and illustrate its use with all three databases. "An on-line interface for exploring *Roget's Thesaurus* in this manner is available at www.roget.org." A similar interface for WordNet will be available at the same site in the near future. An interface for Wikipedia is more difficult because of its size. Rebecca Green and Michael Panzer (United States) discussed "The ontological character of classes in the *Dewey Decimal Classification*." The ontological relationships among topics and classes in *DDC* are examined through a case study. The paper begins with a description of the nature of *DDC* classes as neighborhoods. This identifies implications for the ontological representation of the system and for operations on that representation. The association of topics with classes is examined and the development of neighborhoods is explored. The description of a neighborhood was achieved by expressing class-topic relationships as OWL class axioms. In concluding the experiment, the authors state that there is conflict with the possibility of a *DDC* ontology but they believe that the construction of an ontological model that *reuses* a certain level of knowledge in the *DDC* is feasible. In a paper entitled "Semantic interoperability and retrieval paradigms," Felix Boteram, Winfried Gödert and Jessica Hubrich (Germany) "present a new approach to understanding how indexing strategies, models for interoperability and retrieval paradigms interact" and how it can be used to support semantic navigation in information retrieval systems. A comprehensive interoperability model is presented. Clarification of characteristics and structural qualities are required to implement semantic interoperability on various levels. Some of these levels are introduced and discussed. The concept "theatre" is illustrated and discussed. In the fourth paper in this group entitled "Finding Bliss on the Web" Vanda Broughton (United Kingdom) used the *Bliss Bibliographic Classification*, second edition (BC2) to address "some problems of representing faceted terminologies in digital environments." Bliss, the only example of a fully faceted general classification is maintained and managed in electronic format but the format is not suitable for use in a public interface. What are the problems of achieving such a format? Broughton describes the coding used in BC2, some work done in converting the schedules to a thesaurus format and problems found in vocabulary control. Some degree of manual editing is required and the system cannot yet deal with many of the associative terms. Some of the coding added manually should be incorporated into the mark-up language. The role of encoding is described and various encoding systems such as EXML (Extensible faceted mark-up language), SKOS (Simple Knowledge Organization System) and TEL (Text encoding initative) are considered for representation "of the concepts, their functional roles and the relationships between them. "Integrating these aspects in a coherent and interchangeable manner appears to be achievable but the most

appropriate system is as yet unclear" In the final paper in the session, Kathryn La Barre (United States) in a paper entitled "Facets, search and discovery in next generation catalogs: informing the future by revisiting past understanding" describes a project entitled "Folktales, Facets and FRBR" funded by a grant from OCLC/ALISE. The paper begins with a description of the North American encounters with Ranganathan and faceted classification theory through the years 1950 to 1969. This is followed by the contemporary understanding of facets in North American applications and examines faceted search and navigation in six next generation catalogues which provide such interfaces as guided discovery, mobile browsing and social networking—search refinements which have been called facets. A sample of 200 catalogues was selected and the search term 'folktales' was used. A variety of types of facets were located (e.g. subject/ topic, format, location, call number, genre, etc.). Faculty members were interviewed about their searches and six categories of information tasks identified including exploring, creating, synthesizing, studying, collecting, and searching. Additional facets were found in the interviews. Some, but not all, of the facets were found in existing bibliographic records. Greater understanding of user information tasks is needed. Facet analysis should precede system design. "This paper highlights the ways in which the heritage of facets and facet analysis may continue to inform research and development and spark a dialogue between system designers and facet theorists, thereby enhancing future access and discovery systems."

