ONTOLOGY CONSTRUCTION IN PRACTICE
Experiences and Recommendations from Industrial Cases

Kurt Sandkuhl, Annika Öhgren
School of Engineering at Jönköping University, P.O. Box 1026, 55111 Jönköping, Sweden

Alexander Smirnov, Nikolay Shilov, Alexey Kashevnik
SPIIRAS, 39, 14 line, 199178 St.Petersburg, Russia

Keywords: Ontology construction, ontology engineering, experience report.

Abstract: Significant progress in ontology engineering during the last decade resulted in a growing interest in using ontologies for industrial applications. Based on case studies from different industrial domains, this paper presents experiences from ontology development and gives recommendations for industrial ontology construction projects. The recommendations include (1) using defined roles in a matrix project organisation, (2) perspectives on generalisation/specialisation strategy and ontology lifecycle phases, and (3) aspects of user participation in ontology construction.

1 INTRODUCTION

During the last decade, significant progress was made in development methods, engineering tools and application environments of ontologies, which resulted in a growing interest from industry in applying ontologies for solving various industrial problems. Integration with other technology areas, like knowledge management (Nanoka and Takeuchi 1995), enterprise modeling (Vernadat, 1996), or information systems (Scheer, 1992) is another reason for the increasing use of semantic technologies in industry.

This paper presents experiences and recommendations for ontology construction projects with focus on industrial application contexts. The research approach is conceptual and argumentative based on a number of case studies that were carried out in industrial enterprises from different domains.

In two of these cases, recent trends in the worldwide economy to implement new production and marketing paradigms were clearly visible and among the reasons for starting the ontology construction projects. These trends indicate major shifts of the knowledge-dominated economy: (i) shift from "capital-intensive business environment" to "intelligence-intensive business environment" – an "e" mindset – and (ii) shift from "product push" strategies to a "consumer pull" management – mass customisation approach. (Smirnov et al., 2002). All changes required for implementation of mass customisation approach are connected to changes in information factors: production systems need more knowledge about customers and customers need more knowledge about products (Caddy, 2000).

The remainder of the paper is structured as follows: After a brief description of important concepts and methods for ontology construction (chapter 2), three industrial cases of ontology development from different application domains are introduced (chapter 3). Experiences and recommendations are presented and discussed (chapter 4) related to project organisation, development strategy and user participation. Conclusions and future work are presented in chapter 5.

2 ONTOLOGY CONSTRUCTION

Ontology construction has been subject of work in numerous research activities. This chapter will briefly present two aspects of ontology construction important for the following part of the paper: ontology representation (2.1) and ontology development methods (2.2).

Sandkuhl K., Öhgren A., Smirnov A., Shilov N. and Kashevnik A. (2007).
ONTOLOGY CONSTRUCTION IN PRACTICE - Experiences and Recommendations from Industrial Cases.
In *Proceedings of the Ninth International Conference on Enterprise Information Systems - ISAS*, pages 250-256
DOI: 10.5220/0002384002500256

From the many different definitions of the term "ontology" in computer science related research, Gruber's proposal will be used in this paper:

"An Ontology is a formal, explicit specification of shared conceptualization." (Gruber, 1993).

2.1 Ontology Representation

For case 1 and 3 presented in chapter 3, we use Protégé frames or the W3C recommendation ontology language OWL (Web Ontology language) to represent the ontology. An OWL ontology consists of Individuals, Properties and Classes.

In case 2 an object-oriented constraint network paradigm (Smirnov et al., 2006) is applied, which supports capturing of constraints in a more sophisticated way. The OOCN-based ontology consists of Classes, Class Attributes (Values), Value Domains and Constraints. The Class Instances (Individuals) are stored separately from the ontology. The OOCN representation can be transformed to OWL and vice verse.

2.2 Ontology Development Methods

Previously, we have evaluated existing manual methods for ontology construction, see (Öhgren and Sandkuhl, 2005). The aim was to find a methodology which fits to the requirements in small-scale application contexts (shorten development time; minimize need for ontology expert; etc.). Since none of the existing methodologies fulfilled the requirements, an enhanced methodology was proposed. Relevant parts from different methodologies have been combined into a new methodology, together with some other ideas. The enhanced methodology consists of four different phases: requirements analysis, building, implementation, and evaluation and maintenance.

