

# **Advances in Knowledge Organization, Vol. 10 - Supplement**

## **Compatibility, Media and Ethics in Knowledge Organization**

Proceedings of the 10th Conference of the German Section of the International Society of  
Knowledge Organization

Vienna, 3-5 July 2006

Edited by

H. Peter Ohly, Sebastian Netscher & Konstantin Mitgutsch

ERGON Verlag

Würzburg 2008

### **1. Principles of Knowledge Organization**

#### **1.1 Gerhard Rahmstorf: Determining positions of Knowledge Organization**

The following paper takes position to some fundamental questions for the scientific understanding in the area of 'knowledge organization': What is knowledge? Into which subject does knowledge organization belong? For what is the knowledge organization useful? Which solutions does knowledge organization offer? How should knowledge organization be trained? These questions addresses important topics, which could only be answered with first hints.

#### **1.2 Christine Rabl: Ignorance in Knowledge Organization. Types of Ignorance and their Importance for Knowledge Organization**

This article explores the relevance of ignorance within knowledge organisation. The following forms of ignorance will be discussed: first, ignorance as 'not knowing yet', which is found in connection with problem awareness in respect to specific research questions, which implies a lack of knowledge in certain matters; second, forms of ignorance whose intentionality or lack of intentionality is influenced by political, economic and private factors; and third, ignorance defined as 'impossibility of knowing', which implies that all knowledge is uncertain. Based on this distinction, the following thesis will be discussed: ignorance as 'not knowing yet' can be understood as a means of generating new knowledge, thus underpinning the importance of knowledge organisation, whereas 'impossibility of knowing' emphasizes the relative validity of knowledge (and consequently of its organisation).

#### **1.3 Rosa San Segundo Manuel: From the invalidity of a General Classification Theory to a new Organization of Knowledge for the Millennium to come**

The idea of organizing knowledge and the determinism in classification structures implicitly involve certain limits which are translated into a General Theory on the Classification of Knowledge, given that classification responds to specific parameters and structures more than to a theoretical concept. The classification of things is a reflection of their classification by man, and this is what determines classification structures. The classification and organization of knowledge are presented to us as an artificial construct or as a useful fiction elaborated by man.

#### 1.4 Mikel Breitenstein: **Indexing Models as Social Systems**

The activity of indexing has always had as its central goal the determination of the concepts in a document and the representation of those concepts in vocabulary terms that can be used in systems of document storage and retrieval. Traditionally, careful reading and careful term choice, relating each document to a larger terminological whole, have been the special techniques of the indexer. The involvement of users in term selection (such as in democratic indexing) has changed the assumption that professionals alone could do it best, and the revolution of user-based information posting on the internet has overwhelmed old models, and created challenges for information professionals. In this paper the author reflects on some indexing approaches as social styles, and suggests some new or expanded skills for professional indexers.

#### 1.5 Fulvio Mazzocchi & Paolo Plini: **Refining thesaurus relational structure: implications and opportunities**

In this paper the possibility to develop a richer relational structure for thesauri is explored and described. The development of a new environmental thesaurus -EARTH (Environmental Applications Reference Thesaurus)- is serving as a case study for exploring the refinement of thesaurus relational structure by specialising standard relationships into different subtypes. Together with benefits and opportunities, implications and possible challenges that an expanded set of thesaurus relations may cause are evaluated.

#### 1.6 Ingetraut Dahlberg: **Compatibility and Integration: Problems and Solutions in Knowledge Organization**

The trend in the fifties and sixties of the past century away from the use of universal classification systems such as the UDC towards establishing thesauri in special subject fields for the description of the conceptual contents of documents lead documentalists soon to realize that the necessary common tool for a collaboration among centers of similar subject fields was lacking. Therefore compatibility and integration studies began between the different thesauri of such fields, leading often to more comprehensive thesauri, such as macrothesauri. The paper describes this historic development and also the solutions found at the 1995 ISKO-Conference in Warsaw/Poland on Compatibility and Integration as given in its papers, its recommendations and also in the conceptual frame of its comprehensive bibliography on this topic. In conclusion a new solution is presented oriented toward combining the use of a universal classification system with the new developments of ontologies and their problem of interoperability and heterogeneity.

