Ontology-based Modeling and Inference for Occupational Risk Prevention

Alexandra Galatescu, Adriana Alexandru
National Institute for R&D in Informatics,
8-10 Averescu Avenue, Bucharest, 011455, Romania
e-mail: {agal, adriana}@ ici.ro

Corneliu Zaharia
Stefan Nicolau Institute of Virology
285 Mihai Bravu Avenue, 030304, Bucharest, Romania
e-mail: corneliu.zaharia@virology.ro

Stefan Kovacs
National Research Institute for Occupational Safety
35A Ghencea Avenue, 061695, Bucharest, Romania
e-mail: stefan_agk@yahoo.com

***Abstract*— The paper describes and motivates the use of ontologies and of an ontology-based model in a training system (under development) for the occupational risks prevention. The personalized training (for a specified context, e.g., a given activity, workplace, operator type, work machine, etc.) will be the result of the automatic discovery of the prevention documents and actions that fit the training request. The paper also sketches the basic components of the training system for risk prevention, adapted to the proposed semantic view.**

Keywords- ontologies; ontology-based modeling and inference; occupational risk prevention; e-training

I. Introduction

The paper gives an approach based on ontologies for a web system (under development) aiming at the online, fast and personalized training for occupational risks prevention. Risk prevention is a combination of disciplines necessary to reduce the risk of injuries and fatalities in any work environment. A proactive approach is the early recognition and prevention of the risk factors

Occupational risks are a category of risks that appear in work environments with a high probability of harming people or machines. Occupational risk prevention and management comply with the principles and methodology of the *risk management* (RM) process, a key process within both private and public organizations [1].

The training for occupational risk prevention should advise the operator on the health, safety, security and environmental issues related to his work. He can ask for training before or during the execution of an activity or before the use of a certain machine.

The system presented in this paper is general and adaptable to any domain with major occupational risks (industry, biology, construction, transportation, environment, agriculture, health etc.). It unifies existing methodological rules and standards for risk prevention and provides tools for the personalized training and consulting, by the dynamic and multi-criteria selection of the prevention documents and actions. It will eliminate the need for a periodical training of the employees and will diminish the costs from the poor information on the risks.

Terminology and guidelines for risk prevention. The system relies on the standard terminology proposed with ISO CD31000 [2], combined with the terminology common to several upper-level ontologies and process models.

There are several risk-related standards published by ISO and other standards bodies, as well as many proposals and principles that refer to risk management. In 2005, ISO has initiated a working group to develop a guidance standard on RM, ISO CD31000. In conjunction with this standard, the group has updated the ISO/IEC Guide 73-Risk Management – Vocabulary [3] that gives the basic vocabulary and the definitions of RM generic terms. It encourages a mutual and consistent understanding and a coherent approach to the description of the RM activities.

In Europe, the risk prevention is subject of two directives Seveso I and Seveso II [4] that establish the domain terminology, the obligations and normative documents to be elaborated regarding the large scale industrial hazards.

State of the art in software for risk prevention. There are products for the risk control in industrial environments, and domain-specific standards and software tools for RM in health, environment, insurance, finances, construction, transportation, etc. Risk prevention is automated for the security of computers, Web, networks. Security components are integrated lately with the operating systems. Ontologies are also used mainly for the security management (of assets, networks, information systems, databases, etc.). Some examples are in [5]-[9]. There is no system based on knowledge and semantics for occupational risk prevention and for training and dynamic discovery of prevention information, documents and actions.

However, [10] proposes the risk evaluation and analysis along the life cycle of the construction projects, based on ontologies and a conceptual model. They rely on a simpler reference ontology and model and have a different inference goal. Also, [11] gives an example of Web Ontology Language (OWL) [12] ontology for occupational health. And, [13] confirms the idea that a model of occupational risks is important because it describes relevant data in the context of event occurring and this data can be transformed into knowledge navigated using an intelligent search engine (similarly to the goal of the system presented in this paper).