The first phase is the requirements analysis, in which all the formalities are specified, such as users and uses of the ontology, purpose, knowledge sources, etc. This should also include usage scenarios, competency questions and answers. Also, in order to shorten the development time, integration of already existing ontologies should be elaborated already at this stage.

The building phase uses a middle-out approach and is iterative. The previously defined knowledge sources should be used and each term should also be described in natural language. It is natural in this phase to include domain expert evaluation, since the domain experts probably are involved in this phase.

The implementation phase consists of implementing the ontology in an appropriate ontology editor tool, which also includes choosing the most suitable ontology representation.

Finally the ontology needs to be evaluated and tested according to the requirements stated in the requirements analysis. The ontology should also be evaluated according to criteria such as clarity, consistency, and reusability.

3 INDUSTRIAL CASES

This chapter introduces the industrial cases forming the basis for discussion of experiences in chapter 4. When selecting these cases, the objective was to achieve a wide heterogeneity regarding

- The type of project including research project, applied research and contract development,
- The application domain, which here encompasses automotive industry, media industry and industrial automation
- The purpose of the ontology developed including information structuring, model integration and product codification.

Besides the above cases, experiences from a number of other ontology projects contributed to this paper, like (Smirnov, 2001), (Billig and Sandkuhl, 2002), (Smirnov et al, 2006), (Tarasov, 2006).

3.1 Autoliv Electronics

The first industrial case was taken from automotive industries. Automotive manufacturers and suppliers have to manage a large number of product variations and their integration into a specific car model. In order to manage and control variety, manufacturers and suppliers increasingly recognize the need to manage project entities like models, documents, metadata, and classification taxonomies in such a manner that the integrative usage of these entities is supported.

The application scenario for the ontology developed is integration of different kinds of structures reflecting the artefacts and their interrelations. On the one hand, model hierarchies have to be captured, indicated and implemented by different modelling levels (system, software, hardware, etc.), which furthermore will have model instances (artifacts) to be managed. On the other hand, term networks and taxonomies have to be considered as equally important. These networks represent organizational structures, product structures or taxonomies originating from customers that are closely related to artifacts. Explicit denotation of these relationships are considered beneficial for identification of reuse potential of components or artifacts.

The ontology construction was performed in a Swedish automotive supplier of software-intensive systems. The development process applied is an enhanced version of the METHONTOLOGY process as described in 2.2. Most important knowledge sources were (1) a description of the suppliers internal software development process with defined procedures for all major aspects of software development and software project management and (2) documentation of two example cases for requirement handling, including original customer requirements, system and functional requirements, and (3) interviews and working sessions with members of the software development department were conducted including project manager, software developers and engineers.

The resulting ontology consisted of 379 concepts and with an average depth of inheritance of 3,5.

3.2 Festo

The second industrial case originates from industrial and process automation. For companies with wide assortments of products (more than 30 000 – 40 000 products of approx. 700 types, with various configuration possibilities) it is very important to ensure that customers can easily navigate among them. One solution is to provide a codification system that can produce easily recognizable and at the same time relatively short codes. In this section an overview of the ontology-based approach to designing product codes is presented. The approach has been implemented at the industrial company Festo that has more than 300.000 customers in 176 countries. Festo has more than 52 subsidiaries worldwide with more than 250 branch offices and authorised agencies in further 36 countries. A detailed description of the approach can be found in (Oroszi et al., 2006).

The developed approach is based on the idea that knowledge can be represented by two levels. The first level describes the structure of knowledge. Knowledge represented by the second level is an instantiation of the first level knowledge; this knowledge holds object instances. The knowledge of the first level (structural knowledge) is described by a central ontology of the company's product families (classes). In this particular case the entities are product families. Usage of product families enables defining product platforms that can be reused across whole families of similar products.

In Festo case the goal was to build a problem-oriented ontology for the given, specific purpose. Hence it was more reasonable to build a new ontology using the formalism that met the requirements than to try to adapt other existing ontology models like CYC or SENSUS. The creation of the ontology described above was done automatically based on existing documents and defined rules of the model building. The resulting ontology consists of more than 1000 classes organized into a 4 level taxonomy, which is based on the VDMA classification.

