Interactive Learning Environments
Vol. 19, No. 1, January 2011, 25–44

Semantic annotation of video fragments as learning objects: a case study with *YouTube* videos and the Gene Ontology

Elena García-Barriocanal[a], Miguel-Angel Sicilia[a]*, Salvador Sánchez-Alonso[a] and Miltiadis Lytras[b]

[a]*Information Engineering Research Unit, University of Alcalá, Polytechnic Building, Ctra. de Barcelona, Km.33,6–Alcalá de Henares, Spain;* [b]*American College of Greece, Gravias 6 St., Aghia Paraskevi, Greece*

(*Received 10 July 2009; final version received 11 January 2010*)

Web 2.0 technologies can be considered a loosely defined set of Web application styles that foster a kind of media consumer more engaged, and usually active in creating and maintaining Internet contents. Thus, Web 2.0 applications have resulted in increased user participation and massive user-generated (or user-published) open multimedia content, some of which is potentially useful for education. In this context, the problem from the educator's perspective is selecting and annotating existing content so that software applications can filter fragments that were previously marked as useful for particular learning needs. This article discusses a solution for this problem that is non-intrusive to existing applications. This solution fits in the philosophy of multiple metadata profiles, allows for expressing fine-grained learning needs, and leverages the growing mass of contents by reusing well-established domain ontologies. A description of the technical aspects required for the infrastructure supporting such solution is first provided. Then, the solution is contextualized with a case study using a knowledge model of gene-related elements – the Gene Ontology – to semantically annotate videos in *YouTube* that could be used in teaching biology and bioinformatics.

Keywords: Web 2.0; video clips; multimedia; learning objects; ontologies; semantic annotation; *YouTube*; gene ontology

1. Introduction

Web 2.0 is a term, which refers to several things, including behaviors, technologies, and ideals. It considers services and activities that foster a new kind of media consumer who is more engaged and active in creating and adding value to the content which is the basis for using the Internet (Allen, 2008). Thus, the key of Web 2.0 applications from the viewpoint of digital resource production is that they enable mass participation in social activities structured around the contents. User-generated content (UGC) consists mostly of micro-contents, which can be considered fine-grained learning objects (LOs), provided they have the minimal accuracy and reliability requirements to be appropriate for education in participatory pedagogies

*Corresponding author. Email: msicilia@uah.es

ISSN 1049-4820 print/ISSN 1744-5191 online

DOI: 10.1080/10494820.2011.528879
http://www.informaworld.com

(Collis & Moonen, 2008). In some cases, users do not create the contents themselves, but play the role of publishers, annotators and reviewers of contents already available. Therefore, Web 2.0 applications become massive repositories in which users share the results of their use of Web resources. A good example of this is the hundreds of video clips about deoxyribonucleic acid (DNA) in *YouTube*, which can be annotated, commented or rated.

It is extremely attractive to reuse the growing mass of UGC, selecting content subsets that are known to be accurate and useful for particular educational needs. These resources are publicly available right now at negligible development costs. Considering that designing and implementing quality materials is one of the most expensive parts of e-learning courses' budget, teachers would greatly benefit if they did not have to develop new materials – a very costly choice when we speak of multimedia resources – but only to adapt pre-existing resources to each specific learning context or situation. In the following pages, we will detail a form of reusing user-generated videos through the provision of metadata information in the form of semantically meaningful annotations.

There are several approaches to the annotation of digital video in the literature. Feng, Manmatha, and Lavrenko (2004) introduced an automatic image and video annotation technique for retrieval based on textual queries, where the images forming the sequence are partitioned into regions. Their motivation was to improve the conventional approach (i.e. keyword-based manual annotation of all the images carried out by librarians), considered as too expensive. Their model was based on a fixed word vocabulary representing some informal specialization hierarchy, e.g. "face, male face, male news subject" could be a branch in the tree. Then, a probabilistic model is used to automatically assign words in the vocabulary to media files, depending on the objects automatically recognized through existing algorithms. Such kind of models exhibit reasonable performance in fairly homogeneous collections, but as long as they rely on a training process, which uses previously annotated media files, the quality of the approach critically depends on the coverage of the training set used. This often produces good results for closed domains, but it is less likely to be useful in open collections like Web 2.0 sites storing UGC. Miyamori and Iisaku (2000) proposed an automatic annotation method of sports video for content-based retrieval, using models of the behavior of players, e.g. characterizing what in tennis are "smashes" or "over-the-shoulder swings". Such models are again well equipped for homogeneous collections in which the characteristics of behaviors of the objects (e.g. players, balls) can be pre-categorized and defined in terms of image or video frame properties of certain regions in the media. Another interesting research direction that in contrast integrates well with Web 2.0 applications was pointed out by Bargeron, Gupta, Grudin, and Sanocki (1999), who introduced MRAS, a prototype of a collaborative video annotation system. This system was intended for personal use – annotations were not intended to serve for future processes of information management – but revealed the possibilities of annotating video contents. A similar effort worthy of comment is the one by Volkmer, Smith, and Natsev (2005) who "describe and evaluate a web-based system for collaborative annotation of large collections of images or temporally pre-segmented videos". Just to mention a few other significant contributions to the annotation of video, those about speech (Brown et al., 2001) and text (Lienhart, 1996) deserve recognition.

Our main motivation here is to describe a way of adapting and transforming existing video sequences, not specifically created for learning, into ready-to-use learning resources. But to achieve the goal of reusing UGC, we first need to study to what extent UGC is reliable and accurate enough to be used for learning. It is reasonable to assume that part of the UGC in Web 2.0 applications is about information potentially useful for educational purposes. However, the information available is often diverse, and in some cases problematic or even contradictory. The following are relevant examples:

- According to a recent study, which examined a total of 146 unique *YouTube*[1] video clips about human papillomavirus (HPV) vaccination, only 74% portrayed HPV vaccination in a positive manner (Ache & Wallace, 2008).
- For some topics, such as contraception, most scientific claims in *YouTube* have been considered as unsubstantiated (Luttrell, Zite, & Wallace, 2008).
- In contrast, for other topics such as surgical procedures, the information in Web 2.0 sites is mostly accurate. Devgan, Powe, Blakey, & Makary (2007) identified that about 86% of 35 procedures in the Wikipedia were deemed to be appropriate "for patients".
- Regarding the information categorization typical of Web 2.0 systems, there is evidence that ambiguity and polysemy in the structure of the folksonomy tags are indeed important problems according to Spiteri (2007). Scattered analyses by this author portray a very diverse landscape of quality and accuracy in the contents and organization of Web 2.0 applications.