In the session on *Knowledge Organization Systems Construction* three papers were presented. Two of papers deal with specialized subject areas – gold and neurosurgery. Elena Cardillo, et al. (Italy) presented a paper entitled "GoldThes a faceted thesaurus for goldsmith handcraftsmanship in a regional context." The authors present the construction of a very specialized thesaurus for a very precise subject area. Their intention is to demonstrate how classifying and organizing information into multi-dimensional hierarchies makes it more accessible than using a single taxonomy, that is a unique hierarchical dimension." The paper first sets up the background, citing *AAT* (*Art & Architecture thesaurus*) as an example of the type of thesaurus required for their own particular approach. The project is carried out in two macrophases—the creation of a knowledge base and the construction of the thesaurus. Construction follows ISO 2788 (1986) and ANSI NISO Z39.19 (2005). The facets were chosen from those defined by the Classification Research Group The second paper by K.S Raghavan and Chathoth Sajana (India) was entitled "NeurOn: modeling ontology for neurosurgery." The authors report on the initial results of an ongoing experiment in building an ontology using concepts extracted from the patient records in a large hospital. The process of building the ontology is described, including the nature of the domain concepts and the creation of a small query library to be used in defining the classes as well as properties to be included in the ontology. The classes and subclasses are described, as is the terminology. The authors hope that the final product will provide a usable decision support system for health care personnel. A paper on "Development of thesaurus structure through a work-task oriented methodology" was presented by Asam Sanatjoo (Iran). It describes an empirical study which investigated a mixed set of methods and developed a prototype thesaurus to evaluate the potential of a work-task oriented methodology (WOM) for constructing a more enriched thesaurus. The result was evaluated for usability and performance against a conventional thesaurus (specifically Agrovoc). The methodology and construction of the thesaurus is outlined. The research design and findings are described. From the research result it was concluded "that an enriched work-task oriented thesaurus inspires information searchers by offering enhanced conceptual content in contrast to the classic thesaurus in the traditional format." However the author states that "the method cannot be a stand-alone one and must be combined with other construction methods. The challenge is to combine the methods in harmony with the context and purpose of the thesaurus in such a way that the advantages of each method are exploited optimally."

Under the topic *Knowledge Organization Systems Maintenance, Updating and Storage* one paper was presented. Joseph Tennis (United States) spoke on "Measured time: imposing a temporal metric to classificatory structures." For purposes of understanding and evaluating classificatory structures, he divides time into three units: long time (versions and states of classification schemes); short time (the act of indexing as repeated ritual or form) and micro-time (where stages of the interpretation process of indexing are separated out and inventoried). As professional practice, the act of classification has inherited the assembly-line work ethic of early twentieth century scientific management. This suggests assignment of a work to an appropriate location, and assigning only once. That is, in time an act of permanence. However as

time goes by new subjects appear and new documents are added, some of which fit or do not fit and require a new location. Thus the scheme has to be revised and permanence is called into question. Given the state of impermanence. "This paper takes as its purpose the identification and characterization of impermanence (temporal metric) of classmarks in schemes." Each of the units is discussed. The author concludes that "the assembly line approach is not art because it removes the authorial or artist's presence. Yet it is still a major component of the institutions of long term classification schemes."

Under the rubric of *Compatibility, Concordance, Interoperability between Indexing Languages* two papers were presented. Barbara Kwasnik and Mary Grace Flaherty (United States) discussed "Harmonizing professional and non-professional classifications for enhanced knowledge representation." They compared two separate but related classification schemes in the area of medical information in order to understand how they might be used together and support each other. Used in the experiment were *Medical Subject Headings (MESH)* and a "naïve" scheme used by the consumer health website WebMD.com. The term 'autism' was used to compare the strengths and weakness from the perspective of vocabulary, syntax, classificatory structure, context and warrant. The paper describes previous work, the method used and the results. In the conclusion, the authors recognize the many differences between the two and make some suggestions as to how they might be used to support each other, including, the use of *MESH* "to informally update and keep things current, harvesting current usage and concepts from the site and giving them temporary or 'pending' status to bridge the gap between the scientific and lay perspective." They have provided some first steps in developing some guidelines for mapping of disparate classification. A presentation by Jan-Helge Jacobs, Tina Mengel and Katrin Müller (Germany) addressed "Benefits of the CrissCross project for conceptual interoperability and retrieval." The paper discusses the goals, methods and benefits of the conceptual mapping approach of this particular project in which the topical headings of the German subject headings' authority file (SWD) are being mapped to notations from the *Dewey Decimal Classification*. The purpose of the project is to create crosswalks between the two systems to improve retrieval processes, at the same time ensuring the continuous use of already existing indexing data. The two systems are briefly described and the methodology set out. To date, three applications of the system have been worked out: enhancing access to *DDC* class and *DDC*-indexed documents, structuring document sets, and conceptual exploration. These are described in the paper and examples given.

In the category *Theory of Classing and Indexing,* two papers were given but only one was available at the time of printing. Carlos Alberto Corrêa and Nair Yumiko Kobashi (Brazil) described "A hybrid model of automatic indexing based on paraconsistent logic." Methods of automatic indexing are based on different theoretical assumptions. This paper aims to argue "the theoretical potential for the use of hybrid models of automatic indexing, specifically the "paraconsistent logic, a non-classical logic with capacity to handle situations that involve uncertainty, imprecision and vagueness." The type of system is described.