Taxonomical relationships support inheritance that makes it possible to define more common attributes for higher level classes and inherit them for lower level subclasses. The same taxonomy is used in the company's PDM and ERP systems.

For each product family (class) a set of properties (attributes) is defined, and for each property its possible values and their codes are defined as well. The lexicon of properties is multilingual and ontology-wide, and as a result the values can be reused for different families. Application of the central single ontology provides for the consistency of the product codes and makes it possible to instantly reflect incorporated changes in the codes.

3.3 Jönköpings-Posten

The third application case was taken from media industries and is focusing on a local newspaper specialising on news from the region of Jönköping in Sweden. In order to meet the demands of their readers, the newspaper aims at providing not only news and background information about events, activities, politics or businesses in the region, but also if local organisations or individuals are involved in events outside the region, like a local politician having a car accident in Stockholm or the ice hockey team finishing on 2nd place at a tournament abroad.

Various information sources are used for detecting relevant news, like different news agencies, archives, own reporters or e-mails containing hints from readers. The intended use of the enterprise ontology developed is to support evaluating, ranking and assigning the incoming information to the reporter in charge for the respective subject area. (Uschold and King, 1995) name several intended uses for an enterprise ontology, including being “communication medium” between different people, people and computational systems, and different computational systems. From these intended uses, the enterprise ontology developed for the newspaper primarily aims at supporting communication between people. Furthermore, Uschold and Kings recommendation to use enterprise ontologies for integrating and relating available information is also an important motivation for constructing the ontology.

The development process followed the description in section 2.2 and started by a number of interviews at Jönköpings-Posten to get a feeling of which were the most important terms and what should be the focus of the ontology. In the initial stage it was decided that the ontology should capture regional and organisational aspects, and different terms and concepts related to sports. Old issues of the newspaper were analysed in order to try to extract and find important terms and concepts and the relationships between them.

The resulting ontology has 457 concepts, but has not yet been evaluated thoroughly, and might be subject for further revisions.

4 EXPERIENCES

Experiences and recommendations presented in this chapter were based on the industrial cases introduced in chapter 3, findings from other research and development projects applying ontologies and teaching ontology construction in university courses.

4.1 Project Organisation

General project management principles, like those introduced in (Sommerville, 2004, chapter 5) should also be applied for ontology development projects, i.e. there should be clear and unambiguous description of the purpose of the ontology to be developed, a project plan including resource allocation, milestones and deliverables should include the complete development process, quality control should be established independently from the project management and be guided by a quality plan, and resource allocation for the project should be adequate. In addition to these recommendations, which are valid for most project types, focus of this section will be on project organisation for ontology development projects.

The critical resources during the development of ontologies for industrial purposes are from our experience the domain experts in the companies. Typical situation in companies is that these experts are involved in several projects or activities and that agreeing on meetings or workshops with them is a difficult endeavour. An appropriate resource allocation to the overall project solves this problem only partly because it still is depending on the other activities running which priority the ontology development will be given by the superior of the domain expert. Our recommendation is to additionally establish a matrix organisation for the ontology development project, i.e. for the runtime of the project an organisation unit should be created reflecting the project. For the required share of their work the domain experts should then be released from their position in the line organisation and moved to the projects' organisation unit, which gives the project manager in the company better control on how to use their time. (Jenny, 1997) discuss additional aspects of matrix and line organization in IT projects.

In the Autoliv case, the matrix organisation was not established. Although both the company and the department in question committed to support the project, the availability of the domain experts was not as expected due to other activities with higher priority. The Festo case had a dedicated project team following the matrix organisation principle resulting in a good availability of the resources. This difference between the projects was to some extent of course due to the fact that one project was applied research not financed by the company (Autoliv) whereas the other case was contract research with Festo as the paying customer.

