Ontology Engineering Methodology for Intelligent System for Global Monitoring, Detection and Identification of Threats[1]

Kamil Gleba, Joanna Śliwa

Military Communication Institute, ul. Warszawska 22A, Zegrze, Poland

Wojciech Chmiel, Piotr Szwed, Andrzej Głowacz

AGH University of Science and Technology, Kraków, Poland

Abstract: The article describes ontology engineering methodology designed for building ontology for Intelligent System for Global Monitoring Detection and Identification of Threats. It is based on the software development methodology and has been developed after the analysis of a few well known ontology engineering methodologies. The article discusses several approaches for building ontologies for complex systems and finally proposes a methodology for building ontology within the INSIGMA project – IOEM (INSIGMA Ontology Engineering Methodology).
Keywords: methodologies for building ontology, ontology engineering methodology, approaches for building ontologies for complex systems, ontology requirements, ontology life cycle

1. Introduction

Intelligent System for Global Monitoring Detection and Identification of Threats will be the result of the INSIGMA project carried out by four Polish academic, research and commercial bodies (AGH – the consortium leader, MCI, WAT and WSTKT). The project is supported by the funds from European Ministry of Regional Development within one of the Polish National Strategic Reference Frameworks – Innovative Economy. This programme is directed mostly to all entrepreneurs who want to implement innovative projects connected with research and development, modern technologies, investments of high importance for the economy or implementation and the use of information and communication technologies.

The objective of the INSIGMA project is to develop and implement heterogeneous information system for complex detection, identification of threats, monitoring

and identification of mobile objects. The project is to propose innovatory solutions in 5 areas related to:

- Monitoring and identification of vehicles and people,
- Dynamic monitoring and identification of threats within the traffic,
- Analysis of traffic in terms of optimizing routes for external users,
- Identification of suspicious people and threats,
- Discovery of data and multimedia with the watermarking technology.

The general idea is to develop a complex monitoring system that will allow identifying objects in the monitored environment and, based on the stored information and advanced algorithms, identifying threats related to both the traffic and suspicious behaviour of people. The system can be also used for traffic management and route planning for individual users and for the public safety services. The route planning will also take into account complex parameters that provide the possibility of selecting the route in special circumstances, e.g. after a road accident or a natural disaster, in difficult weather conditions, etc.

The INSIGMA system will store and process large amounts of data of various types. Some of them will be raw data flowing from the sensors, some will have to be fused and processed in order to provide additional information. Moreover, the identification of threats requires methods for automatic recognition and classification of events in the system. One of the main tasks within the INSIGMA project is thus the definition of ontology that would help to manage the information and, at the same time, would help in automatic identification and classification tasks.

2. The role of ontologies in INSIGMA system

One of the key assumptions about the INSIGMA system is that it will be developed as an ontology driven information system. The role of ontologies for such systems can be classified according to two dimensions: temporal and structural. [1] The temporal dimension is related to a stage in the software lifecycle: *development time* or *run time*. The structural dimension concerns the usage of ontologies in particular components: databases, user interface and business logic layer.

Ontologies developed within the INSIGMA project fall into two categories: domain ontologies, specifying concepts of a particular domain (e.g. routes, vehicles, elements of a monitoring infrastructure, threats, weather) and task ontologies related to tasks, activities, processes (e.g. threats detection, services configuration, route planning). A particular software component can use both types of closely related domain and task ontologies.

Ontologies are used during the development of the INSIGMA system in the following areas corresponding to RUP (Rational Unified Process) [2,3] disciplines: Business Modelling, Requirements Specification and Design.

Domain ontologies equivalent to RUP business object model are primary artefacts of the Business Modelling. Due to the size of models (containing several

hundred classes linked by relations), the authors have decided to model it in OWL (*Web Ontology Language*) language [4] instead of depicting it on UML (*Unified Modelling Language*) diagrams [5] that rarely provide enough consistency and precision for such large specifications.

The requirements are usually documented in the form of use cases. The specification is mapped onto task ontologies specifying actors and activities in which they participate. These ontologies are planned to be used for system configuration at the deployment stage.

