Towards knowledge organization with Topic Maps

Alexander Sigel M.A.
Informationszentrum Sozialwissenschaften, Bonn, Germany
sigel@bonn.iz-soz.de
http://index.bonn.iz-soz.de/~sigel/

Abstract:

The exciting Topic Maps (TMs) are an ideal catalyst for mutual learning experiences for proponents from the partially overlapping communities of Knowledge Organization (KO), Knowledge Management (KM) and Information Technology (IT). A long-term goal would be a tutorial white paper on the relationship between KO, KM and TMs, together with free reference software. KO is interested in optimizing the organization (the conceptual access structure) of knowledge repositories to support easier retrieval, creation and sharing of knowledge for user communities. TMs can indeed play an important role within KO: Together with related technologies, they have made it easier to provide innovative KO services. With TMs you can define arbitrarily complex knowledge structures and attribute them as metadata to information resources. Decentrally creating, maintaining and exchanging even more heterogeneous metadata is a powerful basic service of high interest for a broad range of applications. However, sooner or later you have to cope with the new semantic heterogeneity and come up with strategies to achieve better semantic interoperability. How could TM-based services alleviate the pressing KO problem of how to reorganize, enhance and semantically integrate heterogeneous subject data? Dedicated to this question, this talk takes a KO perspective: By sketching three typical scenarios in which heterogeneous metadata occur, it shows how classical KO challenges reappear with TMs, but also that TMs may be of value. Because the authors of the TM standard were right in not prescribing the application semantics of the structured link network, the widespread use of large-scale TMs will aggravate the well-known problem of the comparability and compatibility of KO schemata. A closer co-operation between the communities could aid the potential of TMs for KO/KM. Fortunately, the TM community has already started the fruitful exchange by discussing KO-relevant topics. Because of the flexible orientation of TMs towards usage contexts, especially user-oriented indexing should benefit from TMs. Approaches for achieving semantic interoperability within a layered model of decentral information provision are briefly presented as background against which further directions of KO with TMs can be discussed. One consequence for KO is that its methodology must be partially redesigned to take collaborative knowledge building activities on distributed resources more into consideration. This article also asks about the relationship between TMs and other means to computationally handle semantics in next-generation ontology- and agent-based knowledge services. In the end, possible further research towards this vision is suggested.

Introduction

Browsing through the advance material of the Metastructures 1999 conference , I stumbled across TMs (Topic Maps) and immediately got hooked. My background in information science and conceptual knowledge organization made me, of course, wonder about the potential of this technology for KO (Knowledge Organization) and KM (Knowledge Management) . After all, KO is much about knowledge structures and their interrelations, and TMs promise to provide a standardized technical means to build and share such. Naturally, I wanted to learn more and to discuss implications with my KO colleagues after having read all the TM material I could get hold of.

However, at that time none of them had even heard about the TM standard, despite its long history. Consequently, my request for discussion initiated no response. This is probably just another instance of the widespread phenomenon of separated communities. What are some of the obstacles in this case? First, it always takes some time for an innovation to diffuse to other fields. Second, the TM (Topic Map) standard

is not really intended to be read and understood by scholars of the humanities. Third, differing terminology is most hindering (cf. e.g. "facets"). Forth, there existed no tutorial material that was excellent, affordable and easily available. Furthermore, no large-scale real-world applications and examples could be reviewed. (I want to add: In spite of the rising marketing hype, there is still much room for improvement).

To overcome this situation, I imagine an interdisciplinary team working towards a high-quality white paper on the relationship between KO/KM and TMs (working title: "How to manage knowledge with TMs"). It may eventually evolve into a tutorial booklet. In my opinion, it should:

shortly set the necessary background on KO/KM,

present the basic idea, possibilities and shortcomings of TMs and how they fit into this background,

identify diverse KO/KMareas in which TMs could be or have been successfully applied (ideally illustrated by some success stories and business opportunities),

guide through the process (e.g. the analysis of user information needs, the semantic (ontological) modelling of controlled vocabulary, the design of the TM types, typical usage and problems).

develop one complex, continuous example,

explain the essential technical jargon and the implicit suppositions as clear as possible, and

have a kind of cookbook part which maps from a problem or question to technical advice in recipe form.

