Semi-automatic Semantic Annotation Tool for Digital Music

Fazilatur Rahman¹ and Jawed Siddiqi²

Sheffield Hallam University

Faculty of Arts, Computing, Engineering and Sciences, Sheffield, S1 2NU, UK

f.rahman@shu.ac.uk, j.i.siddiqi@shu.ac.uk

The Worldwide Web/Internet has changed the music industry by making huge amount of music available to both music publishers and consumers including ordinary listeners or end users. The Web2.0 tagging techniques of music items by artist name, album title, musical style or genre (technically these are termed as syntactic metadata) have given rise to the generation unstructured free form vocabularies. Music search based on these syntactic metadata requires the search query to contain at least one keyword from that vocabulary and it must be an exact match. The semantic Web initiative by W3C proposes machine process-able representation of information but does not stipulate how that can be applied to music items specifically. In this paper we present a novel approach that details a semi-automatic semantic annotation tool to enable music producers to generate music metadata through a mapping between music consumers' free form tags and the acoustic metadata that are automatically extractable from music audio. The proposed annotation tool enables onotology guided annotation process and uses MPEG-7 Audio compliant music annotation ontology represented in dominant semantic web standard OWL 1.0.

Semantic, Annotation, Music Ontology, MPEG-7.

1. INTRODUCTION

Both syntactic and semantic metadata are proven approaches of tagging resources to make them readable/ accessible to content management and search engines. Tim Berners-Lee's vision for the next stage Semantic Web requires not just tagging information with syntactic metadata, but annotating or 'enhancing' information with semantic metadata to enable machine understand-ability of the full context of what that information means.

Music search based on syntactic metadata requires the search query to contain at least one keyword from that vocabulary and it must be an exact match. But, keywords used by ordinary music fans (i.e. consumers) often contain diverse range of description of the song that differs from the publishers categorization tags. As a result, consumers are left with unsatisfactory search results.

From a technical perspective music is also expressed by another type of syntactic metadata which is actually audio signal level statistics and is only understood by machines but not by ordinary end users. As a result signal level metadata are not

included in the end-users vocabulary. But, signal level metadata are often used to automatically detect musical style or acoustic properties to detect musical instruments, key or structure. Such an approach could be used to yield satisfactory search results but the acoustic metadata is not suitable to be used for search and tagging purpose as meaning of those metadata are not perceived by ordinary human users while searching/annotating songs of interest.

Based on the idea - if music items can be annotated semantically then search and retrieval may be enhanced - in this paper we will present a semi-automatic semantic music annotation tool that will utilize the following basic features:

- (i) *Use of MPEG-7 audio based standardized acoustic metadata:* The MPEG-7 Audio metadata generated by automated tools (e.g. Crysandt, 2005) are not suitable for annotation by ordinary music users (both producers and consumers). The proposed semantic annotation tool will enable music producers to generate semantic metadata that will provide automatic mapping of the

MPEG-7 acoustic features with end-users metadata.

- (ii) *Semantic Metadata vocabulary for music:* At present, available semantic music metadata vocabularies are mainly for use by music domain only without any mapping to the acoustic metadata. On the other hand, prominent semantic multimedia metadata (Hunter, 2003) provides only upper level multimedia (audio) ontology. But none of these provides any means how they can be utilized for music annotation by ordinary music users as do not stipulate how to annotate/categorize music using structured metadata (Gutierrez et al., 2011). The proposed semi-automatic annotation tool has been designed to demonstrate how music producers may be able to annotate music items using a music annotation ontology (detail of the ontology has been presented in (Rahman and Siddiqi, 2011) and at the same utilize MPEG-7 acoustic features in a seamless manner.

The rest of this paper is organized as follows: section 2 will detail on the state of the art of music tagging and annotation tools. In section three we will present the detail of the annotation process featured by the proposed annotation tool as well as highlight evaluation of the tool. Section 4 summarizes the contribution made by this paper.

2. BACKGROUND

2.1 State of the art: Music Tagging and Annotation

The present limitations of music tagging and annotation that have been discussed in Turnall et al., (2008) leads to several challenges faced by music discovery, search and retrieval systems. The main reason behind the limitations of current music search and tagging techniques are that these techniques as mentioned previously operate on a syntactic level rather than a semantic one (Yang, 2007). To overcome this limitation the W3C Semantic Web Activity (Herman, nd) establishes standard ways to represent machine process-able web contents to allow machines to follow links and facilitate the integration of data from many different sources. When the relationships among data are fully accessible to machines, machines will be able to enable us browse those relationships and interpret the data as well as assess the appropriateness of the data for the intended purposes of users (Miller and Swick, 2003). Thus W3C plays a leadership role in both the design of specifications and the open, collaborative development of technologies focused on representing relationships and meaning and the automation, integration and reuse of data. One of

the standard ways of representing data semantics through explicit relationships among different concepts is using 'ontology' that may be used by automated tools to empower more accurate web search. To that aim the Web Ontology Working Group's standard efforts (Connolly et al., 2004) has designed another language to build upon the RDF language (Calvanese et al., 2011) namely Web Ontology Language – best known as OWL (Herman, 2007), for defining structured, web-based ontologies. The Semantic Web really started from metadata that provides a syntactic description of data represented in XML. With the development of a resource description framework or RDF that models things and relationships it provides semantic (meaningful description of) data; which is much more reusable and enable machine processability (Hardin, 2005). With the advent of OWL on top of RDF has brought richer representational capability to model knowledge to further inference on the meaning of data.