Articles on the use of Web 2.0 technologies in education are increasingly frequent in the literature, even though there is a need to discriminate sound research from hype (Rollet, Lux, Strohmaier, Dösinger, & Tochtermann, 2007). In any case, the role of these technologies in participatory pedagogies has been profusely discussed (Collis & Moonen, 2008; Eijkman, 2008). Web 2.0 sites like *YouTube* have been found to have a considerable potential in education (Milliken, Gibson, O'Donnell, & Singer, 2008), so that they have been used as the point of departure for novel trends in educational innovation (Jenkins, 2007). In this article, we focus only on the potential of UGC as a large base of contents from which educationally meaningful resources can be selected. However, the educational use of the mass of openly available UGC is not straightforward, for some reasons including the following:

- Web 2.0 applications are by definition not quality controlled, except by the social assessment of the contents. Some basic control includes regulations, but these are not related to quality but to law enforcement.
- As learning needs are diverse, the same resource may be useful for different contexts and under different usage conditions. This idea of multiple uses is central to the concept of metadata profile (Downes, 2004) and compatible with the philosophy of many educational repositories, which only store the metadata but not the resources themselves, so different repositories can hold different metadata descriptions for the same resource.

Therefore, there is a need to build filters for information in Web 2.0 systems, which enable the general public to select the appropriate resources for a particular use. In the case of education, instructors would benefit of filters set by others for explicit educational purposes. This kind of peer-review mechanism is currently practiced in LO repositories that only store metadata (not the contents themselves) such as Merlot (Cafolla, 2002). The idea is to link metadata annotations to existing resources in a non-intrusive way, i.e. external systems managing the resources. This filtering requires the provision of metadata storage facilities connected to classification systems offering an unambiguous way of describing learning needs. Existing formal ontologies provide such descriptive framework, e.g. ontology-based approaches to storing LO metadata like the one reported by Gaševic, Jovanovic, and Devedžic (2007). In the case of continuous media such as video clips, it is also a need to formally annotate resources (e.g. a fragment of a video clip) and link them to existing multimedia synchronization languages to combine them, which can be achieved using multimedia metadata ontologies (Hunter, 2006). This way, ontologies can be employed as a filtering mechanism for the increasing amount of UGC available in Web 2.0 applications, whereas educational considerations can be used in the filtering process by means of formally representing some pedagogical concerns.

This article reports on an approach to ontology-based annotation of media fragments. This approach allows not only multiple descriptions but also links from those descriptions to any arbitrary domain ontology, including a knowledge basis about educational considerations. We focus on video clips as can be found in popular Web 2.0 applications like *YouTube, Vimeo*, or *MySpace*, which can be potentially reused in pedagogies considering multimedia (Mayer, 2005). Our approach builds on existing efforts in semantic LO metadata repositories and the possibilities offered by service-oriented architectures (Monceaux et al., 2007). The case of video clips on molecular genetics is used as an illustration of the main elements of the approach.

The rest of this article is structured as follows. Section 2 provides some background information on multimedia and ontologies in learning technology. Section 3 describes the general approach and technical elements, and Section 4 describes the case study. Finally, conclusions and future work are provided in Section 5.

2. Background and related work

The use of video clips for the design of LOs has been described in some scattered literature. For instance, Cochrane (2007) reports on the use of action research for devising pedagogically sound interactive LOs using QuickTime technology. However, the challenge of ontology-based annotation of media fragments has not yet been explicitly addressed.

There are several languages and metadata schemas specific for multimedia objects. For example, the *Synchronized Multimedia Integration Language* (SMIL),[2] is an HTML-style language which enables the authoring of multimedia presentations. Even though such markup languages enable loose composition of media objects, nowadays they are rarely used in Web 2.0 community sites like *YouTube*. In contrast, these sites use specific video formats such as FLV, optimized for quick rendering. SMIL provides some media control through the `clipBegin` and `clipEnd` attributes, along with the mechanisms to pass parameters to media players. Then, existing clips annotated for educational purposes can be integrated in synchronized SMIL presentations if there is information available about both timing and educational properties. In consequence, the techniques described below can be integrated with authoring environments in which they could play the role of semantic content search mechanisms.

The discussion below assumes concepts on LOs and their relationship to continuous media, as well as some basic knowledge on formal ontologies. These elements are briefly surveyed in the rest of this section.

2.1. LOs, granularity, and media objects

The concept of LO has been subject to considerable debate. In consequence, several definitions have been proposed (McGreal, 2004), some of them referring the need of metadata records associated to the resources. In any case, the application of Semantic Web to LOs requires them to be annotated with rich metadata expressed with a formal, logics-based language and referring to shared domain ontologies. However, there is a lack of tools that allow instructors or cataloguers to formally describe fragments of media assets.

Current LO specifications like SCORM (Advanced Distributed Learning Initiative, n.d.) allow the composition of units of instruction. However, single files are dealt with as atomic assets, which make it impossible to reference parts of such assets for the purpose of sequencing or selecting particular parts. In other words, video clips are considered atomic pieces, irrespective of their granularity (in this case, related to the length and conceptual density of the video). A possible solution could be splitting the continuous stream in smaller fragments, but this would entail redundancy because different pedagogical usages may lead to different fragmentations of the same media file. Also, the philosophy of reuse in LO fosters the reuse of resources for learning that originally did not have an educational intention. Consequently, there is a need to specify idioms or extended metadata elements that complement SCORM and other standards to deal with multimedia fragments. For example, ways of specifying segments of a continuous media as SCORM assets would enable enhanced SCORM players to display these fragments without the need to break the original files into several pieces.

2.2. Ontologies and the semantic annotation of learning contents

The "Semantic Web" vision described by Berners-Lee Hendler, and Lassila (2001) has resulted in a considerable amount of research and development initiatives to extend the current Web technology with machine-understandable metadata. Formal ontologies (Gruber, 1993) play an essential role in the Semantic Web paradigm, providing the shared conceptualizations expressed in logics-based form that can be used by software agents to act on behalf of humans in search processes or distributed activities. In other words, ontologies provide metadata with shared semantics so that interoperable intelligent agents can act, reason and make decisions according to the information in the metadata.

Although simpler mechanisms exist such as mappings to RDF (Nilsson, Palmér, & Brase, 2003), several initiatives are working to represent the IEEE standard for LO metadata, IEEE LOM,[3] in ontological form. Some of them go further than just mapping the original IEEE LOM to an ontology language like OWL (McGuinness & van Harmelen, 2004) or WSML (de Bruijn et al., 2005). In fact, these efforts implement different mechanisms of referencing domain ontology elements inside metadata elements (Dodero, Díaz, Aedo, & Sarasa, 2005; Sánchez-Alonso, Sicilia, & Pareja, 2007). Several proposed ontological schemas for LO metadata allow us to describe LOs in terms of any available ontology (Gaševic et al., 2007; Sicilia, Lytras, Rodríguez, & García-Barriocanal, 2006;) in such a way that specialized software can be used to exploit the relationships, rules and axioms in the ontologies for (1) navigating repositories, (2) creating tentative LO compositions, or (3) searching for learning resources. Known applications related to the research presented here include ontology-based composition to build exercises (Fischer, 2001) and compositions tailored to personalized learner needs (Jovanovic, Gaševic, & Devedžic, 2006).