Moreover, we need more free reference software like ’tmproc: A Topic Maps implementation (in Python). I am strongly convinced that the availability of such open source tools will increase the number of critical minds which contribute. This in turn will probably lead to the discovery of more flaws and alternative interpretations in the standard, and to additional free software, perhaps even to innovative GUI (Graphical User Interface) concepts. In addition, it will give people the chance to acquire hands-on experience of what it is to work with TMs.

It would be fruitful to arrange for some intellectual cross-over, maybe in the form of a small workshop. The exciting TMs are indeed an ideal catalyst for mutual learning experiences for proponents from the partially overlapping communities of KO, KM and IT (Information Technology) . (The experts on KR (Knowledge Representation) have been subsumed under those communities). KO people need more explanation and hands-on training, KM people require a holistic, interdisciplinary methodology, and IT people may welcome suggestions about related work in areas based on a different epistemological horizon.

Even though this is very ambitious indeed, as the TM movement gained momentum, we now may be closer to that goal. Several companies became TM-aware (e.g. by Holger Rath’s article or by the TM tutorials at this conference). The new XTM-WG (Working Group on TMs for the Web) will hopefully push in the same direction.

In this technical TM session I - a KM apprentice - will share my limited insights on TMs from a KO perspective in order to learn from you. My main message is: Yes, you should seriously consider TMs if you plan for applications on top of digitally organized intellectual assets. But make sure to include strong KO expertise in your KM methodology and team applying TMs, since the classical challenges of KO will inevitably haunt you in a new disguise, even with this promising technology.

After having explained the nature of KOin general and the instrumental role of TMs for KO in particular, I will shortly sketch three typical KO scenarios in which TMs could be fruitfully applied. This should deepen your understanding of the nature of KO tasks and of the TMs’ potential to better accomplishing those tasks. The core of this talk is about the most interesting aspect of the relationship between KO and TMs: The possible role of TMs within the classical KO challenges of how to:

1. order knowledge in a principled way,
2. achieve sound comparability and compatibility of knowledge structures and of heterogeneous subject data.

Thereafter, I will briefly discuss some concepts for achieving semantic interoperability within a layered model of decentral information provision. I will pose some questions about the relationship between TMs and other means to computationally handle semantics in knowledge services. In the end, I will suggest possible further research directions toward the vision of next generation ontology- and agent-based knowledge services.

What is KO about? How do TMs relate to it?

KO is interested in optimizing the organization (the conceptual access structure) of knowledge repositories in order to support easier retrieval, creation and sharing of knowledge for user communities. (You may find a wealth of information about KO via the homepage of ISKO (the International Society for Knowledge Organization) , especially in its quarterly journal "Knowledge Organization" , and in the ISKO conference proceedings). The aim of KM (especially in knowledge-intensive enterprises) resembles this optimization to a certain extent, because KM has to ensure strategically that all important knowledge assets and flows are known, utilized and enhanced according to their respective long-term contribution to the business value.

A basic conviction of KO is that there exist important domains where some form of vocabulary control is beneficial. Hence the interest of KO in principles of classifications, thesauri, and ontologies.

The aim of an optimal conceptual access structure requires the careful design and steady maintenance of additional knowledge (meta-)structures. In order to provide the most useful subject access points for various user communities, indexers have to judge the potential subjects of an item within a collection from various viewpoints. This costly intellectual operation creates added value and obviously results in metadata. One is acquainted with such metadata from database indexing in the form of entries on library records, but also from subject-oriented clearinghouses (subject gateways,) on the internet. Another form is sophisticated back-of-the-book indexing of scientific and technical writings (see e.g.).

At that point we can already see the type of relation between KO and TMs: TMs are instrumental to KO, since they constitute (yet another) basic technology which KO may employ inter alia within a broader methodology in order to provide improved information retrieval and collaboration services. Therefore, TMs can indeed play an important role within KO, if they (together with other technologies) make it easier to provide innovative KO services. On the other hand, the requirements of KO help to spot what TMs are good for and what they do not achieve.