The common form of creating association in the context of digital multimedia (in this context it is digital music) is often referred to as annotation (Ruvane, 2006). There are different ways of creating annotation (Fu et al., 2005). Apart from methods of creating annotations, depending on the level and granularity, annotations might be of two types. Content annotation considers using (e.g.) text selection, emphasis and adding notes and Structural annotations may be achieved by (e.g.) creating a logical structure that is different from the physical structure or linking from one place to another place in a book. In the context of digital music, both content and structural annotations take a different view than that of text documents and images. Music unfolds in time and information conveyed by different musical dimensions is quite distinct from text and images. For example, creating annotations at the composite level refers to many audio segments of a single music file partitioned by time dimension.

Use of ontological semantics in different domains has increasingly become popular (e.g. [Zang et al. 2008]) and use of annotation ontology by extending features needed for annotation is also not new (Dedek et al., 2008). Similarly, in a novel way, this paper will present a general approach for annotating music files that considers both MPEG-7 objective data about the music audio as well as extends multimedia ontology for the purpose of annotating music files in particular. Most of the related works in creating multimedia ontologies have only attempted to create upper level multimedia ontologies and only few of them have focussed on creating customized vocabulary for a specific category (e.g. Greek music (Tsekeridou et al., 2006)) of music files. None of them addressed the problem faced by music producers while

annotating music items that they want to publish in general. The detail of how the music annotation ontology known as *mpeg-7Music* may be found in (Rahman and Siddiqi, 2011).

The Web 2.0 (Zhang, 2007) technologies present applications for developing folksonomies. It is true that folksonomies did not provide techniques to develop meaningful annotations but folksonomies can be utilized as a means for cheap and quick alternative for automatic annotation but the created annotations are not designed for machine processing (Bateman et al., 2006). Though Web 2.0 applications appear to be an excellent opportunity to create semantic resource, in the field of Semantic Web technology it is still suffering from few shortcomings in the context of developing semantic web applications (Gruber, 2008) - user intervention requirements it is usually found that manual annotation tools do not seem to be effective because of error-prone annotations and the laborious effort needed to perform the annotation task makes it an unattractive option. On the other hand, fully automatic annotation tools rely mainly on huge training data needed for effective machine learning algorithms and such tools are unable to capture the high-level meaning conceived by human users. Consequently, this dictates us to

design a semi-automatic annotation tool for music items as a trade-off to bring the best of both worlds of Web2.0 and Semantic technologies (Damme et al., 2007).

Again the dominant industry standard to representation structured machine level acoustic information of music (standardized by MPEG-7 (Martinez et al., 2002)) is not understood by most of the end users though those perceptual feature based representation could lead to finding of that music media easily. So, it would be very useful if we could enrich the semantic interpretation of musical pieces with those acoustic/perceptual features following standard representational techniques. At present, ordinary users metadata vocabulary to describe music is not only unstructured but also is not defined with any standard interpretation mechanism that can be linked with acoustic metadata that can be automatically derived from music audio. The semantic annotator tool presented in this paper will demonstrate how music producers may create a mapping between music consumers' metadata and the underlying music's acoustic metadata in a meaningful (semantic) way that is actually termed here as 'semantic annotation of music'.

Table 1(a): Survey of existing Simple Annotation Tools

	Annotation Tool	Content format (Text/Image/Audio/ Video)	Standard representation of Annotation
Simple Syntactic Annotation Tools	<i>Armadillo</i>		<i>RDF(S)</i>
	<i>KnowItAll</i>		<i>HTML</i>
	<i>SmartWeb</i>		<i>RDF, RDFS, OIL</i>
	<i>PANKOW</i>		<i>HTML</i>
	<i>WiCKOffice</i>	<i>Microsoft Office</i>	<i>Microsoft Smart Documents</i>
	<i>AktivDoc</i>	<i>HTML</i>	<i>HTML, RDF</i>
	<i>Semantic Word</i>	<i>Word</i>	<i>DAML + OIL</i>
	<i>MagPie</i>	<i>HTML</i>	<i>HTML, OCML</i>
	<i>Lixto</i>		<i>Wrappers</i>
	<i>Mangrove</i>	<i>HTML</i>	<i>RDF</i>
	<i>M-Onto-Mat-Annotizer</i>	<i>MPEG-7</i>	<i>XML, RDFS</i>
	<i>SemTag</i>	<i>HTML</i>	<i>RDFS</i>
	<i>Vannotea</i>	<i>JPEG2000/MPEG-2/Direct3D</i>	<i>XML</i>
	<i>Ontolog</i>	<i>QuickTime, JMF</i>	<i>RDF</i>

