# Retrieving Film Heritage content using an MPEG-7 Compliant Ontology

Y. Cobos, C. Sarasua, M.T. Linaza, I. Jiménez, A. García
*Dept. of Tourism, Heritage and Creativity- Visual Communication Technologies VICOMTech*
*{ycobos, csarasua, mtlinaza, ijimenez, agarcia}@vicomtech.org*

## Abstract

*This paper presents the design and implementation of an MPEG-7 based Multimedia Retrieval System for Film Heritage. The multimedia content has been indexed using an Annotation Tool based on MPEG-7 standard. Moreover, an MPEG-7 Compliant Ontology in OWL DL, which is briefly explained in this paper, has been specially developed to fulfil the requirements of the CINeSPACE project. This ontology has been instantiated so that the retrieval process can be handled. The system has been assessed during the validation of the CINeSPACE project.*

## 1. Introduction

The availability of huge amounts of multimedia documents requires a careful design and an efficient implementation of multimedia retrieval systems that facilitate storage, retrieval and browsing of not only textual, but also image, audio and video files. The need for a high-level representation that captures the true semantics of a document has led to the development of the MPEG-7 standard for describing multimedia documents using machine-consumable metadata descriptors.

On the other hand, multimedia retrieval systems have always forced humans to describe their query in terms of a written language. Systems that are aware of multimedia semantics have already flourished in the multimedia information retrieval community. These systems allow the user to specify a set of keywords or concepts, which are thus used to search for multimedia contents containing those concepts. This is already a big step towards more semantic search engines from previous approaches, but it still may be too limiting in some cases.

This paper presents the design and implementation of an MPEG-7 based Multimedia Retrieval System for Film Heritage. Multimedia content is indexed using an Annotation Tool based on MPEG-7 standard. Moreover, an MPEG-7 Compliant Ontology in OWL DL has been specially developed to fulfil the requirements of the CINeSPACE project. This ontology has been instantiated so that the searching and retrieval process can be handled. The system has been assessed during the validation of the CINeSPACE project.

The paper is organized as follows. Section 2 introduces some basic concepts and a brief state of the art. The CINeSPACE project is outlined in Section 3. Section 4 explains the methodology developed for the instantiation of the CINeSPACE MPEG-7 Compliant Ontology. In Section 5, the Multimedia Retrieval System is briefly outlined. Section 6 deals with the evaluation of the implemented system. Finally, Section 7 presents conclusions and future work.

## 2. Related work

Multimedia ontologies model the multimedia domain, especially media structure descriptions related to still images and videos. Structures and semantics are carefully modelled to be largely consistent with existing multimedia description standards like MPEG-7 [1]. From 2001 until the present time, four main ontologies that formalize the MPEG-7 standard using Semantic Web languages have been proposed.

Hunter [2] provided the first attempt to model some parts of MPEG-7 into RDFS. This ontology covers only the part of MPEG-7 related to high-level features and has been integrated with the ABC model [3] into an OWL Full ontology. In 2004, Tsinaraki et al [4] start from the core of this ontology and extend it to cover the complete Multimedia Description Scheme (MDS) part of MPEG-7 in an OWL DL ontology.

A complementary approach was explored by Isaac and Troncy [5], who proposed a core audio-visual ontology inspired by several vocabularies (MPEG-7, TV Anytime, ProgramGuideML). Garcia and Celma [6] produced the first complete MPEG-7 ontology by automatically generating a generic mapping from XSD to OWL. The definitions of the XML Schema types and elements of MPEG-7 are converted into OWL Full axioms.

Simou et al [7] proposed an OWL DL Visual Descriptor Ontology (VDO) based on the visual part of MPEG-7 and used for image and video analysis. Finally, the European SIMAC [8] project has implemented a complete and automatic mapping of the whole MPEG-7 standard to OWL. In this project, a MPEG-7 OWL ontology is generated by XSD2OWL, which constitutes the basic ontological framework for semantic multimedia metadata integration.

Regarding ontology instantiation, a number of approaches have been taken to speed up ontology instantiation of a Web using a variety of techniques. An example of such a tool is Knote which creates the forms directly from the ontology [9]. As in CINeSPACE, all the instances have to be created and maintained manually. The Artequakt system [10] is concerned with automating the extraction of knowledge about the life and work of artists from Web documents, and using it to generate tailored bibliographies.

