POVOO – Process Oriented Views On Ontologies

Eva Gahleitner, Wolfram Wöß

Institute of Applied Knowledge Processing (FAW)
Johannes Kepler Universität Linz, Austria
A-4040 Linz, Altenbergerstr. 69
{egahleitner, wwoess}@faw.uni-linz.ac.at

Abstract. Ontologies still lack of including and considering the dynamic aspects of business processes. Therefore existing ontology-based information systems provide only static information which does not suit the actual working context of a user. In this project we extend information retrieval techniques with ontologies through a Process Oriented View On Ontologies (POVOO). The purpose is to satisfy a user with information that depends on the current process the user is working in. Due to a context aware approach it is possible to dynamically adapt the information to the user's current working situation. We introduce a methodology for generating views on ontologies and we illustrate how an application can use them to query highly specialized knowledge bases.

1. Introduction

Ontologies are widely used in the area of computer science, but did not really reach the step into the area of commercial business engineering. Whereas in the field of information retrieval (IR) ontologies emerged as a major support for improving the recall and precision of search mechanisms, they only play a subordinate role in process modeling[1]. Nevertheless, in daily work business processes are often the starting point for software development and define requirements for software systems. Research and industry have addressed the alignment of business processes and information technology (IT) only marginally. This leads to separate modeling areas: one for information management and retrieval and one for business engineering.

Ontology-based query techniques suffer from a number of disadvantages which have major impacts on their usage in business process modeling:

- Ontologies provide a single monolithic structure, splitting them up into small units is hardly possible.
- Ontologies do not consider dynamic aspects. A process is typically characterized by a dynamic sequence of events and operations. The need for knowledge may change according to these different process events and operations. For example, a technician who designs a new car engine needs information which is different from

[1] Within business process modeling ontologies are used to represent explicit formal specifications of the terms in the entire process management domain and relationships among them.

the information a worker at the assembly line or a car dealer requires for the customers.

- As the size of the ontology raises so does the complexity of its structure and therefore the complexity for a user to find the right concepts (cf. highly specialized ontologies in medicine like UMLS [24]).
- The context in which a user (an employee in a department, a user of a software application, etc.) works determines the user's view on the available knowledge. Much work has already been done in the field of context-based ontologies for certain users or user groups, but little for particular views on knowledge in the context of business processes (cf. [31]).
- One structure does not fit all: information can not easily be categorized within a single (tree) structure, so that users will always find what they are looking for. This is due to the multi-dimensional nature of the information. Any piece of information can be categorized according to one or more facets. Such a multi-facet categorization better reflects the different viewpoints one can have on a single piece of information.

In this work we introduce our approach to integrate views on ontologies in the information retrieval process with specific consideration of business processes. The acronym POVOO stands for *Process Oriented Views On Ontologies*. The purpose of POVOO is to satisfy a user with information that depends on the current process the user is working in. We propose a context aware solution which considers a user's working process and the corresponding information required by a user during certain tasks in this process. For example, when office workers are working on a specific task they are working in a certain context, thus only specific information is necessary to get the work done. Working contexts differ according to the required information and the involved people. This characteristic of work is exploited in the IR mechanism of our approach where views on ontologies represent the working contexts. In this way the system can search and present relevant information in the current context.

For highly specialized knowledge bases which can be found e.g. in medicine or biology, we assume that the information itself, which is relevant to execute a certain task (the documents in a knowledge base), stays the same, whereas the relationships, the various specialization and generalizations, the integration of various concepts in a new one, the ordering of the concepts etc. may differ depending on the user or the actual business process step. We therefore emphasize an approach which uses ontologies not only as simple vocabulary to define a lingua franca in business process engineering but rather as a way to structure the knowledge for particular processes. Within POVOO we develop a methodology for generating views on ontologies (which we call *ontology views*) and we demonstrate how applications can use them to query ontology-based knowledge bases.