Although there are works to annotate the temporal behavior of multimedia audio or video objects with ontologies (Hausmanns, Zerry, Goers, Urbas, Gauss, & Wozny, 2003), these were used basically to index high granularity resources and not specific fragments of pre-existing media. Gahegan, Agrawal, Banchuen, and DiBiase (2007) reported on a system to exploit the richness of relationships inside ontologies integrating several facets, namely, technology, tasks, interactions, learning approaches, techniques, and learning outcomes. Multimedia metadata ontologies usually provide a description of entire media files, even though the practicalities of annotating fragments are specified only at a high level. An example of this is the `AudioVisualSegment` concept in the MPEG-7 ontology defined by Hunter (2006), a concept that could be used for defining a fragment of a video clip. Other ontologies on multimedia include structural features such as shape recognition (Simou et al., 2005), which are not relevant to our research, although they might be useful for describing recognized elements inside media.

Regarding the existence of similar applications to the one we present in this article, Bagdanov, Bertini, Bimbo, Serra, and Torniai (2007) reported on the implementation of a software tool for semantic annotation of video in digital format using multimedia ontologies. This tool allowed to perform higher-level annotation of the clips to generate complex queries that comprise actions and their temporal evolutions and relationships and to create extended text commentaries of video sequences. VIA is a similar tool created for the BOEMIE (http://www.boemie.org) project. Its user-friendly graphical interface allows users to manually annotate both image and video files, using descriptors that are formalized as ontologies. However, these tools are not aimed at carrying out annotations for educational purpose, nor provide the possibility of allowing the user to create "new" virtual video sequences from the temporal delimitation and metadata annotation of existing resources.

Given that the representation of segments in existing media ontologies is not present in LO ontologies, we identified the need for combining educational resources with the media descriptions in a single formal framework. The semantic clip annotation described in this work complements existing ontology-based annotations of learning resources by enabling the reuse of pre-existing media in Web 2.0 applications. As it integrates this kind of media in LO ontologies in a non-intrusive way, it does not affect applications relying on them, e.g. the semantic web services offered by advanced LO repositories such as *ont-space*[4] or *SLOR*[5]. In this article, we will not use any of the mentioned IEEE LOM mappings in a strict sense but will instead provide the main ideas on how to seamlessly take profit of the educational and the media descriptions in a unique knowledge representation model.

3. Semantic annotation of existing continuous media

An approach to semantically annotate resources as generated in the current Web 2.0 social environment should meet a number of requirements specific of the nature of

Web 2.0 content (Kolbitsch & Hermann, 2006). Concretely, the following requirements are important in this environment:

R1: Each uniquely identified resource (e.g. a video clip, a picture, or a blog post) should be allowed to be associated with several descriptions, with different interpretations, following the idea of *resource profiles* (Downes, 2004). For example, a video clip could be described in two different ways, one capturing its possible use as a resource in a K-12 context and other completely different documenting its potential use in Higher Education as a source for research. This extends the plurality of uses of micro-contents to metadata, and allows us to define and evaluate educational contexts, which is pivotal to evaluate reusability (Sicilia & García-Barriocanal, 2003). In other words, a single resource could be described in several ways for different purposes, and technology should allow such plurality of views.

R2: Metadata should be stored externally to the resources themselves, making the annotation non-intrusive to the original applications. Early approaches to annotation of Web pages embedded the annotations in their HTML code (Corcho, 2006). However, this poses several problems including increased download time for the resources, interference with the evolution of the resource, and conflict when several potentially conflicting descriptions for the same description exist.

R3: Annotations could use different domain ontologies, including existing ontologies specific to pedagogical models. A single metadata record could combine links to several domain ontologies.

Regarding R1, some form of permanent, unique identifier for each resource is required. Fortunately, this requirement is somewhat included in most of current Web 2.0 applications, as most of them use permanent URIs or other unique identification mechanisms for their resources. Regarding R2, approaches using separate annotation servers are already in place. A relevant example is Annotea (Kahan, Koivunen, Prud'Hommeaux, & Swick, 2002), included in Amaya, a reference Web browser provided by the W3C.[6] Ontologies expressed in languages like OWL provide namespaces as a way to allow their combination, as required by R3. R2 entails that resources need to be referred inside separate storage systems, and a way of tackling this situation is described in Section 3.1. The rest of the following subsections deals with how to address R3.

3.1. Representing LOs as ontology instances

The main aspect of the annotation approach presented herein is to represent LO metadata as ontology instances. This approach is similar to that of representing tags and tagging in formal terms as used in the *TagOntology* (Gruber, 2007) and has the benefit of enabling reasoning about the LOs by using description logics reasoners or existing inference engines. The IEEE LOM standard can be translated to ontology description languages by mapping or repurposing elements. The main elements of the adaptation of IEEE LOM to an ontology language used are described elsewhere (Sánchez-Alonso et al., 2007). When using such ontology translations, it is not only possible to add rules and logical axioms to the metadata elements but also to link metadata elements to instances in domain ontologies. Regarding persistency, the storage of metadata in ontological form can be carried out by using *ont-space*, a purposefully built open source framework used to include any number of domain ontologies in OWL. This framework was developed as part of the EU project LUISA[7] and has been successfully tested in different educational environments.

Such external metadata repository infrastructure fulfills requirement *R2* above, since any number of metadata repository instances can be setup. This can be used as a semantic-based extension of the approach of distributed annotation servers chosen in Annotea annotation servers (Kahan et al., 2002). In formal terms, we should speak about resource profiles, as a resource such as a video clip identified by a URI is represented by a number of potentially distributed metadata records.

3.2. Mapping continuous media fragments

The annotation process starts by creating a `LearningObject` instance in the ontology for a permanent identifier (if not yet existing). Subclassing allows for combining through subsumption the concepts `AudioVisual` – a subclass of `MultimediaContent` from the MPEG-7 ontology (Hunter, 2006) – and `LearningObject` into a new concept labeled `AVLearningObject`. Instances of this concept are any video sequences in digital format which can be (re)used in different educational contexts, and which include metadata descriptions in a standardized form. An alternate approach could be separating LOs and multimedia contents, but the MPEG-7 ontology does not require a single atomic physical media asset for each multimedia content object, so that the representation described is valid.

Our first problem is that the just mentioned mapping assumes that a metadata record describes an entire media resource: the one identified by URI in the `Identification` section of IEEE LOM. This is in many cases inappropriate, especially in the case of continuous media, where segments of the full content can be considered learning resources. MPEG-7 ontologies provide the `Segment` concept to model fragments of multimedia content, and the `decomposition` property relates segments to media contents. However, it is not possible to specify the starting and ending time points of a segment. Consequently, the following necessary and sufficient condition can be used to automatically classify segments (and thus LOs describing segments) as multimedia contents that are part of others and are determined by some start and end offsets:

```
ContinuousSegment ≡ MultimediaContent and
                    exactly partOf.MultimediaContent and
                    exactly starts.float and
                    exactly ends.float
```

The `partOf` property has the same meaning that the homonymous property in the IEEE LOM relationship category. It refers to strict semantics based on physical containment as related to continuous entities. In this way, the identity of a part is simply determined by the extent relative to the containing content, i.e., equality in `start` and `end` (which should determine a valid, non-empty segment). This simple representation is sufficient for video clips or audio files. As `AudioVisual` is a subconcept of `MultimediaContent` in the MPEG-7 ontology, `AVLearning-Object` declarations with concrete extents would be classified as continuous segments.