1. Content- and quality-rating agencies issuing SOAPs ("Seals of Approval") (e.g. quality-controlled subject gateways) and higher level content providers (e.g. abstracting and indexing agencies) crucially depend on the feature that additional knowledge structures can be expressed (as superimposed views) on top of, but independent from the original resources. No problem: After all, this is what TMs were designed for.
2. Considering the times when thesauri were updated and distributed in printed form by a central agency every few years, KO has advanced in big steps towards the decentral, collaborative creation and maintenance of knowledge structures. Instead of focussing on the statical and monolithical product, it is much more natural to view this effort as a dynamically evolving process in which domain-oriented experts constantly make small improvements. However, KO must more fully explore what this liberation from the traditional limits in non-collaborative environments implies for its methodology. A

few examples: Is voluntary co-operation a substitute for central authority if we want to maintain the same quality? Do we have to reindex more often? Is every view useful? Will the ability of TMs - to "consistently" support master indices of sets of documents with different owners and maintainers - have a positive effect on the consistency at the conceptual level?

3. Highly developed KO builds complex knowledge structures with model-based indexing languages possessing a grammar. For such applications it is highly useful that with TMs you can define arbitrary structures as indexing models.
4. With the growing number and size of internetworked information resources, scalability is vital. Although, e.g., the metadata for both a research organization's library and an important literature database in one specialty is certainly large, TMs can still handle it.
5. TMs (i.e. user-editable views on information assets) further aid KO because of their flexible orientation towards usage contexts. Newer KO theory is explicitly interested in the multiple alternative views and indices which different user communities tend to build for the same item or collection. Hence, especially user-oriented indexing should benefit from TMs. This approach will exploit the scoping and filtering mechanisms of TMs to achieve adaptibility to target groups.
6. Tightly interwoven with the collaborative creation, sharing and maintenace of conceptual knowledge structures (a key activity of all knowledge workers, not only of documentalists) is the need to search and navigate in the resulting semantical structures. While this will remain an open field for decades to come, a more principled structure will always further aid usability. It is, e.g., not important if the interface is hyperbolic but rather if the underlying knowledge structure is natural and predictable.

In sum, if it is really true that TMs are

*"**the** solution for organizing and navigating large and continuously growing information pools" (, p. 18),*

then KO would be well advised to use that tool.

I think, it is now obvious why TM, KO and KM experts should talk to each other: Although a fool with a tool still remains a fool, a tiger team with the right approach and tool set may accomplish the breakthrough.

A short sketch of three typical KO scenarios in which TMs may be usefully applied

All three KO scenarios make typical KO/KM tasks more vivid and illustrate the point that TMs may have a great potential for KO tasks as a basis technology, but that a more comprehensive approach and much further co-operation is needed.

Scenario 1: Knowledge structures in the social sciences: database indexing, a specialized information service and a clearinghouse

The Informationszentrum Sozialwissenschaften, Bonn, Germany , is (inter alia) responsible for the national research databases of the social sciences (SOLIS (Information system on publications in the social sciences) , FORIS (Information system on research projects in the social sciences)). In database indexing, all documents are assigned an abstract, controlled descriptors and classifications. From this pool, specialized information services (so-called soFids (specialized SDI services in the social sciences)) on 28 topics are compiled intellectually. There is some overlap between the topics. Each such service (view) superimposes its own conceptual structure (its table of contents) and filtering upon the already existing structures in the databases. One topic is "migration and ethnic minorities". In this case, a subject gateway linking to external, online information about that topic has recently been started . Again, it has its own conceptual structure.

We are looking for a means to:

1. consistently maintain those views without having to radically change the production procedures and systems,

2. provide users (e.g. of the soFid CD-ROM (Compact Disk read-only Memory/Media) on migration) with an experimental browsing and searching interface to enable them to explore those rich knowledge structures.

Scenario 2: A virtual reference room for digital cultural heritage

We take the mission and the objective of the MMI (Maastricht McLuhan Institute) from its homepage :

"[ ]o study and develop methods for knowledge organisation and knowledge management in a digital, distributed, multimedia world. These methods will be applied to cultural heritage, the design of learning technologies and new electronic services for business. Research will also explore the implications thereof.