## 2. Compatibility and Heterogeneity of Knowledge

### 2.1 Harm Paschen: **About Epistemic Integration and Integratedness in Heterogeneous Domains of Knowledge?**

The most important resource of a knowledge controlled society are its ideas of forms and kinds of knowledge as well as of its development. This seems especially important for long-term investments ( as for productions and education) and also – in a rapid, complex, wide and deep structured development of information – for stable interpretations of the further developments safeguarding identity and coherence. We can demonstrate the history of differentiating of domains of knowledge and an early (aesthetic) presentation of their re-integration with a painting of Rembrandt. Besides indicating further forms of integrations and their topical meanings I want foremost

demonstrate (again with a paradigmatic example) the phenomenon of a primary integration of heterogeneous domains of knowledge. For example there can be detected effective entanglements of cognitive and social domains of knowledge with their differentiations.

Such an understanding of knowledge seems to be helpful for the modern tasks of interdisciplinary, holistic, networking (often more or less still rather programmatic slogans) and for a search of physical, psychic, and mental forms of differentiated unity in multicultural societies.

2.2 Michael Panzer: **Semantic integration of heterogeneous und heterolingual Knowledge Organization Systems: CrissCross and beyond**

The usefulness of controlled vocabularies for retrieval purposes remains largely untapped in current systems. This is in part because the integration of different terminologies for unified access to resources is a challenging process due to lacking standards and competing approaches.

The DFG-funded project "CrissCross" builds links between the topical terms of the German subject headings authorities SWD, the American LCSH, the French RAMEAU and the Dewey Decimal Classification to facilitate verbal access to collections indexed by Dewey numbers. The paper describes the theoretical difficulties of integrating heterogeneous systems into a common mapping framework and discusses the semantic difficulties that arise because of shortcomings in thesaurus and classification construction. In particular, the semantic implications for mapping relationships with regard to sense, meaning and polysemy of the mapped terms are analyzed.

2.3 Maximilian Stempfhuber: **Enhancing Access to Heterogeneous Information**

With the integration of distributed information sources into information systems the semantic heterogeneity between the vocabularies used for content indexing of the individual databases becomes apparent to the user – despite the ongoing activities towards standardization in knowledge organisation. This paper presents two examples where semantic heterogeneity between thesauri and nomenclatures has been treated with cross-concordances, bi-lateral mappings between the controlled vocabularies. They can be used to help users in formulating cross-database queries or to automatically transform their queries onto the thesauri used for different databases, yielding a higher recall. Cross-concordances have been applied in the field of market research for the retrieval of statistical data and with reference databases from pedagogics, social sciences and psychology.

2.4 Anne-Kathrin Walter & Philipp Mayr: **Mapping Knowledge Organization Systems**

The integration of information systems and databases in the field of scientific information neglects the aspect of compatibility and concordance between controlled vocabularies (semantic heterogeneity). Especially the semantic relations, mappings or cross-concordances between the underlying Knowledge Organization Systems (KOS) play a crucial role for the subject access of heterogeneously indexed collections. This paper describes the adoption and examples of cross-concordances in the project "Competence Center Modeling and Treatment of Semantic Heterogeneity" (KoMoHe) and the net of the terminology mappings established. The cross-concordances established at the Information Centre can be used in the future through a terminology web-service, which will be introduced by example.

## 2.5 Francisco Javier García Marco: **Compatibility & heterogeneity in Knowledge**

### **Organization: some reflections around a case study in the field of Consumer Information**

A case study in compatibility and heterogeneity of knowledge organization (KO) systems and processes is presented. It is based in the experience of the author in the field of information for consumer protection, a good example of the emerging transdisciplinary applied social sciences. The activities and knowledge organization problems and solutions of the Aragonian Consumers' Information and Documentation Centre are described and analyzed. Six assertions can be concluded: a) heterogeneity and compatibility are certainly an inherent problem in knowledge organization and also in practical domains; b) knowledge organization is also a social task, not only a logical one; c) knowledge organization is affected by economical and efficiency considerations; d) knowledge organization is at the heart of Knowledge Management; e) identifying and maintaining the focus in interdisciplinary fields is a must; f) the different knowledge organization tools of a institution must be considered as an integrated system, pursuing a unifying model.

## 2.6 Maja Žumer & Genevieve Clavel-Merrin: **TEL-ME-MOR: investigating subject access tools and practices of European national libraries**

TEL-ME-MOR is funded by European Commission, Directorate E - Content, as a result of the third call for proposals published in 2004, under the Sixth Framework Programme (FP6). As part of Workpackage 3 the national libraries of the 10 New Member States responded to a questionnaire to enable the TEL-ME-MOR project to establish an overview of subject access tools and practice in their institutions. Results are analysed and areas for further study are identified, in particular in the field of cross language access.