Figure 1 shows a possible configuration of annotations with different options. Profile 1 represents a conventional annotation of the entire asset, so it is not describing a `ContinuousSegment`. Profiles 2 and 3 are continuous segments and

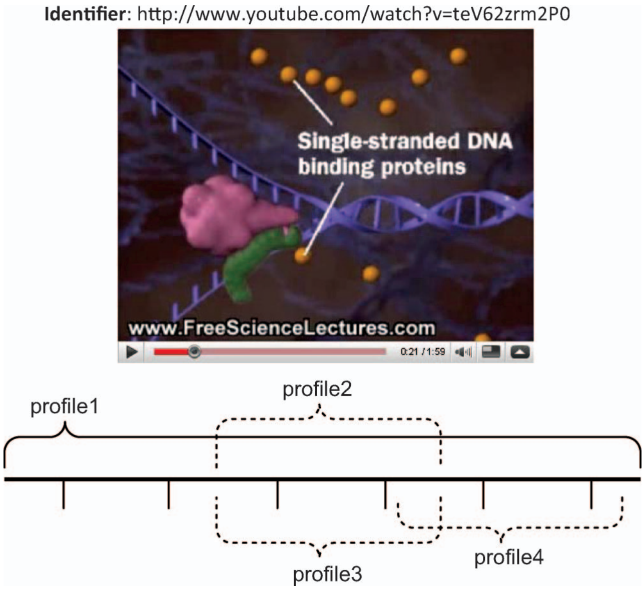

Figure 1. Annotation cases for a given resource.

represent two different educational perspectives on the same resource (the same segment also in this case). For example, it may be that the first one is annotated for the context of secondary education, while the second one is annotated for bioinformatics courses, so that the educational uses and properties are widely diverging. Using an abstract syntax to define it, profile 2 can be represented by using the following collection of assertions:

```
AVLearningObject(?profile2)
partOf(?profile2, ?profile1)
MetadataRecord(?profile2metadata)
describes(?profile2metadata, ?profile2)
educationalContext(?profile2metadata, ?secondary-education)
starts(?profile2, ?13:01)
ends(?profile2, ?25:34)
```

The instance `profile2` would be classified as a continuous fragment automatically, according to the definition provided above, and this could be used to trigger the appropriate software functions that play the specific part of the content. Profile 4 is an additional perspective with a subsegment overlapping some other profiles. Each of the profiles in this example would be represented by an instance of the concept `AVLearningObject`. At the same time, profiles 2–4 would be related to profile 1 with a `partOf` predicate. It should be noted that with this representation, different or even the same fragments of the same multimedia file can be represented as different LOs, and described with different and even inconsistent values in the different IEEE LOM metadata elements, as mentioned earlier. For example, `profile3` will be an instance different from the `profile2` instance in (at least) the value of the `educationalContext` property, as described earlier. The `MetadataRecord` concept is required so that different descriptions for the same segment or the same entire clip can be created and dealt with separately.

3.3. Linking to domain ontologies

In our case, the main purpose of reusing domain ontologies is to express learning objectives mainly through the use of the IEEE LOM element 9 "Classification". Although other forms of linking LOs to ontology elements could have been devised – e.g. defining a property for linking video fragments to concepts of domain ontologies instead of using IEEE LOM elements – we chose this alternative because unlike other options, this decision has the advantage of producing IEEE LOM conformant metadata records. Besides, IEEE LOM Category 9 is a multi-purpose metadata element aimed at describing the contents of the learning resource, so we can use any ontology translation of IEEE LOM for this purpose. To proceed with this form of annotation, a two-step process will be followed:

- Specify the `purpose`: This requires the compilation of a number of purposes of the classification that combine the content of the resource with the type of interaction, as will be described in what follows.
- Select the `TaxonPath`: To be set depending on the domain and the purpose of the classification.

As an example, let us consider the purpose of visualizing processes. This could be represented as a specific purpose from which a logical rule can infer that the associated `taxonpath` is a process. SWRL,[8] a high-level abstract syntax for rules in OWL, can be used to represent such rule:

```
LearningObject(?x)∧
MetadataRecord(?r)∧
describes(?r, ?x)
classification(?r, ?c)∧
purpose(?c, visualizing-process)∧
taxonpath(?c, ?tp)
->Process(?tp)
```

This kind of formalization serves two purposes. First, if the associated taxon path for the classification was something incompatible with a process, an inconsistency would be detected. Second, if the inference is successful, then it will be possible to use knowledge about it if such knowledge is formally represented in the ontology. For instance, if there is a formal description of processes and their constituents, it is possible to link them by using some sequencing technique as in elaboration theories (Reigeluth, 1999). Assuming that, formal definitions of processes (such as the processes of replication of DNA, or the photosynthesis process) can also be used to select related resources.

The notion of `Process` as "sequence of steps" is recurring in different high-level ontologies such as OpenCyc (Lenat, 1995), DOLCE (Gangemi, Guarino, Masolo, Oltramari, & Schneider, 2002) or SUMO (Niles & Pease, 2001). Analyzing some high level concepts can lead to a number of stereotypes in element `purpose` that can be defined to enable concrete kinds of reusable reasoning. For example, stereotypes regarding visualization of processes entail a very concrete learner cognitive activity which is considered in theories of multimedia learning. Another similar reusable purpose is `visualizing-composition`, which in this case refers to any kind of formal mereology, as may be the parts of the cell, or the parts of the digestive system of humans.