Finally, in the Mondeca portal, instances of ontology concepts can be added in various ways (forms, linguistic tools). To fill the knowledge base with data, the system imports OWL files using XTM as an export format.

On the other hand, semantic search has become very popular recently. Many systems have been implemented, such as QuizRDF [11], SCORE [12] or OWLIR [13]. A shortcoming of the available approaches is that they do not exploit the connections within the ontology, but use the ontology only as a kind of thesaurus. Only some of them use the labels of the ontology elements, and some of them exploit the concept of taxonomy.

ERIC7 [14] is a software test-bed that implements Content-Based Image Retrieval (CBIR) compatible with the MPEG-7 multimedia standard. The visible outcome is a web site that allows video retrieval using a proprietary XQuery-based search engine, based in a XML native database.

There are also others approaches which query directly to the MPEG-7 XML files. For example, the MP7QF [15] framework is based on MP7QF XML schema types. This framework uses the MPEG-21 digital item declaration language (DIDL) for exchanging MP7QF. Finally, the iFinder [16] system is an MEPG-7 based interesting framework in which XML based metadata is the basis of the multimedia management process.

## 3. Application scenario: the CINeSPACE project

CINeSPACE project [17] aims at designing and implementing a mobile rich media collaborative information exchange platform, scalable, accessible through a wide variety of networks and therefore, interoperable and Location-based for the promotion of Film Heritage, going beyond the current state of the art.

CINeSPACE will enable users to interact with Location-based multimedia contents while navigating a city. Audiovisual information will be delivered through a small low-cost wireless Head Mounted Display with a high definition screen situated near-the-eye, and audio phones. CINeSPACE also comprises a camera able to record or send what the user is "seeing". This information can be uploaded to a database through a WLAN hot spot or a 3G connection in order to create collaborative experiences with other end users.

The system has been designed to be client and platform independent. Taking into account these requirements, Web services have been chosen as the interface of the system in order to achieve interoperability between different platforms and programming languages.

Multimedia content is stored in a Rich-Media Server in each city (Glasgow, Venice and San Sebastian). Although not focused specifically in optimizing annotation, CINeSPACE will provide the cities with tools to manually annotate the multimedia contents that will be available for the prototype. Once the content is indexed with the CINeSPACE annotation tool, the metadata will be used to query the system.

On the other hand, the user of the system will retrieve information. The main inputs received by the system will be IDs of the users, locations and profiles. With this information, the system will maintain the state of each user and deliver the content that best fits the requirements of the user. Search by geo-reference means that the CINeSPACE will show multimedia information around the current location of the user. Customer personalization will further filter the information taking into account the profiles defined.

## 4. The CINeSPACE MPEG-7 Compliant Ontology and Knowledge Base

The methodology for the creation and instantiation of the CINeSPACE MPEG-7 Multimedia Compliant Ontology includes seven steps (Figure 1). First, the metadata required regarding the CINeSPACE scenarios [18] have been selected. Three types of descriptions

have been considered: structural description, classification description and semantic description.

Several concepts coming from the specific requirements of the CINeSPACE project have also been taken into account, such as Point of Interest (*POI*) to inform about the closest POI of the multimedia content. A POI can be considered as any part of the city with some tourist and/or Cultural Heritage attractions, e.g. La Concha beach in San Sebastian or Piazza San Marco in Venice.

Most of the descriptors of the MPEG-7 standard have been used as the standard specifies, such as *PrivateIdentifier*, which is included in the *Description Metadata Scheme*. Moreover, other descriptors from the standard have been modified, such as the *thumbnail* of the multimedia content or *TitleMedia*, which have been included in the definition of the *Creation Description Scheme*. The addition of these new concepts has extended the MPEG-7 CINeSPACE Compliant Ontology, without modifying the basic descriptors coming from the standard.

```
<?xml version="1.0" encoding="UTF-8" ?>
- <Mpeg7 xmlns="urn:mpeg:mpeg7:schema:2001" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance">
  - <DescriptionMetadata>
      <PrivateIdentifier>0000000001</PrivateIdentifier>
      <PDA>TRUE</PDA>
      <Binoculars>TRUE</Binoculars>
      <BlackandWhite>FALSE</BlackandWhite>
      <Thumbnail>file:/C:/CINESPACE_WS_CONTENT/CONTENT/SanSebastian/Image/0000000001.jpg</Thumbnail>
    - <POI>
        <id>0000000001</id>
        <idCity>0000000001</idCity>
      - <GeographicPoint>
          <Longitude>1.985556</Longitude>
          <LongOrientation>W</LongOrientation>
          <Latitude>43.321114</Latitude>
          <LatOrientation>N</LatOrientation>
        </GeographicPoint>
      </POI>
    </DescriptionMetadata>
  - <Description xsi:type="UserDescriptionType">
    - <UserPreferences>
      - <FilteringAndSearchPreferences>
        - <CreationPreferences>
```