Furthermore, the project team should include several roles:

- Process owner: the owner of the development project who is responsible for establishing the project in the company, assigning the right personnel resources, arranging meetings, etc.
- Planner: the person responsible for proposing the way of working and establishing a consensus between all participants, coordinator of different tasks, moderator of meetings, etc.
- Method expert: provides expert knowledge in ontology development process and method to the project
- Facilitator: is experienced in using the selected ontology management tool and facilitates ontology construction
- Domain expert: provide knowledge about the domain under consideration, which is basis for ontology development

Persons assigned to these roles from company side, i.e. project manager and domain experts, should be members of the projects organisation unit.

4.2 Development Process

The development process as such is governed by the method applied (see chapter 2). Experiences with these methods have been published earlier (see 2.2).

Additional experiences contributed by this paper concern the identification of relevant concepts, relation and properties or constraints. One aspect to discuss is whether to work top-down, bottom-up or middle-out. Our impression from ontology development projects indicates that experience from enterprise modelling (e.g. (Vernadat, 1996)) concerning these strategies can be applied as rule of thumb for ontology projects:

Top-down approaches should be used in application domain well-known to the project team where the complexity in terms of required level of detail and the scope of the development is clearly defined. An example from our background would be the Festo case, where the existing codification system, number of products and potential variation limited the complexity of problem at hand. In cases with unclear or unknown complexity, there is the danger of consuming the resources allocated to the project before reaching the goal of having developed an ontology.

Also in projects with bottom-up development strategy, the danger of exceeding the given time and budget of a project is existing, if required granularity of the ontology and complexity of the overall area are not clear. Bottom-up can result in a number of thoroughly defined parts of an ontology, which are not very well interlinked and do not cover the intended scope of the ontology. These “solution islands” often contain more details than required for the purpose of the ontology. Our recommendation is to always test suitability of the bottom-up approach by using it in a pre-study with limited scope and clearly defined evaluation criteria.

The middle-out approaches is from our experience suitable to explore both, complexity of the problem at hand and required level of detail, in application fields unknown to the ontology expert. At the same time the middle-out way in most cases creates results that can be used for the final ontology. The middle-out approach was for example used in the Jönköpings-Posten case in order to capture sports related concepts in combination with regional concepts. What level of detail of regional information was needed in order to describe the sports selection in this region in sufficient detail became only clear during the ontology development process.

In addition to this generalisation/specialisation strategy, we recommend to also have different lifecycle phases of an ontology in mind during the development process, like

- The conceptual stage where the main elements, structures, relations and constraints of an ontology are identified based on the knowledge of the domain experts and other knowledge sources. This stage should be independent from the actual ontology representation or ontology engineering tool to be used in order to avoid unnecessary dependencies from implementation technology
- The implementation stage coding the result from the conceptual stage in appropriate representation with a suitable tool. Separation of conceptual and implementation stage allows for selection of the implementation technology based on the lessons learned from the conceptual stage
- The application stage concerning the pruning and optimisation of the implementation for the application purpose, which for example can include additional instances or axioms for consistency purposes.

To distinguish between these different phases is part of several methods, like METHONTOLGY or the method proposed by Staab et al. (2001), who use feasibility study/ontology-kickoff and refinement instead of conceptual stage and implementation stage. However, many other methods do not make this strict distinction, which is why this paper emphasizes the importance of clear separation.

In the Autoliv case, the decision to use Protégé was made quite early in the development process, i.e. the conceptual and the implementation stage were not clearly separated. As a consequence, the enterprise ontology developed and represented in Protégé did require a number of compromises and work-arounds in order to represent feature models in this ontology. Feature models to a substantial part include “subsumes” relations between the features, which could not adequately be captured by the “is a” relations in Protégé. The users of the ontology were not satisfied with the solution and demanded a rework. Separation between conceptual and implementation stage would have avoided the unnecessary iteration.

In the Festo case, the application stage is of particular importance. Since the company continuously introduces new products, the built ontology has to be modified accordingly. For this purpose the company experts were given a tool that they successfully use. The tool is domain oriented: the experts do not even necessarily know that they work with the ontology.