3.4. Exploiting ontologies for instructional purposes

There is a considerable amount of research that points to the appropriateness of animations and videos as a resource in education. Some of them highlight the importance of having adequately short videos, because it is less likely that learners will watch the complete clip if this is too long (Korakakis, Pavlatou, Palyvos, & Spyrellis, 2009). In addition, there is a need to express the role of multimedia elements in the delivery of learning resources. Current theories of multimedia learning (Mayer, 2005) provide a number of empirically validated guidelines on how to combine visualizations with verbal representations, emphasizing at the same time the use of two modes of representation rather than one when explaining a concept (Mayer & Moreno, 1997). In the case of devising full instructional sequences, the role that multimedia presentations can play is supporting concrete, fine-grained concepts that are embedded in the activity sequence. A checking mechanism would be that of marking those LOs aimed at teaching something about a process, which do not include any audiovisual representation.

```
(1)
LearningObject(?lo)∧
MetadataRecord(?r)∧
describes(?r, ?lo)
classification(?r, ?c)∧
purpose(?c, competency)∧
taxonPath(?c, ?p)∧
upper:Process(?p)∧
hasPart(?lo, ?lo2)∧
AVLearningObject(?lo2)∧
MetadataRecord(?r2)∧
describes(?r2, ?lo2)
classification(?r2, ?c2)∧
purpose(?c2, visualizing-process)
->AVProcessLO(?lo)
(2)
LearningObject(?lo)∧
MetadataRecord(?r)∧
describes(?r, ?lo)
classification(?r, ?c)∧
purpose(?c, visualizing-process)∧
taxonPath(?c, ?p)
->AVProcessLO(?lo)
(3)
(not AVProcessLO)(?lo)∧
LearningObject(?lo)∧
```

```
MetadataRecord(?r)∧
describes(?r, ?lo)
classification(?r, ?c)∧
purpose(?c, competency)∧
taxonPath(?c, ?p) )∧
upper:Process(?p)
->LOMissingVisualization(?lo)
```

The above rules examine whether the LO is about a process according to a high-level ontology generically denoted with the namespace `upper`. If this is the case, then the rules check whether it is itself a visualization or contains a visualization. The resulting classification can be used to trigger a search for a visualization of the process.

Rule (2) classifies as process-related those LOs that are described as process visualizations. Rule (1) serves the purpose of classifying as process-related those LOs that contain a visualization of a process in its internal part-whole structure and are related to process as a whole. It should be noted that each LO can be composed of others forming a tree of contents (as it is common with SCORM packages for example). The `hasPart` predicate has transitive semantics, so that the rule is valid for multi-level composite LOs. Finally, Rule (3) classifies those LO that are missing visualizations but are related to processes, so that they can be detected by tools to issue a recommendation to instructional designers about the possibility of adding a visualization somewhere in the structure (it can also be triggered if the metadata is not describing that the LO actually contains a visualization of the process it is about). This is a basic example of rule-base support for instructional guidelines.

In our formal representation, this can be translated into constraints on the IMS Learning Design (Koper, Olivier, & Anderson, 2003) structure. For example, the requirement "*Provide the learner control of the sequence when lengthy instructional sequences must be completed by the student in no specific order*" (Stemler, 1997) can be implemented as a constraint on a learning activity. Using the `lessThan` SWRL built-in, which satisfies if and only if the first argument `a1` is less than the second argument `a2` according to a given ordering – temporal sequence in this case – and the IMS-LD ontology by Amorim, Lama, Sánchez, Riera, & Vila (2006), the constraint would look like:

```
Learning-Activity(?a1)∧
Learning-Activity(?a2)∧
Activity-Structure(?as1)∧
execution-order(?a1, ?o1)∧
execution-order(?a2, ?o2)∧
execution-entity-ref(?as1, ?a1)∧
execution-entity-ref(?as1, ?a2)∧
lessThan(?o1, ?o2)
->COMP_showsBefore(?a1, ?a2)
COMP_showsBefore(?a1, ?a2)∧
AVLearningActivity(?a1)∧AVLearningActivity(?a2)∧
LengthyActivity(?a1)∧LengthyActivity(?a2)
```

```
∧objectives-not-related(?a1, ?a2)
->not-recommended-sequencing(?a1, ?a2)
```

The first rule aforementioned determines the sequencing of individual activities according to the IMS LD model. Then, the second checks whether a given lengthy learning activity is based on audiovisual material (which can be defined similarly to `AVLearningObject`) and whether it is sequenced. If so, the rule recommends removing the sequence. The rules determining what a `LengthyActivity` is simply checks the `TypicalLearningTime` IEEE LOM element related to the media elements in the activity, while the `objectives-not-related` predicate can be determined by checking the `objectives` of the IMS LD activities, provided they are expressed through ontology terms (the details are omitted as they do not add significant material to the present discussion). This kind of checks and constraints based on pedagogical guidelines can be used to develop reusable libraries of guidelines. In that case, instructional designers have two options. On the one hand, they can check that their design complies with a subset of the library. On the other hand, they can select some of the formally represented guidelines so that an intelligent authoring environment could help them in the process of composition of units of learning.

The rules sketched so far are just examples aimed to illustrate the ideas introduced in this section. As such, they do not form part of any available set of rules we have developed. However, they show the path for the development of a larger, organized set of rules. This would of course imply a wider effort, which would include full development and evaluation over selected video fragments.

4. Case study

The case study reported here is an application of the semantic annotation concepts described in the preceding sections to the *YouTube* website. As widely known, *YouTube* is one of the largest and most successful Web 2.0 applications, where users upload, watch, and share video clips. Among the millions of videos available, *YouTube* contains many clips on the dynamics of molecular biology processes, which can be used as resources in teaching biology and bioinformatics for very diverse learner profiles. Even though this is a particular domain, it serves as a proof of concept on how the mass of resources shared in current community sites could be reused in a non-intrusive way.

In this section, some background on the gene ontology (GO) is provided. The main elements of the design of the annotation tool are described along with a concrete example of pedagogical annotation.

4.1. The GO and biological processes

The GO[9] project is a collaborative effort to address the need for consistent descriptions of gene products in different databases (Gene Ontology Consortium, 2004). Each entry in the GO has a unique numerical identifier of the form `GO:nnnnnnn`, and a term name. The GO includes a detailed representation of biological processes, a series of events accomplished by one or more ordered assemblies of molecular functions. Examples of broad biological process terms are *cellular physiological process* or *signal transduction*. However, a biological process is not equivalent to a pathway. At present, the GO does not try to represent the dynamics or dependencies that would be required to fully describe a pathway.

The statement of relationships between processes in the GO is specified by a straightforward rule on the degree of overlap, which avoids ambiguous interpretations. To determine whether a process term should be an `is_a` or a `part_of` child of its parent, we need to check whether an instance of the child process is also an instance of the entire parent process, i.e. whether the whole process takes place from start to finish. If answered in the affirmative, then we have an `is_a` child; otherwise, if it is only a portion of the parent process, it will be a `part_of` child.

4.2. Example annotation

As a case of annotation, we will deal with the video clip "*From RNA to Protein Synthesis*".[10] This video visualizes an abstract view of protein synthesis, which can be identified with the GO `biological process` coded GO:0006412 and labeled "`translation`". This process has the following parts in the GO:

- Translational initiation (GO:0006413), the process preceding formation of the peptide bond between the first two amino acids of a protein.
- Translational elongation (GO:0006414), the successive addition of amino acid residues to a nascent polypeptide chain during protein biosynthesis.
- Translational termination (GO:0006415), the process resulting in the release of a polypeptide chain from the ribosome, usually in response to a termination codon (UAA, UAG, or UGA in the universal genetic code).