The ontology engineering process described in the next section is in many aspects similar to the process of object oriented analysis [6] and requirements management [7]. It is a natural consequence of the fact that UML specifications and ontological models play similar roles in the software development lifecycle.

The INSIGMA system will comprise several databases containing data necessary to support various system functionalities (e.g. dynamic maps, traffic parameters, monitoring records, event logs). The database structure is to be defined based on domain ontologies with the use of tools supporting automatic translation of OWL constructs into database schemas. Similar approach will be taken to design the static parts of interfaces between components.

Ontologies at the run time are used in the situations where the domain model cannot be fully elaborated during the system development (some domain aspects are unknown or uncertain) or a kind of reasoning is required. In case of the INSIGMA system main uses are: classification of threats and their properties and reasoning about incidents, dynamic configuration of the system architecture to fit particular needs of an end-user, providing semantic interoperability between components that cannot be defined during the design stage and integration with external systems.

3. Ontology engineering process

Ontology development is a difficult and complex task which includes many steps beginning with the most fundamental ones, such as definition of its purpose. This is very important because ontology development is usually dependent on the context and objective of the project it is designed for. For this reason the most popular and best known ontology engineering methodologies result from the project they have been developed in.

Thus, several approaches to the development process can be defined, e.g.:
- Induction – the ontology is developed as generalization of a particular case.
  - Generally low costs;
  - Close relation to the considered case;
  - Possibility of utilization in small domain ontologies.
- Deduction – the ontology is based on general, widely accepted rules and assumptions derived from the analysed domain.

- ▫ Can lead to the development of a very big ontology;
- ▫ Requires great amount of work on the analysis of the domain knowledge.
- – Inspiration – is based on individual approach of the engineer to the modelled domain.
  - ▫ Usually leads to quite interesting results, however, the ontology can be unaccepted by the domain experts;
  - ▫ Can lead to revolutionary changes in the way the domain knowledge is perceived.
- – Synthesis – the created ontology is a synthesis of a few existing ones.
  - ▫ Requires analysis of the existing ontology models and their background.
- – Cooperation – the ontology is created in a group work process and is incremental.
  - ▫ Uses experience of engineers and domain experts and their cooperation in gathering the domain knowledge;
  - ▫ Requires close cooperation of the task groups.

The above listed approaches sometimes are overlapping, which means that using one of them does not forbid utilization of another one. They show, however, general methods for building an ontology that must be further expanded to provide ontology engineering methodology that would describe it as a defined process comprising usually initial actions, gathering domain knowledge, terminology definition, model production as well as many management and maintenance actions.

4. Related work

The IEEE defines methodology as [8]: "*a comprehensive, integrated series of techniques or methods creating a general systems theory of how a class of thought-intensive work ought be performed*".

Methodologies for ontology engineering have been subject of research for a number of years. In general, ontology development depends on the context and the purpose of the particular project. Therefore, each project team is very often trying to build its own methodology. In order to develop a methodology for the INSIGMA system, we assumed the analysis of the most known existing approaches in that field as a starting point.

There are many different ontology engineering methodology proposals. We analysed typical methodologies used to build ontologies from scratch or by reusing other ontologies. In particular, the approaches dealt with were:
- • On-To-Knowledge Methodology [9],
- • Ontology development proposed by Noy and McGuinness [10],
- • Enterprise Ontology [11],
- • TOVE (TOronto Virtual Enterprise) [12],
- • HCONE (Human Centered ONtology) [13],

- UPON (Unified Process for ONtology building) [14],
- Holsapple [15],
- METHONTOLOGY [16].

Each methodology describes the ontology development process that identifies which activities are to be performed. This development process is called ontology life cycle. *"Life cycles identifies when the activities should be carried out, that is, it identifies the set of stages through which the ontology moves during its life time, describes what activities are to be performed in each stage and how the stages are related (relation of precedence, return, etc.)"* [17].

