

## Automatic generation of tests from domain and multimedia ontologies

Andreas Papasalouros<sup>a\*</sup>, Konstantinos Kotis<sup>b</sup> and Konstantinos Kanaris<sup>b</sup>

<sup>a</sup>Department of Mathematics, University of the Aegean, Karlovassi, Greece; <sup>b</sup>Department of Information and Communication Systems Engineering, University of the Aegean, Karlovassi, Greece

(Received 1 July 2008; final version received 10 May 2009)

The aim of this article is to present an approach for generating tests in an automatic way. Although other methods have been already reported in the literature, the proposed approach is based on ontologies, representing both domain and multimedia knowledge. The article also reports on a prototype implementation of this approach, which automatically creates tests using the Semantic Web standard technology OWL (Ontology Web Language) as well as proper annotations of images. The proposed approach is independent of specific domain characteristics, since question items are generated according to generic ontology-based strategies. In the presented prototype implementation, simple natural language generation techniques are used to project the items in the tests.

**Keywords:** ontologies; multimedia; computer-aided assessment

### 1. Introduction

Electronic tests are a very popular means of assessment and self-assessment in both traditional and electronic learning settings. They are appealing to the examinees, they can be automatically graded and they provide the capability of frequent testing, almost immediate feedback on the performance. The most common type of electronic tests comprises of text-based multiple choice questions (MCQ). Such a type of a test comprises a number of questions named *items*. Each item consists of a short text describing a question or a sentence to be tested, called *stem*, and a number of alternative *choices*. In single-response MCQ, one of the choices is the *correct answer* and the wrong alternatives are called *distractors*.

Typically, the number of items constituting a test must be large enough for a test to provide credible evaluation. Furthermore, the construction of question items requires specialized knowledge. Thus, the creation of MCQ tests is a resource-consuming task in terms of time and human effort, which would be benefited by automation.

Automatic creation of MCQ tests can be considered as a specialized application of natural language generation (NLG), which is based on the following:

- (1) The existence of a knowledge base, expressed in a knowledge representation language, which contains a set of facts about a specific domain. From the

---

\*Corresponding author. Email: andpapas@aegean.gr

existence of a knowledge base, expressed in a knowledge representation language, which contains a set of facts about a specific domain. From these facts, question items together with correct answers are extracted for the test.

- (2) The use of the semantic relationships between various elements in the knowledge base to assert “false” statements. These statements are used for generating the distractors in question items.
- (3) The application of NLG techniques for actual sentence generation (Bateman, 1997; Bontcheva & Wilks, 2004) based on the semantic descriptions generated in the first two steps.

Beyond text-based MCQ, other types of question items exist, which incorporate different types of media, such as images, sound, and video. Tests comprising these types of question items can be used in both formal education, in domains with emphasis on audio-visual information, as well as in informal educational activities combined with entertainment (edutainment). In order for tests of this type to be automatically generated, the above process must be enhanced by enriching the knowledge base with media-related semantics.

This work proposes a method for utilizing ontologies as a knowledge representation formalism to support automatically generating tests for assessment and self-assessment. Ontologies provide well-defined domain conceptualizations, which formalize facts as well as definitions and rules about a particular subject. Both text-based and media-rich tests are generated, by using domain ontologies in OWL (W3C, 2004a) format, as well as associated annotations on multimedia content. Multimedia tests are restricted to items with images, but the approach can be generalized to sound and video media as well. For experimental purposes, the article reports on results produced with a number of domain ontologies for text- and image-based tests. Based on this approach, a prototype tool was developed. Certain strategies are used for selecting the correct answers in question items, as well for selecting distractors. These strategies are analytically presented and constitute the main contribution of this article. Benefits, with respect to existing literature, that motivated this work are domain independence, the creation of different types of questions through a systematic utilization of ontology semantics and its support for media-rich question generation.

The rest of this article is organized as follows. In section 2, related work is presented. In section 3, the proposed framework is described. In sections 4 and 5, the strategies for text and image question generation are discussed, while section 6 describes the generation of questions and insights of the implementation of the prototype system. In section 7, an evaluation of the approach is presented, and conclusions are outlined in section 8.

## **2. Related work**

Tsumori and Kaijiri (2007) propose a methodology similar to ours for MCQ generation based on ontologies. These ontologies are used as a knowledge base for item generation based on predefined templates. The ontology used relates each concept (term) with a synonym, a definition, a description, a figure, a superconcept, and a potential concept with a part of relationship. This ontological description is not as expressive as OWL, used in our approach. It focuses on estimation of question difficulty and student assessment. Furthermore, although MCQs with images are supported, more advanced questions with images are supported by our approach, by using richer image annotation semantics, as discussed in section 6.

The methodology presented by Mitkov, Ha, and Karamanis (2006) generates MCQ based on text corpora in a specific domain by employing ontologies, such as WordNet (Miller et al., 1995). The proposed system functionality can be divided in three steps: term extraction concerning frequent concepts inside the text, stem generation, and distractor selection. The extraction of the terms is implemented by shallow parsing of scanned text corpora. Afterward, a frequency measure is applied so the noun or noun phrase with occurrence frequency above a customizable threshold is selected. The stem generation actually filters the clauses and transforms the selected ones to the stem of an item. This is done by utilizing a simple set of rules that are assisted by WordNet. The last part, namely the distraction selection part, is dictionary-based and uses mostly WordNet to obtain the candidate distractors. Apart from the fact that our approach can deal with media tests, OWL ontologies are used for representing the domain knowledge, which can provide richer semantics/axioms than Wordnet or any other lexicon, but more importantly, they support the reasoning on the existing knowledge for the production of new (inferred) knowledge.

Holohan, Melia, McMullen, and Pahl (2005) describe the *OntoAWare* system, which provides a set of tools useful for educational content authoring, management, and delivery. It exploits the semantic web technology along with knowledge-representation standards and knowledge-processing techniques. Moreover, the authoring environment that is introduced by the system concerns the semi-automatic generation of the learning objects (standard e-learning and courseware elements). To generate the learning objects, one can customize existing ontologies or even create a new one from scratch. One of the features of the presented system is the generation of tests from ontology elements, which is the focus of our article. However, the focus of *OntoAWare* is on adaptivity and personalization. Thus, the generated tests are based on subsumption axioms (class-subclass relationships). No further strategies are applied, and thus, the expressiveness of OWL-DL in describing domain knowledge is not utilized for evaluation. The delivery environment of this system can also be configured and can vary from free navigation to learning technology standards-based course delivery in the form of Simple Sequencing specification (IMS, 2003).

Holohan, Melia, McMullen, and Pahl, (2006) present an advancement of the *OntoAWare* system toward generating assessments for problem solving skills in the domain of relational databases. These assessments are produced by utilizing an ontology that describes the domain in question. Students may customize the system to produce personalized problems. While this approach goes one step further by assessing higher order skills, such as problem solving, it is confined in a single knowledge domain.