The remainder of the paper is organized as follows: Section 2 overviews related work concerning ontologies in IR and process modeling, view based search and process oriented ontologies. Section 3 describes the characteristics of highly specialized knowledge bases, presents POVOO’s three ontology levels and the concept of ontology views. Section 4 gives an explanation of our query mechanisms and how ontologies are connected with process modeling techniques. Finally, section 5 describes further research and concludes the paper.

2. Related Work

A business process defines the sequence of activities and the kind of resources (machine or human) which a process or an activity needs for its execution [3]. In recent years various (business) process modeling techniques have been introduced: Starting from the well known Petri nets [4] or high-level Petri nets [5], over UML activity diagrams [6] and object behavior diagrams [7] to more enterprise and business related techniques such as Event-Driven Process Chains (EPC) [8], UML Profile for Enterprise Distributed Object Computing (EDOC) [6] or the Business Process Modeling Language (BPML) [9].

All these techniques have in common that they describe the *behavior* of a system. By contrast, ontologies describe the *knowledge* of the system. Ontology is an explicit specification of a conceptualization [10], it captures the knowledge of a certain domain. But ontologies are not limited to the description of domain knowledge. They can also be used to define problem-solving knowledge (so-called task knowledge or task ontologies). In business engineering task ontologies create an ordering over sets of tasks and subtasks and are therefore defined as hierarchically ordered task ontologies [11] [12] [13].

Whereas in business engineering ontologies are simply used as a common vocabulary for processes and tasks, ontologies in information engineering are applied in various ways. E.g. in information retrieval (IR) ontologies have been commonly used to improve recall and precision [2]. Their main advantage relies in their ability to organize information into hierarchically ordered taxonomies of concepts, and to define attributes and relationships between these concepts. Two approaches are used in IR: *query expansion* and *conceptual distance measures*. The former expands the user query by adding terms semantically related to those used in the original user's query and therefore documents that do not necessarily contain the queried terms may be retrieved [14]. The latter uses a conceptual distance measure to calculate the similarity between terms in a query and terms in a document [15].

An extension to these IR methods is the concept of view-based or multi-faceted search methods [16] [17]. Here the idea is to organize the terminological keywords of the underlying knowledge base into various hierarchies which help the user to better formulate queries. For example, the keywords of a knowledge base can be ordered according to different aspects, e.g. "Time" or "Place". Such hierarchies are often called facet or views. The facets provide complementary views on the content along different dimensions.

However, existing multi-facet search tools use simple subclass-taxonomies [18]. They do not consider various relations between the concepts of an ontology (they are built for database querying). The Ontogator [19] approach combines the usage benefits of multi-facet search with the answer quality benefits of ontology-based search. But Ontogator does not support automatic querying; the users have to define the queries on their own.

3. Views on Ontologies

Ontologies describe those parts of knowledge, which are interesting for a certain domain. If a user likes to tailor ontologies to specific aspects of the phenomena of interest (e.g. to implement a certain application) he/she has to create different versions of the same ontology. Ontology versioning is a well known research area in the field of ontology engineering [1]. Unfortunately, these approaches only take care of the changes in the ontology *itself*, they do not deal with different *views* somebody may have when working with the ontology within a given process. In that case the ontology does *not* change, only the parts which are relevant to a certain user query change.

In general, views create virtual schemas and resource descriptions reflecting only the users' (applications') conception of a specific application domain. There is a large body of work on views for the relational data model. For example, the commonly used structured query language SQL [25] serves as a view definition language. By contrast, in the semantic web ontology views have been regarded only marginally until now. One known adoption is the use of scopes within topic maps [20] [21]. Currently, there are only two semantic web view languages for ontologies, both of them are built upon RQL query language and are aimed at RDF(S) data models: RDF View Language (RVL) [22] and the ontology view language proposed by Volz et al [23].

3.1 3-Level Architecture of Ontologies

Views (also called facets) provide complementary views on the content along different dimensions. They are widely used by database management systems (DBMS). A prominent role in DBMS plays the ANSI 3-schema architecture [26] that describes the different views on a database. In the center of the 3-schema architecture is the logical schema, which represents a complete business-oriented view on the information model. The underlying physical schema reflects the physical representation of data according to the requirements of the database. The external schema on top of the architecture represents specific views on the logical schema from the perspective of an individual application.