**Fig. 1. MPEG-7 file describing an image.**

As a second step, multimedia contents are indexed using the CINeSPACE Annotation Tool. The output of the Annotation Tool consists on pairs of XML files and multimedia contents, that is, there is one XML file description for each multimedia item. Each XML file includes three different parts: description metadata, where all specific CINeSPACE descriptors are included; user description type to include descriptors related to user preferences; and content entity type that describe multimedia items (Figure 2).

The third step is the generation of the Java classes related to the selected metadata of the MPEG-7 standard. The Castor framework [19] has been selected, an Open Source data binding framework which allows the combination of Java objects, XML documents and relational tables. Among the four main features of Castor, the CINeSPACE project has mainly focused on two of them: Castor XML and Castor XML Code. The former is a XML data binding framework to bind XML artefacts to Java objects and vice versa. The latter generates Java source code from XML Schema information.

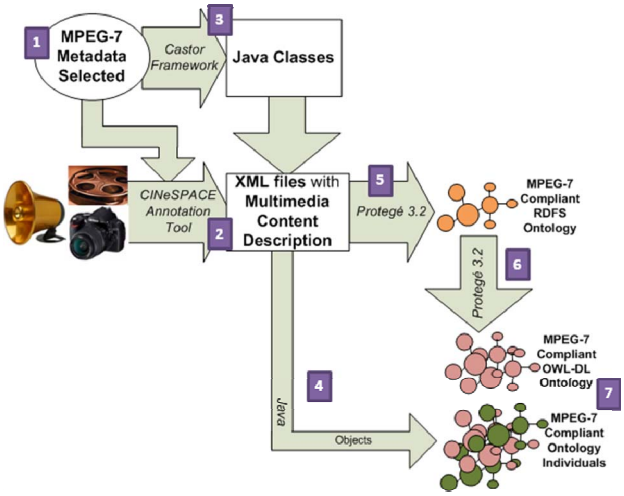


**Figure 2. Methodology to instantiate the CINeSPACE MPEG-7 Compliant Ontology.**

The Java classes created are then instantiated with the corresponding XML file description. Java objects are the instances of all the selected MPEG-7 concepts mapped to Java classes.

In a further step, the CINeSPACE MPEG-7 Compliant Ontology has been designed on the basis of the MPEG-7 descriptions in XML generated by the Annotation Tool. The concepts within the MPEG-7 descriptions are connected in a plain tree structure. Therefore, several relationships have been established among the concepts in order to get an MPEG-7 Compliant Ontology.

Written in RDF, this ontology has been updated into a MPEG-7 Compliant OWL-DL Ontology (Figure 3). The ontology has been developed using the version 3.2 of the Protégé Ontology Editor [20] with the XML structure of the multimedia content as the input, while the output is the MPEG-7 compliant OWL-DL ontology.

Finally, the combination of the ontology with the Java objects produces the final CINeSPACE MPEG-7 Compliant Ontology with Individuals. The most significant metrics of this OWL DL model are 45 classes, 41 object-type properties, 33 data-type properties and 63 individuals.