4.3 User Participation

Since more than a decade, participative modelling is recognised as valuable and practicable instrument contributing to solving design problems in particular in organisational contexts (see e.g. (Nilsson et al, 1999)). As opposed to the traditional approach of gathering facts by interviewing stakeholders in an organisation and afterwards developing a solution without stakeholder involvement, the participative way of working includes development of the intended solution with direct involvement and contribution of the future users, like modelling in facilitated group sessions.

Experiences from ontology development projects like the cases presented in chapter 3 indicate the value of user participation even for ontology development projects. The main recommendations are to thoroughly prepare participation and to concentrate on the conceptual stage of development.

To prepare user participation should start with the key persons at the industrial company, e.g. the head of the organisation unit in charge and the process owner, who should be introduced to the potentials and limits of ontologies. In many organisations there exists the myth of ontologies as the silver bullet for all knowledge representation problems; in other they are considered as yet another modelling technique. By clearly defining purpose of the project, intended use of ontologies and known limits, the expectation of the industrial partner should be adjusted to realistic possibilities. This should preferably happen before the project starts.

After sufficient management information and attention, the intended participative steps of the ontology development should be prepared by individual discussions with the participants. Each participant should be informed about the purpose of the ontology development project and the intended way of working. However, main purpose of these individual discussions is to start identifying existing knowledge sources in the organisation relevant for the ontology development, to build up trust to the participating users, and to increase their commitment to the project.

During the participative parts of the ontology construction, focus should be on the conceptual stage of the ontology development and on use of techniques like card sorting or pencil and paper sketches. Main reason for this is to not put the burden of learning and understanding the formalities of an ontology language or the functionality of an ontology engineering tool on the domain experts and end users participating. A notation that everyone understands should be used, otherwise to too much attention is lost when the participants try to understand the notation used.

Furthermore, the role distribution proposed in section 4.1 should also be observed during the participative part. In particular, the ontology expert and the facilitator should lead the overall process based on the selected construction methods. Too much focus on the method or even training the participants in method knowledge will based on our experience distract the participants from solving the problem at hand. The facilitator should also make sure that the process owner does not dominate the sessions. An important point of participatory working is extending the view by including a wider group of stakeholders and giving them an equal possibility to contribute.

5 CONCLUSIONS

Based on the discussion of three industrial cases and on results from earlier ontology construction projects, this paper presented a number of experiences and recommendations for ontology construction in an industrial context. The recommendations can be summarizes as follows:

Project organisation:

- establish a matrix organisation for ontology construction projects
- Use the roles process owner, planner, facilitator, method expert and domain expert

Development process

- Use top-down development for well-known domains with clearly defined scope
- Use middle out development for unknown application fields
- Always perform a pre-study before deciding to use bottom-up development
- Clearly separate between conceptual, implementation and application stage

User participation

- Create realistic expectations on company management side by introducing potentials and limits of ontology use
- schedule individual discussions with every participant to prepare the participative parts
- use participative work mainly in the conceptual stage, i.e. avoid details of ontology representation
- do not train users in development method

Although the work presented in this paper is based on quite a few industrial projects, the main limitation of the research is that the empirical grounding should be improved by an increased number of cases. The recommendations presented are considered useful, but they can not be expected accurate for all industrial cases.

In the Festo case, the main limitation is that the ontology built and the methodology used are not universal but oriented to the particular problem. However, due to usage of the OOCN formalism they still can be used for different purposes, e.g., those mentioned in (Oroszi et al., 2006). Besides, compatibility between OOCN and OWL makes it possible to use the developed ontology in other company projects.

Future work will include further elaboration on the recommendations. We consider it interesting and important to investigate, whether and how the different application domains, development purposes or project types affect project organisation, development process and user participation, i.e. is it possible to recommend – based on extended empirical grounding – project organisation and development process for enterprise ontology development in for example automotive industry?

Furthermore, there are a number of experiences from industrial projects not discussed in this paper because they were just based on a single case, like the integration of ontologies with existing IT-systems in the companies under consideration.

ACKNOWLEDGEMENTS

Some parts of the research presented here were done under the joint project "New Order Code" between Festo AG & Co. KG and SPIIRAS. Elements of the ontology-driven architecture were developed under project # 16.2.35 of the research program "Mathematical Modelling and Intelligent Systems", project # 1.9 of the research program “Fundamental Basics of Information Technologies and Computer Systems” of the Russian Academy of Sciences, and project funded by grant # 05-01-00151 of the Russian Foundation for Basic Research.