The quest is to create comprehensive strategies for searching, structuring, using and presenting digital resources more coherently and efficiently; to integrate past knowledge and to produce ordered knowledge that leads to new understanding and insights."

One of the pending research projects applies the concept of a digital (virtual) reference room to selected cultural heritage resources, e.g. in the Limburg region. The material is virtually combined to support specialized usage contexts and tasks (e.g. learning about a certain painting technique). Such views put additional requirements on the subject metadata. This poses the challenge of how to cope with the semantic heterogeneity of those diverse resources and their metadata. Of course, in this case, the original sources cannot be changed, either.

Scenario 3: KM: Innovative information services based on controlled vocabulary

The CKO (Central Knowledge Officer) of a large technical consultancy sees the necessity to replace the existing, autonomous, redundant and decentrally maintained keyword lists with a centralized repository of structured vocabulary. In his eyes, the homogenization of the vocabulary and its synchronization with global classifications and translation aids is an absolute precondition for efficient search engines, push and pull services, interest profiles, portals, yellow pages, etc. His main concern is the bottleneck of developing and maintaining appropriate vocabularies, not the availability of technology or applications. To reduce the expected high efforts, he wants to buy and re-use as much knowledge structures as possible. A large part of the integration effort would be easier if those structures were formally defined, and more context information was available. The provision of knowledge structures marked up as TMs will open up business opportunities, but only if interested parties will agree on application-oriented TM semantics.

TMs and classical KO challenges: How to organize knowledge in a principled way? How to make order systems interoperable? How to deal with heterogeneous subject data?

Once we are faced with more than one content provider and stakeholder, with more than one rather homogeneous user community, and once we move from homogeneous databases to the plethora of independent internetworked information resources, searching and navigating by subject requires that we (at least virtually) bridge the gap between the numerous schemata which organize knowledge by subject. Hence we are faced with the well-known KO problem that it is extremely difficult to achieve comparability and compatibility between various schemata which organize knowledge by subject . All the more this extends to the problem of mapping, merging and integrating the corresponding subject metadata referring to those different schemata.

Let’s assume that all your data is "XML (Extensible Markup Language) -ified", each data source has a predefined basic order structure, and your objects have been associated with metadata derived from these structures. Current best practice for Web (WWW - World Wide Web) resources suggests to use a RDF (Resource Description Framework) -embedded cousin of Dublin Core , which points to a machine-readable version of your order system.

Only recently have documentalists recognized the full power of XML/RDF as a basis for information systems . If there exists an important difference between TMs and RDF, this may have far reaching implications. Given we already possess a vocabulary: Which standard shall we adhere to? While RDF is more general and it may be possible to write a converter between both , RDF is rather centered around the resources instead of around the cross-resource knowledge structures. Moreover, it is important to note that only a TM processor can make useful interpretations, since RDF does not deal with the semantics (, , pp. 21ff.).

As most of this knowledge is implicit, the computer has no clue of the underlying concepts and is limited to string processing. TMs allow us to formally define such schemata. The standard designers, however, wisely did not prescribe the user semantics of the structured link network. (A fact which we enjoy in lengthy threads on the TMmailing list).

Given that the key players decide that TMs are the preferable alternative, especially for the Web: Then we still have to resolve:

1. How can we define the KO structures in TMs in a principled way?
2. Which semantic relation types should we use and standardize in TMs?
3. What approaches exist to tackle the scalable interrelating of TMs in order to achieve layered semantic interoperability?

How can we define the KO structures in TMs in a principled way?

We would like to know if it was possible to create knowledge structures in a principled way in order to ease their conjoint use. How can we plan today for future merging? (Note that this is more than asking for the architectural "organizing principle" of the topic paradigm itself). The answer is: While KO offers some guiding principles (cf. e.g. ,), there are more open problems than solutions.