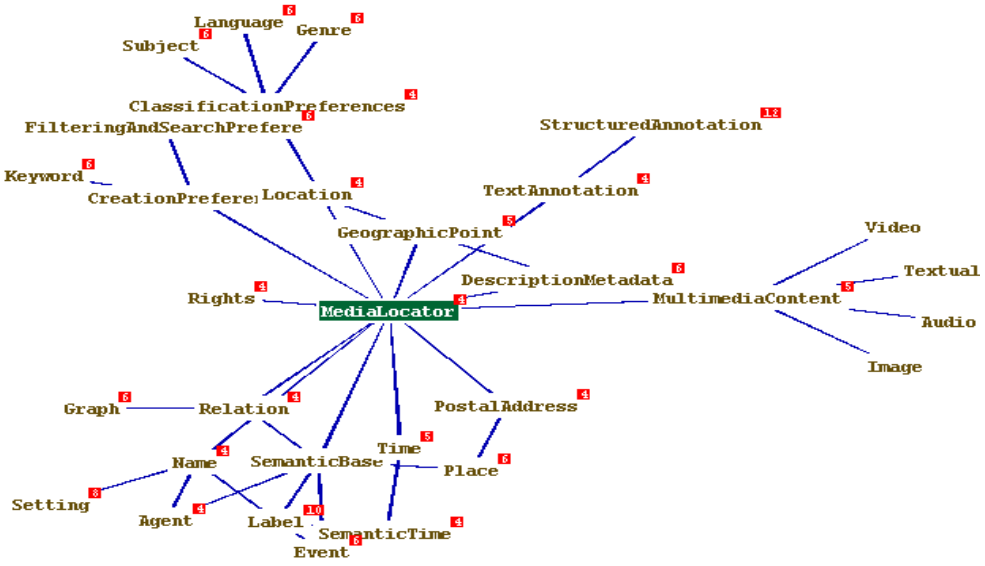


**Fig. 3. MPEG-7 Compliant RFDS Ontology.**

## 5. CINeSPACE Multimedia Retrieval System

Once multimedia content has been indexed and stored, it is necessary to search and query over the ontology. Queries are made over the CINeSPACE MPEG-7 Compliant Ontology that is used to semantically retrieve content. This ontology located at the server part is the core of a Web Service developed for this Multimedia Retrieval System, which has been implemented in Java 2.

The Web Service has been developed with Eclipse 3.2 [21], an Open Development Platform. The client sends all the queries to the server, while the server provides the results of the queries from the ontology through the Service. The interface of the client application is depicted in Figure 4.

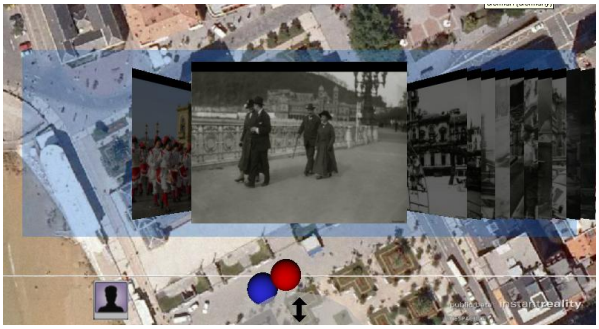

**Fig. 4. Client interface in the Multimedia Retrieval System of CINeSPACE.**

Due to the Location-Based nature of the project, queries are mostly related to geo-reference information stored in the CINeSPACE MPEG-7 Compliant Ontology. Although not the only ones, these types of queries can be defined as follows in a natural language way:

Give me all the multimedia content…

(i) … around this geo-reference data.
(ii) … in black and white colour.
(iii) … visualized by the $user_x$.
(iv) … created yesterday by $user_x$.

Among all of the possible queries related to the Location-Based contents, a complex set of test queries, composed of six queries [22] has been chosen to test the proposed architecture as shown in Table 1.

$$Q = \{q_1 \cup q_2 \cup q_3 \cup q_4 \cup q_5 \cup q_6\}$$

**Table 1.** Simple input queries.

| |
|---|
| $q_1$ = CLASS *Geographic Point* with the PROPERTY *Latitude* GREATER A |
| $q_2$ = CLASS *Geographic Point* with the PROPERTY *Latitude* SMALLER B |
| $q_3$ = CLASS *Geographic Point* with the PROPERTY *Longitude* GREATER C |
| $q_4$ = CLASS *Geographic Point* with the PROPERTY *Longitude* SMALLER D |
| $q_5$ = CLASS *Geographic Point* with the PROPERTY *LatOrientation* EQUALS E |
| $q_6$ = CLASS *Geographic Point* with the PROPERTY *LongOrientation* EQUALS F |

$\beta_1$=0.0035, the radius in degrees around the current latitude

$\beta_2$=0.0035, the radius in degrees around the current longitude

x, queried latitude

y, queried longitude
$A<x<B,\ A=x-\beta 1,\ B=x+\beta 1$
$C<y<D,\ C=y-\beta 2,\ D=y+\beta 2$
E={North, South}
F={West, East}

## 6. Evaluation of the system

We now discuss the user experiments with the CINeSPACE multimedia content for Film Heritage. The pilot user study was conducted by over 30 users in Glasgow, San Sebastian and Venice. The trials were conducted to demonstrate the effectiveness of the location-based retrieval.

