SECURITY ONTOLOGY CONSTRUCTION AND INTEGRATION

Tomasz Boiński, Piotr Orłowski, Julian Szymański and Henryk Krawczyk
Department of Computer Architecture, Faculty of Electronics, Telecommunications and Informatics, Gdask University of Technology, Poland
{*tomasz.boinski, julian.szymanski, henryk.krawczyk*}*@eti.pg.gda.pl, cmsptcp@gmail.com*

Keywords: ontology design, ontology integration, security ontology.

Abstract: There are many different levels on which we can examine security. Each one is different from others, all of them are dependent on the context. Hence the need to bear additional knowledge enabling efficient utilization of the knowledge by the computers. Such information can be provided by ontologies. The paper presents gathered requirements needed to be taken into account when creating an onthology. The method of ontology creation and the criteria for keywaords selection are presented. Ontology created in such way should provide means for interoperability with other systems.

1 INTRODUCTION

The Universal Polish Language Dictionary (Dubisz, 2008) defines security as "state of non-danger, tranquility, confidence." Experts related to international relationships (Żukrowska and Grącik, 2006) and military (Nowakowski, Z. and Szafran, H.,) all agree with that statement. Experts from other fields, like (Anderson, 2005) or (Borgosz-Koczwara and Herlender, 2008) usually provide a definition that is similar but just more detailed.

The meaning of the term security is greatly influenced by the context it is used in. Security will be understood differently when we are talking about a personal computer, going on a vacation to some unstable country and when we are talking about a nuclear plant. There are many different layers on which we can examine security, each differs with possible risks and outcomes.

(Żukrowska and Grącik, 2006) distinguishes two different types of security dependent on the source of threat:

- external security – security related to external factors,
- internal security (safety) - security related to internal factors.

(Kim et al., 2005) marks that the creation of annotations describing the security of network resources allows better service allocation. (Herzog et al., 2009) adds that accepting one, common set of terms and relations between those terms is good for any type of organization, as any misunderstanding in terms of security can be costly and time consuming. Unfortunately many terms in security have vague definitions. It is also worth mentioning that in 2003 (Donner, 2003) raised a need of creation of single, common ontology for network resources description. Such appeal was later raised again by European Network and Information Security Agency (ENISA) (ENISA, 2006).

Such common ontologies start to emerge but often specific aspects of a given problem makes it unusable without some changes or additions or even creation of new ontology. It is however important to perform such modifications or creation of a new ontology in a way that will allow future integration with other solutions. Such process will be described in further chapters of this paper.

2 SECURITY ONTOLOGY DESIGN

The method used for the creation of the presented ontology is a combination of the best practices from the solutions proposed by researchers working in this field. Creation of an ontology characterized by good quality, extendability and usability in applications and research, like OCS project (Boiński et al., 2009)(Boiński et al., 2010a), was a priority. The method was based on works of (Noy et al., 2001), mainly extended by aspects from NeOn (Suárez-Figueroa et al., 2009b)(Suárez-Figueroa et al., 2009a) and UPON (De Nicola et al., 2009) methodologies.

In initial phase gathering and analysis of require-

ments for the ontology was made. As a result Ontology Requirement Specification Document was created. Before implementation requirement analysis was performed and available solutions were checked.

Ontology development started with analysis and selection of basic concepts as it influences further interoperability (Boiński et al., 2010b). Using selected basic concepts a model in descriptive logic was created. After this step further development was performed.

2.1 Ontology Requirement Specification Document

Creation of this document was based on a template proposed for NeOn project (Suárez-Figueroa et al., 2009b). Some elements were taken from (De Nicola et al., 2009) and second edition of the book Handbook on Ontologies (Sure and Studer, 2009). The documents goals were:

- describe the reason for creating the ontology,
- describe its boundaries,
- find its applications,
- specification of requirements that need to be realized by the ontology.

Additionally it contained:

- list of technical requirements for the ontology,
- list of quality requirements,
- list of other requirements needed by the client of the ontology,
- proposition of list of knowledge sources about the ontology domain,
- list of other ontologies that can be used,
- list of questions that should be answered by the final ontology.