Translational initiation starts at second 62 in this clip and lasts until second 102. From that point to second 134, some examples of elongation are provided and the rest of the video illustrates the termination phase. This would result in four LO metadata records, one for the full clip and the other three declared as `partOf` of that one. The annotations for the `purpose` would be set to `process-visualization`, while the `taxonpath` would match the corresponding GO term identifier. The GO does not contain information on the sequence of sub-processes in its current version. It however requires every process definition to have a starting and end points, so that in this case it is possible to derive the sequence. The video clip does also contain the visualization of a fragment of another process in the beginning, concretely corresponding to the process of `mRNA transcription` (GO:0009299). In this way, different parts of the video can be used for different learning objectives.

An interesting feature of the multiple annotation approach is that different aspects can be considered when describing LOs. For example, the video "DNA translation animation"[11] starts with a description of amino acids preceding the description of the translation process. Furthermore, it is two-dimensional and introduces additional concepts such as *codons, anti-codons*, and *peptide bonds*, including details of its chemical structure, providing a more detailed level. These elements can also be described as annotations in a simple way, in this case, using classification `purpose` equal to `describing-concept` and specifying the corresponding ontology concepts.

Regarding relationships, it is important to point out that the `partOf` GO relationship can be used in several ways in the context of learning resource search. These include the following:

- Expanding a query using a term, to its constituent parts (which are in turn terms in the ontology). In this manner, each part of a process might be visualized by a fragment of a different video clip. This process can be further exploited through the several levels of traversal of the `partOf` hierarchy.
- Aggregating pre-requisites. In the case of intermediate or final parts of whole processes, knowledge on preceding sub-processes could be used either to check the pre-requisites on the learner side or to find additional background material for understanding the complete process.

There are other interesting relationships in the GO that could be exploited for this case. Existing `isA` relationships can be used to implement "examples first" tactics in sequencing instruction, which would typically go bottom-up from the leaves of the `isA` hierarchy. Also, domain-specific relationships are a source of interesting knowledge for instruction. Following the preceding example about the video *From RNA to Protein Synthesis*, learning resources annotated with positive regulations of translation (GO:0045727) could be used for instruction, having molecular-level regulation of processes as an objective. These regulations are represented in the GO explicitly with relationships like `positively_regulates` or `negatively_regulates`.

In addition to relating the resources to the elements in the ontology, a different source of knowledge that annotators could easily provide are relations between different resources covering the same content. Following the example of video clips visualizing translation, the first one stays at a less detailed level than the second (which even provides chemical formulae for some elements). It is difficult to represent absolute levels of detail in IEEE LOM, since the interpretations of some metadata elements such as `DifficultyLevel` are tied to subjective interpretation. An alternative approach is to state a relationship like `(A moreDetailedThan B)` in the ontology representation. This relationship between LOs is useful for loosely specifying partial orders of difficulty when searching resources for a given educational level.

4.3. *Designing the annotation tool*

Devising an annotation tool to capture different usages of media fragments requires capturing starting and ending offsets in existing video clips, along with a flexible way for creating metadata records connected to the ontologies of interest.

The prototype shown in Figure 2 is an Adobe Flex 3 graphical interface for *YouTube* videos. As users paste the identifiers of clips on the left hand side of the tool, they are automatically played. On the right hand side, users can create a new LO metadata record for the video in the left part of the tool. Clicking on the left buttons, the user can set the exact start and end points of the segment being annotated. It is also possible to select previously created metadata records, as well as creating contained records which will be implicitly related by `partOf` relationships to the full video. Then, the rest of the metadata can be edited using an

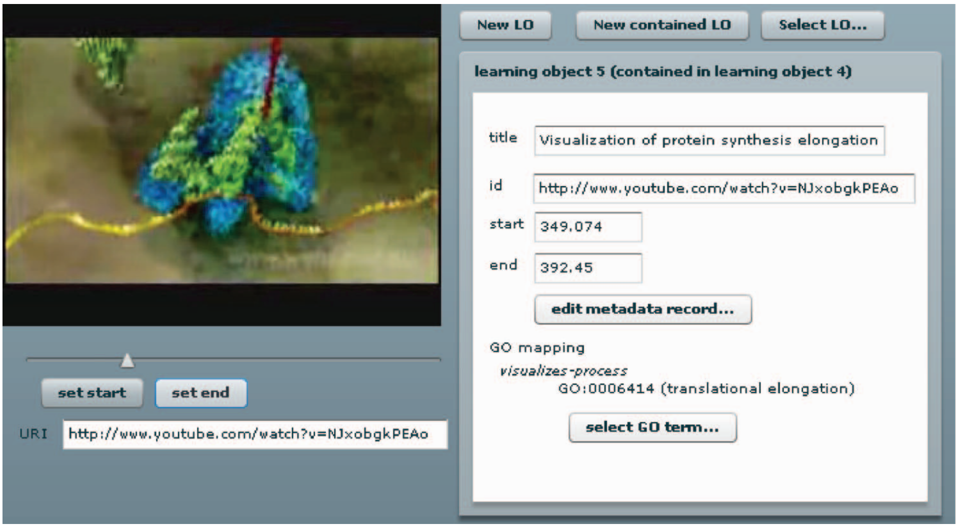

Figure 2. A simplified annotation tool for process annotations.

ontology-enabled metadata editor such as SHAME.[12] As long as editors like SHAME can be set up to work with existing ontologies by means of annotation profiles, the annotations related to the GO can be integrated in the *ont-space* by using the OWL version of the ontology.

5. Conclusions and future work

The exponential growth of user-generated content available through Web 2.0 applications represents an opportunity for educators seeking for quality resources on the Web. However, neither sites collecting such resources apply quality control procedures nor educational value is one of their objectives. In consequence, educators would need to select and annotate those resources deemed appropriate for particular learning needs. This can be performed by creating metadata pointing to the original contents, an approach which results easy to use and almost cost-free. Considering economic reasons as the central target of our effort, we aim at making new educational resources available with negligible development costs. Adapting pre-existing materials through semantic annotation provides the ability of solving this issue in a fairly quick and effective manner.

The use of existing mature domain ontologies for creating that metadata together with repository technology supporting their storage enables reuse and plays the role of a filter. Furthermore, such technology enables the composition of multimedia presentations combining them, and the semi-automated composition of resources based on instructional and domain knowledge represented in ontologies. This article has described an approach for such purpose that can be used for any kind of Web 2.0 application that (1) provides permanent identifiers, e.g. permanent URIs, and (2) allows multiple descriptions of the same or overlapping fragments and their combination. A case study in the domain of molecular genetics has been provided as a proof of concept of the possibilities of the technology developed.

Future work will investigate the connection between user reviews, comments and the structure of internal bookmarking in applications like *YouTube* with the educational quality of the resources. At first glance, there are many motives that lead users to positively rate or comment on a piece of content. Even if these reasons are almost not related to its potential educational usage, these ratings may serve as a first filter for finding potentially useful user-generated content. Another interesting research direction for future work is to examine how the semantic coherence of social nets around content (Paolillo, 2008) could be exploited for educational purposes.

