Semantics-Driven Multimedia Data Access

Marian Dardala, Adriana Reveiu, Ion Smeureanu
Academy of Economic Studies Bucharest, Romania
dardala@ase.ro, reveiu@ase.ro, smeurean@ase.ro

Abstract. The multimedia data are more present in the actual informatics systems. Because of characteristics of multimedia data, their management and querying techniques are unlike than those of traditional data. We can increase the potential use of multimedia data across various applications by assuring semantics-driven access to multimedia data. In this paper we propose a solution, based on MPEG-21 standard, to describe the multimedia content and to integrate it with the semantic web.

Keywords. Data access, MPEG-21 standard, multimedia data, ontology, semantic web

1. Introduction

The development of multimedia applications from the last years, due to the exponential growth of the Internet had as consequence a great usage of multimedia data.

Unfortunately, the more data become available, the harder it is to identify and to find what you want, and the more difficult it becomes to manage the data.

People looking for content typically use web browsers with moderate retrieval performance. The web engines rely on human operators to manually describe the multimedia content using keywords. This solution is unacceptable because it is a costly process, and because these descriptions are inherently subjective.

Accessing multimedia data based on its semantics add powerful features to multimedia applications.

The semantic web is an extension of the current web where, in addition to being human-readable, multimedia files are annotated with metadata. Metadata definitions are the base for multimedia content description. Metadata are data about data. Metadata are used to describe and to store characteristics of multimedia resources, and are helpfully to decide quickly how can be processed the multimedia content. The usage of metadata is suitable for multimedia resources.

The representation of metadata, accompanied by semantic theories, will enable a web that provides a qualitatively new level of services. Metadata can exist at two levels: they may describe a video sequence and on the other hand, they may describe entities within the video sequence, for example the key frames of the video sequence.

The metadata add semantic to multimedia data, which tells us about the content of a multimedia file like its subject matter or relationship with other files, or about an entity within the media file.

There are a number of additional services which this metadata can enable [5].

Using semantics we can improve the solutions to present the information. Instead of a search, providing a linear list of results, the results can be clustered by meaning. Using semantics the informational systems can understand where words, or phrases, or multimedia content are equivalent.

Ontologies are a key enabling technology for the semantic web. They interweave human understanding of symbols with their machine processability.

There are many applications of semantics-driven data access approach like automated information processing, information integration and knowledge management.

2. Multimedia Data Features

Because of multimedia data characteristics, their management and querying techniques are unlike than those of traditional data.

Multimedia data are heterogeneous from many points of view:
- Some data are time dependent and the others are time independent,
- Multimedia data use different formats for data representation,
- Some data are structured and the others are represented as unstructured or as semi-structured streams of data,
Some data can be transferred remotely in a short time and the others need a large period to be transferred. In contrast with text-based systems, multimedia applications must include powerful descriptions of human thinking in video, audio, animations and images [2]. But it is a big gap between how the users think about the reality and the physical representation of multimedia data. Users want to access multimedia data at the object level, but multimedia-retrieval systems tend to represent multimedia data based on their low-level characteristics such as colour, patterns, textures and shapes, with no way to automatically represent the human thinking issue. The representation of multimedia data according to human’s perspective is difficult to be realized. No system can provide automated identification or classification of objects from multimedia data streams [6].

Multimedia data collections have various types and are poorly indexed. In a distributed environment, because of the autonomy and diversity of data sources, multimedia objects are often represented in heterogeneous data types. The difference of data formats leads to the difficulty to integrate multimedia objects within a unique indexing framework.

3. Semantics for Multimedia Data

At the heart of all semantic web applications is the use of ontologies. A commonly agreed definition of ontology is: ‘An ontology is an explicit and formal specification of a conceptualization of a domain of interest’ [3]. The use of ontologies has also become widespread in fields such as intelligent information integration, cooperative information systems, information retrieval, electronic commerce, knowledge management and multimedia applications.

Ontologies describe semantic web-based knowledge management architecture and a suite of innovative tools for semantic information processing.

To enable multimedia content to be discovered and exploited by services, agents and applications on the semantic web it needs to be described semantically.

Generating descriptions of multimedia content is inherently problematic because of the volume and complexity of these kinds of data, its multidimensional nature and the potentially high subjectivity of human-generated descriptions.