A number of authors made reference to special subject taxonomies in geography, pathology and practical medicine, and psychology, in their papers focused on other topics. One author focused specifically on *Taxonomies in Communications Engineering.* Michiko Tanaka (United States) presented a paper entitled "Domain analysis of computational science: fifty years of a scientific computing group." Bibliometric and historical methods were used to study the domain of the Scientific Computing Group at the Brookhaven National Laboratory over a period of fifty years from 1958 to 2007. The methodology and data analysis is described and statistics included. In his results the author noted "the growing emergence of interdisciplinarity" and "identified a strong and consistent mathematics and physics orientation" within the group.

Similarly, there were a number of presentations that briefly referred to knowledge organization systems from special subject areas (e.g. biology, agriculture and horticulture, food science and technology, sociology, social aid and social politics, and general economics) while their papers focused mainly on broader topics. Three papers focused specifically on *Special Knowledge Organization Systems in Literature*. Two were printed in the proceedings. Pauline Rafferty (Wales) described "Genre theory, knowledge organisation and fiction." The author is concerned with "the epistemological assumptions underpinning fiction categorisation, explores current genre theory and argues for an approach to the understanding of genre, and ultimately the description of genre, that is based on a cultural-materialist, historical world-view." There are sections on access to fiction in historical terms; also on genre theory and mapping generic history. Finally she proposes a fiction retrieval tool through a user-based website. Francisco-Javier Gracia-Marco (Spain), Joâo-

Batista-Ernesto de Moraes (Brazil), et. al. described "Knowledge organization of fiction and narrative documents." They were dealing with the "challenge in the age of multimedia." The paper examines the key facets for knowledge organization in the field of fiction building on literature theory and faceted classification theory. The research focuses on the integration of two fields—library and information science and research from the researchers in literature theory. Subtopics addressed include the nature of narrative and fictional narrative, specificity in fiction and the theory of literature and subject indexing, fictional documents, levels of complexity and intertextuality and the problem of canonical order. A model is proposed, recognizing that there is no single classificatory approach to fiction. In their conclusion, the authors recognize that much work still needs to be done in clarifying the big levels of analysis proposed in the paper. Also there was one paper on *Special Knowledge Organization System in Cultural Sciences.* Carol Tilley and Kathryn La Barre (United States) proposed "New models from old tools: leveraging an understanding of information tasks and subject domain to support enhanced discovery and access to folktales." Their paper provides an introduction to an ongoing research project, the purpose of which is to provide users with a method of enhancing the effectiveness and efficiency in discovering and accessing folktales. In general, the research combines task analysis with facet analysis and plans are to develop an enhanced bibliographic record type. This paper describes the first phase of the project, specifically the information tasks to be used, the information seeking obstacles and the desired features of the project. It also includes some of the bibliographic, cultural, and intellectual facets derived from a sample of folktale resources. Finally, it proposes a model for enhanced bibliographic records. The methods and findings are described and the model is illustrated.

The very broad category entitled *General Problems of Natural Language, Derived Indexing, Tagging* was popular among participants. Five papers were given in this category. Marianne Lykke ((Denmark), Susan Price and Lois Delcambre (United States) explained "Using semantic components to represent and search domain-specific documents: an evaluation of indexing accuracy and consistency." The authors developed a semantic component model to supplement the existing representations of documents. Then they conducted a comparative indexing study using a national health portal to assess the feasibility of semantic component indexing. Findings suggest that accuracy and consistency might be higher for semantic component indexing (SC) than conventional indexing. Additional study is needed. Future analysis will evaluate the nature and number of indexing facets. Jungran Park, et al. (United States) described "Locally added homegrown metadata semantics: issues and implications. The authors used data from a nationwide study carried out by cataloguing and metadata professionals to assess the current state of metadata elements used in digital depositories. The homegrown elements included local notes and description, local personal and place names and local subjects, as well as administrative, technical and preservation data that had been added locally to records. The additions are seen as examples of perceived needs of local users. Currently there is a lack of a common data model of this kind of data used in records. The aim of the study was to examine records to find out the answers to three questions: What homegrown elements are added? What were the criteria for adding such data? How are local metadata practices documented and shared? An overview of previous studies is given; methodology used in the study is provided and conclusions drawn. The results indicated that "widespread use of homegrown metadata elements may present a potential challenge to the effective reuse and sharing of metadata in the networked environment." Further research is needed in which other research methods are used and more varied data sources would provide a fuller picture. Maria Aparecida Moura and Juliana Assis (Brazil) investigated "Social networks, indexing languages and organization of knowledge: a semiotic approach." They conducted a theoretical discussion on semiotic categories and their application in information organization. This was followed by an experiment on the performance of the Gemet and Eurovoc thesauri using the subject "sustainable development" and comparing with folksonomies and classification systems. The result was a proposal for a semiotic approach to design of indexing languages. In the fourth paper Pertti Vakkari (Finland) described "How specific thesauri and a general thesauri cover lay persons' vocabularies concerning health, nutrition and social services." The aim of the study was two-fold: 1) to compare the semantic structures in lay persons' questions addressed to an expert service in the areas of health, nutrition and social services; and 2) to determine to what extent lay persons' vocabularies are covered by a general thesaurus and a specific thesaurus in each of the three fields. Questions were for tests in each of the three areas. The results show that the overlap between general controlled vocabulary and a specific one was most ex-