Other parts of this work were financed by the Swedish Knowledge Foundation (KK-Stiftelsen), grant 2003/0241, project “Semantic Structuring of Components in Model-Based Software Engineering for Dependable Systems” and by Hamrin Foundation (Hamrin Stiftelsen), project “Media Information Logistics”.

Cooperation between SPIIRAS and Jönköping University was supported by the Swedish Foundation for International Cooperation in Research and Higher Education under grant # IG2003-2040.

REFERENCES

Billig, A. and Sandkuhl, K., 2002. *Match-Making based on Semantic Nets: The XML-based BaSeWeP Approach.* Proceedings XSW 2002, Springer Verlag.

Caddy, I., 2000. *Moving from Mass Production to Mass Customization: the Impact on Integrated Supply Chains*, Proceedings of the Workshop on Mass Customization Management (MCM 2000), University of Wollongong, Australia, 2000. (Electronic Proceedings).

Gruber, T., 1993. *A translation approach to portable ontology specifications*. Knowledge Acquisition 5: 199-220.

Jenny, B., 1997. *Projektmanagement in der Wirtschaftsinformatik.* 2. Edition, 1997, vdf Verlag ETH Zurich, ISBN 3728123552.

Nilsson, A., Tolis, C. and Nellborn, C. (Eds.), 1999. *Perspectives on Business Modelling: Understanding and Changing Organisations*, Springer-Verlag.

Nonaka, I., Takeuchi, I., 1995. *The Knowledge Creating Company.* Oxford University Press.

Öhgren, A. and Sandkuhl, K. 2005. *Towards a Methodology for Ontology Development in Small and Medium-Sized Enterprises.* In IADIS Conference on Applied Computing, Algarve, Portugal.

Oroszi A., Jung T., Smirnov A., Shilov N., Kashevnik A., 2006. Ontology-Driven Codification Based on Product Families. Blecker Th., Friedrich G., Hvam L., Edwards K. (Ed.): *Customer Interaction and Customer Integration*. Series on Business Informatics and Application Systems, Vol. 2, Gito Verlag, Berlin 2006, pp. 383-396.

Scheer, A., 1992. *Architecture of Integrated Information Systems: Foundations of Enterprise Modelling.* Springer-Verlag New York.

Smirnov, A, 2001. *Rapid Knowledge Fusion into the Scalable Infosphere: A Concept and Possible Manufacturing Applications.* Proceedings of the International NAISO Congress on Information Science Innovations. U.A.E., Dubai, March 17-20.

Smirnov, A., Pashkin, M., Levashova, T., Chilov, N., 2006. *Ontology-based support for semantic interoperability between SCM and PLM.* International Journal of Product Lifecycle Management, Inderscience Enterprises Ltd., Vol. 1, No. 3, 2006, 289-302.

Smirnov, A., Pashkin, M., Chilov, N., Levashova T., 2002. *Knowledge Fusion in the Business Information Environment for e-Manufacturing Pursuing Mass Customisation*, Moving into Mass Customization. Information Systems and Management Principles. (eds. C. Rautenstrauch, R. Seelmann-Eggebert, K. Turowski), Berlin: Springer, 2002. pp. 153 – 175.

Sommerville, I., 2004. *Software Engineering*. 7th Edition. Addison-Wesley, 2004.

Staab S., Studer R., Schnurr H.-P., Sure Y.: „Knowledge Processes and Ontologies.” IEEE Intelligent Systems, 1094-7167/01, 2001.

Tarassov, V., Sandkuhl, K., Henoch, B., 2006. *Using Ontologies for Representation of Individual and Enterprise Competence Models.* IEEE RIVF 2006 conference, Vietnam, February 2006.

Uschold M., King M., 1995. *Towards a Methodology for building Ontologies.* Workshop on Basic Ontological Issues in Knowledge Sharing. International Joint Conference on Artificial Intelligence.

Vernadat, F., 1996. *Enterprise Modelling and Integration.* Chapman & Hall.