Soldatova and Mizoguchi (2007) describe an ontology-based approach for test composition. There is a threefold use of ontologies: a test ontology provides a detailed specification of test items; a student model describes the level of understanding of topics comprising the domain model under consideration; finally, a set of rules governs the test construction process. In this approach, emphasis is given to constructing tests delivering items of appropriate difficulty with regards to a particular user model. As exemplified in (Soldatova and Mizoguchi (2003), for some types of knowledge, question items are automatically generated based on facts, events and terms contained in the domain model. Our approach does not require the use of a predefined ontology, but rather uses the standard semantics of OWL for assessing the knowledge of a domain.

AquaLog (Lopez, Uren, Motta, & Pasin, 2007) is an integrated system for answering questions by users expressed in natural language. Questions are mapped to triples representing queries applied to an OWL-based knowledge base. Query results are also triples, which must be converted to natural language sentences, just as in our approach. While AquaLog provides a sophisticated methodology for mapping from natural language queries to ontology triples, there is no mechanism for doing the opposite, i.e. conversion from query triples to natural language, which is of interest in our approach.

Li and Sambasivam (2005) provide a method for generating questions that involve students in solving equations of multiple variables. These variables correspond to quantities related to terms that are specified in appropriate ontologies of specialized domains. While this is an advanced method for problem solving questions generated from ontologies, its use is confined in specific domains, our method is domain independent and furthermore can take advantage of media annotations.

### 3. The proposed framework and architecture

Ontologies represent domain knowledge in the form of definitions of terms, individuals belonging to these terms and relationships between these terms and individuals. The above constitute the asserted knowledge, i.e. explicitly defined facts within the ontology.

OWL (W3C, 2004a) is the standard Web ontology language with well-defined semantics. OWL is based on Description Logics knowledge representation formalism (Baader, Calvanese, McGuinness, Nardi, & Patel-Schneider, 2003) for expressing concept definitions and basic relationships between concepts, which is considered of adequate level of expressiveness. Furthermore, a number of software tools are available for ontology management and reasoning. Thus, OWL was chosen for the description of domain ontologies. The approach presented in this article follows specific ontology-related conventions:  $A, B, C, D$  are names of *concepts* (also known as *classes*);  $R, S$  are names of *roles* (also known as *relationships* or *properties*); and  $a, b, c$  are names of *individuals* (also known as *instances*). Based on these conventions, we produce the following statements:

- $A(a)$ : states that  $a$  is an individual of class  $A$ .
- $R(b,c)$ : states that individuals  $b$  and  $c$  participate in binary role  $R$ .

In ontologies, individuals are members of collections named concepts or classes. These classes are organized in subsumption (class/subclass) hierarchies, i.e. “*is a*” relationships. Properties (roles) are relationships between individuals in the ontology. A property has a *domain*, which is the “class of individuals to which this property can be applied” and a *range*, which is the “class of individuals that a property can have as its value” (W3C, 2004a). There are two kinds of properties in OWL: object properties, which are relationships between individuals and datatype properties, i.e. relationships between individuals, and basic types, e.g. numerical or string. In terms of OWL,  $R$  is an *object property* and  $b, c$  are individuals, which are related by this particular property.

Annotation of content with multimedia semantics refers to assigning proper descriptions to multimedia documents, as a whole, as well as to specific segments in these documents. Such descriptions can be either structural, providing spatial and

temporal information related to segments, frames and moving regions of videos, audio segments and still regions in images, or conceptual, formalizing entities depicted at multimedia content and their relationships by assigning domain-related semantics (ISO, 2004). MPEG-7 (Martínez, 2002) is the prominent multimedia description standard covering media-related aspects of annotated content. Latest efforts such as COMM<sup>1</sup> (Core Ontology for Multimedia) are based on MPEG-7 and are proposed as formal descriptions compatible with existing semantic web technologies (Arndt, Staab, Troncy, & Hardman, 2007). In these efforts, domain ontologies are used to describe the semantics of the domain associated with content.

The proposed framework utilizes domain ontologies and associated annotations on media items, currently images, to generate text and media-rich texts. The overall architecture of the proposed framework is illustrated in Figure 1. A domain ontology in OWL format is provided as input.

Ontologies incorporate a reasoning mechanism to derive facts from explicitly defined knowledge (Baader, Calvanese, McGuinness, Nardi, & Patel-Schneider, 2003). These facts, not explicitly defined in the initial ontology, constitute the inferred knowledge. In the approach presented in this article, reasoning is applied *before* question generation and thus, generated questions are based on both asserted and inferred knowledge. As a result, a student performing a test not only is assessed on the recalling of factual knowledge, but also is expected to apply inferencing based on domain specific rules to answer questions. Skills related to this type of inferencing are referred by Gagné, Briggs, and Wager (1992) as *concrete* and *defined concepts* and are related to the ability to *identify* and *classify* specific individuals as members of particular concepts. Nevertheless, domain ontologies are capable of specifying declarative knowledge only and thus they cannot be used alone for assessing higher order cognitive skills such as problem solving (Holohan, Melia, McMullen, & Pahl, 2006). Thus, reasoning is applied for class subsumption relationships identification and classification of individuals. Text-based strategies are applied for identifying the semantics of both correct answers and distractors for each question item. Then, simple NLG techniques are applied for creating the actual items.

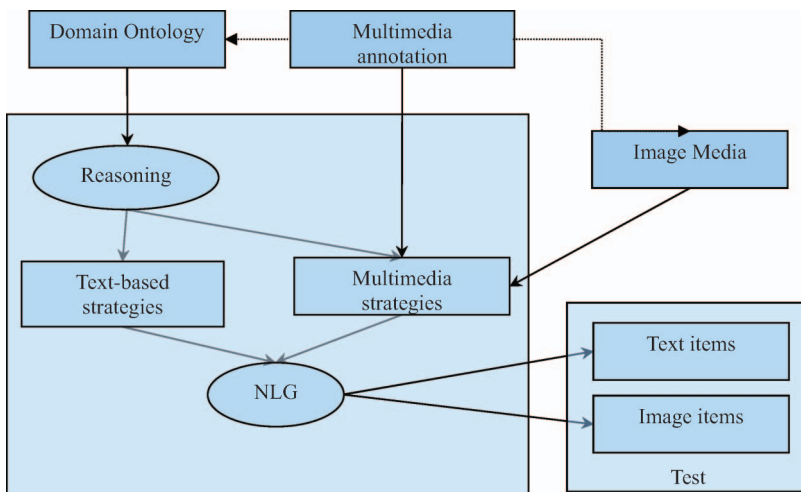


Figure 1. Tests generation architecture.

Media-rich questions are generated from a knowledge repository, which contains the following:

- a domain specific ontology which is composed by an ontology scheme, i.e. classes and relationships between those classes in a subject domain and a set of individuals;
- a set of images which are related to the aforementioned domain;
- annotations that relate images, as well as particular image regions to individuals of the domain ontology.