tensive in health (32%) and least extensive in social services (9%). It seems that in all fields tested there are limited links between the general and specific vocabularies from the point of view of users. In the case of nutrition and social services, the match was low and the need for enrichment from the specific tool is very great. In the final paper in this section Isto Huvila (Sweden) researched "Aesthetic judgements in folksonomies as criteria for organising knowledge." Using Flickr photosharing service as an example the folksonomies were examined as a "potential source of collective judgements of a large group of people with a special focus on everyday life aesthetics." Visual analysis of clusters of photographs was carried out using a system of the tags.

One presentation fell in the category of *Automatic Language Processing.* Klaus Lepesky et al. (Germany) provided a paper on "Metadata improvement for image information retrieval." It discusses the goals and results of the research project Perseus-a. This project attempts to improve image retrieval by automatically connecting the images with text-based descriptions. The project uses the image collection of *Prometheus*, a distributed digital image archive for research and studies. In order to connect the works with related texts a matching process for images and texts had to be developed. Art historical terminological resources, classification data and an open source system for linguistic and statistical automatic indexing called lingo were used. It was concluded, that while the principle idea of the project was successfully demonstrated, there needs to be much more research on the underlying algorithms,

Under the category *Online Retrieval Systems and Technologies* seven papers were given. Six were available for printing. Margherita Sini, et al. (Italy) provided a paper on "Smart organization of agricultural knowledge: the example of the AGROVOC concept server and Agropedia." The authors noted the importance of the use of the computer in disseminating information in the food and agriculture field. This paper analyses projects developed by two such organizations, "aiming to make use of a concept-oriented approach, while describing agricultural topics." The two projects are described with respect to their innovative aspects, their benefits and the technology used. The authors conclude that "the work undergoing by FAO and other AOS partners for making better use of traditional thesauri is in line with the current strategies of making data more processable." Similarly, the Agropedia project opens the road to the representation of agricultural knowledge in the form of concept based maps. It is noted that there is still much to be done. For example "for the AGROVOC Concept Server investigations on the role of OWL2 and OWL rules should be carried out as well as the completion of the collaborative tool to maintain the data pool." Further work is planned. A discussion by Currado Di Benedetto, et al. (Italy) fcocused on a "Semantic approach to bioethics in the Ethicsweb project." Specifically, the authors describe "building a semantic architecture for a European documentation system." The purpose of the paper is to present the activities of the European project referred to as Ethicsweb. The project has four general objectives: 1) to facilitate access to information on ethics in science using an integrated infrastructure; 2) the development of sophisticated tools, technical and semantic, to establish the infrastructure; 3) the creation of a European Reference Center for Bioethics; and finally 4) the development of multilingual tools (thesauri and ontologies) for searching of documents in the bioethics field. The content of the paper focuses on the steps taken up to now. Maria Teresa Biagetti (Italy) discussed "Pertinence perspective and OPAC enhancement. A starting point for her paper is the ongoing "debate on OPAC enhancement and the necessity to design OPACS based on search engine features." The previous work done on OPAC enhancement is outlined; relevance/pertinence is defined; the semantic perspective addressed and an improved model using traditional semantic indexing strategies is proposed. Anna Nosek, et al. (Poland) reported on research on "Multidimensional analysis of the information structure of public libraries' websites in the Podlasie region (Poland)." It includes the results of quantitative surface research, covering the contents of library websites and a detailed analysis of three subjects: information about literature, borderline knowledge and formal website quality assessment. The introduction discusses the general nature of website use in small public libraries in Poland. The methodology and scope of the study is set out and quantitative research results are outlined. The results of the study were both revealing and disappointing. They indicated that these small libraries are most connected only within the area in which they exist. There is little connection between the main regional library and the county libraries. Some of the libraries do not have websites of their own and there are few links between library websites. Thus "an essential regional information network of libraries is practically non-existent." Many libraries see their websites as virtual bulletin boards giving practical information such the library's address, telephone number, hours, etc.