The CINeSPACE database included the following multimedia content for the system evaluation: 8 videos from Glasgow; 17 videos and 12 images from Venice; 4 videos and 22 images from San Sebastian. All the multimedia content has been previously indexed using the CINeSPACE Annotation Tool.

The CINeSPACE Multimedia Retrieval System has been tested in January 2008 in a Windows XP SP 2 computer with an Intel Celeron CPU 3.33GHz with 512 MB of RAM.

A simplified use case of the first evaluation test would be the following. The user logins into the CINeSPACE system, and he/she starts moving around the city. The Application Layer receives a query informing that the user is physically located at the position 43.3205º N, 1.9883º W (Playa de la Concha, San Sebastian), with a CINeSPACE compliant device.

The system checks the multimedia contents that can be rendered at that location, searching for the closest POI near the geographic data. Finally, the Application Layer retrieves the multimedia content (images and videos) and renders them on the CINeSPACE device. It must be mentioned that users do not formulate queries over the ontology.

The evaluation in each of the cities led to several comments [24]. The main retrieval inaccuracies were due to the limitations of the GPS sensors as some POI specially in Venice were sometimes incorrect especially in the area near the bell tower. This was mainly due to the loss of GPS coverage at some points of the field trial. Furthermore, the refresh of the GPS position was too slow in some areas of the field trial. However, users were satisfied with the multimedia content provided by the three cities.

Further comments were related to the quantity of the multimedia content available for the evaluation and retrieved (enough or not enough); the problems of achieving the same contents at close POI due to the uncertainties in the position estimation; or the waiting time because of the bandwidth restrictions of the wireless networks of the cities.

## 7. Conclusions and future work

This paper presents the design and implementation of an MPEG-7 based Multimedia Retrieval System for Film Heritage. A methodology for the design and instantiation of the CINeSPACE MPEG-7 Compliant Ontology has been developed. First, the metadata required for the CINeSPACE scenarios have been selected. Furthermore, multimedia contents selected fro the evaluation have been indexed using the CINeSPACE Annotation Tool.

The third step is the generation of the Java classes related to the selected metadata using the Castor framework. The Java classes created are then instantiated with the corresponding XML file description.

The CINeSPACE MPEG-7 Compliant Ontology has been designed mainly based on the MPEG-7 descriptions generated in XML by the CINeSPACE Annotation Tool. In a further step, the MPEG-7 Compliant RDFS Ontology has been updated in order to get the MPEG-7 Compliant OWL-DL Ontology. Finally, the combination of the ontology with the Java objects produces the final CINeSPACE MPEG-7 Compliant Ontology with Individuals.

This methodology has been assessed during the validation of the CINeSPACE project. Queries have been made over the CINeSPACE MPEG-7 Compliant Ontology that is used to semantically retrieve content. This ontology located at the server part is the core of a Web Service developed for this Multimedia Retrieval System, which has been implemented in Java 2. Due to the Location-Based nature of CINeSPACE, queries are mostly related to geo-reference information stored in the CINeSPACE MPEG-7 Compliant Ontology.

Regarding future work, the current Multimedia Retrieval System will be extended in several ways. On the one hand, new criteria for queries will be added to personalize content retrieval taking into account the profile of the user. On the other hand, the system will include more advanced retrieving facilities, that is, high-level semantic queries support.

Finally, the CINeSPACE MPEG-7 Compliant Ontology will be mapped to other standard domain ontologies such as the CIDOC Conceptual Reference Model (CIDOC-CRM) or the Geoconcepts ontology.

## 8. Acknowledgements

This work has been partially funded under the 6th Framework Programme of the European Union within the IST project "CINeSPACE" (IST FP6- 034990, http://www.cinespace.eu). Authors would also like to thank the Basque Government for the funding provided through the ETORTEK strategic project eTourgune. This work has been done within the COST Action 292.

## 7. References

[1] MPEG-7, ISO/IEC JTC1/SC29/WG11 Coding of Moving Pictures and Audio. http://www.chiariglione.org/mpeg/standards/mpeg-7/mpeg-7.htm.