This ANSI reference model can be adapted to the area of ontologies (cf. Figure 1). The ontology concepts and their interrelationships, which are described according to the terms and principles of the domain (the semantic), represent the logical schema. We call this level the *semantic level*. The physical schema represents the syntactic specification of the ontology (e.g. built-in constructs given in RDF(S) or OWL) (*syntactic level*). The external schema is a mapping between the ontology schema and the schema the application is using. In the simplest case this is just a subset of the concepts, attributes, and relations of the ontology. For more complex applications, views are arranged in the way how the ontology concepts and relationships are viewed by an agent (human or software agent). More precisely, creating such a view over some data on the semantic web essentially consists of the creation of virtual metadata schemas and descriptions consistent with the agent's perception of those data. We call this level the *application level*.

Whereas the semantic and syntactic level is well discussed in the semantic web, views on ontologies are only marginally regarded. The application level has a major impact on the usability of existing ontology based knowledge systems. For example, ontology change management could be based on views where each view represents a major change in the knowledge model. Different versions of the same ontology could be specified in different views on that ontology. Additionally, working with views makes maintenance of ontologies easier.

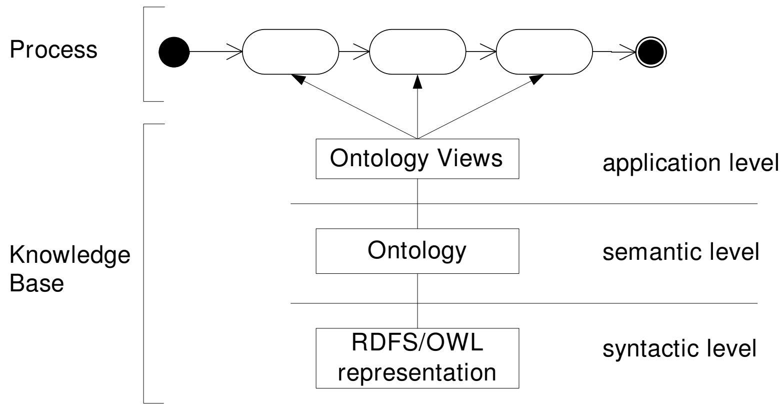

Fig. 1. Ontologies and ontology views analogous to the ANSI 3-level architecture

Another advantage of ontology views is that they describe information according to different contexts. This characteristic can be used to align ontology management systems with process oriented approaches. The kind of working process and its different conditions and dependencies between the single process steps have a major influence on the kind of information a worker requires. This information is described through various ontology views (compare Figure 1). The ontology views are based on the underlying ontology of the system (of the semantic level).

3.2 Specialized Domain Ontologies

Views on ontologies can only be built for relatively static and constant processes. Such processes can be found e.g. in medicine (e.g. diagnostic processes in medicine have a common structure). Ontologies belonging to specialized fields of long academic and professional tradition show a high degree of stability. Although disciplines such as medicine or biology have experimented drastic changes, this does not however mean that these novelties completely invalidate earlier conceptual organizations. An ontology about oncology may be affected by scientific advances but it is much less likely that it will be reformed in its totality. This relies on the high level of international consensus that some of these disciplines demonstrate.

Another feature of specialized field ontologies is the high granularity of their content. For example UMLS [24] provides a content base with highly specialized terms and documents. Additionally, the sources used for constructing specialized domain ontologies are well structured. Proof-reading and controlled communication leaded to a high formality of the sources. This high level of granularity makes it easier to split up knowledge in small, coherent knowledge pieces.

These properties (stability, high granularity and formality) make specialized domain ontologies a reliable resource for the retrieval of information, as well as a more effective one than its counterparts of non restricted fields and those used for common language. In other domains, which are not that structured and well defined the building of views on ontologies and the alignment of views to processes may lead to modeling problems.