Let me mention some **difficulties**:

> In contrast to public belief, the conceptual recognition of an item’s subjects cannot be determined objectively once and forever, because this is an interpretive, hermeneutic process which is always dependent on the social situation, purpose and context, including culture. (The "totality of the epistemological potential" is unknown and infinite). Thus it is rather difficult to find a criterion to judge "correctness" of a given subject assignment. This in turn means that two different ordering systems for the same specialty may be equally valid. The CKO (from "Scenario 3: KM: Innovative information services based on controlled vocabulary ") may find that user communities will resist his approach of central control because the variants which he regards as unnecessary and redundant are possibly grounded in different social praxis.
>
> Most existing ordering structures were not primarily made for computerized usage. Hence they lack exact definitions and a more formal specification which could be utilized by automated means. In addition, the context of their social genesis and their underlying suppositions are only implicit and cannot always be reconstructed.

In business contexts, stakeholders segment the market and secure their position by differentiation on order schemata. Only voluntary co-operation can help here.

As you cannot simply throw order systems (or their metadata) together and inappropriately use them out of their context, someone seriously working with TMs will also need strong expertise in how to create useful conceptual structures to organize knowledge in a principled way. His business may depend on the state of research in such a methodology.

If we expect large-scale TMs to find widespread application, it is self-evident that the physical sharing of TMs per se will not alleviate the problem of incompatible KO schemata. TMs can be a vehicle for semantic integration, but, on the contrary, the heterogeneity will increase, because it has become easier to create such schemata.

I now turn to a few **principles**:

Fortunately, the TM community has already started the fruitful exchange by discussing KO-relevant topics (such as constraint mechanisms like schemata or templates), relationship types, validation and inferencing services, or the principles of analytico-synthetic (faceted) KO schemata). It is always helpful to check with a good handbook on thesaurus construction and maintenance, or on the subject indexing process. Regrettably, books concentrating on conceptual issues are rare, and KO will have to integrate the new requirements and possibilities into new textbooks.

It is currently debated whether a constructive way to guide user-oriented depth indexing exists at all . I believe that a thorough, domain-oriented analysis of the types of user requests with the relevant answers, together with their embedding into the social praxis of this special community, will ultimately uncover pragmatically relevant core knowledge structures.

The next step is to find out how these complex structures can be broken down and be expressed as a combination of simpler constructs. This brings us to the most prominent way to design principled KO structures: Facetted classification: Every compound subject can be synthesized from a set of elementary, independent building blocks, using the grammar of a powerful knowledge order language. The definition of concepts as specialized composition of faceted (mutually-exclusive) sorts which are subsumed by postulated, very general basic categories avoids enumeration. However, it needs a great deal of expertise to find the right building blocks! Nota bene: Apart from the name, the advanced TM concept "facet" is not related to that method, and calling your TM topics faceted does not guarantee a useful KO structure.

Much like object-oriented models, faceted knowledge structures are advantageous if you want to merge them: The structure of the building blocks is clearer, and you only have to handle fewer and more abstract elements.

Which semantic relation types should we use and standardize in TMs?

KO has at length dealt with the question of which types of relationships are needed. This is more than part-of/has-a, is-a and instance-of. The "related terms" relationship has been semantically differentiated by various specialties, including pragmatism in linguistics and rhetoric structuring theory. However, the KO community did not achieve consensus on which relations to standardize and therefore still lives with the outdated thesaurus standards. Some programs for the structuring of vocabulary offer up to 30 relation types, but only very few are actually used in documentary practice.

The reason may be that the very differentiated relation types were not directly more useful, since they were not supported by retrieval software. In addition, the less predictable the assignment of a specific relation, the more errors happen. In sum, the extra work did not pay off. But with ontological engineering this situation may change, because inference and validation services need fine-grained relations. A lesson for the standardization of TM templates may be that consensus on such sophisticated templates can only be achieved in specialized domains.

What approaches exist to tackle the scalable interrelating of TMs in order to achieve layered semantic interoperability?

The original requirement that had led to the development of TMs was that of a publisher who wanted to merge indices in technical documentation. But how can TM applications merge topics if even similar topics within the same scope may have different extentions? With TMs one can express conceptual structures, but, of course, TMs do not come up with valid fusion strategies. Without some background in KO or comparable experience, naive merging will result in a big pile of rubbish in which all context will be lost.

Because the problem of how to reorganize, enhance and semantically integrate heterogeneous subject data will persist with TMs, I will very briefly discuss concepts for achieving semantic interoperability within a layered model of decentral information provision. Here the focus is on semi-automatic methods which depend on intellectually maintained schemata.