2.2 The reason for creating the ontology

The goal for the ontology is creation of common, unambiguous semantic model of terms from security domain, providing means for easy extension and usability in research projects. The ontology will be developed in OCS ontology editor. Currently OCS supports all languages provided by OWL API 2.1.1 (Horridge and Bechhofer,), like RDF, RDFS and OWL 1.1. OWL 2.0 will be supported in future versions of OCS. Because of this limitation OWL 1.1 was selected as language of the ontology in its DL dialect. As file representation rdf/owl was selected, what ensures its portability. Also the ontology need to have one namespace, as owl:imports are not yet supported by OCS.

2.3 Boundaries of the ontology

The created ontology in general should contain subjects from the following areas:

- general meaning of security and safety,
- domains closely related with security but in limited scope.

In detail the ontology should contain:

- basic and general terms in the field of security,
- terminology from the domain of information safety and security,
- the most important terms from other fields of security, e.g.:
  - road traffic,
  - national and international,
  - energetic, etc.

To further limit the wide scope of the ontology it was decided that it will contain general security and safety terminology and detailed terminology describing security and reliability of computer systems. Information security is close to other fields of security and is generally understood by people from field of computer science.

Following knowledge bases were selected as basic sources of knowledge:

- NIST Glossary (Kissel, 2006),
- ENISA risk management glossary (Enisa, 2010),
- Ian Sommerville book "Software Engineering" (Sommerville, 2006).

Such selection allows capture of different approaches to security and safety. First of them shows American approach to security, second shows European aspects of the problem and the third one shows approach represented by software engineers.

Additionally to the above sources, it was decided to extend the ontology by the following taxonomies:

- IEEE computer security taxonomy (Avizienis et al., 2004),
- Firesmith security requirements taxonomy (Firesmith, 2005a)(Firesmith, 2005b).

Those two publications are well formalized so moving contained in them information to the ontology should be easily doable. Unfortunately there are no publicly available ontologies based on those sources.

2.4 End users

Target users of the created ontology are:

1. researchers and students interested in topics regarding security, safety, ontologies or just are in need of security ontology,
2. software developers that need security ontology to create semantic annotations to their web services or other applications,
3. agent systems or search systems utilizing data previously annotated by the security ontology.

2.5 Intended usage

The ontology was created for following usages in mind:

1. scientific targets, including testing and extension of OCS system,
2. creation of a WWW page simplifying learning by the students topics connected to security,
3. applications created with ubiquitous programming in mind,
4. applications related with developed by Gda nsk University of Technology system for traffic management,
5. search engines and agent systems,
6. other applications related to Semantic Web.

It is also recommended that created domain ontology should be so called general purpose ontology so that it will be usable in other applications.

2.6 Nonfunctional requirements

1. The ontology should support both polish and English language.
2. Concept and properties definitions should come from renown sources or standards.
3. Knowledge sources should be clearly provided.
4. The ontology should be expandable.
5. The ontology should be adaptable.
6. Concepts and properties should be described in a way that should be human readable to allow their easy understanding.
7. The ontology should be consistent.
8. The ontology should be complete.
9. The ontology should be portable.
10. The ontology should work with OCS system.
11. Classification operation should last no more than one second.

2.7 Functional requirements

Functional requirements, as suggested by (Suárez-Figueroa et al., 2009b), were presented in form of competency questions (Table 1).

Table 1: Functional requirements.

Question	Expected answer
What is a risk?	Probability of a loss.
What type of attacks can be performed against computer systems?	DoS, unauthorized access, etc.
What is an internal safety?	State of internal threats.
What are attributes of external security?	Accessibility, integrity, confidentiality.
What is an attack?	Violent usage of force against somebody.

2.7.1 Ontology architecture

It was decided that the ontology should be divided into modules according to ODP (ODP Portal, 2011). Due to lack of owl:imports support by OCS system target ontology must be merged into one file with common namespace. Three modules are planned: basic core module, security and reliability module and security requirements module.

2.7.2 Naming convention

Naming convention proposed by (Schober et al., 2007) was applied to the ontology. It forces usage of context free and human readable names. It forbids usage of names that are negations. Additionally, independently from the naming convention, it was decided to use camel case for multi word names. Names of classes and individuals will start with uppercase (e.g. SampleClassName) and names of properties will start with lowercase (e.g. samplePropertyName).