After analyzing the above methodologies it was possible to distinguish some common activities (processes) for all methodologies which form an ontology life cycle [16, 18]:

- Management process: consists of scheduling, control and quality assurance;
- Ontology development process: consists of environment study, feasibility study, specification, conceptualization, formalization, implementation, maintenance and use;
- Support process: consists of activities run in parallel to the ontology development process. It includes knowledge acquisition, evaluation, integration, documentation, merging, configuration and alignment.

Each methodology defines its individual approach to carrying out the complete ontology life cycle.

5. Requirements for INSIGMA ontology

Generally, it was assumed that the ontology for the INSIGMA system should be:

- explicit: an ontology should efficiently communicate intended meaning of defined terms. Definitions of terms should be objective and insusceptible to unintended interpretations. If it is possible, definitions should be expressed in formal axioms. Additionally, a complete definition is preferred over a partial definition. The richer and more formal the restriction placed on use of words in the ontology, the less likely an unintended interpretation is caused by a mistake;
- coherent: an ontology should be coherent and allow for drawing meaningful inferences that are consistent with definitions and axioms. Axioms should be logically coherent. Coherence should be related not only to formal axioms, but also to the terms informally defined. For instance, the ontology is incoherent, if there is a contradiction between a sentence inferred from the axioms and definition or example given informally;
- extensible: an ontology developing should take into account future extensions of created ontology without the need for revising definitions. New terms in ontology should be able to be defined based on the existing vocabulary. It should also be open to other existing ontologies.

6. INSIGMA ontology engineering methodology

The approach

The broad range of domains the knowledge of which is represented in INSIGMA makes it necessary for the IOEM to be based on a mix of ontology definition approaches. The deduction and cooperation approaches should be used for the domain knowledge representation and cooperation with domain experts. Without them it is impossible to provide useful target ontology. For domains that are not strictly related to INSIGMA-specific field like e.g. weather conditions and time, there is an opportunity to use existing ontologies. For this purpose it is necessary to involve the synthetic approach. The inspiration approach could lead to innovatory solutions, however, there is a risk that in case of complex project led by geographically distributed teams may also quickly direct to missing the objective of ontology engineering. It is though not required as a main approach.

The ontology development for such a complex system as INSIGMA may be similar to the development of a complex information system. Following this idea it seems well founded to use methodology based on the software development techniques. This process can be referred to the activities that are performed when building ontologies.

IOEM stages

The proposed process model for ontology building consists of the following stages:
- ontology vision – general description of a system architecture and ontology content. This stage should be concluded with creating approved vision of the ontology which will be used and defining ontology development milestones,
- ontology construction planning – definition of which steps and when should be done,
- validation and acceptance of created ontology plan and milestones,
- identification of people responsible for each task within ontology building process and reviewers,
- ontology engineering (includes documentation),
- ontology stabilization and acceptance by the end-users,
- ontology maintenance.

Additionally, the authors assumed that during the process of ontology engineering living documents will be used, providing the possibility of sharing information about the developed ontology and its version with engineers (domain experts) who should have the ability to modify it. Moreover, the process of ontology versions acceptance with all interested parties will be used during the ontology lifecycle.