In my view, ideally one would improve all schemata involved towards faceted schemata and reclassify the items. However, limited resources render this approach rather unfeasible. In this situation, TMs could be helpful, because they allow to define structures independent from and across the original documents, they support a more formal definition, they are open for alternative views, and they make collaborative work on evolving structures possible.

Thus TMs could be one apt IT that fits into Krause's layered model of information provision in which no longer a central agency exerts its authority in subject indexing and vocabulary control upon agencies located lower in the hierarchy, but in which a group of partners co-operate. Such a strategy does not result in uniform metadata, but leads to layers of heterogeneous metadata with different quality control procedures. Intellectually controlled high-quality subject schemata lie in the heart of those layers. Intelligent transfer components are sought which can improve on subject data on outer layers by using the structure of inner layers. The main methods are: The compilation of cross-concordances which map between entries, and a combination of quantitative-statistical with qualitative-deductive approaches. The right mix seems to be domain-dependent and is hitherto unknown. Personally, I am convinced that qualitative methods and case analysis will yield rich material and exploitable ideas for transfer strategies.

The ongoing research project ViBSoZ (Virtual Library in the Social Sciences) (,) explores how to cope with heterogeneous subject data in the social sciences. The SIMS (School of Information Management Systems,

UC Berkeley) project "Search Support for Unfamiliar Metadata Vocabularies" by Michael Buckland et al. is a related approach which also includes issues of scalability and information agents. Several projects have tried to automatically gather and experimentally classify HTML (HyperText Markup Language) documents in one specialty according to existing knowledge structures. E.g. Koch and Ardö have thoroughly compared the results both with intellectual classifications and with expert judgements. At best, 2/3 of the results match.

In sum, we dispose of no overall convincing strategy to achieve semantic interoperability, but of a broad range of necessarily heuristic methods. Theory does not say much about this "repair case" in which most systems to be integrated are not principle-based and much context has already been lost. This stresses that it is worthwile to put great effort in the meticulous, intellectual maintenance of conceptual structures, since such key assets are at the heart of the layered model. All transfer components (including automatic ones) will depend on the quality of the innermost schema. TMs could be one tool with which knowledge structures could be maintained more easily, and thus more time could be dedicated towards better quality. Such high-quality knowledge structures will be needed anyway by clever strategies in next-generation ontology- and agent-based knowledge services.

Questions about the relationship between TMs and other means to computationally handle semantics in next-generation knowledge services

The vision of high-level ontology- and agent-based knowledge services is not new. Likewise, at first sight TMs seem to be nothing more than a new format to mark up what formerly was expressed as assertions and rules in AI (Artificial Intelligence) databases. As there have been other formats and languages before which did not receive that much attention, the purpose of this short section is to ask, whether they are already superseded, or are just variants, or whether they constitute a welcome complement to TMs (e.g. in order to validate the semantics of a TM application according to a TM schema (or a similar mechanism)). This question is of relevance, since an innovative information service provider is interested in estimating the survival power of a technology before making huge investments.

Before TMs, you may have thought about introducing the computational semantics needed for metadata fusion by equipping information agents with clever heuristics based on AI tools. After all, that's what the validation and inferencing services of terminological logic's subsumption is good for. You are right if you object that this is overkill, that one cannot make everything explicit, and that it is computationally demanding. So which other ontological tools did you use instead to express your knowledge structures? XML-encoded SHOE (Simple HTML Ontology Extensions) ? Its relative CKML (Conceptual Knowledge Markup Language) ? Stanford's KQML (Knowledge Query and Manipulation Language) /KIF (Knowledge Interchange Format) (, ,)? Why not? Now with TMs: Would the higher TM services differ from general tell-ask-performatives of information agent languages like ACL (Agent Communication Language) ?

I would like to learn more about the relation between this research and TMs. Maybe someone can point me in the right direction? I know of no demanding TM validation service. Is it possible to convert between the formats or to communicate between applications? How can we achieve that information agents exploit knowledge structures expressed in TMs?