The proposed method is anticipated to be feasible for test generation in real educational settings due to the fact that domain-specific ontologies can be available in the following ways:

- The manual summarization of a domain by ontology engineers and domain/pedagogic experts, as performed in this work;
- The manual summarization of a domain in the form of concept maps, generated either by the teacher, or typically, collaboratively by the students themselves. Concept maps can be converted to ontologies by using appropriate tools (Simón, Ceccaroni & Rosete, 2007);
- The automatic ontology generation from text using appropriate tools. Although this constitutes an open research problem, it is possible to create domain ontologies from text corpora such as textbooks, manuals and tutorials (Zouaq & Nkambou, 2008).
- The reuse of ontologies created for specific educational technology purposes, such as educational content organization (Boyce & Pahl, 2007), searching and planning, (Dicheva & Dichev, 2006; Karampiperis & Samson, 2004) and knowledge representation for intelligent tutoring systems.
- The reuse of existing domain ontologies for educational purposes. These ontologies are typically created by domain experts, e.g. as annotations of historical archives and museum digitized resources (Trant, Bearman, & Richmond, 2002). Assessment and entertainment activities can be supported by automatically generating tests based on such existing ontological descriptions.

Furthermore, multimedia ontologies are becoming available for the annotation of digital libraries, archives, and multimedia databases to enhance human user and agent access to multimedia content (Garcia, & Celma, 2005; Dasiopoulou, Tzouvaras, Kompatsiaris, & Strintzis, 2008; Hunter, 2001; Tsinaraki, Polydoros, & Christodoulakis, 2007). Such ontologies are mostly based on the MPEG-7 standard and can also be reused for educational purposes.

#### **4. Text-based multiple-choice question generation**

Questions are generated by the application of specific strategies. Strategies presented in this section deal only with the semantics and not with the syntactic aspects of text-based question formation. Sentence generation is discussed in section 6. All strategies were selected, so as to provide distractors semantically as similar as possible to the correct answers, so that they successfully mislead students not

knowing the correct answer. In the following subsections, these strategies are presented together with examples taken from a domain ontology in the Greek ancient history domain called “Eupalinos Tunnel,” which is partially illustrated in Figure 2. In this figure, a rectangle denotes a class or individual, a line with an *io* symbol denotes an *instance\_of* relationship between a class and an instance, a line with an *isa* symbol denotes a subclass relationship between two classes and finally a dashed line with a role attached denotes a binary relationship between two classes/instances. In the examples presented, individuals are typeset in italics, while classes and properties are typeset in larger fonts, e.g. *Eupalinos* and *EupalinosTunnel* are both individuals, **Engineer** and **Tunnel** are both concepts that describe the type of the related individuals, respectively (i.e. *Eupalinos* is an **Engineer** and *EupalinosTunnel* is a **Tunnel**), and **constructed** (*Eupalinos*, *EupalinosTunnel*) is a property that relates two individuals (binary role).

Strategies are distinguished into three major categories, depending on the elements in the ontology that are used to extract the appropriate knowledge for sentence creation.

#### 4.1. Class-based strategies

This category contains strategies that generate distractors based on classes and their individuals. For all strategies, correct answers are of the following type: “Instance *a* is a class *A*,” e.g. “*Eupalinos* is an **Engineer**.” In ontology engineering terms, the

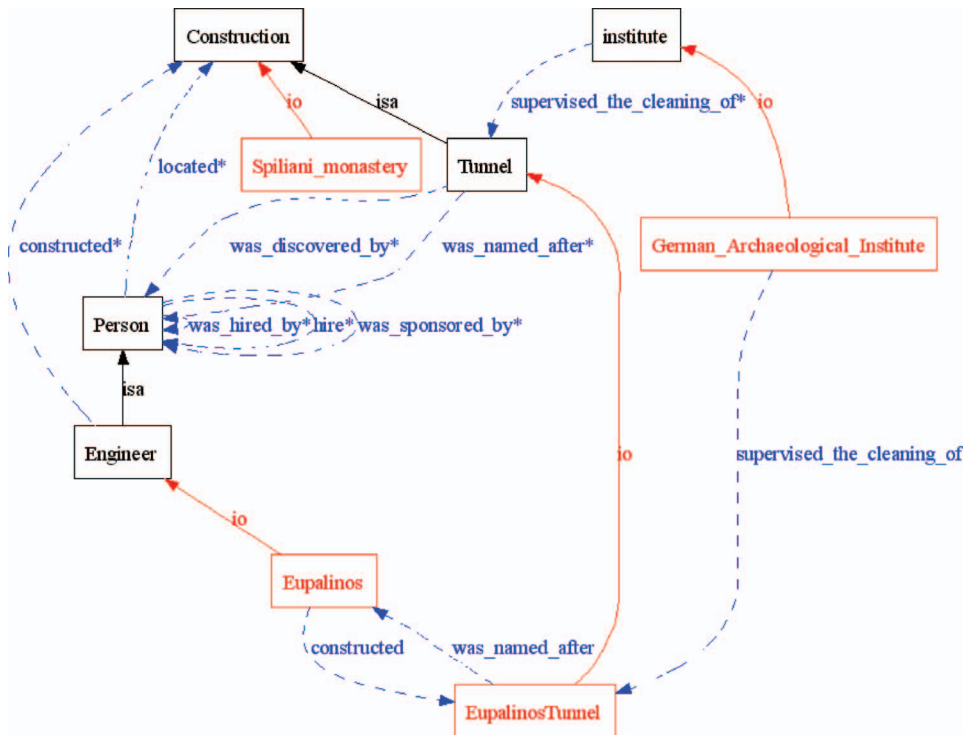


Figure 2. Example snapshot from the *Eupalinos Tunnel* domain ontology.

above sentence means that *Eupalinos* is an instance of **Engineer** or that *Eupalinos* is of type **Engineer**.

Distractors are formed by creating sentences in the same format as the correct answer, by choosing proper individuals or classes different than those that appear in the correct answer.

*Strategy 1.* Choose individuals, which are not members of a given class, provided that they are members of one of its superclasses. More specifically, if  $A(a)$  for some  $a$ , then correct answer is  $A(a)$ . For the distractors selection, we assume that  $B$  is a superclass of  $A$ . Then, if  $B(b)$ ,  $b \neq a$  and  $b$  is not an individual of  $A$ , then  $A(b)$  is a distractor.

Example: “*Ampelos Hill* is a **Mountain**” is the correct answer, since *Ampelos Hill* is an instance of class **Mountain**. “*North Opening of main tunnel* is a **Mountain**” is a distractor, since *North Opening of main tunnel* is an instance of concept **Location**, which is a superclass of class **Mountain**. As shown in the example, distractors formed by this strategy differ from the correct answer in the name of individual used as subject.