Table 1(b): Survey of existing Ontology Guided Semantic Annotation Tools

	Annotation Tool	Content format (Text/Image/Audio/ Video)	Standard representation of Annotation
Ontology guided semantic annotation Tools	<i>OntoMedia</i>	<i>QuickTime, JMF</i>	<i>RDF</i>
	<i>Amaya</i>	<i>HTML/XML</i>	<i>RDFS, XLink, XPointer</i>
	<i>OntoMat</i>	<i>HTML</i>	<i>DAML + OIL, OWL, SQL</i>
	<i>SHOE-knowledge Annotator</i>	<i>HTML</i>	<i>SHOE</i>
	<i>SMORE</i>	<i>HTML/Text/Image</i>	<i>RDFS</i>
	<i>Open Ontology Forge</i>	<i>HTML/Image</i>	<i>RDFS, XML, XLink</i>
	<i>COHSE Annotator</i>	<i>HTML</i>	<i>DAML+OIL</i>
	<i>MnM</i>	<i>HTML/Text</i>	<i>RDFS, DAML+OIL, OCML</i>
	<i>Melita</i>	<i>HTML/Text</i>	<i>RDFS, DAML+OIL</i>
	<i>Parmenides</i>	<i>HTML</i>	<i>XML (CAS)</i>
	<i>AeroSwarm</i>	<i>HTML</i>	<i>OWL</i>
	<i>KIM</i>	<i>HTML</i>	<i>RDFS, OWL</i>
	<i>Rainbow Project</i>	<i>HTML</i>	<i>RDF, WSDL/SOAP</i>
	<i>h-TechSight</i>	<i>HTML</i>	<i>DAML+OIL, RDF</i>
	<i>Thresher</i>	<i>HTML</i>	<i>RDF</i>

2.2 Existing Semantic Annotation Tools

The available surveys on annotation tools vary in the criteria they adopted to assess the tools. Uren et al. (2006) presented a survey of ontology based annotation frameworks and tools that helps us to categorize the tools based on standard formats and ontology support. To design a semantic annotation tool for music items it requires us to identify the tools based on the various content formats they allow annotating. Another survey done by Schroeter et al. (2006) focussed on content type (text, image, audio and video) the support but emphasizes the tool's ability to provide collaboration and sharing among different communities of users. Based on these observations we focus on existing annotation tools on the basis of following *three* requirements -

- (i) Ontology based annotation for achieving interoperability
- (ii) Standard representation of annotations
- (iii) Type of content of heterogeneous multimedia resources

Table 1(a) & 1(b) provide an overview of different tools, systems and projects according to above mentioned three requirements. From table 1(a), it is

clear that tools listed under simple syntactic annotation tools presented here fail to satisfy all three of these requirements. The rest of the tools as listed in table 1(b), though they all support ontology based annotations but were limited by their ability to provide annotations for diverse content formats (for music audio files). Most of them supported textual content (mostly HTML), very few of them could enable annotation for other types such as MPEG, JPEG, QuickTime, JMF etc.

But, in practice there are a huge variety of digital music media of different types such as audio (wav, mp3, real audio etc.), and the musical object that comes as embedded in video resources(avi, mpeg etc.); for which none of these tools is able to provide generalized support. So, the need for a generalized semantic annotation platform (that satisfies all three of the requirements mentioned above) is evident.