A semantics-based approach for risk prevention. In order to benefit from semantics, the system relies on:

- A *reference ontology* that gives the basic types of taxonomies and structures for the classification and description of the risk factors, of the relationships among them, of the consequences and actions for their prevention, etc. This ontology represents the background of a *reference model* for occupational risk prevention (represented in Fig. 2 and detailed in Section 3).
- *Domain-specific ontologies and knowledge bases,* built by the specialization of the generic concepts and relationships in the reference ontology and model.
- A *query language* and *editor* for risk prevention based on the domain ontologies and reference model. The framework for the query composition based on ontologies is given in Fig. 3.
- Automatic and semantics-based inference for search and discovery of prevention documents and actions based on the risk context and requestor's preferences.

The semantics will support the interoperability of the organizations with respect to risk prevention, by common vocabularies and model. The ontology-based inference will increase the precision of the search algorithm. Also, the vocabularies and model can be dynamically extended and used in further inferences or they can be reused in other applications.

The constructs in the reference and domain ontologies comply with a subset of the constructs proposed for OWL in [12]. The constructs in the reference model comply with the basic constructs in the entity-relationship model, adapted to the use of ontologies instead of entities.

Structure of the paper. Section 2 sketches the basic components of the training system. Section 3 describes the semantic and modeling layers for the representation of the occupational risks and of the context for their occurrence. It also enumerates the basic types of inferences that will be implemented in the system and exemplifies the composition of a training query based on ontologies.

II. COMPONENTS OF THE ONTOLOGY-BASED TRAINING SYSTEM FOR OCCUPATIONAL RISK PREVENTION

The intended system will have components distributed on two platforms (see Fig. 1):

- *Platform for the risk design,* i.e., for the risk identification, description and analysis (Fig. 1, left);
- *Platform for the risk evaluation and decision-making* on the prevention documents and actions (Fig. 1, right).

The two platforms share the repository composed of ontologies, rules, queries and documents.

The *components of the platform for the risk design* are:

- *Ontology Editor* to build (specialization or composition) ontologies or list structures in the model given in Fig. 2;
- *Rule Editor* to define or customize rules for risk prevention in the designer's organization;
- *Query Editor* to predefine or customize queries for risk prevention training.

The c*omponents of the platform for the risk evaluation and decision-making* are:

- *Query Composition and Submission Engine* to dynamically compose the training queries.
- *Model Navigator* to navigate the ontologies in the reference model in order to compose the query, as exemplified in Fig. 3.
- *Inference Engine,* called automatically after the submission of the query, in order to perform the automatic discovery of the training documents registered in the system and of the appropriate prevention actions. Besides the conditions and constraints in the query, the discovery will also rely on rules previously defined by the risk designer.
- *Query Result Generator*, called automatically by the Inference Engine, after the document discovery, in order to arrange the results.

Discovered documents can be stored either in the system repository or in the repositories/ Web servers of the organizations registered in the system. The documents can be in different formats: Web pages, Word, .pdf, .xls, etc.

The system is developed using Microsoft Visual Studio 2008 and obout Suite for ASP.Net [14]. It integrates the expression evaluator given in [15] and the interface for the rule and query editors is inspired from [16].

III. AN ONTOLOGY-BASED MODEL FOR RISK PREVENTION

The system integrates three layers representing the occupational risks and the context for their occurrence and prevention: semantic, modeling and execution layers.

The ***semantic layer*** is composed of the *reference ontology* and the *domain ontologies* that give the basic vocabularies for domains with potential risks. The domain ontologies are populated by domain experts (risk designers) using the ontology editor. They are represented by:

- domain-specific *taxonomies*, i.e., hierarchies composed of concepts connected, in this system, by relationships like: *specialization or synonymy* or *composition* (part-of) or *list*-like relationships;
- *attributes* of and *constraints* upon the concepts and relationships in each ontology.

The concept attributes in any ontology can refer to external ontologies. For example, the "domain" attribute of an "activity" in Activity ontology can be selected from Domain ontology.