*Strategy 2.* Choose individuals belonging to disjoint siblings of a given class, to generate distractors. If  $A(a)$ ,  $b \neq a$ ,  $B(b)$  and  $B$  is a sibling class (disjoint or not) of  $A$ , then  $A(b)$  is a distractor. Note that if  $B$  is not a disjoint sibling of  $A$ , the strategy still applies, but in this case, there is a possibility that a distractor is a correct answer, which leads to an invalid distractor.

Example: “*Eupalinos* is an **Engineer**” is the correct answer, since *Eupalinos* is an instance of class **Engineer**. “*Polykrates* is an **Engineer**” is a distractor, since *Polykrates* is an instance of concept **Politician** and **Engineer** and **Politician** are disjoint subclasses of class **Person**.

*Strategy 3.* Choose individuals belonging to a class that has a non-empty intersection with a given class. If  $A(a)$ , there exist an individual  $b$  different than  $a$  and a class  $B$ , such that  $B(b)$  and  $A(b)$  and there exists individual  $c$  such that  $c \neq a$  and  $B(c)$ , then  $A(c)$  is a distractor.

Example: “*Eupalinian Aqueduct* is a **Remarkable Achievement**” is the correct answer, since *Eupalinian Aqueduct* is an instance of class **Remarkable Achievement**. “*Dyros* is a **Remarkable Achievement**” is a distractor, since *Dyros* is an instance of class **Tunnel** and *Eupalinian Aqueduct* is a member of both classes **Tunnel** and **Remarkable Achievement** (**Tunnel** and **Remarkable Achievement** have a non-empty intersection).

*Strategy 4.* Choose sibling classes (disjoint or not) to a given class. If  $A(a)$ ,  $B$  is a sibling of  $A$ , disjoint or not, then  $B(a)$  is a distractor. Distractors differ from correct answers in the name of the class. This strategy is dual to Strategy 2.

Example: “*Eupalinos* is an **Engineer**” is the correct answer, since *Eupalinos* is an instance of class **Engineer**. “*Eupalinos* is a **Politician**” is the distractor, since **Engineer** and **Politician** are disjoint subclasses of class **Person**.

*Strategy 5.* Choose subclasses of a given class. If  $A(a)$  and,  $B$  is a subclass of  $A$  and  $a$  is not a member of  $B$  then  $B(a)$  is a distractor. Again, distractors differ from correct answers in the name of the class in generated sentence. This strategy is dual to Strategy 1.

Example: “*Aristrachus* is a **Person**” is the correct answer, since *Aristrachus* is a member of class **Person**. “*Aristarchus* is an **Engineer**” is a distractor, since **Engineer** is a subclass of **Person**.



#### 4.2. Property-based strategies

This category contains strategies that generate distractors based on properties. Correct answers are generated from property instances in the ontology,  $R(a,b)$ , i.e. individuals  $a,b$  related with property  $R$ .

*Strategy 6.* Choose individuals from a class equal or subclass of the domain of a given property. If  $R(a,b)$ ,  $c \neq a$  and  $c$  an individual of a class which is equal to or subclass of the domain of property  $R$  then  $R(c,b)$  is a distractor.

Example: “*Polykrates hired Eupalinos*” is the correct answer, since the domain of this property is class **Person**. “*Herodotus hired Eupalinos*” is a distractor, since *Herodotus* is an instance of class **Historian** which is a subclass of class **Person**.

*Strategy 7.* Choose individual members of a class which is equal or subclass of the range of a given property to generate distractors. If property  $R$ ,  $R(a,b)$ ,  $c \neq b$  and  $c$  an individual of a class equal or subclass of the range of property  $R$  then  $R(a,c)$  is a distractor.

Example: “*Polykrates hired Eupalinos*” is the correct answer, since the range of this property is class **Person**. “*Polykrates hired Herodotus*” is a distractor, since *Herodotus* is an instance of class **Historian**, which is a subclass of class **Person** (the property range).

*Strategy 8.* Choose a property having both domain and range equal or subclass of the domain and range of the property of the correct answer. More formally, if a property  $S$  has a domain and a range that are equal or subset to the domain and range of property  $R$  correspondingly and  $R(a,b)$  is a correct answer, then  $S(a,b)$  is a distractor.

Example: “*Eupalinian Aqueduct brings water to ancient city of Samos*” is the correct answer. “*Eupalinian Aqueduct leads to ancient city of Samos*” is a distractor, since property *leads to* has range **Location** and domain **Tunnel**.

*Strategy 9.* Choose a numeric *datatype property* value by taking multiples and submultiples of a given property value. This strategy is based on numeric datatype properties, i.e. on properties that relate individuals to numerical values.

Example: If “*Eupalinian Aqueduct years spent for completion 10*” is the correct answer, then “*Eupalinian Aqueduct years spent for completion 16*” is a distractor.

#### 4.3. Terminology-based strategies

Strategies in this category are based solely on concept/subconcept relationships, without dealing with ontology individuals at all.

*Strategy 10.* Choose sibling classes of a given class to substitute the *subject* of the correct answer sentence. If class  $A$  is a subclass of  $B$ , then  $A$  is a  $B$  is the correct answer. If  $C$  is a sibling of class  $B$  then a distractor is  $C$  is a  $B$ . Distractors differ from correct answers in the class name used as subject.

Example: “*Sovereigns are politicians*” is the correct answer, since **Sovereign** is a subclass of **Politician**. “*Monks are politicians*” is a distractor, since **Monk** is a sibling class of **Politician**.

*Strategy 11.* Choose sibling classes of a given class to substitute the *object* of the correct answer sentence. If class  $A$  is subclass of  $B$ , then, again,  $A$  is a  $B$  is a correct answer. If  $C$  is a sibling of class  $B$ , then a distractor is:  $A$  is  $C$ . Distractors differ from correct answers in the class used as object. Strategies 10 and 11 use the same

technique for substituting the subject and object of a correct answer, correspondingly, for producing distractors.

Example: “Sovereigns are politicians” is the correct answer, since *Sovereign* is a subclass of *Politician*. “Sovereigns are monks” is a distractor, since *Monk* is a sibling class of *Politician*.

## 5. Multimedia question items generation

Beyond text-based multiple-choice question items, our approach handles the generation of multimedia items. In the following paragraphs, we discuss additional strategies, to support the integration of multimedia content in questions based on simple multimedia semantics. As an example, we illustrate the semantic annotation of a sample image available in COMM web site<sup>1</sup> depicting a picture from Yalta Conference illustrated in Figure 3. A multimedia annotation file is also used as example, containing definitions of image regions illustrated in the picture, i.e. SR1, which corresponds to the face of Winston Churchill. This file also contains a mapping between image regions and individuals in the domain ontology, also illustrated in Figure 3, described in OWL format. In the domain ontology, “*Winston\_Churchill*” is an individual associated with the above still region.