Significant progress has been made in recent years on automatic segmentation and structuring of multimedia content and the recognition of low-level features within such content. However, comparatively little progress has been made on machine-generation of semantic descriptions of audiovisual information.

The use of semantic metadata are crucial to integrating information from heterogeneous sources, whether within one organization or across organizations. Typically, different schemas are used to describe and classify information, and different terminologies are used within the information. By creating mappings between the different schemas, it is possible to create a unified view, and to achieve interoperability between the processes which use the information.

There are some technologies available for ontology development like: OWL (Web Ontology Language) and RDF (Resource Description Framework) but most of them are not adapted to multimedia data characteristics.

The OWL language provides mechanisms for creating all the components of ontology: concepts, instances, properties (or relations) and axioms.

Resource Description Framework is essentially a data modelling language, graph-based, but it consists of triples: subject, predicate and object.

4. Multimedia Content Description

Metadata definitions are the base for multimedia content description. Metadata are used to describe and to store some characteristics of multimedia resources and is used to decide quickly how can be processed the multimedia content. The usage of metadata is suitable for multimedia resources.

There are a lot of standards and research projects having as a goal the multimedia content description, the most successful being: Dublin Core Standard, Metadata Dictionary SMPTE (Society of Motion Picture and Television Engineers), MPEG-7 and MPEG-21. These are general standards; there are not dedicated to some application topologies, with applicability in various areas and having the support of some prestigious international organizations. [4]

Dublin Core standard uses as support for multimedia resources description a set of metadata. This metadata facilitate the electronic resources detection. The goal of Dublin Core
standard is to describe multimedia resources for some TV archives.

**Metadata Dictionary SMPTE** project defined a large collection of names and data types developed for TV and video industry. The aim of this project was to propose a universal method for resources management and a hierarchical solution for data structures. It uses metadata to describe attributes for environment description. The main disadvantage of this system is that it does not use the semantic annotations.

**MPEG-7** is a standard for multimedia metadata descriptions, based on XML opened standard. MPEG-7 proposes descriptive elements for the whole succession of multimedia processing: from multimedia data acquisition, i.e. devices classifications, devices characteristics description, to data analyze and filtering, to delivery, i.e. environment descriptors, and till the interface with the final users, i.e. descriptors for users’ preferences. [9]

**MPEG-21** standard defines a robust and flexible multimedia opened framework, which identify a distributed architecture for multimedia elements interactions. The aim of the MPEG-21 standard is to define an opened framework that allows transparent usage and multimedia resources adaptation, using various networks and terminal devices.

### 5. MPEG-21 Standard Presentation

MPEG-21 is an ISO/IEC 21000 standard of Moving Picture Experts Group (MPEG) that defines an open framework for multimedia. The power of MPEG-21 is demonstrate in the following situation: there are a lot of multimedia resources which could be used to develop an infrastructure for multimedia content delivery and consume, but there is no architecture to describe the interactions between their elements.

The goal of MPEG-21 standard is to define an open framework for multimedia, which allows the transparent usage of multimedia resources, networks and peripheral devices, assuring good performances in the multimedia resources controlling process. MPEG-21 offers an open support to deliver and to use the multimedia data. [4]

MPEG-21 covers the whole flow of multimedia content processing from the delivery channels point of view. MPEG-21 refers to the solutions to produce, to consume, to present and to sell the multimedia content, having the goal to personalize the multimedia applications.

MPEG-21 standardizes the stream of multimedia information and services from the content development until the delivery to the final users. To realize these stages it is necessary to identify, to describe, to manage and to protect the multimedia content. The multimedia content transport and delivery could be perform throw various networks and between different types of terminals.

MPEG-21 is based on two important concepts: the definition of a fundamental unit for distribution and transaction, namely the digital element, and the users that interact with digital elements.

The digital element is a digital object with a standard structure that has metadata associated with it. The digital elements contain multimedia resources, the content, and metadata associated with the resources or with the digital element.

In the MPEG-21 standard, a user is any entity that interacts with the MPEG-21 environment or that uses a digital element, like a multimedia data consumer, an organization or another standard that use multimedia resources. A user can consume multimedia content by publishing or by delivering it and can have specific rights and responsibilities.