 ISBN: 978-1-61208-104-5

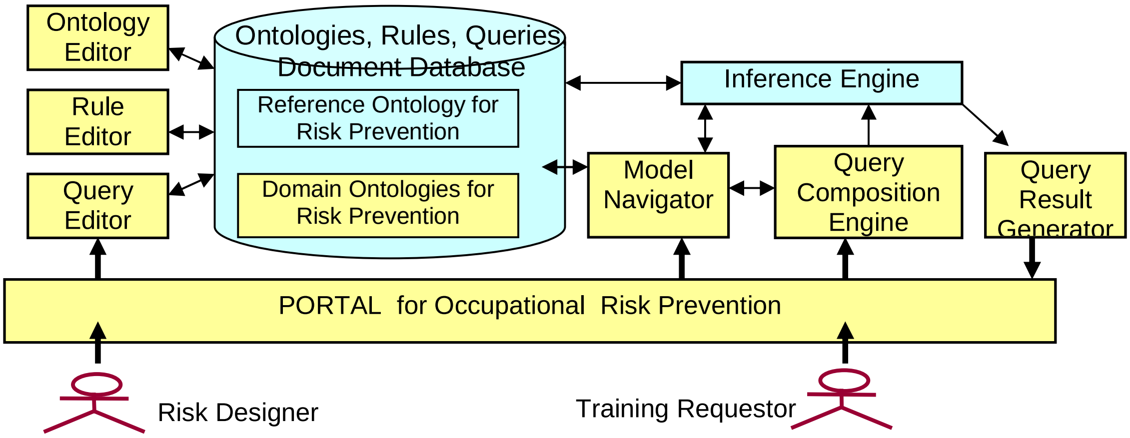

Figure 1. Components of the training system for occupational risk prevention

In this system, the ontology editor treats separately the specialization, composition and list-like ontologies, because each type of ontology has its specific features. For example, the *attribute inheritance* is implicitly implemented only in the specialization ontologies. For the composition ontologies, it can be requested by the user for attributes of the ontology or of certain concepts. For list-like ontologies it is useless.

The ***modeling layer*** is needed in addition to the semantic layer in order to represent the application-specific *relationships* and *constraints* between concepts in different ontologies. In this system, the inter-ontology relationships are defined according to the reference model represented in Fig. 2. An *ontology-based model* is seen as a union of relationships between concepts in different ontologies, along with their attributes and constraints.

The ***execution*** **(*technological)*** ***layer*** represents the ontologies and models, the documents, rules, constraints and queries in a format interpretable by the software.

A. Semantic Layer for Occupational Risk Prevention

The concept types connected as in Fig. 2 and described below root ontologies based on specialization, composition or list relationships.These ontologies have been proposed to help for the identification and classification of the risk factors, of the consequences and preventive measures, of the dangerous activities and of the processes they compose, etc.

Risk**:** combination of an event probability and its consequences [3]. The event can take place in a certain workplace, during a certain activity/ task or resulting from a material source action (e.g., water, a substance, gas, etc.).

Executant **(**or Starter or Operator**)**: the (human or material) agent that, during an activity, can cause unexpected events and also can be injured by them or can get professional diseases.

Process: a sequence of activities/ operations/ tasks in a certain domain and workplace. The activities in a process can be executed by different executants, at different moments and in different places.

Activity (operation/ task): atomic operation executed independently or during a process inside the organization.

Event: occurrence or existence of a particular set of circumstances. An unpredictable event is called "incident" [3]. It can be the consequence of the executants' action using a certain instrument and acting on a certain object.

Workplace**:** location in the organization where unexpected events can occur and affect/ destroy it.

Consequence**:** outcome of an event or change in circumstances affecting the achievement of objectives [3]. An event may lead to a range of consequences. A consequence can have positive or negative effects. For the occupational risks, only the negative effects are considered.

Work_Instrument: tool/ machine/ substance/ etc., used by the operator during an activity/ task. It can determine an event or be damaged by it.

Work_Object: object existing at a workplace. It can determine an event or an event may impact on it. It can be material (e.g., a substance) or human (e.g., an infected patient in a hospital).