### 5.1. Multimedia-based strategies

In this section, we describe strategies for multimedia-based question item generation. In order to generate questions, a combination of techniques from processing of

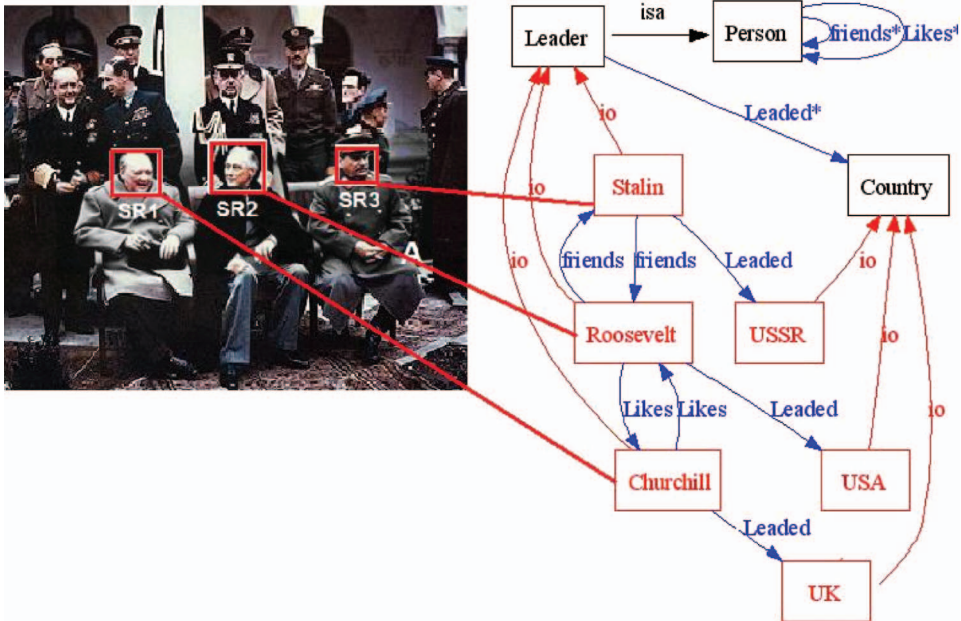


Figure 3. Yalta Conference example image<sup>1</sup> with three annotated regions associated with a snapshot of the domain ontology representing real world knowledge about the Yalta Conference event.

multimedia ontology files, basic image processing and NLG are incorporated. More specifically, we specify *image regions*, which are associated with specific domain ontology individuals. The questions generated belong to various types of image-related question items. In the following, we have adopted the terminology of the IMS QTI specification (IMS, 2005) for question types.

#### 5.1.1. MCQ with image

MCQ with images are types of questions presenting an image and prompting learners to select among one or more options concerning information related to the image. The following strategy is related to this kind of questions:

*Strategy 12.* Highlight an image region which is associated with an individual  $a$  of a class  $A$  in the domain ontology to indicate the area in question to the learner. Highlighting can be performed by an image filter algorithm, e.g. by changing the color of the region in question. Possible distractors are different individuals of class  $A$ , which are also illustrated in the picture.

Example: in the Yalta Conference picture, a question item generated with this strategy is “*The person highlighted in the picture is*”. If the highlighted person is “*Stalin*,” then possible distractors are “*Franklin Roosevelt*,” “*Winston Churchill*.”

#### 5.1.2. Hotspot interaction questions (single and multiple answers/responses)

In this type of questions, an image is presented to the learner, who is asked to identify a specific region in the presented image. One or multiple selections may be correct. Each region can be marked by a hot spot, which is a visual sign, indicating an image region, to facilitate selection. Alternative false hotspots can be given as distractors. The following strategies apply to this kind of questions:

*Strategy 13.* Identify an image region instantiated in the multimedia ontology, which is associated with an individual  $a$  of class  $A$  in the domain ontology. Then, construct an item with stem: “*Identify  $a$  in the picture.*” Distractor hotspots correspond to individuals of class  $A$  or a superclass of  $A$ , different than  $a$  in the picture. An example output of this strategy is the stem “*Identify Winston Churchill in the above picture.*” Possible distractors are “*Stalin*” and “*Franklin Roosevelt*,” as illustrated in Figure 3.

*Strategy 14.* Identify an image region instantiated in the multimedia ontology. Find an individual  $a$  in the domain ontology, which is related with an individual  $b$  through an object property  $R$ , i.e.  $R(a,b)$  specified into the domain ontology. Distractor hotspots correspond to individuals different than  $a$ , belonging to the domain of property  $R$ .

Example: If *led* is an object property in the Yalta Conference domain ontology, and sentence *led* (*Stalin, USSR*), then “*Identify a person which led USSR in the picture*” is an example output of this strategy. Distractors are, again, the hotspots of Churchill and Roosevelt, individuals of class Person, depicted in the picture.

*Strategy 15.* Identify an image region instantiated in the multimedia ontology. Find an individual  $b$  in the domain ontology, which is an individual of class  $B$  and is related with an individual  $a$  through an object property  $R$ , i.e.  $R(a,b)$  is specified into the domain ontology. Distractor hotspots correspond to individuals different than  $b$  in the range of property  $R$ .

Example: If *likes* is an object property in the domain ontology describing Yalta Conference, and sentence *likes (Roosevelt, Churchill)* is included in the domain ontology, then “*Identify a person, which Roosevelt likes in the picture*” is an example output of this strategy.

Strategies 14 and 15 differ in the individual participating in the relationship. In Strategy 14, it is the first participant (role) in the relationship that is associated with a region in the picture, while in Strategy 15, it is the second participant in the object property. It thus leads to a difference in the form of generated sentence between the two strategies.

### 5.1.3. Other types of interaction

Graphic association interaction question items are types of questions prompting the learner to connect two regions of an image, which are associated with some semantic relationship. In a question delivery environment, a learner typically connects these regions by linking together appropriate hotspots, as these depicted in Figure 4. A strategy for generating questions of this type involves the generation of a graphic association item for connecting hotspots that correspond in image regions associated with a specific property. Distractors are formed by choosing image regions or hotspots referring to individuals belonging to the domain or the range of the object property.