2.7.3 Evaluation and verification of the ontology

To verify consistency, completeness and adaptability of created ontology tests need to be performed. For that reason Protégé Ontology Tests, included into version 3.4.4 of Protégé editor (Knublauch et al., 2004), were used. Tests should be performed for each module separately as well as for the whole combined ontology and they were based on questions defined in Table 1. To find answers for that questions Protégé plugin called DL Query was used. Achieved results were compared with the ones that

were expected. Completeness tests required involvement of domain experts and was performed manually as there are currently no means of automatic coverage checks (Krawczyk, 2007)(Tartir and ATHENS,).

2.8 Basic concepts

Basic concepts are necessary for creation of ontology core. Their selection influences further interoperability (Boiński et al., 2010b) of the final ontology. To choose proper set of basic concepts external sources were checked for common terms and important topics.

2.8.1 The most important concepts in security engineering

When describing security (Anderson, 2005) concentrates on targets and attributes of security: confidentiality, integrity and availability. For description of security itself he uses the following concepts: system, subject, participant, identity, trusted, reliable, privacy, secrecy, anonymity, authenticity, vulnerability, threat, security breach, security, security profile.

The terms system, subject and participant can be connected with organization's assets. Terms vulnerability, threat and security were mentioned both in ISO model as well as in Fenz and Herzog ontologies. Rest of the concepts in their ontology are treated either as security attributes (confidentiality, integrity, availability) or as not significant.

2.8.2 Security in software engineering

As a main subjects connected with internal safety Sommerville (Sommerville, 2006) states: incident, threat, harm, level of threat, probability of risk, risk. As main subjects connected with external security the author states: exposure, vulnerability, threat and surveillance. Additionally he connects security with reliability, availability and credibility.

2.8.3 Basic concepts in field of international security

Żukrowska (Żukrowska and Grącik, 2006) defines basic concepts in context of international relations. Proposed lists differs substantially form previously presented and includes: defense (readiness to repel the attack, and a synonym for security guarantees), neutrality (not to engage in situations and organizations, which can cause conflicts), budget (military), national interest, raison d'etat, strategic sectors, strategic reserves, interdependence, cooperation, globalization.

Defense and neutrality can be connected to remedy measures. Military budget, strategic sectors and strategic reserves are directly connected with national treasury.

2.8.4 Basic concepts inf field of energetic security

Czerpak (Żukrowska and Grącik, 2006) as main concepts related to energetic security states the following: availability, threat, security policy.

All concepts from above sources were combined with each other and as a result following basic concepts were selected: security, safety, security attribute, vulnerability, security policy, risk, harm, safeguard, threat, protected subject.

3 ONTOLOGY CONSTRUCTION

Target ontology was implemented in an iterative and incremental manner. The design decision was that the ontology will be implemented in three modules, which in the final stage will be merged into one monolithic ontology.

The core module was created form three small ontologies each counting less than 100 classes. Those ontologies were designed during research on basic concept selection (Boiński et al., 2010b). Those ontologies were combined into single OWL file. Security and reliability module contains one average sized ontology based on Avižienis taxonomy (Avizienis et al., 2004). Security requirements module also contains one ontology created in respect to Firesmith taxonomy (Firesmith, 2005a)(Firesmith, 2005b). Each of those ontologies was created according to procedure described in previous sections of this paper using Protégé 4.0.2 editor (Gennari et al., 2002)(Noy et al., 2000).

(Noy et al., 2001) methodology after extension with UPON and NeOn aspects contained the following steps:

- lexicon creation – this step involves selection of concepts form chosen knowledge sources (both glossaries and taxonomies),
- concept selection – each subject included into the lexicon automatically becomes a member of concept set. Additionally from definition of those concepts proper names and more significant nouns were selected. Such set of concepts was converted into OWL classes. When available those classes were annotated with their description taken form the glossary or taxonomy,

- concept hierarchy creation – each occurrence in glossary or taxonomy of statements similar to "expression A is of type B" was converted into OWL subclassOf relation. In taxonomies such relations usually are presented graphically as inheritance which speeds up its conversion to the ontology,
- selection of disjoint concepts and synonyms – in case of classes which names are clearly disjoint like AccidentalBreakdown and NonAccidentalBreakdown disjointWith relation was introduced. In other cases disjointness was added manually basing on human experience and descriptions in chosen glossaries. Same procedure applies in case of synonyms,
- relations identification and selection – as base for relations verbs were selected from concept definitions. If such verb connects any two selected concepts than it is transformed into a relation between those concepts. In taxonomies often aggregation relations are presented graphically. Such relations were converted into "has part" relation,
- creation of relation hierarchy – relations similar in respect of verbs in their names were grouped.
- refining of the relations – domains and counter domains of all relations was added to the ontology. Where possible relations were marked as (non)functional, (non)transitive etc.
- ontology integration – ontologies were combined first within modules and later into one monolithic ontology.