Document: a document containing prevention/ protection/ control instructions, regulations, rules or measures for risk prevention.

Prevention_Action: management action preventing the unexpected events or diseases. An example is the training of the operators in workplaces with potential risks.

B. Modeling Layer for Occupational Risk Prevention

Figure 2 shows how the semantic and modeling layers for risk prevention are integrated from the conceptual point of view. The modeling layer represents the relationships between the ontologies defined on the semantic layer. These relationships have been selected depending on the needed reasoning on them and on the context for the risk identification, analysis, evaluation and prevention, identified at this moment. The model can be dynamically enhanced with new ontologies, relationships, attributes and constraints that will be used in future rules, queries and inferences.

Risk modeling for their prevention and control is today mainly a mathematical modeling complemented with formal methods to assess or measure the risks and to help the

 ISBN: 978-1-61208-104-5

decision-making for their prevention. Also, this modeling is usually a domain-specific one for health/ financial/ insurance/ economic/ business/ etc. risks.

The *benefits from an ontology-based model* in general and, in particular, for risk prevention are:

- The types of concepts and the relationships between them in the model, as well as the reasoning on them, are explicit (external to the application code) and independent of the application tools;
- The ontologies can be shared by different diagrams or models (e.g., for risk monitoring and control, in addition to risk prevention).
- The separation of the ontology-based model from the representation of the ontology content (the domain-specific hierarchies) makes it flexible, adaptable and extensible. The tools for ontology editing and navigation may differ from the tools for the model editing and navigation. Also, the reasoning on the model can be separately implemented from the reasoning on ontologies.

The basic *relationships in the reference model* are enumerated below.

Activity->Process relationship is a "*part-of*" relationship between the component activities and the process they belong to. In a process, the activities might be executed by operators in different departments and even in different organizations. The risks should be tracked for each activity, but also for each process in/ cross organizations.

Process->Process relationship is a "*part-of*" relationship between a process and its sub-processes with potential risks that should be tracked.

Activity->Workplace and ***Process->Workplace*** relationship "*executed_IN*" are necessary to track the risks per activity, process and workplace at the same time.

Executant->Activity relationship "*is_agent_of*" and also ***Activity->Work_Instrument*** and ***Executant->Work_Instrument*** relationship "*acts_WITH*" are necessary for reasoning on an *operator-activity-machine* sub-model.

Executant->Event relationship "*causes*" helps for the identification of the events that an operator might determine by his work. Inverse relationship "*acts_ON*" between ***Event->Executant*** helps for the identification of the executants that can be injured after certain events.

Work_Instrument->Event relationship "*causes*" is necessary to identify the events determined by the inappropriate use of a certain instrument. The inverse relationship "*acts_ON*" between ***Event->Work_Instrument*** is necessary to identify the instruments that can be damaged after certain events.

Besides the executants and the work instruments, the unexpected events or diseases might be caused by other objects existing at the workplace. These events can be found by the relationship "*causes*" between ***Work_Object->Event***. Also, the objects damaged by certain events can be found by the relationship "*acts_ON*" between ***Event->Work_Object***.

Event->Workplace are correlated by the relationship "*acts_ON*" in order to associate the events to the workplaces they can damage.

Risk->Event relationship "*has_effect*" associates the identified risks to the events they may produce.

Event->Consequence relationship "*has_effect*" associates the events with their consequences.

Document->Risk relationship "*describes*" associates to the identified risks the documents and the actions necessary for their prevention.

The reference model in Fig. 2 also associates the elements with potential risks that should be tracked (types of concepts like Activity, Executant, Workplace, Work_Instrument, Work_Object) with their specific risks (in Risk ontology), by the relationship "*has_risk*".