Table 2: Digital Content Representation Standards

Content Type	Existing Standards	Formal Representation	Available Mapping to RDF/OWL	Workflow	Domain	Industry
Still Image	Visual Resource Association (VRA)	VRA-RDF/OWL - non-XML	No-commonly accepted mapping	Publish	Culture	Archive
	Exchangeable Image File Format (EXiF)	EXiF-RDF/OWL non XML	Few available	Capture, Distribute	Generic	Digital Camera
	DIG35	DIG35-RDF/OWL	Available	Publish	Archives	Consumer
	PhotoRDF	RDF		Capture, Distribute	Personal media	Photo
Audio	ID3	non XML		Distribute	Generic	Music
	MusicBrainzMeta-DataInitiative 2.1	RDF		Production	Generic	Music
	MusicXML	XML		Production	Generic	Music
Still Image, Audio, Video	MPEG-7	XML, RDF/OWL	Available	Archive, Publish	Generic	Generic
	Advanced Authoring Format (AAF)	non-XML		Production	Content Creation	Broadcast
	MXF-DMS-1	non-XML		Production	Content Creation	Broadcast
General Purpose	Synchronized Multimedia Integration Language (SMIL)	XML		Publish, Distribute, Presentation, Interaction	Generic	Web, Mobile Applications
	Scalable Vector Graphics (SVG)	XML		Publish, Presentation	Generic	Web, Mobile Applications
	Dublin Core	XML, RDF		Publish	Generic	Generic
	TVAnytime	XML		Distribute	Electronic Program Guide (EPG)	Broadcast
	MPEG-21	XML, non-XML		Annotate, Publish, Distribute	Generic	Generic
	XMP/IPTC	XML, RDF		Annotate, Publish, Distribute	Generic	Generic
Generic (Text, Still Image, Audio, Video)	Flash	XML, non-XML		Publish	Distribution, Presentation & Interactivity	Broadcast
	MPEG-4 BIFS	XML, non-XML		Publish	Distribution, Presentation& Interactivity	Broadcast
	MPEG LASER	XML		Publish	Distribution, Presentation& Interactivity	Broadcast
	NewsML	XML		Publish	News	News Agencies
Video, Audio	EBU P/Meta	XML, non-XML		Publish	Generic	Broadcast

2.3 Goals for Music Annotation

Referring to table 2, there are lots of standards available to represent general purpose digital content and content type. In relation to audio materials most of the standards are concerned about publication and distribution of the content for commercial exploitation of the content. Very few of them did actually paid attention to musical objects and different levels of information content conveyed by those objects. None of them provided any further direction for enabling the music producers to facilitate annotation for further search and retrieval. Most of the existing audio content representation standards were designed to satisfy the requirements for specific domains for which they were created to document the workflow that support for a particular industry's purpose.

For interoperable multimedia systems, it is necessary to use a standard set of type definitions (schemas) and standardized values in the description (Smith and Schirling, 2006) following five discriminating characteristics¹:

- (i) Representation: The primary (official) serialization format for the MM description standard. e.g. XML, RDF/OWL
- (ii) Content Type: The type of media, a certain MM standard is capable to describe. e.g. Still image, audio, video, text.
- (iii) Workflow: understood in terms of the *Canonical Processes of Media Production* (Hardman, 2005) e.g. publish, process, creation etc.
- (iv) Domain: The realm in which a MM standard is intended to be used in the industry. e.g. news, sports, music etc.
- (v) Industry: The (main) branch of productive (commercial) usage e.g. broadcast, music etc.

3. THE SIMPLE SEMANTIC ANNOTATOR TOOL

3.1 Ontology based Annotation of Digital Music

Annotation framework needs guidance from the ontology to allow sharing of knowledge and newly created annotations must be consistent with a community's ontology. If

newly created annotations are created by instantiating arbitrary classes and properties the semantics of these properties remains void. So, the ontology is important in order to guide annotators towards creating relational metadata. Without the ontology it will be hard to establish more cues for assigning relationships between class instances (Hands Schuh et al., 2001).

Advantages of utilizing ontology as an annotation schema are many folds (Kosch et al., 2005). The formalization of annotation schema as an ontology represented in standard language such as RDF/OWL meets the interoperability requirement of different conceptualization. Thus the use of standard encoding of annotation schema enables the reusability of the schema across different tools making the created annotations completely independent of the annotation tool actually used. Such RDF/OWL based ontological annotation model offers a general framework for the task annotation that may be broadly applied to diverse contents. The fact that annotation is performed with respect to an ontological hierarchy offers annotators the possibility to choose the appropriate level of annotation detail. Moreover, ontology based semantic framework for annotation helps to constrain the possible relations between two concepts, thus reducing the amount of errors in the annotation process.

Most annotation tools make use of schema which specifies what can actually be annotated. These schemas can be understood as a formal representation of the conceptualization underlying the annotation task (Cimiano and Hands Schuh, 2003). As ontologies are formal specifications of a conceptualization (Gruber, 1993) it seems straightforward to formalize annotation schemes as ontologies.

Based on the intended methodology for annotating multimedia several tools and annotation frameworks have been developed and they fall in three general category. The first category is focused on the direct exploitation of signal level features for the purpose of annotation supporting a limited collection of content format (Schroeter et al., 2006) while the second one emphasizes th

¹<http://www.w3.org/2005/Incubator/mmsem/wiki/Vocabularies>

association of the resource as a whole with ontological concepts encoded in several standard representation and formats (Uren et al., 2006). Thirdly, the M-OntomatAnnotizer tool (developed under Acemedia project) (Petridis et al., 2006a) that enriches domain ontologies by using visual descriptor ontology (spatio-temporal ontology) represented in RDF. But the created ontology is limited to Image and visual feature descriptors only.