Comparing the security and reliability and security requirements modules the basic core module one can see that to much greater extent they resemble taxonomies. They describe their area more strongly at the same time. During construction process each ontology got it's own namespace and URI address.

4 ONTOLOGY INTEGRATION

Ontology integration using PROMPT, Ontology Module Composition orContent Map didn't provide satisfactory results so hybrid, partially manual approach was undertaken. Ontologies were combined in parts of two. Given pair was compared using Falcon-AO tool (Jian et al., 2005). Result of such comparison was than manually implemented using Protégé editor by means of subClassOf or subPropertyOf relations. Following rules were applied during integration:

- if concept definitions allows statement that one of the concepts is more general than the other one than subsumption relation is used,
- if both concepts have identical definitions and classes (properties) related to them in both ontologies does not have subclasses (subproperties) than one of them was either removed from integrated ontology or equivalentClass (equivalentProperty) was used, when there were subclasses (subproperties) related to them than equivalentClass (equivalentProperty) was always used,
- when concepts are different than they were connected by closes, common super class (super property), when necessary such generalizing concept/property was created.

When applicable combined classes were sufficed by a token identifying source of their definition as advised by Ontology Design Patterns (ODP Portal, 2011). As a result one big ontology was created. It contains 757 classes and properties. After merging those ontologies the namespace and URI was unified. Such combined ontology proved to implement all planed requirements.

5 ONTOLOGY SHARING

According to basic definition of ontology (Gruber et al., 1993), for a highly formalized data model has become the ontology, it must also be shared. For this reason the created ontology was placed into the OCS portal. In the future the ontology can became part of one of the services developed using that system. Each registered user will be able to track and influence changes in the ontology. Changes will be however applied only by specified experts. OCS also supports versioning of ontologies so every change, can always be reversed or the ontology can be developed as many, concurrent versions of the same core idea.

6 CONCLUSIONS

Way how ontologies are created is very important. Both the process itself and even selection of basic concepts can influence how and even if the ontology will be usable outside its primary application or will it be suitable for integration with other ontologies. In this paper we shown that an ontology can be designed and created in a way that will make it suitable for interoperability and integration. In near future we would like to extend our works to include automatic ad hoc ontology integration eliminating manual aspects form presented steps in ontology design.

ACKNOWLEDGEMENTS

This work was financed by National Science Center.

REFERENCES

Anderson, R. (2005). Inżynieria zabezpieczeń.

Avizienis, A., Laprie, J., Randell, B., and Landwehr, C. (2004). Basic concepts and taxonomy of dependable and secure computing. *Dependable and Secure Computing, IEEE Transactions on*, 1(1):11–33.

Boiński, T., Budnik, Ł., Jakowski, A., Mroziński, J., and Mazurkiewicz, K. (2009). OCS – Domain Oriented Ontology Creation System. In *SMI'09, 4th International Conference 'Congress of Young IT Scientists'*. HARD Olsztyn.

Boiński, T., Jaworska, A., Kleczkowski, R., Kunowski, P., and Szamański, J. (2010a). Zespołowa budowa ontologii z wykorzystaniem systemu OCS oraz edytora Protégé. *Zeszyty Naukowe Wydziału ETI Politechniki Gdańskiej*, 19:101–105.

Boiński, T., Orłowski, P., Szpryngier, P., and Krawczyk, H. (2010b). Influence and selection of basic concepts on ontology design. In *Proceedings of KEOD2010 of the 2nd International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management*, pages 364–369. INSTICC.

Borgosz-Koczwara, M. and Herlender, K. (2008). Bezpieczeństwo energetyczne a rozwój odnawialnych źródeł energii. *Energetyka*, pages 194–197.

De Nicola, A., Missikoff, M., and Navigli, R. (2009). A software engineering approach to ontology building. *Information Systems*, 34(2):258–275.

Donner, M. (2003). Toward a security ontology. *IEEE Security and Privacy*, pages 6–7.