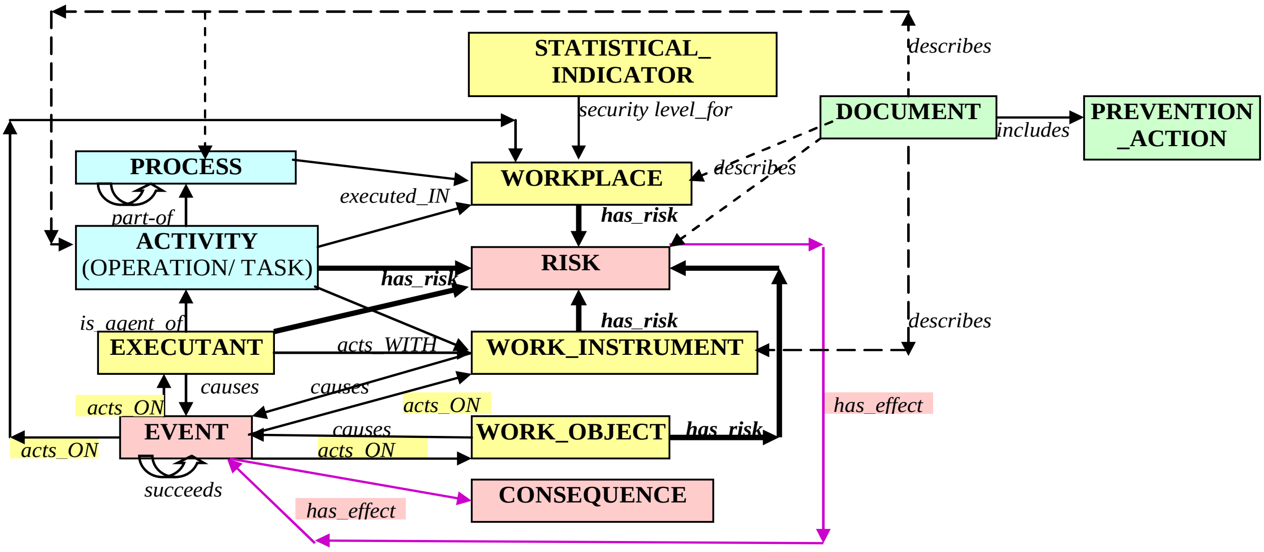

Figure 2. Basic ontologies and the relationships between them in the reference model for occupational risk prevention

The generic concepts above, their attributes, relationships and constraints in the reference model are specialized and instantiated by the risk designers, resulting in *domain models* (e.g., for biological or industrial risks). For any domain concept, the designer instantiates the concept attributes defined in the reference or domain ontology (either implicit, for the concept unique identification, or inherited or concept specific attributes). By their instantiation, the generic relationships and constraints in the reference model become domain-specific relationships between concrete concepts in ontologies.

C. Ontology-based Inference and Query Composition for Risk Prevention

The basic *knowledge and inference* for risk prevention will include (see details in the case study given in [17]):

- Rules for query formulation and for the verification of its syntactic and semantic correctness;
- Rules for the semantic completion of the search query: for each concept in a specialization ontology, its subtypes and synonyms are added in the query, as search alternatives;
- Rules for navigation on domain ontologies and models;
- Rules for risk evaluation and for search and discovery of documents and actions for risk prevention;
- Inheritance rules for both concept attributes and attribute values. In this system, for the specialization ontologies, the values of certain attributes can be inherited by concepts from their parents, at the designer's request. But, in the concept description, there are identification attributes with concept-specific values that cannot be inherited (e.g., ID, concept author, creation date).
- Inheritance rules for concept relationships. For instance, the relationship between a risk and a certain activity or workplace can be inherited by the subtypes of the respective risk.
- Rules for ordering the query results, depending on the conditions and constraints on the involved concepts.

The *query semantic completion* for a concept C is performed by the navigation in the ontology the concept C belongs to and by the extraction of its subtypes and synonyms. They are correlated with the initial concept C by the logical operator OR. Hence, the search algorithm does not use the concept C and its subtypes/ synonyms simultaneously, but successively, even if the search with the initial concept C is successful. The benefit is that more appropriate results are obtained than using only the initial concepts.