3.3 Creating Ontology Views in POVOO

Existing approaches in view based ontology management [19] have a number of disadvantages:

- Views are only built on taxonomies: the multi-facet search just regards concept hierarchies (*subclass_of* or *part_of* relationships) not the entire semantic relationships of an ontology.
- The taxonomies are built on hierarchy rules which tell how to construct the taxonomies. In the mentioned approaches the taxonomies are built on simple Java applications and are therefore hardly to maintain.
- The GUIs are not suitable for large ontologies with various views on the ontology.

In POVOO we regard ontology views not only as a set of simple taxonomies. In our approach a view consists of various concepts, attributes and relationships which themselves build an ontology. The querying is therefore not restricted to a set of hierarchical ordered concepts but considers the entire semantic dependencies of concepts. Based on these modeling conventions, in POVOO views on ontologies are created in two ways:

- Manually by using an editor: this editor allows the integration of different classification schemas into the ontology as well as various relationships between the concepts. The editor will be integrated in the Protégé ontology editor framework[2].
- With the help of semantic web querying languages, e.g. RDQL (RDF Query Language) [27] and OWL-QL (OWL Query Language) [28].

The connection of process modeling and ontology views disburdens the user from choosing the right facets in the querying process. POVOO automatically identifies the necessary ontology views for the given process step, expands the user query and displays the result set according to this view. For example, when searching for medical reports one gets an anatomical ordering during the anamnesis process, whereas the same reports are relevant in a temporal ordering when preparing a surgery. Due to traceability reasons a user can always switch to a non-view-based search. The result set is then presented according to the underlying ontology of the information system.

[2] http://protege.semanticweb.org/

4. Integrating Process Models and Ontologies

When integrating structural and behavioral system aspects into an information system it is necessary to know a) which processes should be performed, b) who is responsible for certain tasks, c) which kind of information is needed and d) which resources are used. These different viewpoints are regarded in various business process engineering models. For example, ARIS (ARchitecture of integrated Information Systems [29]) a well known method in the German speaking part of Europe for analyzing processes distinguishes between a workflow model, functions, data and data flows as well as organizational units. In addition, INCOME/WF [30] follows a very similar approach. It supports four kinds of workflow views, an information object view, a view on existing resources, and a management view.

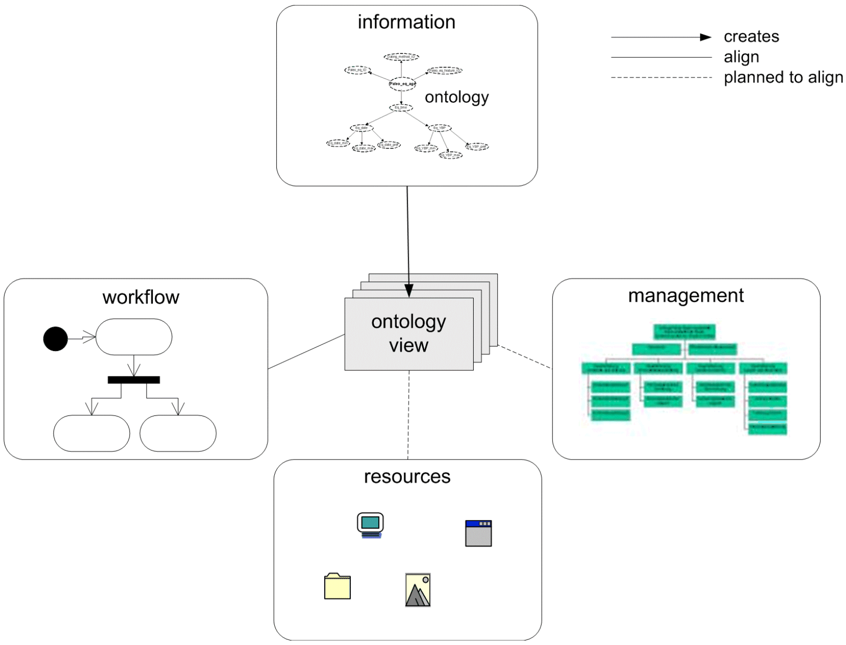

Fig. 2. Ontology views for integrating information, workflow, management and resource model

POVOO maintains the mentioned four viewpoints and connects them with the help of ontology views. In POVOO ontologies are used to represent the information model of a company's knowledge base. The information model (the ontology) is the central model for which various ontology views are built in order to connect all other models together (compare Figure 2). The workflow or process models are based on existing process modeling languages. Analogously this is true for the management and resource models.