Acknowledgements

This work has been partially supported by project PERSONAL (Personalizing the learning process through Adaptive Paths based on Learning Objects and Ontologies), funded by the Spanish Ministry of Education – Project code TIN2006-15107-C02-01 and by project LUISA (Learning Content Management System Using Innovative Semantic Web Services Architecture), code FP6-2004-IST-4-027149.

Notes

1. *YouTube* is a video sharing website on which users can upload and share videos. The company is based in San Bruno, CA, and uses Adobe Flash Video technology to display a wide variety of user-generated video content. Retrieved at: http://www.youtube.com
2. http://www.w3.org/AudioVideo/
3. http://ltsc.ieee.org/wg12/
4. http://sourceforge.net/projects/ont-space/
5. http://slor.sourceforge.net/
6. http://www.w3.org/Amaya/
7. http://www.luisa-project.eu
8. http://www.w3.org/Submission/SWRL/
9. http://www.geneontology.org/
10. Identified in *YouTube* with the URI: http://www.youtube.com/watch?v=NJxobgkPEAo
11. Identified in *YouTube* with the URI: http://www.youtube.com/watch?v=ncjIiD51hQk
12. http://kmr.nada.kth.se/shame/

Notes on contributors

Elena García-Barriocanal obtained a university degree in Computer Science from the Pontifical University of Salamanca in Madrid (1998) and a PhD from the Computer Science Department of the University of Alcalá. In 1999, she joined the Computer Science Department of University of Alcalá and in 2001 she became an associate professor. She is a member of the Information Engineering Research Group. Her research interests focus on topics related to the role of knowledge representation in human–computer interaction and learning technologies where she actively works on ontological aspects both in e-learning and in usability and accessibility.

Miguel-Angel Sicilia obtained an MSc degree in Computer Science from the Pontifical University of Salamanca, Madrid (Spain) in 1996 and a PhD degree from the Carlos III University in 2003. He worked as a software architect in e-commerce consulting firms, being part of the development team of a Web personalization framework at Intelligent Software Components (iSOCO). Currently, he is an associate professor and leads the Information Engineering Unit at the Computer Science Department, University of Alcalá. His research interests are primarily in the areas of adaptive hypermedia, learning technology and human–computer interaction. He leads the AIS SIG on reusable learning objects (www.sigrlo.org).

Salvador Sánchez-Alonso obtained a university degree in Computer Science from the Pontifical University of Salamanca in 1997, and a PhD in Computing from the Polytechnic University of Madrid in 2005. He worked as an assistant professor at the Pontifical University of Salamanca for 7 years during different periods, and also as a software engineer at a software solutions company during 2000 and 2001. From 2005, he is an associate professor at the Computer Science Department of the University of Alcalá. His current research interests include learning objects reusability, metadata, Object-Oriented technologies and Software and Web Engineering.

Miltiadis Lytras is a Research Professor in the American College of Greece – Deree College. His research focuses on semantic web, knowledge management and e-learning, with more than 100 publications in these areas. He has co-edited 25 special issues in international journals and has authored or co-edited 25 books. He is the founder and officer of the Semantic Web and Information Systems Special Interest Group in the Association for Information Systems (http://www.sigsemis.org). He serves as the (Co) Editor in Chief of 12 international journals and he is associate editor or editorial board member in seven more.

References

Ache, K., & Wallace, L. (2008). Human papillomavirus vaccination coverage on *YouTube*. *American Journal of Preventive Medicine*, *35*, 389–392.

Advanced Distributed Learning Initiative. (n.d.). Sharable content object reference model (SCORM). *Advanced Distributed Learning*. Retrieved from http://www.adlnet.org.

Allen, M. (2008). Web 2.0: An argument against convergence. *First Monday*, *13*.

Amorim, R., Lama, M., Sánchez, E., Riera, A., & Vila, X.A. (2006). A learning design ontology based on the IMS specification. *IEEE Journal of Educational Technology Society*, *9*(1), 38–57.

Bagdanov, A.D., Bertini, M., Bimbo, A.D., Serra, G., & Torniai, C. (2007). Semantic annotation and retrieval of video events using multimedia ontologies. *International conference on semantic computing, 2007. ICSC 2007* (pp. 713–720). Los Alamitos, CA: IEEE Press.

Bargeron, D., Gupta, A., Grudin, J., & Sanocki, E. (1999). Annotations for streaming video on the web. *CHI'99 extended abstracts on human factors in computing systems* (pp. 278–279). Pittsburgh, PA: ACM.

Berners-Lee, T., Hendler, J., & Lassila, O. (2001). The semantic web. *Scientific American*, *284*, 34–43.

Brown, W., Srinivasan, S., Coden, A., Ponceleon, D., Cooper, J.W., & Amir, A. (2001). Toward speech as a knowledge resource. *IBM Systems Journal*, *40*, 985–1001.

Cafolla, R. (2002). Project Merlot: Bringing peer review to Web-based educational resources. In *Proceedings of the USA Society for Information Technology and Teacher Education International Conference*, pp. 614–618. Chesapeake, VA: AACE.

Cochrane, T. (2007). Developing interactive multimedia learning objects using quicktime. *Computers in Human Behavior*, *23*, 2596–2640.

Collis, B., & Moonen, J. (2008). Web 2.0 tools and processes in higher education: Quality perspectives. *Educational Media International*, *45*, 93–106.

Corcho, O. (2006). Ontology based document annotation: Trends and open research problems. *International Journal of Metadata, Semantics and Ontologies*, *1*, 47–57.

de Bruijn, J., Lausen, H., Krummenacher, R., Polleres, A., Predoiu, L., Kifer, M., & Fensel, D. (2005). The Web Service Modeling Language WSML. *WSML Final Draft* D, *16*.

Devgan, L., Powe, N., Blakey, B., & Makary, M. (2007). Wiki-Surgery? Internal validity of Wikipedia as a medical and surgical reference. *Journal of the American College of Surgeons*, *205*(3 Suppl 1), S76–S77.

Dodero, J.M., Díaz, P., Aedo, I., & Sarasa, A. (2005). Integrating ontologies into the collaborative authoring of learning objects. *Journal of Universal Computer Science*, *11*, 1568–1578.

Downes, S. (2004). Resource profiles. *Journal of Interactive Media in Education*, *5*.

Eijkman, H. (2008). Web 2.0 as a non-foundational network-centric learning space. *Campus-Wide Information Systems*, *25*, 93–104.

Feng, S., Manmatha, R., & Lavrenko, V. (2004). Multiple Bernoulli relevance models for image and video annotation. *Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004. CVPR 2004* (Vol. 2, pp. II-1002–II-1009).

Fischer, S. (2001). Course and exercise sequencing using metadata in adaptive hypermedia learning systems. *Journal of Educational Resource in Computing*, *1*, 5.