Although not implemented in current version of our prototype, we provide an example of this strategy:

If *likes (Roosevelt, Churchill)* is a fact in the domain ontology, as presented in previous examples, then a stem for an item of this type can be “*Connect Roosevelt*

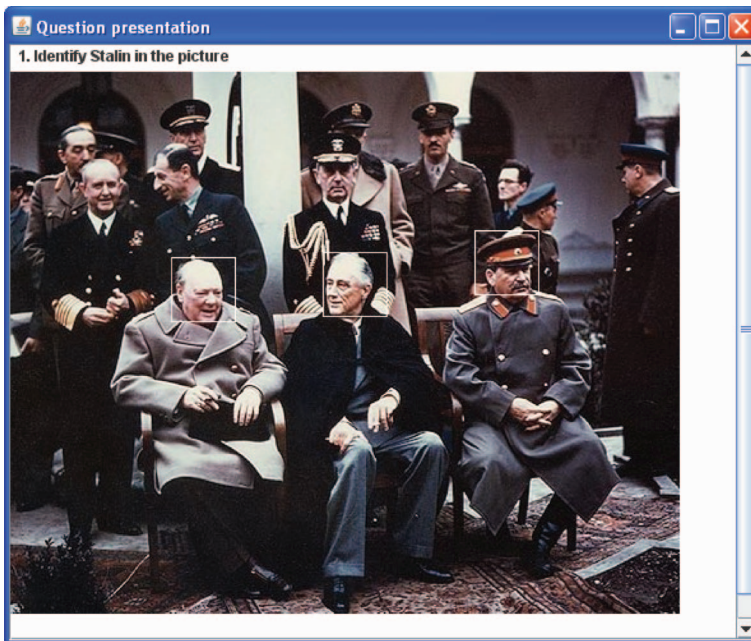


Figure 4.<sup>1</sup> An example of Strategy 13. Hotspots are displayed as rectangles.

with a person which Roosevelt likes.” This strategy is similar to Strategy 15 above in the sense that they share the same semantics in terms of both domain and multimedia ontologies. They differ only in the type of generated item.

Spatio-temporal semantics of multimedia annotations in combination with domain ontologies can be used for utilizing other types of media, such as video and audio, toward more sophisticated ways of assessment. This can be achieved by generalizing the above strategies, so that they use other types of media segments apart from still regions. Thus, questions can be generated using video or audio segments associated with domain classes/individuals to be recognized in a video or audio sample, respectively. As an example of generalizing Strategy 12, after displaying a properly annotated video about dance positions, a question can have the form: “*After arabesque, the dancer performed.*” with possible answers “*Fouetté,*” “*En pointe,*” and “*Jeté.*” The other strategies can also be generalized. In this case, a question delivery system enabling real time interaction with video needs to be utilized.

## 6. Automatic question generation prototype

A prototype tool<sup>2</sup> was developed which accepts as input OWL documents and multimedia annotations and generates tests using the above strategies. The format of questions is illustrated in Figure 5.

In all types of text-based MCQ questions, the stem is “*Which of the following sentences is correct?*”. In question items based on instance-of relationships, i.e. for class-based strategies, declarations in the form **Concept (Individual)** in the ontology are transformed in sentences of the form “*Individual is a(n) Class*”. As an example, “*Eupalinos is an Engineer.*” For strategies based on subclass relationships, i.e. for questions generated by terminology-based strategies, class names appear in plural number. For questions generated by property-based strategies, sentences are

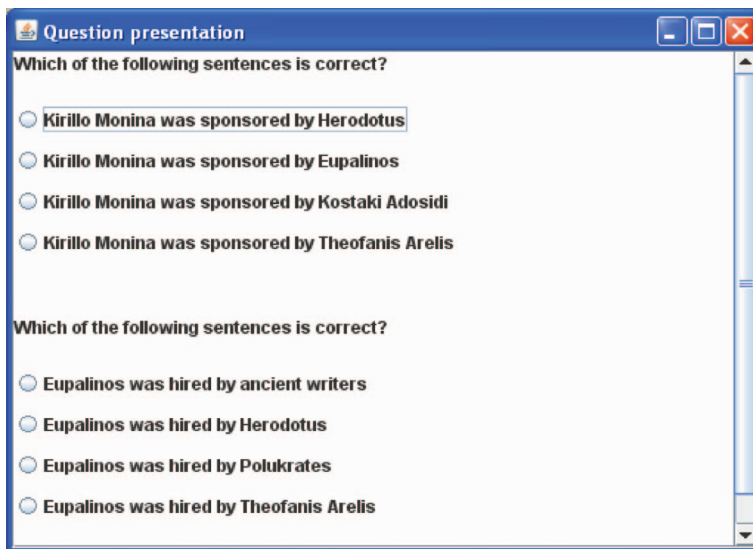


Figure 5. The MCQ generation output for a text MCQ item.

generated in the form “*Individual* propertyName *Individual*.” “was\_sponsored\_by” is the name of an object property and thus, an example of a generated sentence is “*Kirillo Monina was sponsored by Kostaki Adosidi*.” The name of the property is tokenized according to simple rules, e.g. underscore is recognized as a separating character. Sentence generation is performed using the SimpleNLG framework.<sup>3</sup>

Input ontologies should adhere to certain conventions to generate syntactically correct sentences, with the NLG techniques adopted. Properties’ names should be written as verbs or verb-like phrases. The names of classes, individuals and properties must contain words connected with underscores, hyphens, or multiple concatenated words with first word letter capital, such as “hasValue.”

The MCQ presentation application is implemented in Java. JENA Semantic Web framework<sup>4</sup> is used for OWL ontologies management and storing and thus for the implementation of strategies presented in previous section. Pellet open source Description Logics reasoner<sup>5</sup> is used as an inference engine, i.e. for subsumption and automatic classification of individuals in the ontology. Both text and multimedia (image) question items are presented using the Java Swing graphics toolkit. This solution was preferred to exporting tests in QTI format, due to the fact that we were not able to find an (open source) QTI player implementation to support images.

Regarding image question generation, we have been experimenting with the COMM API, as well as with other semantic annotation multimedia content solutions, such as PhotoStuff<sup>6</sup>, to semantically annotate multimedia objects (currently images) using both domain and multimedia knowledge represented in domain and multimedia ontologies respectively. Although COMM API is a modern approach toward standardization of semantic multimedia annotation, it was not used eventually in the extended approach for multimedia-based question items construction due to incompatibility with the ontology management framework<sup>4</sup> that we were using for the domain ontologies. Thus, we have annotated the example image with the same domain ontology using the PhotoStuff tool. Although the MPEG-7 standard is not fully supported in the multimedia ontology of this tool, there is adequate knowledge represented for experimenting with spatial semantics. Other MPEG-7 based descriptions can be supported as well, though not in the present version of our tool.

## 7. Evaluation

A number of ontologies from different domains were used for evaluating the proposed approach for text-based questions (Strategies 1–11). Some metrics pertaining to the number of classes, individuals, and properties contained in these ontologies are presented in Table 1. From the example ontologies, *Eupalinos Tunnel*

Table 1. Example ontologies used for evaluating the proposed approach.

	Eupalinos tunnel	MSc program	Travel v.1	Travel v.2	Grid resources	Food ontology
Individuals	40	100	145	38	21	12
Classes	29	25	72	21	54	63
Object prop.	25	38	22	13	26	4
Datatype prop.	16	23	13	0	24	0

ontology was developed by domain experts and ontology engineers as a test-case for the method presented in this article.

The generated tests were evaluated in three dimensions: Pedagogical quality, linguistic/syntactical correctness, and number of questions produced. These dimensions were considered for each strategy category.