## 3. Ontologies in Knowledge Organization

### 3.1 Winfried Schmitz-Esser: **Ontologien – What are they good for, and do they help us much?**

Modern ontologies, what are they good for? How did they take us in KO theory and application, and what can we expect from them in the future? These were the focal questions in an afternoon workshop initiated and run by three passionate knowledge organizers from the front of ontology research, Alexander Sigel (Cologne), Winfried Schmitz-Esser (Hamburg), and Roberto Poli (Trento), at the International ISKO 06 Congress in Vienna.

### 3.2 Winfried Schmitz-Esser: **Formalizing terminology-based knowledge for an ontology independently of a particular language**

Last word ontological thought and practice is exemplified on an axiomatic framework [a model for an Integrative Cross-Language Ontology (ICLO), cf. Poli, R., Schmitz-Esser, W., forthcoming 2007] that is highly general, based on natural language, multilingual, can be implemented as topic maps and may be openly enhanced by software available for particular languages. Basics of ontological modelling, conditions for construction and maintenance, and the most salient points in application are addressed, such as cross-language text mining and knowledge generation. The rationale is to open the eyes for the tremendous potential of terminology-based ontologies for principled Knowledge Organization and the interchange and reuse of formalized knowledge.

### 3.3 Roberto Poli: **Upper Ontologies Hold It Together**

After presenting some of the basic features of upper ontologies, the thesis is defended that all the relations needed by any concrete application can be generated by a small set of general relations, by adding proper ontological constraints to the general relations' arguments. This procedure provides an explicit and verifiable grounding to all forms of knowledge managements, including acquisition, interchange, integration, reuse, merging, aligning and updating knowledge. Upper ontologies therefore provide cues for developing both unification and decomposition methods. Finally, upper ontologies pave the ground for enhancing automatic reasoning and other machine-oriented procedures. I conclude by mentioning a difficulty in the theory of semantic fields.

## 4. Knowledge Management and Knowledge Access

### 4.1 Kerstin Zimmermann: **A Research Ontology for Telecommunications**

In this paper we discuss the main issues of an ontology creation process. We show the underlying concepts, the methodology and the player involved. The life cycle and iteration loops are described as well as the coding aspects. First we mention the goal, define then the domain and end up with the evaluation in the wider context of information management.

### 4.3 Jörn Sieglerschmidt: **Convergence of internet services in the cultural heritage sector – the long way to common vocabularies, metadata formats, ontologies**

Since several years it has been observed that information offered by different knowledge producing institutions on the internet is more and more interlinked. This tendency will increase, because the fragmented information offers on the internet make the retrieval of information difficult as even impossible. At the same time the quantity of information offered on the internet grows exponentially in Europe – and elsewhere - due to many digitization projects. Insofar as funding institutions base the acceptance of projects on the observation of certain documentation standards the knowledge created will be retrievable and will remain so for a longer time. Otherwise the retrieval of information will become a matter of chance due to the limits of fragmented, knowledge producing social groups.

### 4.4 Cornelia Dippold: **Knowledge Management and Knowledge Research by Integration of a Metaconcept into FDZ-RV**

The Research Data Centre of the German Statutory Pension Insurance (FDZ-RV) provides data (cross sectional and longitudinal data sets) of the German pension insurance for scientific research. These data products are based on management data, regularly published by DRV Bund in statistical reports. Besides the generated data products the scientific community should also be provided with relevant extra knowledge concerning the data products, independent of the knowledge's appearance in digital, paper or what ever form. The richness of this additional knowledge represents in form and content an heterogeneous data and document landscape. To handle this plenty of information we use a metadata concept seeing the data product – the micro data set – as the centre of our study and present a two-stage metadata model. First we identify the so called “wide sense metadata”, the relevant set of information. Second we enlarge the information items building the “wide sense metadata” by adding extra content to each item, the so called “narrow sense metadata”. Finally, for realizing an efficient knowledge research we associate the

model of knowledge representation with the model of interaction based on the previously developed meta data model.

4.5 H. Peter Ohly: **Mining: Added Value from Document Analysis and Retrieval**

Bibliometrics is understood as statistical analysis of scientific structures and processes. The analyzed data result from information and administrative actions. The demand for quality judgments or the discovering of new structures and information means that Bibliometrics takes on the role of being exploratory and decision supporting. To the extent that it has acquired important features of Data Mining, the analysis of text and internet material can be viewed as an additional challenge. In the sense of an evaluative approach Bibliometrics can also be seen to apply inference procedures as well as navigation tools.