Figure 1: High level System view of the Semantic Annotation Tool

Figure 1 shows the high level system view of the tool that uses *mpeg7Music* as backbone to guide the annotation process. MPEG-7 feature was extracted from the music file (to be annotated) automatically and mapped to the ontology through the properties defined in the ontology. Then the user was presented to associate higher level concepts with these auto generated features. So, a part of the annotation was automated and to complete the process of annotation it required user intervention. Hence the designed annotation tool was classified to be semi-automatic.

The high level system view of the simple semantic annotation tool as shown figure 1 has two distinct phases of workflow: Lifting MPEG-7 metadata by associating it with annotation ontology (figure 2) and the actual annotation process.

3.2 Lifting MPEG-7 Metadata

When a music file is loaded into the annotation tool several background steps are executed without any user intervention as displayed by figure 2. At first, MPEG-7 feature extraction output is generated automatically for that particular music file using *MPEG-7 Feature Extraction Tool*. A freeware MPEG-7 audio

encoder tool (Crysanadt, 2005) was utilized to generate MPEG-7 XML output from the music file. Secondly, MPEG-7 datatype values were extracted from that XML output using the custom built *XML Data Source Processor*.

These data type values were then mapped to the *mpeg7Music* ontology using the melodic and timbral rules in the ontology. To perform the mapping process, a web service was designed to process OWL ontology using Protege OWL API². The reason to use web services was to create interoperable service oriented architecture (Mahmoud, 2005) for the proposed annotation system. Having designed the web service (OWL Ontology Processor - Web Service) to fit the need, a web service client was also created to interface the annotation tool with the web service. The web service client was used to map the data types extracted from the XML data source processor with properties and concepts of the ontology. Association rules for linking MPEG-7 XML tags defined by the annotation ontology preserved the connection of the created annotations with MPEG-7 description of the underlying audio of music files. After the data type values were mapped then the assertions generated were made available for the graphical user interface step 3 with the help of the *mpeg-7Music* music annotation ontology.

Once the generated assertions have been displayed annotation workflow relied on user intervention by music producers.

3.3 Actual Annotation Process

The actual annotation process of semantic annotation tool was based on a simple five steps process; those will now be explained with reference to an illustrative example as shown by figure 3 (1-e) that depict screen shots been generated using Adobe FireWorks³ for clarity.

² <http://protege.stanford.edu/plugins/owl/api/>

³ <http://www.adobe.com/products/fireworks/web/>

Figure 2: Workflow for lifting MPEG-7 Metadata

Step1- Loading and naming the Resource: When a user loads a resource (i.e. music file) using the annotator interface a Globally Unique Identifier⁴ (GUID) is created in order to provide a unique reference number to the loaded resource and that reference is unique in any context. For example, one can load the resource and create an identifier *X* to denote it.

Step2-Identifying the resource type: In order to associate semantic concepts to *X*, the annotator application first determines the type of content (whether the content is mp3, wav or in real audio format) from the resource's file format with respect to the ResourceFormat class as defined in the *mpeg7Music* ontology (The *mpeg7Music* ontology contains a classification of the content type (ResourceFormat class) depending on the file formats).

Figure 3(a): Simple Semantic Annotator- Step 1

Step 3-Displaying relations: In this step, concepts and individuals from the domain ontology is displayed for the user to associate with the resource's content through the relation that the user has chosen in step 3.

Figure 3(b): Simple Semantic Annotator- Step 2

⁴http://en.wikipedia.org/wiki/Globally_Unique_Identifier

Figure 3(c): Simple Semantic Annotator- Step 3

Step 4-Assigning MusicalConcepts from the instances of the annotation ontology: Users may click on the individual of choice to relate it with the resource. Following the example of resource X, a user may create annotation in this step as follows:

"X *relatesTo* (TimbralExpression-Sharpness) *acoustic*"

Figure 3(d): Simple Semantic Annotator- Step 4

Figure 3(e): Simple Semantic Annotator- Step 5

After creating a single annotation the user goes back to step 3 where s/he can choose another relation/property to associate concepts from the property hierarchy of the ontology displayed in step 3 again otherwise if the user chooses to finish the process of associating domain ontology concepts then s/he goes to step 5.

Step 5-Confirmation of Annotation: In step 4, the user may have a final check of the created annotations and confirm the results through saving/storing annotations functionality that stores the results of the annotation process and finally completes the annotation process.

Step 1 starts by providing browsing facility for the user to upload music file on the annotator tool. For example, the file that the user chooses is *example1.wav* and the user clicks the 'upload' button.

Then Step 2 prompts the user to enter a title for the uploaded music resource. Here the title parameter is a mandatory input for annotation. Let the arbitrary title input be 'Divine' and the user clicks the 'next' button and before step 3 is displayed few process is executed at the back end.