Gahegan, M., Agrawal, R., Banchuen, T., & DiBiase, D. (2007). Building rich, semantic descriptions of learning activities to facilitate reuse in digital libraries. *International Journal on Digital Libraries*, *7*, 81–97.

Gangemi, A., Guarino, N., Masolo, C., Oltramari, A., & Schneider, L. (2002). Sweetening ontologies with DOLCE. *Lecture Notes in Computer Science*, *2473*, 166–181.

Gašević, D., Jovanovic, J., & Devedžic, V. (2007). Ontology-based annotation of learning object content. *Interactive Learning Environments*, *15*, 1–26.

Gene Ontology Consortium. (2004). The gene ontology (GO) database and informatics resource. *Nucleic Acids Research 32*, D258–D261.

Gruber, T.R. (1993). A translation approach to portable ontologies. *Knowledge Acquisition*, *5*, 199–220.

Gruber, T. (2007). Ontology of folksonomy: A mash-up of apples and oranges. *International Journal on Semantic Web and Information Systems*, *3*(1), 1–11.

Hausmanns, C., Zerry, R., Goers, B., Urbas, L., Gauss, B., & Wozny, G. (2003). Multimedia-supported teaching of process system dynamics using an ontology-based semantic network. *Computer Aided Chemical Engineering*, *15*, 1453–1459.

Hunter, J. (2006). Adding multimedia to the semantic web – Building and applying an MPEG-7 ontology. *Multimedia content and the semantic web: Methods, standards and tools*. New York: Wiley.

Jenkins, H. (2007). From *YouTube* to YouNiversity. *Chronicle of Higher Education*, *53*, B9.

Jovanovic, J., Gašević, D., & Devedžic, V. (2006). Ontology-based automatic annotation of learning content. *International Journal on Semantic Web and Information Systems*, *2*, 91–119.

Kahan, J., Koivunen, M.R., Prud'Hommeaux, E., & Swick, R.R. (2002). Annotea: An open RDF infrastructure for shared Web annotations. *Computer Networks*, *39*, 589–608.

Kolbitsch, J., & Hermann, M. (2006). The transformation of the web: How emerging communities shape the information we consume. *Journal of Universal Computer Science*, *12*, 187–213.

Koper, R., Olivier, B., & Anderson, T. (2003). *IMS learning design information model*. IMS Global Learning Consortium technical report.

Korakakis, G., Pavlatou, E., Palyvos, J., & Spyrellis, N. (2009). 3D visualization types in multimedia applications for science learning: A case study for 8th grade students in Greece, *Computers & Education*, *52*, 390–401.

Lenat, D.B. (1995). CYC: A large-scale investment in knowledge infrastructure. *Communications of the ACM*, *38*, 33–38.

Lienhart, R. (1996). Indexing and retrieval of digital video sequences based on automatic text recognition. *Proceedings of the fourth ACM international multimedia conference*. Boston, MA: ACM Press.

Luttrell, K., Zite, N., & Wallace, L. (2008). Myths and misconceptions about intrauterine contraception on *YouTube*. *Contraception*, *78*, 183.

Mayer, R. (Ed.). (2005). *The Cambridge handbook of multimedia learning*. New York, NY: Cambridge University Press.

Mayer, R.E., & Moreno, R. (1997). *A cognitive theory of multimedia learning: Implications for design principles*. New York, NY: Wiley.

McGreal, R. (2004). Learning objects: A practical definition. *International Journal of Instructional Technology and Distance Learning*, *1*.

McGuinness, D.L., & van Harmelen, F. (2004). OWL web ontology language overview. *W3C Recommendation*, *10*, 2004–2003.

Milliken, M., Gibson, K., O'Donnell, S., & Singer, J. (2008, May). User-generated online video and the Atlantic Canadian public sphere: A *YouTube* study. *Proceedings of the international communication association annual conference* (ICA 2008), Montreal, Quebec, Canada. NRC 50362.

Miyamori, H., & Iisaku, S. (2000). Video annotation for content-based retrieval using human behavior analysis and domain knowledge. *Proceedings of fourth IEEE international conference on automatic face and gesture recognition* (pp. 320–325). Los Alamitos, CA: IEEE Press.

Monceaux, A., Naeve, A., Sicilia, M.A., García-Barriocanal, E., Arroyo, S., & Guss, J. (2007). *Targeting learning resources in competency-based organizations. The Semantic Web: Real-world applications from industry* (pp. 143–167). New York, NY: Springer.

Niles, I., & Pease, A. (2001). Towards a standard upper ontology. *Proceedings of the international conference on formal ontology in information systems-Volume 2001* (pp. 2–9). New York, NY: ACM.

Nilsson, M., Palmér, M., & Brase, J. (2003). *The LOM RDF binding – Principles and implementation*. Paper presented at the Third Annual ARIADNE conference, Leuven, Belgium.

Paolillo, J.C. (2008). Structure and network in the *YouTube* core. *Proceedings of the 41st Annual Hawaii International Conference on System Sciences* (p. 156). Los Alamitos, CA: IEEE Press.

Reigeluth, C.M. (Ed.). (1999). *Instructional-design theories and models, Volume II: A new paradigm of instructional theory*. Mahwah, NJ: Lawrence Erlbaum Associates.

Rollett, H., Lux, M., Strohmaier, M., Dösinger, G., & Tochtermann, K. (2007). The Web 2.0 way of learning with technologies. *International Journal of Learning Technology*, *3*, 87–107.

Sánchez-Alonso, S., Sicilia, M.A., & Pareja, M. (2007). Mapping IEEE LOM to WSML: An ontology of learning objects. *Proceedings of the ITA'07 – Second International conference on internet technologies and applications* (pp. 92–101). Wrexham: North East Wales Institute.

Sicilia, M.A. & García-Barriocanal, E. (2003). On the concepts of usability and reusability of learning objects. *International Review of Research in Open and Distance Learning*, *4*.

Sicilia, M.A., Lytras, M., Rodríguez, E., & García-Barriocanal, E. (2006). Integrating descriptions of knowledge management learning activities into large ontological structures: A case study. *Data and Knowledge Engineering*, *57*, 111–121.

Simou, N., Saathoff, C., Dasiopoulou, S., Spyrou, E., Voisine, N., Tzouvaras, V., . . . Staab, S. (2005). *An ontology infrastructure for multimedia reasoning*. Paper presented at the international workshop VLBV, Sardinia, Italy.

Spiteri, L.F. (2007). Structure and form of folksonomy tags: The road to the public library catalogue. *Webology*, *4*.

Stemler, L. (1997). Educational characteristics of multimedia: A literature review. *Journal of Educational Multimedia and Hypermedia*, *6*, 339–359.

Volkmer, T., Smith, J.R., & Natsev, A.P. (2005). A web-based system for collaborative annotation of large image and video collections: An evaluation and user study. *Proceedings of the 13th annual ACM international conference on multimedia* (pp. 892–901). New York, NY: ACM.