In the first phase of our project we will place emphasis on the integration of ontologies and workflow models, in later stages we will also integrate of the used resources (e.g. lexicons) and the responsibilities of the users in the process (their position in the organization).

4.1 POVOO Querying Mechanisms

With POVOO a user has the opportunity to search for relevant information in two ways. Firstly, by using a simple keyword based search mechanism, and secondly, with the help of views on ontologies. The search mechanism then regards the certain role a user is playing when acting in a process. For example, in a medical environment a user may be a surgeon, an internist, a nursery, etc. who plays a certain role in the process. In the first phase of a process one may need more generic information including only generic knowledge bases, whereas in subsequent process steps one may need a more specialized view on the ontology, including more specialized knowledge bases.

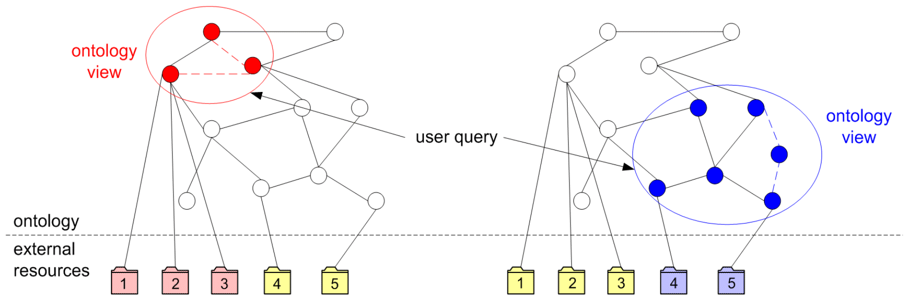

Fig. 3. View based search in POVOO

Figure 3 shows a possible scenario where different views provide different result sets for the same user query. The views are built on the same ontology. Within a view the structure of the ontology may be changed, e.g. a new or existing concept is added or deleted, or the relationship between concepts are changed, etc. Whereas on the left side documents 1, 2 and 3 are in the result set of the query, documents 4 and 5 are in the result set of the same query based on the view on the right side. Additionally, the result entries are organized according to the structure of the current ontology view. For traceability reasons the user can at any time switch to the underlying ontology.

4.2 Connecting Ontologies with Process Modeling Techniques

In order to connect ontologies with process modeling techniques we rely on existing process modeling languages described in section 2. It is more efficient to use established de facto standards instead of introducing another new process modeling language, which then may be perfectly suitable for aligning process models with ontologies, but which has no acceptance and no support in existing management tools.

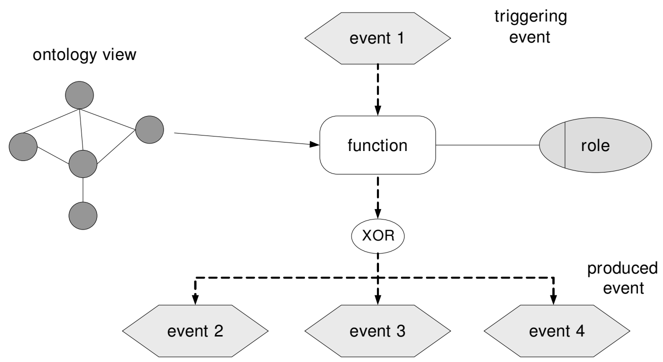

Fig. 4. Example connection EPC with Views

Figure 4 shows a connection between processes and ontology views. In this figure the modeling language EPC (Event-driven Process Chains) is connected with a certain ontology view for a certain process step (in EPC this is modeled by using "functions"). EPC is an important aspect of the ARIS model and connects all other views and describes the dynamics of the business process. Therefore it is possible to identify a user's role, the triggering event of the process step and the generated events. If the user acts in this specific role for this specific function the querying mechanisms are based on the given ontology view.