After having analysed existing methodologies and approaches for building the ontology, the authors proposed the INSIGMA ontologies engineering methodology, which consists of 6 basis phases (see Figure 1):
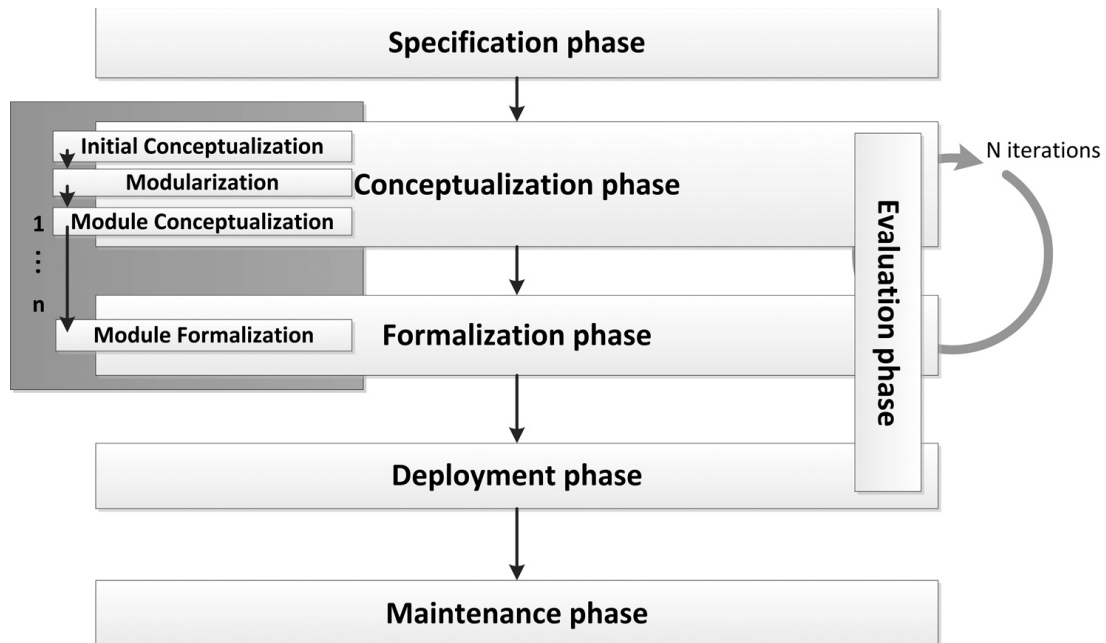

Figure 1. IOEM stages

a. Specification phase

Specification phase consists of the following actions:
– Determine the domain and scope of the ontology by answering several questions, i.e.:
  - What is the purpose of creating the ontology?
  - What is the domain of the developed ontology?
  - For what types of questions should the ontology provide answers?
  - Who will be the end-users of the ontology?
  - Who will be responsible for the ontology maintenance?
  - What are ontology use cases in the INSIGMA system?
– Provide competency questions and motivating scenarios including some use cases.

The questions that an ontology should be able to answer are called competency question. The competency questions, after the ontology is built, enable verification if the ontology is complete.

Motivating scenario describes requirements that the created ontology should satisfy expressed as use cases. They provide examples for ontology utilization in the system.
– Consider reusing existing ontologies.

Large number of ontologies have been developed and published. Using them (importing) into the INSIGMA ontology may accelerate the work and reduce costs.

The potential candidates should be evaluated to check their validity and possibility for refinement and extension for particular domains and task.

For the Specification phase, we proposed a questionnaire, which was sent to persons responsible for particular tasks in the project, who could be interested in application of ontologies in their areas (see Figure 2). The received answers were very useful in determining the role of ontology in INSIGMA project and its range in terms of necessary domain ontologies as well as possible use cases. It also prepared the foundation for the Conceptualization phase.

INNOWACYJNA GOSPODARKA NARODOWA STRATEGIA SPÓJNOŚCI	UNIA EUROPEJSKA EUROPEJSKI FUNDUSZ ROZWOJU REGIONALNEGO		
QUESTIONNAIRE No 1 **REQUIREMENTS FOR INSIGMA SYSTEM ONTOLOGY**			
Personal details	Name and surname: Tasks performed in INSIGMA: E-mail address: Phone number:		
PRELIMINARY REQUIREMENTS FOR INSIGMA ONTOLOGY			
No	**QUESTION**	**Yes**	**No**
1.	**Are you interested in using semantic descriptions (ontologies) in the task you perform?**		
2.	**Are there some existing ontologies you want to use?**		
3.	**Do you find it necessary to use the following ontologies?**		
	Events ontology road situation, threats on the road, threads on public safety, event notifications		
	Objects ontology (high and low-level attributes) vehicles, person, roads (roads classification and attributes, dangerous objects and substances, sounds, monitored devices, data being monitored		
	Dynamic map objects ontology Urban spacer, Road, places, cross-roads, etc.		
	Physical quantity ontology distance, location, duration, temperature, air-pressure, frequency, concentration		
	Security ontology		
	System user profile ontology		
	System users and user privileges ontology		
	Ontology of reaction to events detected by the system		
	Other:		
	Weather ontology		
	Human behaviour ontology		
	Biometric characteristics ontology		
	Interaction between objects ontology		
	System services ontology		