Outlook and conclusion

My personal experience with KO is that because knowledge structures are a socio-cultural product, AI modeling is only of limited help. During the process of detecting and exploring emerging knowledge structures, a tool is needed that allows to start less formally. Thus I recommend to investigate:

1. How useful could TMs be during domain analysis (,)?

2. What effect does principle-based KO have on the quality of the knowledge structures, if the work is done within a TM application?
3. How might an architecture for intelligent transfer components within a layered model of information provision look like, if next-generation information agents worked on TMs?

Altogether, I like to see more information agents which rework subject data with their information strategies and which are informed by improved versions of ontological models.

In conclusion, TM-based services may alleviate KO tasks, but strong KO expertise is indispensable. The main implication of TMs for KO is not that KO thesauri and classifications can (trivially) be defined and maintained as TMs, but that - like with hypertext - there is a paradigm change: The KO methodology must be partially redesigned for collaborative knowledge building activities on distributed resources. This paper is a first attempt to stimulate co-operation between the specialties, but much more work is necessary. What about the tutorial white paper and the reference software? What about joint projects?

Acknowledgments

I am greatly indebted to Steve Pepper for his encouragement.

Bibliography

[1] 6th Annual Metastructures Conference 1999, Montréal, Québec, Canada
http://gca.org/attend/1999_conferences/meta_99/meta_99_program.htm

[2] Sigel, Alexander (July 1999): Request for discussion (RfD): "Knowledge organization and management of heterogeneous subject data with Topic Maps and ontologies"
http://www.isko.org/topic-maps.html

[3] tmproc: A Topic Maps implementation (in Python)
http://www.infotek.no/~grove/software/tmproc/index.html

[4] Rath, Holger (1999): Mozart oder Kugel: Mit Topic Maps intelligente Informationsnetze aufbauen, in: iX 12/1999, pp. 149-155 (in German). The material is similar to /11/).
http://www.topicmaps.com/content/resources/ix9912/hhr.pdf

[5] XTM-WG, hosted by IDEAlliance:
http://www.idealliance.org/ http://www.egroups.com/list/xtm-wg

[6] ISKO Homepage: **http://www.isko.org/**

[7] Journal "Knowledge Organization":
http://www.ergon-verlag.de/contents/06ainfobibl.htm

[8] All proceedings volumes have been publishes with Indeks, Frankfurt/Main, or with Ergon, Würzburg. See also the Call for Papers for the current conference:
http://www.ergon-verlag.de/contents/06ainfobibl.htm

[9] Koch, Traugott (2000): Quality-controlled subject gateways: Definitions, typologies, empirical overview, in: Online Information Review, vol. 24:1, Feb 2000 (Secial issue on "Subject Gateways"). **http://www.lub.lu.se/~traugott/OIR-SBIG.txt**

[10] Fugmann, Robert (1990): Systematical index to the Proceedings of the 1st International ISKO Conference on Knowledge Organization. Darmstadt, Germay, 1990. **http://www.isko.org/KO-Darmstadt90-Index/**

[11] Rath, Hans Holger & Pepper, Steve (1999) Topic maps at work. Chapter 1 in: XML Handbook, 2nd ed. **http://www.topicmaps.com/content/resources/xmlhb2/hhr-stp.pdf**

[12] Informationszentrum Sozialwissenschaften: **http://www.bonn.iz-soz.de/**

[13] Subject Gateway on migration issues: **http://www.bonn.iz-soz.de/themen/migration/**

[14] MMI - Maastricht McLuhan Institute: **http://www.mmi.unimaas.nl/**

[15] International Society for Knowledge Organization, Polish Librarians Association and Society for Professional Information (1996): Compatibility and Integration of Order Systems. Research Seminar Proceedings of the TIP/ISKO Meeting. Warsaw, 13-15 September, 1995. Warszawa: SBP, 1996.

[16] W3C Resource Description Framework: **http://www.w3.org/RDF/**

[17] Dublin Core Metadata: **http://purl.oclc.org/dc/**

[18] Koehne, Michael. Personal Communication 2000-03-09. kraehe@copyleft.de, **http://www.xml-edifact.org/** .