Regarding the inference for the verification of the query semantic correctness, the system will achieve:

- Verification of the *semantic compatibility between each concept type C and its concept-like instance* (value) specified in the query (when the value is a concept, not numeric). This verification is fully automated only when the value concept belongs to the same ontology as the concept type C. Otherwise, the system involves the requestor to confirm their semantic compatibility. For instance, the concept 'Laboratory_Procedure' is compatible with its instance 'p1' only whether p1 is a laboratory procedure as well, not a concept with another meaning.
- Verification of the *semantic relationships between the concepts in the query*. Suppose that the query includes two concepts Ci and Cj that belong to the ontologies Oi and Oj. Also, suppose that, previously, the designer has defined a generic relationship R between the ontologies Oi and Oj. The system checks if the designer has also instantiated the relationship R for the concepts Ci and Cj. If he did not, the occurrence of both concepts in the query might be a semantic contradiction.

  For instance, suppose that the query includes the activity *A* and the work instrument *I*. Also, suppose that between the ontologies Activity and Work_Instrument there is a generic relationship *executed_WITH*. If there is no concrete instance of this relationship between *A* and *I* (*A executed_WITH* I), it is possible that the instrument *I* is incorrectly associated with the activity A in the same query. The requestor will be notified before the request execution in order to review his request. Otherwise, he can receive results about the instrument *I* that are not correlated with the results regarding the activity *A*.

Figure 3 gives an example of query composition based on ontologies. The query has three parts:

- *Search query:* a Boolean expression with concepts as operands. For example, the user asks for the prevention rules and the prevention measures to be selected from the ontology Document and for the physical risks at the workplace to be selected from the Risk ontology);
- *Search condition:* an expression with known concepts as operands. They indicate the work context where the risks and events can occur (e.g., the activity “Laboratory_Procedure” selected from Activity ontology and the work instrument “Substance_with_microorganisms” selected from Work_Instrument ontology).
- *Concept restrictions:* expressions with concept attributes as operands. For example, the search should find the physical risks with high gravity and that can occur frequently.

After the query submission, it is analyzed and semantically completed, the conditions and constraints are syntactically and semantically analyzed and, then, the search algorithm is executed.

IV. CONCLUSIONS

The paper has described the conceptual and semantic framework of a system for training on occupational risks. It relies on a dedicated reference ontology and model, on domain specific ontologies and on reasoning on them, basically, for the search and discovery of registered prevention documents and actions.

Although the importance of the ontologies and of a model for risk prevention has already been revealed in the literature, there is no general software for on-line training, the goal of the system presented in this paper.

The system architecture was adapted to a semantics-based view on the risk prevention. Its *interface dedicated to the domain experts* moves the work for ontology editing and risk design, from IT experts to the domain experts. The system and its portal will contribute to a *knowledge repository for risk prevention* inside and cross organizations. It will be accessible from Web and will gradually replace the periodical training in organizations.

The *risk prevention model* described in this paper *can be dynamically extended* with new ontologies, relationships, constraints and rules, when necessary. They will be automatically considered in future inferences on the model.

The main *benefits* in this system from ontologies and from the ontology-based model for risk prevention are:

- *organization interoperability,* by common vocabularies and models on risk prevention represented in the reference ontology and model and in the domain ontologies and rules. They can be reused in other applications.
- *semantics-based inference,* by ontology and model-based verifications and executions of the rules and queries. They increase the search precision, completeness and correctness;
- *personalized queries* for training, dedicated to domain experts.

The system is partly implemented, as follows:

- The platform for the risk design is already implemented. Besides the *rule editor* and *query editor*, the designer can use general tools for:
  - o *ontology editing* (for specialization, composition and list-like ontologies), with automatic inheritance of the attributes only for the specialization ontologies; and, for
  - o *reference and domain model editing and instantiation.* These tools help the designers to add to the reference model: new ontologies, new inter-ontology relationships, new attributes for ontologies and relationships. And, to add to the domain model and ontologies: new concepts, new concept instances, new relationship instances. They also provide the graphical view of the models.
- The platform for risk evaluation and decision-making is partly implemented: the query composition engine and model navigator are finished; but, the inference engine and query result generator are under development.