The main elements that define the MPEG-21 architecture are: [10]
- The **digital element** is a hierarchical container for heterogeneous resources like video, audio, text and so on, metadata and other digital elements. A digital element is an elementary unit for delivering and transitioning;
- Digital element declaration (DED) defines a set of terms used to declare the digital elements;
- The **identification** and the declaration of the **digital element** (DEA) supposes to identify and to declare the digital element, it’s nature, it’s type or it’s granularity;
- Content management and usage offers the interfaces and protocols that allow to create, to manipulate, to search, to access, to store, to deliver and to reuse the multimedia content over the distribution and consume channels;
- The **intellectual property protection and management** allows the multimedia content management and the digital elements protection inside networks and devices;
- Terminals and networks offer the tools that allow the transparent access to the content throw networks and terminals, with the service quality control;
- **Content representation** is responsible for media resources representation;
- **Events reporting** contain metrics and interfaces that allow users to evaluate the performances of all events reported inside the system.

- The **Component** realizes the link between a resource and all its relevant descriptors and contains information about all parts of one resource;
- The **Resource** is an individual element identifiable by an address and it is use to identify the video sequences and the static images of one presentation.

6. **Adding Semantic to Multimedia Content Description**

In Figure 1 we present an architectural solution for integration of semantic multimedia data with the semantic web based on MPEG-21 standard.

We choose to use MPEG-21 standard for digital elements declaration because this standard offers a generic tool for multimedia content description and for standardization in the multimedia applications development.

MPEG-21 components used for digital elements description have the following signification:
- The **Item** is the declarative form of the MPEG-21 digital element. An item is composed from a group of sub-items and/or elements with relevant descriptors;
- The **Descriptor** contains information about the items, like the activity;
- The **Component** realizes the link between a resource and all its relevant descriptors and contains information about all parts of one resource;
- The **Resource** is an individual element identifiable by an address and it is use to identify the video sequences and the static images of one presentation.

We dynamically generate the MPEG-21 document in XML format, at run time, function on user’s preferences or function on the technical conditions, using content-based querying. Based on these elements, we automatically generate the MPEG-21 document, in XML format, to describe the digital elements used in the semantic web.

We use a C# solution to build and to implement the system.

To generate and to format the MPEG-21 document, we used `XmlTextWriter` class, in the following way:

```csharp
XmlTextWriter xw=new XmlTextWriter(Server.MapPath("Presentation.xml"), System.Text.Encoding.UTF8);
xw.WriteStartDocument(true);
xw.Formatting=System.Xml.Formatting.Indented;
```

![Figure 1. Integrating semantic multimedia to the semantic web](image-url)
The previous code creates a new instance of XmlTextWriter class, used to create and to write on the web server a new XML file, Presentation.xml. The elements and attributed added correspond to the MPEG-21 file structure presented hereinafter.

We transform the MPEG-21 document in an HTML document based on a schema attached to the XML document.

We defined a container that stores information about the multimedia elements, the time interval in which the presentation will display each multimedia element. We use this container for synchronization the multimedia elements during the presentation.

The link between the name of elements and the corresponding multimedia resources is made by digital elements description presented in MPEG-21 format. The MPEG-21 descriptions for images components are presented in the following example:

```xml
<Component id="images">
  <Descriptor>
    <Component id="right">
      <Resource ref="pictures/right.jpg" type="image/jpg" />
    </Component>
    <Component id="front">
      <Resource ref="pictures/front.jpg" type="image/jpg" />
    </Component>
    <Component id="left">
      <Resource ref="pictures/left.jpg" type="image/jpg" />
    </Component>
    <Component id="back">
      <Resource ref="pictures/back.jpg" type="image/jpg" />
    </Component>
  </Descriptor>
</Component>
```

In this way we can add semantic information to multimedia data available in a standardized manner to the semantic web.
7. Conclusions

The use of semantic is crucial to integrate information from heterogeneous sources, whether within one organization or across organizations.

The descriptions of multimedia content are inherently problematic because of the volume and complexity of the data and the potentially high subjectivity of human-generated descriptions.

To enable multimedia content to be discovered and exploited by services, agents and applications on the semantic web, it needs to be described semantically.

To standardize the multimedia content descriptions allows adding semantic to multimedia data and facilitate the integration with the semantic web.

The multimedia content description must be realized in a standardised way, so that it can be integrate with the semantic web.

8. References