[2] J. Hunter, "Adding Multimedia to the Semantic Web - Building an MPEG-7 Ontology", *1st International Semantic Web Working Symposium,* California, USA, 2001, pp. 261–281.

[3] J. Hunter, "Enhancing the semantic interoperability of multimedia through a core ontology", *IEEE Transactions on Circuits and Systems for Video Technology*, 2003, pp. 49–58.

[4] C. Tsinaraki, P. Polydoros, F. Kazasis, and S. Christodoulakis, "Ontology-based Semantic Indexing for MPEG-7 and TV-Anytime Audiovisual Content", *Multimedia Tools and Application Journal*, 2005, pp. 299-325.

[5] A. Isaac, and R. Troncy, "Designing and Using a Multimedia Description Core Ontology", *Workshop on Core Ontologies in Ontology Engineering*, Northamptonshire, UK, 2004.

[6] R. Garcia, and O. Celma, "Semantic Integration and Retrieval of Multimedia Metadata", *5th International Workshop on Knowledge Markup and Semantic Annotation*, Galway, Ireland, 2005.

[7] N. Simou, V. Tzouvaras, Y. Avrithis, G. Stamou, and S. Kollias, "A Visual Descriptor Ontology for Multimedia Reasoning", *Workshop on Image Analysis for multimedia Interactive Services*, Montreux, Switzerland, April 2005.

[8] E. Motta, S. Buckingham Shum, and J. Domingue, *Ontology Driven Document Enrichment,* 2000.

[9] H. Alani, S. Kim, D.E. Millard, M.J. Weal, W. Hall, P.H. Lewis, and N. Shadbolt, "Automatic Ontology-based Knowledge Extraction and Tailored Biography Generation from the Web", *IEEE Intelligent Systems*, 18/1, 2003, pp. 14-21.

[10] J. Davies, R. Weeks, and U. Krohn, "QuizRDF: Search Technology for the Semantic Web", *11th International World Wide Web Conference*, 2002, pp. 24.32.

[11] A. Sheth, C. Bertram, D. Avant, B. Hammond, K. Kochut, and Y. Warke, "Inductry Report: Managing Semantic Content for the Web", *IEEE Internet Computing*, 2002.

[12] T. Finin, J. Mayfield, A. Joshi, R.S. Cost, and C. Fink, "Information Retrieval and the Semantic Web", *Proc. of the 38th Annual Hawaii International*, 2005, pp. 113a.

[13] I. Gagnon, S. Foucher, and V. Gouaillier, "ERIC7: A Generic Environment for Content-Based Image Retrieval within the MPEG-7 Standard", *Proc. of the 2004 Winter International Symposium on Information and Communication Technologies*, 2004, pp. 322-327.

[14] M. Doeller, H. Kosch, W. Ingo, and T. Matthias, "MP7QF: An MPEG-7 Query Format", *The International Conference on Signal-Image Technology & Internet-Based Systems*,. Hammamet, Tunisia, pp.36-45.

[15] J. Löffler, K. Biatov, C. Eckes, and J. Köhler, "IFINDER: an MPEG-7-based retrieval system for distributed multimedia content", *Proc. of the 10th ACM International Conference on Multimedia*, 2002, pp. 431-435.

[16] P. Herrera, J. Bello, G. Widmer, M. Sandler, O. Celma, F. Vignoli, E. Pampalk, P. Cano, S. Pauws, and X. Serra, "SIMAC: Semantic interaction with music audio contents", *Proc. of the 2nd European Workshop on the Integration of Knowledge, Semantic and Digital Media Technologies,* Savoy Place, London, 2005.

[17] P. Santos, A. Stork, M.T. Linaza, O. Machui, D. McIntyre, and E. Jorge, "CINeSPACE: Interactive Access to Cultural Heritage While On-The-Move", *HCI International*, Beijing, China, July 22-27, 2007, pp. 435-444.

[18] Y. Cobos, M.T. Linaza, A. García, and I. Torre, "Applicability of MPEG-7 Descriptors to Film Heritage", *Proc. of SMAP 2007,* Uxbrigde, UK, pp. 205-210.

[19] Castor Framework. http://www.castor.org/.

[20] Protégé Ontology Editor and Knowlegde Acquisition System. http://protege.stanford.edu/.

[21] Eclipse. http://www.eclipse.org/.