Firstly, the system then generates a globally unique identifier (GUID) for the resource so that the resource itself and the generated annotations can be linked together.

Secondly, the MPEG-7 feature extractor plug-in generates XML output (a sample may be found at⁵). Thirdly, the XML DataSource

⁵ <http://carus.aces.shu.ac.uk/coreontology/mpeg-7/AppendixC.xml>

Processor uses the Timbral and Melodic rules as mentioned in the **mpeg-7Music** ontology - the detail of the OWL ontology is available in (Rahman and Siddiqi, 2011) to detect which tag from the MPEG-7 output to extract. For example, timbral rule associates SpectralCentroid (SC) parameter with the *Brightness* concept and melodic rule associates AudioFundamentalFrequency (AFFT) parameter with the Motion concept.

As can be seen from figure 3(c) that step 3 of the GUI is divided into two panels – the left panel displays the property hierarchy for the user to choose for further annotation. The right panel displays a list of assertions auto generated from MPEG-7 output with reference to the rules defined in the ontology. Clearly, MPEG-7 assertions presented on the right panel is a static display, its purposes are to provide the user suggesting on the concept to be chosen and the MPEG-7 data types associated with the resource Divine. The left panel displays the list of properties where the user can click on any of those object properties to associate individual concept of choice.

Let's consider if the user clicks on the *relatesTo* property then step 4 GUI is displayed.

Step4 of the GUI is also divided in two panels (figure 3(d)), the left one displays the class hierarchy from where the user can choose a concept and if a concept is chosen then the corresponding list of individuals are displayed on the right panel. Now, if the user does choose the *Brightness* concept then the individuals under motion sub-class is displayed as shown in figure 3(d). From the right panel then the user can select any individual and let's select **gentle**. Then to add more individuals the 'Back' button must be clicked to go back to Step3 and select any object property to come back again to step4 and so on until the user is happy. Let's assume the user chooses 'describedBy' property and arrives at step 4 and clicked on the Motion concept and selected conjunct individual from the right panel at step4.

After adding two individuals if the user goes back to step 3 and clicks the 'Finish' button then s/he arrives at step 5.

The step 5 displays both the auto generated (maroon rectangle) and user created (light-green rectangle) annotations as shown in

figure 3(e). If s/he is satisfied then click of the 'Submit'

3.4 Evaluation

The annotation process of a musical object using the simple semantic annotation tool involves both automatic and semi-automatic association phases. Automatic annotations are generated using the MPEG-7 descriptors mapping that was detailed in section 3.3. Figure 3 (c) shows an example of the auto-generated annotations marked using maroon rectangle. The semi-automatic annotations were carried out (by music producers) as shown in figures 3 (a-e). The user may upload the music file and name it as shown in figure 3 (a-b). The step 3 of the annotation process (figure 3(c)) displayed the property hierarchy of the **mpeg-7Music** from where the user may choose a property by clicking on it and then step 4 of the annotation process appeared. In step 4 (figure 3(c)), the class hierarchy of **mpeg-7Music** ontology was displayed. As user clicked on the class concept of choice then individuals of the corresponding class was displayed and the user could choose from the list of those individuals. Once the 'Done' button was clicked then the uploaded resource was linked by the chosen property (in step 3 of annotation) with the individual phrase (chosen in step 4) structured under **mpeg7Music** ontology. So, this semi-automatic annotation process was based on two fundamental annotation steps – i.e. Step 3 that related the chosen property from the **mpeg-7Music** ontology with the uploaded music segment; Step 4 in which appropriate class's individuals were discovered and then the user could assign the chosen individual to the music segment. These two fundamental steps were supported by display of and interaction with class & property hierarchy as well as individuals – such functionalities were enabled by the web service client interface used by the semantic annotator application. The web service client interface was served by the web service that was developed for processing OWL ontology.

The popularity of social tagging used by Web 2.0 applications was a result of the low effort required for tagging. This high usability encountered in social tagging served as the blueprint for the Simple Semantic Annotator. The Simple Semantic Annotator was designed to enable the music producers to generate annotations, and attach them to a music resource. These annotations were metadata containing references to the **mpeg-7Music**

ontology. The Simple Semantic Annotator was designed following the usability standards in the *Web 2.0* set for the so-called *social tagging* (Bertin-Mahiex et al., 2008). The high level of usability is considered to be a pre-condition for recruiting the end-users' (Music Producers) effort to annotate resources for the use of others (Music Consumers).

The evaluation of the Simple Semantic Annotator tool was carried out based on criteria mentioned in the Table 7 involving participants who were graduate students not necessarily coming from computer science background. The evaluation session started with six participants to assess the usability of the annotation tool based on the criteria mentioned in table 3.