REFERENCES

[1] MethodWare, "ISO 31000: Risk Management Standard," 2009, http://www.methodware.com/iso-31000-risk-management-standard-published [visited, May 14, 2010]

[2] ISO, "ISO 31000: 2009 Risk management —principles and implementation of risk management," 2009, http://www.iso.org/iso/catalogue_detail.htm?csnumber=43170 [visited, May 10, 2010]

[3] ISO/ TMB/ WG, "ISO/ IEC Guide 73:2002 “Risk management — Vocabulary Guidelines for use in Standards,” 2009, http://www.iso.org/iso/catalogue_detail?csnumber=34998 [visited, May 10, 2010]

[4] MAHB (Major Accident Hazards Bureau), "Safety Management Systems - Seveso II," Official Publications of European Communities, Luxembourg, 1998, http://mahbsrv.jrc.it/ GuidanceDocs-SafetyManagementSystems.html [visited, May 11, 2010]

[5] S.-W. Lee, R. Gandhi, D. Muthurajan, D. Yavagal, and G.-J. Ahn, “Building problem domain ontology from security requirements in regulatory documents,” Proc. Intl. WS on Software Engineering for Secure Systems. ACM Press, 2006

[6] B. Tsoumas and D. Gritzalisi, “Towards an ontology-based security management,” Proc. Intl. Conf. on Advanced Information Networking and Applications (AINA), Vienna, Austria, 2006

[7] M. Klemen, E. Weippl, A. Ekelhart, and S. Fenz, “Security ontology: Simulating threats to corporate assets,” Proc. 2nd Intl. Conf. on Information Systems Security (ICISS), Springer, 2006

[8] A. Simmonds, P. Sandilands, and L. van Ekert, “An ontology for network security attacks,” Proc. Asian Applied Computing Conference (AACC), LNCS, vol. 3285, Springer, 2004

[9] P. Mitra, C. Pan, P. Liu, and V. Atluri, “Privacy-preserving semantic interoperation and access control of heterogeneous databases,” Proc. Symposium on Information, computer and communications security, ACM Press, 2006

[10] N. Forcada, M. Casals and A. Fuertes, "The Basis of a Decision Making Tool for Risks’ Evaluation Based on Ontologies," Proc. Intl. Conf. on Information and Knowledge Management - Helping the Practitioner in Planning and Building (CIB), Stuttgart, 2007, http://www.baufachinformation.de/aufsatz.jsp?ul=2008031001259 [visited, April 12, 2010]

[11] J. Kola, B.Wheeldin and A. Rector, "Lessons in building OWL Ontology driven applications: OCHWIZ – an Occupational Health Application," National e-Science Centre, 2007, http://www.allhands.org.uk/2007/programme/download490f.html?id=849&p=paper [visited, April 12, 2010]

[12] W3C, "OWL 2 Web Ontology Language Structural Specification and Functional-Style Syntax Oct. 2009," http://www.w3.org/TR/2009/REC-owl2-syntax-20091027/ [visited, May12, 2010]

[13] P. Swuste, "Qualitative Methods for Occupational Risk Prevention Strategies in Safety or Control Banding safety," Safety Science Monitor, Issue 3, vol. 11, 2007

[14] obout,"obout Suite for ASP.Net", www.obout.com, 2010 [visited, May 20, 2010]

[15] P. Ganaye, "An expression evaluator written in VB.NET," http://www.codeproject.com/KB/vb/expression_evaluator.aspx [visited, April 2, 2010]

[16] V. Abilov, "WYSIWYG rule editor: create and test rules for any .NET type," http://bloggingabout.net/blogs/vagif/archive/ 2009/04/13/wysiwyg-rule-editor-create-and-test-rules-for-any-net-type.aspx [visited, Jan. 15, 2010]

[17] A. Alexandru, F. Filip, A. Galatescu and E. Jitaru. "Using Ontologies in eHealth and Biomedicine", in book A. Shukla and R. Tiwari (Eds) "Intelligent Medical Technologies and Biomedical Engineering: Tools and Applications", IGP Global, May, 2010