No	QUESTION	Yes	No
	INSIGMA system end-users ontology (police, fire brigade, road service, medical service)		
	Ontology for semantic search and data processing support		
	System and data protection ontology		
4.	**Do you think it is worth using other ontologies than specified above?**		
	Other ontologies:		
	CHARACTERISTICS OF DELIVERABLE FOR A GIVEN TASK		
	What information will be used by the created INSIGMA component?		
	What information will be stored in databases? Determine input and output data.		
	Are there any existing databases you have already used? If yes, please give a structure of these databases.		

List of possible usage scenarios:		
List of used objects, attributes and relations between them		
No.	**Object**	**Attributes**	**Relationships**
1.			
2.			
3.			

Figure 2. IOEM Questionnaire 1: Requirements for INSIGMA system ontology

b. Conceptualization phase

Conceptualization is an activity that should be carried out before the formalization process. It provides a domain ontology model in an informal language with terms and their relationships presented in ordered taxonomy. This process consists of the following actions:
- building a glossary of terms,
- building taxonomies of concepts,
- defining relations (between concepts and data types),
- building concept dictionary,
- defining axioms,
- defining rules,
- creating instances of particular classes.

The ontology for such a broad project as INSIGMA is large and complex. In that cases ontologies do not have a monolithic form, they are rather distributed into several modules (compound ontologies). Therefore, it is recommended to divide Conceptualization phase in three subphases: Initial Conceptualization, Modularization and Module Conceptualization.

The goal of the Initial Conceptualization phase is to describe the scope of the developed ontology by establishing a glossary of key concepts, their hierarchy and relations.

The main goal of Modularization phase is to reuse of existing components and logical distribution of the ontology building into a set of partially independent activities. Above-mentioned issues give an opportunity to parallel development of compound modules. This is particularly important for a team being distributed among several geographical locations (as in the case of the INSIGMA project).

The goal of Module Conceptualization phase is to provide ontology model in an informal language for particular modules defined in Modularization phases.

To support the Conceptualization phase the authors proposed the second questionnaire, which helps to build conceptual dictionary and to define rules that describe given domain.

c. Formalization phase

This phase transforms the conceptual model into the formal model. It includes building ontology model in an ontology language. In the case of IOEM, the Formalization phase will use XML-based semantic languages, i.e.: RDF and OWL. They have appropriate expressiveness, are easy to use and supported by available software development tools (e.g. Protégé, Jena library).

d. Deployment phase

Deployment is a phase that provides tools for applying ontology models to systems, software application, etc. Ontologies developed within the INSIGMA project (domain and task ontologies) will be used by software components to support different activities and processes, such as, for instance, threats detection and route planning. Currently, the most advanced tools and libraries have been developed for the Java platform, e.g. OWL API, Jena, Pellet, Jess (as a rule engine). Thus, components using ontologies in the INSIGMA system are usually coded in Java. In the situation where an integration with another platforms is needed, a wrapper in a form of web services can be used.

e. Evaluation phase

Evaluation is a process aimed at validation and verification of an ontology in terms of its scope, consistency and expressiveness. It relates the created ontology to the requirements defined in the Specification phase (possibility of answering competency questions, using cases coverage). There are a few methods, techniques and tools for ontology evaluation. Within the project, there have been analyzed the most-known approaches in that field, however, due to limited size of this article, we will not present the results.