## 5. Media in Knowledge Transmission

5.1 Sabrina Schrammel: **New Knowledge Spaces?! Considerations regarding the Analysis of qualitative New Space Relations as Implications of Knowledge-Organization in the Decade of New Media**

Computer and internet are not only instruments of knowledge organisation that have enhanced and accelerated the way we live, work and learn. The use of these new media has also changed, and is still changing, our social practices. So far there are hardly any considerations on the consequences of pedagogical-didactical computer-based knowledge organisation for our pedagogical practices. In order to reflect on these implications we have to find suitable categories for our investigation. In this article, the author suggests the category "space" as a basis for examining our modified pedagogical practices.

5.2 Konstantin Mitgutsch: **Incompatibility in Knowledge-Organization. On Productive Negativisms in Learning Processes**

A majority of investigations examine the importance of connectivity of knowledge for its organization and acquisition. However, only little research has been carried out in examining the role of irritations and incompatibility of facts, experiences and thoughts for knowledge organization. Especially in the field of learning by new media knowledge is not only affiliated via a linear process, but by confrontation and irritation. In providing insights into circulating and non-linear process of relearning and learning anew, a different concept of learning and of organizing knowledge will be proposed. How incompatibility might gain learning and what role it can play for the organization of knowledge, will be examined and reflected.

5.3 Sylvana Kroop: **Knowledge Organization in the Medium of increasing Complexity. The encyclopaedic Change from the 15th to the 20th Century**

The 18th century libraries were beautiful and manageable. Limited and manageable was also the totality of human knowledge. Today's electronic helpers were not yet necessary. The 18th century marks also an important change how knowledge was organized. In the age of Enlightenment a turning away from an inflexible and creation centered system towards an open and flexible organization of knowledge took place. The changing can be traced very well by looking at the examples of three encyclopedic forms of knowledge organization: the „Margarita Philosophica“ of Gregor Reisch at 1508, the "Encyclopédie ou Dictionnaire raisonné des sciences, des arts et des métiers" of Dennis Diderot and Jean Le Rond d'Alembert, that has been created between 1750 and 1780, and the Memex concept by Vannevar Bush

published in 1945. They all are popular examples of knowledge organization that minted without a doubt the thinking of a whole society.

## 6. Future of Knowledge Organization

### 6.1 H. Peter Ohly: **What Denotes 'Evolution'? Knowledge Organization in Retrospect**

This paper discusses general developments in the area of knowledge organization. A distinction is made between short and long-term developments. These yield a list of expected dimensions of future developments.

### 6.3 Winfried Schmitz-Esser: **Organizing Knowledge and Putting Knowledge Organization to Use**

How can public media utterances of societal rank (“knowledge”) be organized, not just by search words of a specific language as is practice of the day, but by their meaning and environmental intentions, expressed in some language, yes, but irrespective of a specific language, so as to enable to formalize this knowledge mediating cross-language ontologies that in turn support the action of knowledge machines? In such machines, knowledge could be “tamed” and put to use in a consistent, widely linguistically and logic-controlled way. Knowledge, once formalized, would be apt to be merged with (or segmented for) other ontologies, with a chance to make world-wide sharing of reliable knowledge come true. Moreover, formalization of knowledge in ontologies would produce new knowledge which in turn could be used to detect knowledge hidden in hitherto unknown texts and in heterogeneous texts to come in the future. Formalization of knowledge would enable, require, even cry for, responsible, knowledgeable human, and social, control. Educational curricula as known from library science and archival disciplines remain of importance, but will prove largely insufficient in the face of such challenges. Actually, in present new and upcoming ontology-based knowledge machines, the human resource is found addressed by labels such as peers, curators, knowledge and linguistic engineers and the like, names that all are falling short of expressing the most important three elements that have to come together in individuals and teams bound for constructing, maintaining and running cross-language ontologies for knowledge machines: Intimate domain/task expertise, full command of the source and target languages and their respective linguistics, and author-like, editorial responsibility that is up to the challenges of the new media world.

### 6.4 Michael Nentwich: **Technology-based Knowledge Organization in Academia**

When it comes to technology-related knowledge organisation, three trends are noteworthy: (1) what we may call „google-isation“, that is the typical, rather simple full text research in indices of search-engines, in other words the observation that users are not very demanding; (2) the increasing decentralised organisation of thematic portals; and (3) a trend towards the organisation of knowledge resources in a bottom-up or community-driven instead of a top-down manner. In the face of these trends, the author is tempted to advice the knowledge organisation professionals to either concentrate on solutions for specialised applications or to engage in actively contributing to the development of a „Google 2.0“ environment.