The librarians often do not see the importance and benefits of websites as an information service. There is still a lot to be done in the development of a regional information service here. Elizabeth Milonas (United States) tackled "The use of facets in Web search engines." There were four web search engines in the study—two that utilized facets or facet terms, Exalead and Excite, respectively, and two search engines that do not use facets, namely, Google and AltaVista. The two faceted systems are described. Related research studies are identified and the methodology described. Participants were library and information science master's students (LIS) and PhD information studies students (IS) and the search terms were "social networks" and "lymphoma." The analysis looked at three characteristics—ease of the search process, search time, and confusion during the search process. The results provided three significant findings: 1) Facets make the search process easier, whether searching for familiar or unfamiliar topics; 2) when using facets it takes longer when searching familiar topics, than unfamiliar topics; and 3) when searching for familiar topics, facets do not cause confusion for the searcher. Findings 1 and 3 are supported by the literature. Finding 2 is not supported by the literature. Some discrepancies were found. For example, IS students did find that facets made the search process easier but were confused when searching the term "social networks." LIS students found that facets did make the search process easier and were not confused when searching the term "social networks." Marcia Lei Zeng et al. (United States) spoke on "Expressing classification schemes with OWL 2. In doing so they explored "issues and opportunities based on experiments using OWL 2 for three classification schemes." The schemes used were the *Dewey Decimal Classification,* the *Chinese Library Classification* and the Library of Congress *Classification.* The characteristics that OWL 2 and traditional classification have in common were identified. Most important were the issues in presenting various types of classes and their relationships were discussed and included the following: centered entries, synthesis in classification schemes, class-topic relationships, alternative class location, presentation of auxiliary tables, presentation of index entries, presentation order/sequence of sibling classes, the internal structure of notes and the presentation of notation-building rules. The authors continue to explore issues for evidence of possible use of OWL 2 to resolve some classification issues is emerging.

Under the category of *Problems of Terminology* was one paper by Boyan Alexiev (Bulgaria) and Nancy Marksbury (United States) entitled "Terminology as organized knowledge." It explores the possibilities of integration between knowledge organization and terminology based on analyzing and comparing the basic theoretical methodological premises of the two disciplines with the idea of identifying threads that could be used to apply an interdisciplinary approach to a knowledge-oriented terminology. The theoretical and methodological nature of each discipline is analyzed. The commonalities were sought in three ways; semantic similarity in terminology used; similarity in underpinnings and similarity in methodological approaches. In conclusion, the authors point out that in both KO and terminology there is a tendency to move forward to a domain specific approach. "A final conclusion can be drawn that combined KO and terminology research methods would lead to strengthening the collaborative links between specialists in the two fields bringing about the development and improvement of their theoretical, methodological and practical achievements."

In another single paper topic on *Subject-Oriented Terminology Work,* Peter Ohly discussed "Interrelations and dynamics in thematic networks: how to present bibliometric outcome?" In this presentation, network analyses of term and concept co-occurences are examined to demonstrate their potential in combining both in one map. Alternative possibilities are discussed, and examples are taken from German literature. Using as the example, elderly employees a thematic network analysis was demonstrated, shaping both concept specific words as well as broader concepts.

Under the general category *General Problems of Applied Classing and Indexing, Catalogues, Guidelines* two papers were presented. Lynne Howarth (Canada) talked on the topic "Mapping the world of knowledge: cartograms and the diffusion of knowledge." One issue in providing access to knowledge is the use of non-verbal representations such as notations, symbols or icons, or rich visual displays, including topical map to facilitate access to information and warrant more attention. Here, Howarth uses Wordmapper as an example and "examines cartograms – a derivative of the data map which adds dimensionality to the geographic positioning of information." This is "one approach to representing and managing subject content and to tracking the diffusion of knowledge across place and time." The paper discusses applications of information visualization, mapping of data content and context, mapping the diffusion of knowledge and the use of cartograms to represent and manage subject content.