[19] Hirsch, Sven (2000): Aspekte der Entwicklung von Fachinformationssystemen auf der Basis von XML/RDF, in: nfd 51:2 (2000), 75-82. (Aspects of the Development of Information Systems based on XML/RDF, in German) **http://www.darmstadt.gmd.de/dgi/nfd/Ausg200/abstr1x.html**

[20] Topic Maps Frequently Asked Questions: **http://www.infoloom.com/tmfaq.htm**

[21] Pepper, Steve (forthcoming): Navigating haystacks and discovering needles, in: Markup Languages Journal. **http://www.topicmaps.com/content/resources/mlvolx/stp.pdf**

[22] Fugmann, Robert (1999): Inhaltserschließung durch Indexieren: Prinzipien und Praxis. Frankfurt/Main: DGD (in German).

[23] Iyer, Hemalata (1995): Classificatory Structures: Concepts, Relations and Representation. Frankfurt/Main: Indeks.

[24] Sigel, Alexander (2000): How can user-oriented depth analysis be constructively guided?, in: Procs. 6th Int. Conference on Knowledge Organization. ISKO. July 10-13th, 2000, Toronto, Canada. Würzburg: Ergon. (forthcoming). **http://index.bonn.iz-soz.de/~sigel/veroeff/publications.html**

[25] Krause, Jürgen (1996): Informationserschließung und -bereitstellung zwischen Deregulation, Kommerzialisierung und weltweiter Vernetzung ("Schalenmodell"). IZ-Arbeitsbericht Nr. 6, September 1996 (in German). **http://www/publications/series/working-papers/ab6.pdf/**

[26] Project ViBSoz: **http://vibsoz.bonn.iz-soz.de/**

[27] Kluck, Michael (2000): Exploration der Suche in heterogenen Datenbeständen - Bestandsaufnahme und Realisierungsprobleme für ViBSoz. **http://vibsoz.bonn.iz-soz.de/Pub/ViBSozAB2.pdf**

[28] Buckland, Michael et. al: SIMS project: **http://sims.berkeley.edu/research/metadata/**

[29] Koch, Traugott & Ardö, Anders (2000): Automatic classification of full-text HTML-documents from one specific subject area. DESIRE II D3.6a, working paper 2. **http://www.lub.lu.se/desire/DESIRE36a-WP2.htm**

[30] SHOE project: **http://www.cs.umd.edu/projects/plus/SHOE/**

[31] CKML: Conceptual Knowledge Markup Language: **http://wave.eecs.wsu.edu/CKRMI/**

[32] Knowledge Sharing Effort and Ontology Server: **http://www-ksl.stanford.edu/knowledge-sharing/**

[33] KIF: Knowledge Interchange Format **http://logic.stanford.edu/kif/**

[34] KQML: Knowledge Query and Manipulation Language: **http://logic.stanford.edu/sharing/papers/kqml.ps http://www.cs.umbc.edu/kqml/**

[35] Albrechtsen, Hanne (1993): Subject Analysis and Indexing: >From Automated Indexing to Domain Analysis, in: The Indexer 18(4), Oct. 1993, pp. 219-224.

[36] Hj⊘rland, Birger & Albrechtsen, Hanne (1995): Toward a New Horizon in Information Science: Domain Analysis, in: Journal of the American Society for Information Science 46(6), pp. 400-425.

Author

Alexander Sigel M.A.

Researcher
Informationszentrum Sozialwissenschaften
Postal Address:
Lennéstr. 30
D-53225 Bonn
Germany
Telephon: +49 228 2281 170
Fax: +49 228 2281 120
E-mail: sigel@bonn.iz-soz.de
Web: index.bonn.iz-soz.de/~sigel/

Alexander Sigel - Alexander Sigel, an information science researcher, investigates methods and develops tools for improved conceptual knowledge organization and summarization of intellectual assets. He is interested in how to leverage next generation information systems by information agents with indexing heuristics that can adapt their ontology-based subject analysis to the interests of user communities. His working experience includes two national scientific information centers and the GMD National Research Center for Information Technology, Germany. Alex is an active member of several societies in his specialty, most notably the International Society for Knowledge Organization (ISKO). He is also co-chair of the German Perl Workshop and develops his information services in Perl. His published talks include academic conferences within information science, knowledge organization and user modeling (see homepage).