The generated questions from the *Eupalinos Tunnel* ontology were reviewed by two domain/educational experts. All questions were found satisfactory for assessment by the experts. Nevertheless, all questions are not syntactically correct. Especially, *Travel v.1* and *MSc Program* ontologies have particularly poor performance, because in these, the names of various elements are far from the conventions required for NLG, which were described in previous section.

Table 2 depicts the number of syntactically correct items and the total number of questions per strategy for each domain ontology. For the sake of comparison, first row in the above table displays results for the *Eupalinos* ontology without reasoning (automatic classification and computation of instances' inferred types), while second row displays results after applying reasoning. The same comparison is displayed for two more ontologies, *Food* and *Grid Resources*. The application of reasoning has two main consequences. First, the total number of items increases, since inferred individuals and subsumed classes are taken into account. Second, not only direct subclasses/superclasses are taken into account but also the subsumed ones. As a result, more distant classes are used in generating distractors, which are semantically distant from the correct answer. For example, Strategy 4, after applying reasoning, generates distractor "*Pantheon is a subtunnel*," which is much distant from the correct answer "*Pantheon is a unique construction*" and thus is easier to be rejected by the learner. However, there is a trade off between shortage of distractors and availability, albeit of easy ones.

Meaningfulness depends heavily on the ontology structure. If input ontological descriptions are correct, then OWL management mechanisms guarantee the meaningfulness of the question items generated. When items are produced from concepts with large semantic distance, then items are easier, as stated before. Strategies involving individuals seem to produce more difficult questions than terminology-based ones. Property-based are also more difficult than class-based, since they involve the association between concepts or individuals, thus requiring a more thorough understanding of a particular domain.

Property-based strategies may produce a large number of MCQ but are very difficult to manipulate syntactically. Class and terminology-based strategies on the

Table 2. Multiple choice question items per question generation strategy.

	1	2	3	4	5	6	7	8	9	10	11	Total
Eupalinos T.	3/3	5/8	5/5	5/8	1/2	0/0	0/0	3/3	0/6	3/3	3/3	66/88
Eupalinos T-R.	9/12	6/9	6/7	12/12	3/3	32/41	6/6	3/3	0/6	3/3	3/3	83/99
Travel 1		3/17		5/14		1/3	0/39		0/55	19/19	19/19	47/166
MSc Program		3/6		0/5					0/59			3/70
Grid Res.				0/1					0/1	6/6	6/6	12/14
Grid Res.-R.	5/6	14/16		2/18	10/13				0/1	17/22	13/20	61/96
Food				1/2						7/19	6/19	14/40
Food-R.	8/11	6/15		8/13	4/6					31/40	38/52	95/137

Table 3. Ontologies/images for multimedia strategies evaluation.

	Classes	Individuals	Properties	Images	Annotated regions
Eupalinos tunnel	29	40	26	17	21
Yalta conference	18	27	9	13	20

Table 4. Multimedia items per question strategy.

	12	13	14/15
Eupalinos tunnel	8/11	0/2	0/1
Yalta conference	20/25	11/15	5/10

other hand are much easier to handle syntactically but generate fewer questions for ontologies of the same depth and population.

While the proposed approach works well in defining the semantics of questions, the problem of generating syntactically correct question items is only partially tackled. Evaluation of multimedia strategies was based on two domains: *Yalta Conference* and *Eupalinos Tunnel*. A number of related images were annotated using PhotoStuff and the related domain ontologies.

For multimedia strategies evaluation we run experiments with ontologies/images presented in Table 3, using Strategies 12–15. Question items generated were evaluated based on the same criteria as in text-based generated question items. Table 4 illustrates the results per strategy.

As it is shown, some strategies generate more than one question items per image. The users of the application must select items according to the pedagogical needs of a particular test. As previously, some items are not correct, because property or individual names do not conform to certain conventions previously described. Furthermore, there are images in the form of a map in the *Yalta Conference* domain, in Strategy 13, where countries to be identified are already labeled in the map. This kind of problem is expected to occur in other cases as well. Nevertheless, question items are meaningful and adequate for assessment use. The more engaging and difficult strategies are 14/15 (grouped in Table 4) since they engage relationships, followed by Strategy 13, which involves only subsumption relationships and individuals.

## 8. Conclusions

Domain and multimedia ontologies can be used to automatically produce multimedia-rich tests. In this article, a novel approach for automatic generation of tests for self-assessment has been presented. The proposed approach is based on strategies that use ontological axioms and asserted/inferred knowledge (text and multimedia) of a knowledge base developed in OWL. A prototype tool has been developed for evaluation reasons, proving that the proposed approach, when used with semantically rich and fully populated domain ontologies, can provide successful cases.

Future work on this direction includes the pedagogic evaluation of the approach with elementary school students, as well as improvements on the NLG subsystem.



Furthermore, initial experiments have been also started to provide an automatic mechanism for enriching domain ontologies either in the level of concepts/properties (ontology learning) or in the level of individuals (ontology population). Our first experiments on this direction include information sources, such as the Web, using search engines, such as Google, to “fish” individuals from online lexicons or other linguistic resources.

The list of strategies, which are described here, should not be considered as exhaustive. Other strategies can be further implemented, which utilize more expressive constructs of OWL, such as defined classes and cardinality constraints, or even formalisms supporting rules, e.g. SWRL (W3C, 2004b). Moreover, as already discussed, support for spatial and temporal annotations is planned to generate assessment activities incorporating video, music and sound, based on semantically annotated media.

### Acknowledgments

The authors thank Professor George A. Vouros for his comments and suggestions on this work. They appreciate contribution of domain experts and school teachers in the evaluation process as well as in the refinement of the experimental ontologies. Especially, they thank the archeology, museum studies, and educational programs’ specialist Mrs Stamatia Ladikou and the pre-school teacher and educational programs’ specialist Mrs Aikaterini Koti for their contribution on generated tests’ evaluation. They also thank the anonymous reviewers for their valuable comments.

### Notes

1. <http://comm.semanticweb.org/>. This image is a work of the US Federal Government.
2. <http://www.samos.aegean.gr/math/andpapas/mmcq/>
3. <http://www.csd.abdn.ac.uk/~ereiter/simplenlg/>
4. <http://jena.sourceforge.net/index.html>
5. <http://pellet.owldl.com/>
6. <http://www.mindswap.org/2003/PhotoStuff/>

### Notes on contributors

Andreas Papasalouros is a Lecturer at the Department of Mathematics, University of the Aegean, Greece. He holds a BSc in Physics from University of Athens and a BSc and a PhD in Electrical and Computer Engineering from National Technical University of Athens. His research interests include Educational Technology, Hypermedia Design and Web Engineering. He has participated in several European and national research and development projects and is co-author of a number of articles published in scientific journals, book chapters and conference proceedings.