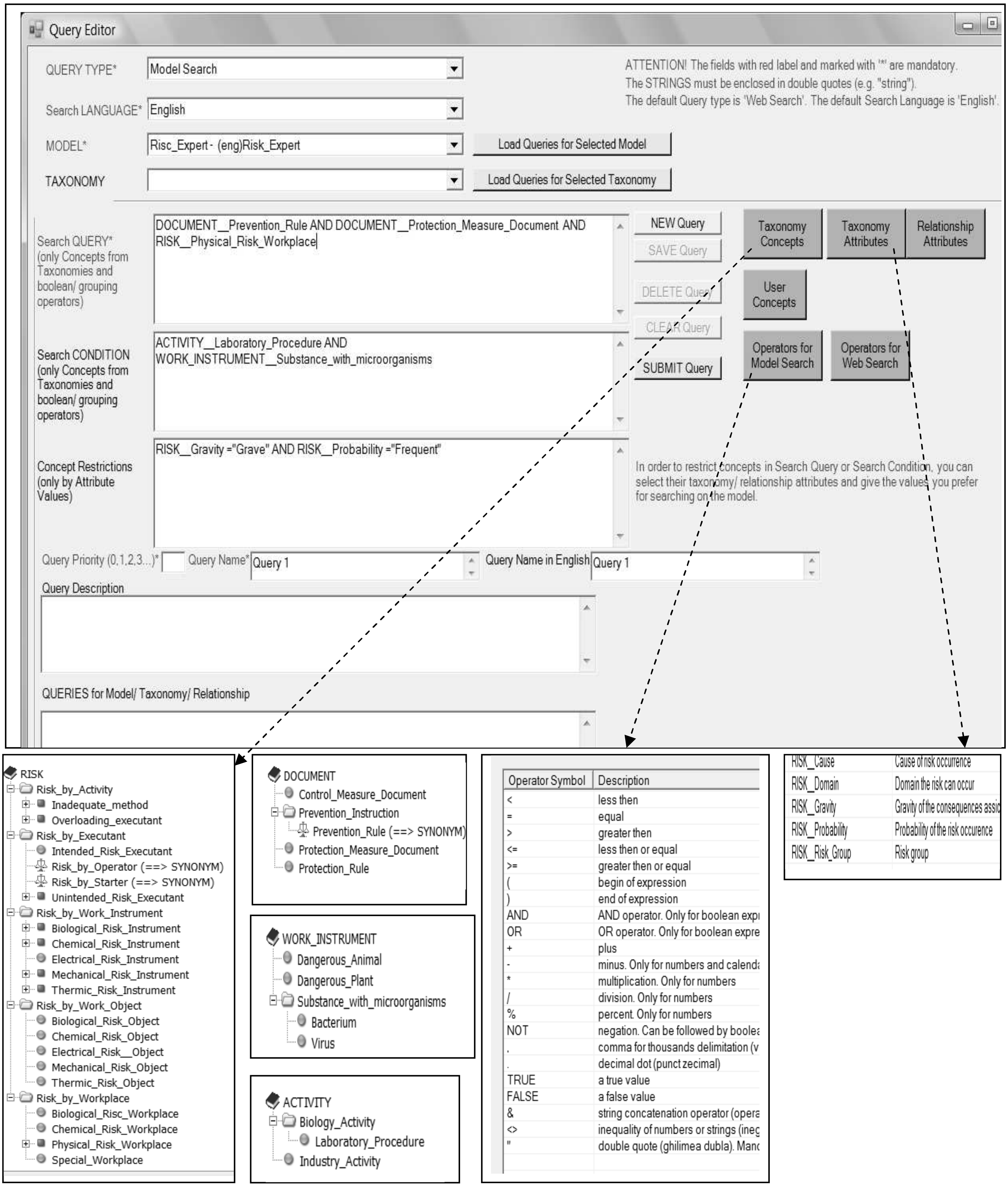

Figure 3. An example for the composition, based on ontologies, of a query for training in risk prevention