Table 3: Evaluation criteria for the Simple Semantic Annotator Tool

Validation criteria	Measurement method	Expected results
Easiness of use (usability)	Average time spent on annotating a resource	Annotating time measured in seconds
	User feedback	Level of satisfaction with the usability

The 5 out of 6 participants took less than a minute to finish the five steps of annotator application to annotate a single music item with single object property. Highly satisfactory feedbacks were received from the participants with regard to easiness of use of the annotator application. The objective of the evaluation session was not only to validate the annotation tool but also to collect recommendation for future improvements.

4. CONCLUSION

The Simple Semantic Annotator application introduced three innovations as specified below. The first two were concerned about introducing a balance between user intervention and automation. The third one is related to the design of the annotation procedure with MPEG-7 features information and structured metadata vocabulary from **mpeg-7Music** ontology. The detail of specific points on three aforementioned innovations is as follows:

The minimal set of mandatory information was defined including:

- (i) Globally Unique Identifier for the annotated resource - URL of the annotated resource (provided as the first step in annotation procedure) – created automatically.
- (ii) Title (provided manually).
- (iii) A list of assertions containing MPEG-7 datatype values from the ontology rules – automatically provided.
- (iv) A hierarchy of the concepts (with individuals) from the **mpeg-7Music** ontology (from the hierarchy the user can manually select).
- (v) The property hierarchy is presented to the user from where the user can choose manually.

From above, it is evident that 2 out of 5 mandatory elements are provided in a full automatic manner. Specifically, presentation of the MPEG-7 features is done automatically. The last two mandatory information required user intervention to choose but the hierarchy of concepts and properties were made available to choose by the user; these two are partially automatic. Finally, depending on the selected annotation concept, the system will attempt to automatically select the matching relation, in order to complete the annotation assertion. Only in case that the selection was ambiguous, the user will be prompted to select the annotation relation manually.

5. REFERENCES

- Bateman, S., Brooks and C., McCalla, G. (2006) Collaborative Tagging Approaches for Ontological Metadata in Adaptive ELearning Systems In the Proceedings of the Fourth International Workshop on Applications of Semantic Web Technologies for E-Learning (SW-EL 2006) in conjunction with 2006 International Conference on Adaptive Hypermedia and Adaptive Web-Based Systems (AH2006), June 20, 2006. Dublin, Ireland.
- Bertin-Mahiex, D. Eck, F. Maillet, and P. Lamere. Autotagger: A model for predicting social tags from acoustic features on large music databases. Journal of New Music Research, 2008.
- Calvanese D., Giuseppe De Giacomo, Maurizio Lenzerini, Riccardo Rosati. (2011) 'View-based query answering in Description