Konstantinos Kotis holds a BSc in Computation from UMIST (University of Manchester), UK (1995), and a PhD in Knowledge Representation and Management from the Department of Information and Communication Systems Engineering, University of the Aegean, Greece (May 2005). Currently, he is a lecturer and research fellow in AI Lab of the Department of Information and Communication Systems Engineering. His research and published work concerns Ontology Engineering and Semantic Web (Grid, Search, and Services) technologies. His published scientific work includes more than 30 peer reviewed articles, including book chapters, journal and international conference articles in the above mentioned themes. He serves as member of organizing and program committees of international conferences on related topics as well as member of editorial boards of international scientific journals.

Konstantinos Kanaris holds a BSc in Mathematics and an MSc in Technologies and Management of Information and Communication Systems from University of the Aegean. He has participated in various projects as software developer. His research interests include Natural Language Processing and Educational Technology.

## References

- Arndt, R., Staab, S., Troncy, R., & Hardman, L. (2007). Adding formal semantics to MPEG-7: Designing a well-founded multimedia ontology for the web (Technical Report No. 4/2007). Koblenz: University of Koblenz.
- Baader, F., Calvanese, D., McGuinness, D., Nardi, D., & Patel-Schneider, P. (eds.). (2003). *The description logic handbook: Theory, implementation, and applications*. New York, NY: Cambridge University Press.
- Bateman, J. (1997). Enabling technology for multilingual natural language generation: The KPML development environment. *Natural Language Engineering*, 3, 15–55.
- Bontcheva, K., & Wilks, Y. (2004). *Automatic report generation from ontologies: The MIAKT approach*. Paper presented at the ninth international conference on applications of natural language to information systems (NLDB'2004), Manchester, UK.
- Boyce, S., & Pahl, C. (2007). Developing domain ontologies for course content. *Educational Technology & Society*, 10, 275–288.
- Dasiopoulou, S., Tzouvaras, V., Kompatsiaris, I., & Strintzis, M. (2008). Capturing MPEG-7 semantics. *Metadata and Semantics* (pp. 113–123). New York, NY: Springer.
- Dicheva, D., & Dichev, C. (2006). TM4L: Creating and browsing educational topic maps. *British Journal of Educational Technology*, 37, 391–404.
- Gagné R., Briggs, L., & Wager, W. (1992). *Principles of instructional design* (4th ed.). New York, NY: Harcourt Brace.
- Garcia, R., & Celma, Ó. (2005). *Semantic integration and retrieval of multimedia metadata*. Paper presented at the fifth international workshop on knowledge markup and semantic annotation (SemAnnot 2005), Galway, Ireland.
- Holohan, E., Melia, M., McMullen, D., & Pahl, C. (2005). *Adaptive e-learning content generation based on semantic web technology*. Paper presented at the workshop on applications of semantic web technologies for e-learning (SW-EL@ AIED'05), Amsterdam, The Netherlands.
- Holohan, E., Melia, M., McMullen, D., & Pahl, C. (2006). *The generation of E-Learning exercise problems from subject ontologies*. Paper presented at the sixth international conference on advanced learning technologies (ICALT'06), Kerkrade, The Netherlands.
- Hunter, J. (2001, July–August). *Adding multimedia to the semantic web – Building an MPEG-7 ontology*. Paper presented at the International semantic web working symposium (SWS), Stanford University, California, USA.
- IMS. (2003). *IMS simple sequencing specification*. <http://www.imsglobal.org/simplesequencing/>
- IMS. (2005). *IMS question and test interoperability implementation guide. Version 2.0 final specification*. [http://www.imsglobal.org/question/qti\\_v2p0/imsqti\\_implv2p0.html](http://www.imsglobal.org/question/qti_v2p0/imsqti_implv2p0.html)
- ISO. (2004). MPEG-7 overview. Retrieved from <http://www.chiariglione.org/mpeg/standards/mpeg-7/mpeg-7.htm>
- Karampiperis, P., & Sampson, D. (2004). Adaptive instructional planning using ontologies. *Proceedings of international conference on advanced learning technologies* (pp. 126–130). Washington, DC: IEEE Computer Society.
- Li, T., & Sambasivam, S. (2005). Automatically generating questions in multiple variables for intelligent tutoring. *The Journal of Issues in Informing Science and Information Technology*, 2, 471–480.
- Lopez, V., Uren, V., Motta, E., & Pasin, M. (2007). AquaLog: An ontology-driven question answering system for organizational semantic intranets. *Journal of Web Semantics*, 5, 72–105.
- Martínez, J. (2002). Overview of MPEG-7 description tools, Part 2. *IEEE Multimedia*, 9, 83–93.
- Miller, G. et al. (1995). *Five papers on WordNet* (CSL Rep. No. 43). Cognitive Science Laboratory, Princeton University.
- Mitkov, R., Ha, L.A., & Karamanis, N. (2006). A computer-aided environment for generating multiple-choice test items. *Natural Language Engineering*, 12, 177–194.
- Simón, A., Ceccaroni, L., & Rosete, A. (2007). Generation of OWL ontologies from concept maps in shallow domains. In D. Borrajo, L. Castillo, & J.M. Corchado (Eds.), *CAEPIA 2007, LNAI 4788*. Heidelberg: Springer.
- Soldatova, L., & Mizoguchi, R. (2003). Ontology of tests. *Proceedings of Computers and Advanced Technology in Education* (pp. 175–180). Calgary, Canada: Acta Press.

- Soldatova, L., & Mizoguchi, R. (2007). Testing and understanding by use of an ontology methodology. *Proceedings of Joint Workshop of Cognition and Learning* (pp. 202–205). Berlin.
- Trant, J., Bearman, D., & Richmond, K. (2002). Open concepts: Museum digital documentation for education through the AMICO Library. *Visual Resources Association Bulletin*. Retrieved from <http://www.amico.org/docs/papers/2002/ArtDoc0302.pdf>
- Tsinaraki, C., Polydoros, P., & Christodoulaki, S. (2007). Interoperability support between MPEG-7/21 and OWL in DS-MIRF. *IEEE Transactions on Knowledge and Data Engineering*, 19, 219–232.
- Tsumori, S., & Kaijiri, K. (2007). System design for automatic generation of multiple-choice questions adapted to students' understanding. *Proceedings of the eighth international conference on information technology based higher education and training* (pp. 541–546). Kumamoto, Japan.
- W3C (2004a). *OWL web ontology language overview, W3C recommendation*. Retrieved from <http://www.w3.org/TR/owl-features/>
- W3C (2004b). *SWRL: A semantic Web rule language combining OWL and RuleML*. Retrieved from <http://www.w3.org/Submission/SWRL/>
- Zouaq, A., & Nkambou, R. (2008). Building domain ontologies from text for educational purposes. *IEEE Transactions of Learning Technologies*, 1, 49–62.

Copyright of Interactive Learning Environments is the property of Routledge and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.