5. Conclusion and Further Research

In this paper a new approach to integrate ontology-based information systems with business process engineering is introduced. To connect and integrate both areas we use process oriented views on ontologies. Within our project we want to identify similarities and differences between ontology and process modeling techniques. We therefore analyze existing process modeling techniques such as EPC, UML 2.0, etc. and compare their possibilities to connect the various viewpoints on business models (information model, resources, management and workflow). Both techniques for process modeling and ontology modeling are then implemented in a prototype using existing tools (e.g. Protégé). In a later phase of this project we try to integrate the various used resources and the responsibilities a user has within a process.

One of our visions is to apply POVOO in a Grid Computing middleware layer, which integrates the underlying information resources and workflows based on grid computing technology and semantic mediation. GRID networks are characterized by a huge number of knowledge bases containing semantically related information that's DBMS make high demands on the used IR techniques. During the Austrian Grid Project [32] the G-SDAM (Grid Seamless Data Access Middleware) prototype is developed. G-SDAM enables electronic data interchange between various distributed and heavily heterogeneous information sources using semantic mediation techniques and allows (authorized parties) to seamlessly bind those information sources for querying and processing data in grid environments. G-SDAM processes queries over multiple data sources and translates data accordingly by applying domain ontologies.

References

1. M. Klein, D. Fensel, A. Kiryakov, D. Ognyanov, Ontology, Versioning and Change Detection on the Web, in 13th International Conference on Knowledge Engineering and Knowledge Management (EKAW02), page 197 ff, Siguenza, Spain, October 1 - 4, 2002
2. E. Mönch, SemanticMinerTM: Ein integratives Ontologie-basiertes Knowledge Retrieval System, Workshop Ontologie-basiertes Wissensmanagement, 2. Konferenz Professionelles Wissensmanagement - Erfahrungen und Visionen, Schweiz, April 2003
3. Workflow Management Coalition, Terminology & Glossary, Document Number WFMC-TC-1011, Document Status - Issue 3.0, Feb 99, http://www.wfmc.org/standards/docs/TC-1011_term_glossary_v3.pdf
4. C.A. Petri, Kommunikation mit Automaten, Ph.D. thesis, Institut für Instrumentelle Mathematik, Bonn, 1962
5. W. Brauer, W. Reisig, Petri Nets: Central Models and Their Properties, Lecture Notes in Computer Science, Vol. 254: Advances in Petri Nets 1986, Part I, proceedings of an Advanced Course, Bad Honnef, Springer-Verlag, September 1986
6. Unified Modeling Language (UML), Object Management Group, http://www.uml.org/
7. P. Bichler, G. Preuner, and M. Schrefl, Workflow Transparency, in Advanced Information Systems Engineering, in Proceedings of the 9th International Conference (CAiSE '97), Barcelona, Spain, pp. 423 – 436, Springer Verlag, June 1997
8. J. Staud, Geschäftsprozessanalyse: Ereignisgesteuerte Prozessketten und objektorientierte Geschäftsprozessmodellierung für Betriebswirtschaftliche Standardsoftware, Berlin, Springer, 2001, ISBN 3-540-41461-4
9. Business Process Modeling Language (BPML), Business Process Management Initiative, http://www.bpmi.org/BPML.htm
10. T. R Gruber, A Translation Approach to Portable Ontology Specifications, in Knowledge Acquisition, Vol. 5(2), pp. 199-220, Academic Press, Stanford University, USA, 1993
11. V. R. Benjamins, L. Nunes de Barros, A. Valente, Constructing Planners through Problem-Solving Methods, in Proceedings of KAW'96 (Banff), pp. 14.1-14.20, 1996
12. A. Tate, Roots of SPAR - Shared Planning and Activity Representation, The Knowledge Engineering Review, Vol. 13(1), pp. 121-128, Special Issue on "Putting Ontologies to Use" (eds. Uschold, M. and Tate, A.), Cambridge University Press, 1998.
13. A. Pease, The Warplan: A Method Independent Plan Schema, in L. de Barros, R. Benjamins, Y. Shahar, A. Tate, A. Valente (edts), in Proceedings of the AIPS 1998, Workshop on Knowledge Engineering and Acquisition for Planning, AAAI Technical, Report WS-98-03, 1998
14. N. Guarino, C.Masolo, G. Vetere, Ontoseek: Content-based Access to the Web, IEEE Intelligent Systems, Vol. 14 (3), pp. 70 – 80, May 1999
15. L. Khan, Ontology-based Information Selection, PhD thesis, Department of Computer Science, University of Southern California, 2000
16. A.S. Pollit, The Key Role of Classification and Indexing in View-Based Searching, technical report, University of Huddersfield, UK 1998
17. E. Mäkelä, Hyvönen, T. Sidoroff, View-Based User Interfaces for Information Retrieval on the Semantic Web, in Proceedings of the ISWC-2005 Workshop End User Semantic Web Interaction, November, 2005
18. A. S. Pollitt, G. P. Ellis Martin P. Smith, HIBROWSE for Bibliographic Database, Journal of Information Science Vol. 20 (6), pp. 413 – 426, 1994, ISSN:0165-5515