### 6.5 Gerhard Rahmstorf: **The Relevance of Language in Knowledge Organization**

Knowledge organization will remain an important task. The knowledge-organizational efforts will not only consist in the development of well-known instruments (Thesauri, classification systems etc.) but predominantly concentrate on the questions, how the

computer-technical representation and use of knowledge are associated. A knowledge representation in the proper sense is however only achieved if sentences and sentence meanings are represented. A machinable, structured knowledge representation requires that one must provide a certain number of record forms for the representation of sentences. In the future complex systems will be able to operate for the knowledge representation, knowledge evaluation question answer and for other functions.

## 7 Ethics in Knowledge Organization

### 7.1 Martin Van der Walt: Normative Ethics in Knowledge Organisation

The paper addresses the problem of whether the information profession needs ethical norms or guidelines specifically aimed at situations that may arise during knowledge organisation processes, and, if so, which specific norms should be included in codes of conduct. To explore this issue the following three specific questions are addressed:

Which forms of unethical conduct actually occur in knowledge organisation?

Which specific guidelines are required for promoting ethical practices in knowledge organisation?

To what extent does existing ethical codes make provision for knowledge organisation practices?

Four categories of unethical conduct in knowledge organisation are identified:

The use of terms with negative connotations

Misrepresentation of the subject

Censorship of “undesirable materials”

Bias in verbal indexing languages, classification schemes, evaluative comments in bibliographic records and subject analysis.

Guidelines in codes of conduct should be aimed at encouraging information professionals to avoid these unethical practices. An examination of a number of existing ethical codes for the information profession shows that, although general ethical statements, that can be seen as applicable to knowledge organisation tasks, do occur in these codes, this is by no means a general trend. It is also clear that very few of the codes give explicit attention to knowledge organisation.

### 7.2 Erwin Lengauer: Analytical Legal Ethics in the Context of Secular Justification Discourses with Respect of Dignity of Biological Entities

The aim of this work is to make the plausibility of secular explanatory discourse concerning the normative status of biological entities more transparent using the method of analytical legal ethics. This attempt tries therefore, first and foremost to explore the vast amount of values and normative assumptions hidden in typical bioethical discourse. The work criticizes the apparently neutral ideological world view of the in disposability of the “dignity of human life” in embryos.

### 7.3 Michael Nagenborg: Privacy – a Human Right or a Problem of Etiquette?

In the first part of the paper I present a short summary of the discussion on privacy in law and ethics. The main focus is on information ethics. In the second part I will discuss some issue with special relevance to knowledge organization like web information and data mining including FOAF as an example of the standardization of data representing persons.

#### 7.4 Karsten Weber: **Codes of Ethics for Knowledge Organization**

In the past, many professions and scientific disciplines decided to create a code of ethics which shall guide the professional activities of their members. The rules in these codes of ethics, sometimes called ethical guidelines, shall provide guidance in situations of moral conflict. Obviously, as other professionals or scholars, persons who are involved in knowledge organization face moral conflicts, too. Therefore, ISKO decided to discuss whether it would be necessary to create ethical guidelines for ISKO. In the paper two options to formulate a code of ethics are discussed: First, it is possible to identify moral values without formulating the way they can be achieved – this option is realised in the ethical guidelines of the German Gesellschaft für Informatik (GI). Second, it is feasible to clearly define morally acceptable professional actions without formulation basic moral values – this option is realised in the ACM and IEEE-CS Software Engineering Code of Ethics and Professional Practice. It is argued that if ISKO should decide to implement an own code of ethics it will be inevitable to choose the second option while it also will be necessary to address the specific needs of knowledge organization and its moral problems, for instance, the conflict of copyright and open access. Additionally, the second option has to be completed by basic moral values that shall underlie the professional actions of knowledge organization.

#### 7.5 Michael Nentwich: **Knowledge Base Law - a Novel Base of Knowledge for the Juristical Domain**

A „Knowledge Base Law“ (KB:Law) is an innovative database system. In this knowledge base legal information is made accessible in a hypertext environment both for laypeople and experts by presenting and interlinking not only basic information, but also comprehensive legal analysis. Richly documented answers are given in multiple languages and with regard to different national legal. The interface allows for browsing between different levels of answers and for switching between languages and countries. KB:Law is at the same time a multidimensional FAQ database, a novel form of publication and an attractive instrument for comparative legal studies. A first implementation of this concept contains questions and answers in the field of copyright law, with a particular focus on issues related to the intellectual property rights in the age of the Internet.