- Logics: Semantics and complexity'. Journal of Computer and System Sciences, In Press, Corrected Proof, Available online 15 March 2011
- Cimiano P. and Handschuh S. (2003) Ontology-based linguistic annotation. Proceedings of the ACL 2003 workshop on Linguistic annotation: getting the model right - Volume 19. July 2003
- Connolly, D.; Hendler, J.; Schreiber, G. 2004. Web-Ontology (WebOnt) Working Group, [online] Available at: <www.w3.org/2001/sw/WebOnt/> [Accessed 22 June 2008].
- Crysanadt, H. Hierarchical Sound Classification using Mpeg-7. 2005 IEEE 7th Workshop on Multimedia Signal Processing, Oct. 30 2005-Nov. 2 2005 Page(s):1 – 4.
- Damme V., C., Hepp, M. & Siorpaes, K. (2007), FolksOntology, An Integrated Approach for Turning Folksonomies into Ontologies. Proceedings of the ESWC 2007 Workshop "Bridging the Gap between Semantic Web and Web 2.0".
- Dedek, A. Eckhardt, L. Galambos, P. Vojtas. Uncertainty Reasoning for the World Wide Web. W3C Incubator Group Report 31 March 2008, available from <http://www.w3.org/2005/Incubator/urw3/XGR-urw3/>
- Xin Fu, Tom Ciszek, Gary Marchionini., Paul Solomon. Annotating the Web: An exploratory study of Web users' needs for personal annotation tools. Proceedings of the American Society for Information Science and Technology Volume 42, Issue 1, Date: 2005, Pages: NA
- Gruber T. R. (1993) "A Translation Approach to Portable Ontology Specifications", Knowledge Acquisition, 5(2), 199-220.
- Gruber, T. (2008). Collective knowledge systems: Where the Social Web meets the Semantic Web. Journal of Web Semantics 6(1): 4-13 (2008)
- Claudio Gutierrez, Carlos A. Hurtado, Alberto O. Mendelzon, Jorge Perez. (2011) 'Foundations of Semantic Web databases'. Journal of Computer and System Sciences, Volume 77, Issue 3, Database Theory, May 2011, Pages 520-541
- Handschuh, S, Staab, S, and Maedche, A . CREAM - Creating relational metadata with a component-based, ontology-driven annotation framework. ACM K-CAP, 2001.
- Steve Hardin. 2005. Tim Berners-Lee: The Semantic Web - Web of Machine-Processable Data. Bulletin of the American Society for Information Science and Technology. Volume 31, Issue 3, pages 12–13, February/March 2005. Willey online Library.
- L. Hardman. Canonical Processes of Media Production. In Proceedings of the ACM workshop on Multimedia for human communication. ACM Press, 2005.
- Herman I., 2007. Web Ontology Language (OWL), [online] Available at: <<http://www.w3.org/2004/OWL/>> [Accessed 14 April 2007]
- Herman I., nd. W3C Semantic Web Activity, [online] Available at: <www.w3.org/2001/sw/> [Accessed 22 June 2009].
- Hunter, J. Enhancing the semantic interoperability of multimedia through a core ontology. IEEE Transactions on Circuits and Systems for Video Technology, Volume 13, Issue 1, Jan. 2003 Page(s):49 – 58
- Kosch, H.; Boszormenyi, L.; Doller, M.; Libsie, M.; Schojer, P.; Kofler, A. (2005) The life cycle of multimedia metadata. Multimedia, IEEE. Volume 12, Issue 1, Jan.-March 2005 Page(s):80 – 86
- Guan-yu Li; Sui-ming Yu; Sha-sha Dai. Ontology-Based Query System Design and Implementation. International Conference on Network and Parallel Computing Workshops, 2007. NPC Workshops. IFIP. 18-21 Sept. 2007 Page(s):1010 – 1015
- Mahmoud, H. 2005. Service-Oriented Architecture (SOA) and Web Services: The Road to Enterprise Application Integration. [online] Available at: <<http://java.sun.com/developer/technicalArticles/WebServices/soa/>> [Accessed 8 November 2008].
- Eric Miller and Ralph Swick. An Overview of W3C Semantic Web Activity. Bulletin of the American Society for Information Science and Technology—April/May 2003.
- Petridis, K.; Bloehdorn, S.; Saathoff, C.; Simou, N.; Dasiopoulou, S.; Tzouvaras, V.; Handschuh, S.; Avrithis, Y.; Kompatsiaris, Y.;

Staab, S. Knowledge representation and semantic annotation of multimedia content. Vision, Image and Signal Processing, IEE Proceedings- Volume 153, Issue 3, 8 June 2006 Page(s):255 - 262

Rahman F. and Siddiqi J. (2011) 'Semantic Annotation of Digital Music'. Accepted for publication in Journal of Computers and Systems Sciences, April 2011. Elsevier Inc. 2011.

Ruvane. B. (2006) Defining annotations: a visual (re)interpretation. Proceedings of the American Society for Information Science and Technology. Volume 43, Issue 1, Date: 2006, Pages: 226

Schroeter, R., Hunter, J., Guerin, J., Khan, I., and Henderson, M. 2006. A Synchronous Multimedia Annotation System for Secure Collaboratories. In Proceedings of the Second IEEE international Conference on E-Science and Grid Computing (December 04 - 06, 2006). E-SCIENCE. IEEE Computer Society, Washington, DC.

Smith, J.R.; Schirling, P. Metadata standards roundup. Multimedia, IEEE, Volume 13, Issue 2, April-June 2006 Page(s):84 - 88

Sofia Tsekeridou, Athina Kokonozi, Kostas Stavroglou, Christodoulos Chamzas: MPEG-7 Based Music Metadata Extensions for

Traditional Greek Music Retrieval. MRCS 2006: 426-433.

D. Turnbull, L. Barrington, G. Lanckriet. Five Approaches to Collecting Tags for Music. ISMIR 2008.

Uren, Victoria; Cimiano, Philipp; Iria, Jose; Handschuh, Siegfried; Vargas-Vera, Maria; Motta, Enrico; Ciravegna, Fabio. Semantic annotation for knowledge management: Requirements and a survey of the state of the art. Web Semantics, v 4, n 1, January, 2006, p 14-28.

Yang, S.-Y., An ontological website models-supported search agent for web services, Expert Systems with Applications (2007), doi:10.1016/j.eswa.2007.09.024

Zang B., Yinsheng Li, Wei Xie, Zhuangjian Chen, Chen-Fang Tsai, Christopher Laing. (2008) 'An ontological engineering approach for automating inspection and quarantine at airports'. Journal of Computer and System Sciences, Volume 74, Issue 2, Web and Mobile Information Systems 2006, March 2008, Pages 196-210

Zhang, J. (2007). Ontology and the Semantic Web. In Tennis, Joseph T., Eds. Proceedings North American Symposium on Knowledge Organization 2007 1, pages pp. 9-20, Toronto, Ontario