19. E. Hyvönen, S. Saarela, K. Viljanen: Ontogator, Combining View- and Ontology Based Search with Semantic Browsing, in Proceedings of the International SEPIA Conference, Helsinki, September 18-20, 2003
20. XML Topic Maps (XTM) 1.0, TopicMaps.Org Specification, http://www.topicmaps.org
21. T. Luckeneder, K. Steiner, W. Wöß, Integration of Topic Maps and Databases: Towards Efficient Knowledge Representation and Directory Services, in lecture notes in Computer Science 2113, DEXA 2001 – 12th International Conference on Database and Expert Systems Applications, Springer-Verlag, pp. 744-753, München 2001
22. A. Magkanaraki, V. Tannen, V. Christophides, D. Plexousakis, Viewing the Semantic Web through RVL Lenses, in Proceedings of the 2nd International Semantic Web Conf., Sanibel Island, USA, pp. 96-112, 2003
23. R. Volz, D. Oberle, R. Studer, Views for Light-Weight Web Ontologies, in Proceedings of the ACM Symposium on Applied Computing, NY, USA, pp. 1168-1173, 2003
24. Unified Medical Language System, http://www.nlm.nih.gov/research/umls/
25. Structured Query Language, http://www.sql.org/
26. D. Tsichritzis, A. Klug, The ansi/x3/sparc/dbms Framework Report of the Study Group on Database Management Systems, in Information Systems Vol 3 (3), pp. 173 – 191, 1978
27. RDQL - A Query Language for RDF, W3C Member Submission 9 January 2004, http://www.w3.org/Submission/2004/SUBM-RDQL-20040109/
28. R. Fikes, P. Hayes, I. Horrocks, OWL-QL - A Language for Deductive Query Answering on the Semantic Web, Knowledge Systems Laboratory, Stanford University, CA, 2003
29. A. W. Scheer, Architektur integrierter Informationssysteme, Springer, Berlin/Heidelberg; 1992; ISBN 3540554017
30. A. Oberweis, R. Schätzle, W. Stucky, W. Weitz, G. Zimmermann, INCOME/WF – A Petri Net Based Approach to Workflow Management, in: H. Krallmann (Edt): Wirtschaftsinformatik '97, Springer-Verlag, pp. 557-580, 1997
31. A. Abecker, G. Metzas, M. Legal, S. Ntioudis, G. Papavassiliou, Business-Process Oriented Delivery of Knowledge through Domain Ontologies, in Proceedings of 2nd International Workshop on Theory and Applications of Knowledge Management, Munich, 3-7 September 2001
32. http://www.austriangrid.at/