Dubisz, S. (2008). *Uniwersalny słownik języka polskiego*. Wydawnictwo Naukowe PWN.

ENISA (2006). Risk management: implementation principles and inventories for risk management/risk assessment methods and tools. Technical report.

Enisa (2010). Enisa: a European Union Agency - Glossary of Risk Management. *http://www.enisa.europa.eu/act/rm/cr/risk-management-inventory/glossary*.

Firesmith, D. (2005a). A Taxonomy of safety-related requirements. In *International Workshop on High Assurance Systems (RHAS'05)*.

Firesmith, D. (2005b). A taxonomy of security-related requirements. In *International Workshop on High Assurance Systems (RHAS'05)*. Citeseer.

Gennari, J. H., Musen, M. A., Fergerson, R. W., Grosso, W. E., Crubzy, M., Eriksson, H., Noy, N. F., and Tu, S. W. (2002). *The evolution of Protege: An environment for knowledge-based systems development*. Stanford Medical Institute, Stanford.

Gruber, T. et al. (1993). A translation approach to portable ontology specifications. *Knowledge acquisition*, 5:199–199.

Herzog, A., Shahmehri, N., and Duma, C. (2009). An ontology of information security.

Horridge, M. and Bechhofer, S. The OWL API: a Java API for working with OWL 2 ontologies. In *Proc. of the 5th Int. Workshop on OWL: Experiences and Directions (OWLED 2009), CEUR Workshop Proceedings, Chantilly, VA, United States, October*, pages 23–24.

Jian, N., Hu, W., Cheng, G., and Qu, Y. (2005). Falcon-AO: Aligning ontologies with Falcon. In *Integrating Ontologies Workshop Proceedings*. Citeseer.

Kim, A., Luo, J., and Kang, M. (2005). Security ontology for annotating resources. *On the Move to Meaningful Internet Systems 2005: CoopIS, DOA, and ODBASE*, pages 1483–1499.

Kissel, R. (2006). Glossary of key information security terms. *Glossary, National Institute of Standards and Technology, US Department of Commerce*.

Knublauch, H., Fergerson, R., Noy, N., and Musen, M. (2004). The Protégé OWL plugin: An open development environment for semantic web applications. *The Semantic Web–ISWC 2004*, pages 229–243.

Krawczyk, H. (2007). *Ontology engineering and its applications*. Department of Computer System Architecture, ETI Faculty, Gdańsk University of Technology.

Nowakowski, Z. and Szafran, H. *Bezpieczeństwo w XXI wieku : strategie bezpieczeństwa narodowego Polski i wybranych państw*. Wydawnictwo Politechniki Rzeszowskiej.

Noy, N., McGuinness, D., et al. (2001). Ontology development 101: A guide to creating your first ontology.

Noy, N. F., Fergerson, R. W., and Musen, M. A. (2000). The knowledge model of Protege-2000: Combining interoperability and flexibility. In *Lecture Notes in Computer Science*. Springer-Verlag.

ODP Portal (2011). Ontology Design Patterns. http://ontologydesignpatterns.org/wiki/Main_Page. [Online; 10-05-2011].

Schober, D., Kusnierczyk, W., Lewis, S., Lomax, J., et al. (2007). Towards naming conventions for use in controlled vocabulary and ontology engineering. *Proceedings of BioOntologies SIG, ISMB07*, pages 29–32.

Sommerville, I. (2006). Software Engineering. 8th. *Harlow, UK: Addison-Wesley*.

Suárez-Figueroa, M. et al. (2009a). D5. 4.2: Revision and extension of the neon methodology for building contextualized ontology networks. *NeOn project. http://www. neon-project. org*.

Suárez-Figueroa, M., Gómez-Pérez, A., and Villazón-Terrazas, B. (2009b). How to write and use the Ontology Requirements Specification Document. *On the Move to Meaningful Internet Systems: OTM 2009*, pages 966–982.

Sure, Y., S. S. and Studer, R. (2009). *Handbook on Ontologies*.

Tartir, S. and ATHENS, G. Ontology-driven Question Answering and Ontology Quality Evaluation.

Żukrowska, K. and Grącik, M. (2006). *Bezpieczeństwo międzynarodowe: teoria i praktyka*. Szkoła Główna Handlowa w Warszawie.