The proposed methodology assumes the possibility of taking a step back and making corrections that will make the ontology tailored to the needs of the project. It is recommended to run at least 2 iterations after which the resulting model is evaluated against semantics, syntactic, coherence, coverage and adherence to the competency questions and use cases.

f. Maintenance phase

Maintenance is a process that corrects and updates the ontology. This phase is usually performed at the stage of ontology utilization in particular system components. Due to the fact that the duration of the system development phase was established to 5 years, several software components will be successively created within this period. Particular domain ontologies will be though maintained by different task groups and in different timeframe. It is necessary for the ontology creator to actively supervise the ontology utilization and support ontology maintenance. Although Specification phase bases on the questionnaire that helps to develop the most appropriate assumptions, they may be not valid after the final system components are developed. The aim of the Maintenance phase is though to react on changes in system functionality, operation and implementation in order to provide the biggest benefit for the project.

7. Summary

The paper presents IOEM, a methodology for ontology development elaborated for the INSIGMA project. Although prepared for a particular use, the methodology is quite general and can be used in a large variety of IT projects requiring ontology components. It is particularly suitable for large and geographically distributed software projects.

The methodology was proposed as a result of analyses of various approaches and known methodologies for building ontologies. The goal of the IOEM methodology is to provide a defined process comprising phases, supporting tools and well determined results. The methodology implementation is a complex process including several activities that involve different experts and work groups. At present, a number of ontologies within the INSIGMA project are concurrently built. In general, their progress ranges from phase 1 to 3. So far, no significant corrections in the methodology were required. At the end of the project, we plan to assess the overall methodology and formulate lessons learned.

REFERENCES

[1] Guarino N., *Formal Ontology and Information Systems*, Proceedings of FOIS'98, Trento, Italy, 6-8 June 1998.
[2] Ambler S.W., *A Manager's Introduction to The Rational Unified Process (RUP)*.
[3] http://www.enterpriseunifiedprocess.com/essays/enterpriseBusinessModeling.html, S.W. Ambler, *The Enterprise Business Modeling Discipline: Scaling Agile Software Development*.
[4] http://www.w3.org/TR/owl-features/, W3C Recommendation 10 February 2004.

[5]   Booch G., Rumbaugh J., Jacobson, *Unified Modeling Language User Guide*, 2nd Edition, 2005".

[6]   Booch G., *OBJECT-ORIENTED ANALYSIS AND DESIGN With applications SECOND EDITION*, Addisson Wesley, 1998.

[7]   Leffingwell D., Widrig D., *Managing software requirements: a use case approach*, Addison-Wesley, 2003.

[8]   IEEE standard glossary of software engineering terminology, 1990. IEEE Standard 610.12-1990, ISBN 1-55937-067-X.

[9]   Sure Y., Studer R., *On-To-Knowledge Methodology – Final Version*, On-To-Knowledge EU IST-1999-10132, 2002.

[10] Noy N.F., McGuinness D.L., *Ontology Development 101: A Guide to Creating Your First Ontology*, Standford University, Standford, CA, 94305.

[11] Uschold M., King M., *Towards a Methodology for Building Ontologies*, The University of Edinburgh, 1995.

[12] Gruninger M., Fox M.S., *Methodology for the Design and Evaluation of Ontologies*, University of Toronto, 1995.

[13] Kotis K., Vouros G.A., *Human Centered Ontology Management with HCONE*, Department of Information and Communications Systems Engineering, University of the Aegean.

[14] De Nicola A., Missikoff M., Navagli R., *A proposal for Unified Process for Ontology building: UPON*".

[15] Holsapple C.W., Joshi K., *Handbook of Knowledge Management*, A knowledge management Ontology, Chapter 6.

[16] Gómez-Pérez A., Fernández-López M., Juristo N., *METHONTOLOGY: From Ontological Art Towards Ontological Engineering*, Universidad Politécnica de Madrid.

[17] Gómez-Pérez A., Fernández-López M., Corcho O., *Ontological Engineering with examples from the areas of Knowledge Management, e-Commerce and the Semantic Web*, Springer-Verlag London Limited, 2004.

[18] Corcho O., Fernández M., Gómez-Pérez A. & López-Cima A., *Building Legal Ontologies with METHONTOLOGY and WebODE*, Springer-Verlag, 2005.