"In this paper, cartograms – emerge as key and opportunistic players" in finding new ways of approaching knowledge. In the second paper in this section Athena Salaba (United States) looked at "Use and users of subject authority data." Here, the author reports on the findings of two surveys of subject authority data and its use by information professionals in the semantic Web environment and in libraries and information agencies. An introduction is provided on authority data and its changing and expanded use outside dedicated information retrieval systems. The two surveys are described—one on use by semantic web professionals and the other on use by information professionals. In the conclusion, the author points out that there would be addition information provided beyond her paper in her presentation, including the implications of the findings and future directions for the research.

In a category on *Classing and Indexing of Non-Book Materials (Images, Archives, Museums)* four presentations were given. Edward Ismael Murguia (Brazil) discussed "Collecting and knowledge organization: a theoretical approach from the material culture studies. Thiago Henrique Bragato Barros and Joân Batista Ernesto de Moraes (Brazil) addressed the topic "From archives to archival science: elements for a discursive construction." They are studying known concepts of archival science. The problem addressed is the identification and analysis of the discourse produced by archival science methodology from its key functions—description, organization, classification (current and intermediate archives) and arrangement (permanent archives). Two manuals were analyzed – the *Manual of Dutch Archivists* and Hillary Jenkinson's *A manual of administration including the problems of war archives and archive making.* The method used was discourse analysis. In their conclusion the authors state that "both manuals are fundamental to the construction of archival science as a discipline … However, though they seek a theoretical approach, as all discourses and productions, their concepts and their approaches are dated by their historical and social space." In the third paper in this category Natalia Bolfarini Tognoli and José Augosto Chaves Guimarâes (Brazil) addressed the topic "Postmodern archival science and contemporary diplomatics" in a search for "new approaches for archival knowledge organization." New information technologies and new forms of document production have lead archivists to rethink the role of archival science in the so-called information age. The authors have chosen to examine two trends with different approaches that have emerged in North America and Europe. These are 1) the reformulation of the basic concepts and the functional analysis method focusing on the process and context of document creation, and 2) incorporation of all the theoretical and methodological models of classic diplomatics. The purpose of this study is "to elucidate the connection points and distinct features between the two trends concerning the organization of archival knowledge." Each method is described in detail. In conclusion, both approaches have important insights to offer in understanding the record and both should be used as interrelated tools. The final paper in category by Hemalata Iyer and Abebe Rorissa (United States) is entitled "Representative images for browsing large image collections: a cognitive perspective." This paper addressed the issue of choice of representative images within categories. A study of the free sorting of 50 images by 75 participants was conducted, in which they sorted the images into categories and selected a representative for the category. They also indicated the prominent feature of the image in the selected image. The authors also found reasonable agreement in the choice of representative images and the identification of prominent features.

In the final major category of these proceedings, *Personas and Institutions in Knowledge organization, Cultural Warrant,* there were three papers. Gloria Origgi and Judith Simon (France) wrote "On the epistemic value of reputation: the place of ratings and reputational tools in knowledge organization." The authors explored epistemological revelance and value of reputation, understood as evaluative social information. They introduced a model of rational concensus and followed with an analysis of different reputational tools on the Web. The nature of the situation is described and caution given of the dangers of using social information for epistemic purposes. In conclusion, it is stated that "a purely epistemologtical or cognitive analysis of using reputation for epistemic purposes will not suffice for KO." Nevertheless, reputational tools open up new possibilities for KO." Suellen Oliveria Milani and José Augusto Chaves Guimarâes (Brazil) examined "Bias in the indexing languages: theoretical approaches about feminine issues." They take as starting point the fact that the process of knowledge representation as well as its procedures or tools and its products are not neutral in value. Instead they imply moral values. In this context, they address the problems of bias in classification and thesauri. Starting from the reflections of earlier writers they propose a preliminary categorization aimed at facilitat-

ing the identification of bias concerning feminine issues in indexing languages. Among other things, they offer suggestions in the form of use of feminine form, insertion of notes and the use of gender qualifiers, to minimize the problems. In the final paper of the proceedings, Carel de Beer (South Africa) described "The troubadour of knowledge: a knowledge worker for the new knowledge age." The author provides a portrait of the "new age" and identifies new qualities or qualities to be reinvented.. The troubadour is described "as the instructed third, competent and very able to link the sciences (the instructed first) with the humanities (the instructed second), while taking him/herself the third position … with the ability to move from one to the other and back again – a kind of traveller or voyager, hence the troubadour." Six characteristics of a troubadour emerge.