

# An Opera Information System based on MPEG-7

Òscar Celma<sup>1</sup> and Enric Mieza<sup>1</sup>

<sup>1</sup>*Music Technology Group, Institut Universitari de l'Audiovisual, Universitat Pompeu Fabra. Ocata 1, 08003 Barcelona, Spain. <http://www.iua.upf.es/mtg>*

Correspondence should be addressed to Òscar Celma ([oscar.celma@iua.upf.es](mailto:oscar.celma@iua.upf.es))

## ABSTRACT

We present an implementation of the MPEG-7 standard for multimedia content description of lyric opera in the context of the European IST project: OpenDrama. The project aims the definition, development and integration of a novel platform to author and to deliver rich cross-media digital objects of lyric opera. MPEG-7 has been used in OpenDrama as base technology for a music information retrieval system. In addition to MPEG-7 Multimedia Description Scheme, different classification schemes have been proposed to deal with operatic concepts as musical form (acts, scenes, frames, introduction, etc), musical indications (piano, forte, ritardando, etc), genre and creation roles (singers, musicians, production staff, etc). Moreover, this project has covered the development of an authoring tool for MPEG-7 standard, namely *MDTools*, which includes segmentation, classification scheme generation, creation and production and media information descriptors.

## 1. INTRODUCTION

An immeasurable amount of digital multimedia material is available today (on the World Wide Web, in broadcast data stream, in digital storage media) and this amount constantly grows. The intrinsic value of this information depends on how easily we can manage, search, retrieve and access to it.

MPEG-7 [1], formally named Multimedia Content Description Interface, aims to create a standard for the description of the multimedia content data that supports some degree of interpretation of the information's meaning. The main goal of the MPEG-7 standard is to provide structural and semantic description mechanism for multimedia content. In that context, we usually talk about metadata (that is data about data) and, in that particular case, data about the multimedia information that is described.

Moreover, the MPEG-7 standard provides content description for audiovisual content, defining normative elements as Descriptors (Ds), Description Schemes (DSs) and a Description Definition Language (DDL). Descriptors are designed for describ-

ing different types of information; low-level audiovisual features, high-level semantic objects, content management and information about storage media. Ideally, most Descriptors corresponding to low-level would be extracted automatically, whereas human intervention would be required for producing high-level Descriptors ([2], [3]). Description Schemes are used to group several Descriptors and Description Schemes into structured, semantic units using the DDL.

European project OpenDrama<sup>1</sup> aims the definition, development and integration of a novel platform to author and to deliver rich cross-media digital objects of lyric opera. Two main services arise from the project: the *Advanced Delivery Client* and the *Open-Opera* web portal. The former offers a new and exciting way of interacting with an opera performance with all its related multimedia information, that is; audio, video and 3D stage representation in real-time. The latter is an interface to explore the operatic world and discover new productions, operas,

---

<sup>1</sup><http://www.opendrama.com>

emotions, etc. Both services show a high degree of interactivity with the user, including transport and track volume control, libretto and plot display along with the multimedia visualization, annotations and links to related material and, finally, an interactive karaoke to sing users' favorite arias.

From the metadata point of view, the main requirements for the described OpenDrama services are: (i) a search and retrieval system to browse database contents and (ii) the streaming of metadata synchronized with multimedia data. For these objectives, the use of the MPEG-7 standard fits the needs. In addition, the use of standards in such a project is a key point to make it a real open, maintainable and durable system. An authoring application based on MPEG-7, namely *MDTools*, has been implemented and it has been used to create metadata information from opera productions. The output of the application, that is XML documents, are introduced into an XML:DB for searching and retrieval purposes.

## 2. OPENDRAMA PRODUCTION WORKFLOW

From a high level point of view, we can distinguish several steps in the production workflow of an opera, in the context of the OpenDrama project.

### 2.1. Opera performance

OpenDrama productions follow a special procedure and method for recording. For instance, a voice has to be recorded apart from the orchestra and the orchestra would be desirable to be recorded in several tracks. That means a multitrack recording for the orchestra and an independent recording session for the singers. Material for special products as "making of" is also needed to be recorded during the performance.

### 2.2. Media files generation and processing

The output of the recording sessions gives a set of media files; basically video —with different camera points of view— and multitrack audio recordings. These files are mixed and processed to produce the desired operatic audiovisual information to show. Then, the material is encoded in several media profiles to enable different type and quality of service, depending on the user's connection bandwidth and service type —simple, multitrack audio, karaoke, video, and so on.

### 2.3. Metadata and content generation

All the necessary metadata of a production (author, singers, characters, producer, abstracts) are created using *MDTools* application. In this process, content is segmented based on operatic forms (acts, scenes and pieces) and edited with information such as actors/characters, instruments present in each segment. More detailed musicological information and musical indications could be also included.

### 2.4. Publication

Content is published introducing the MPEG-7 description into the database system, and the media files into the streaming server. Once this process is done, user can access to the content and browse the database using the web-based search and retrieval interface.

Figure 1 shows the overall explained process.

## 3. TAILORING CONTENT

As mentioned in the introduction, one of the main goals of the project is the creation of a database for all the OpenDrama services. *OpenOpera* web browser is a classical search and retrieval system related to opera information. The other service, the *Advanced Delivery Client*, focuses on real-time visualization of an operatic production that contains audio, video and 3D virtual stage simulation as well as libretto and music score representation. The information in the storage database system is used for both services, and MPEG-7 is a good choice as a base technology.

### 3.1. MPEG-7 Multimedia Description Scheme

MPEG-7 Multimedia Description Scheme (MDS) comprises the set of Description Tools (Ds and DSs) dealing with multimedia entities. MDS contains, among others, the following areas; Basic Elements, Content Management and Content Description [4].

- **Basic Elements:** address specific needs of audiovisual content description, such as the description of time, persons, places and other textual annotation.
- **Content Management:** describes different aspects of creation and production of the process —such as: title, creators, locations, dates—

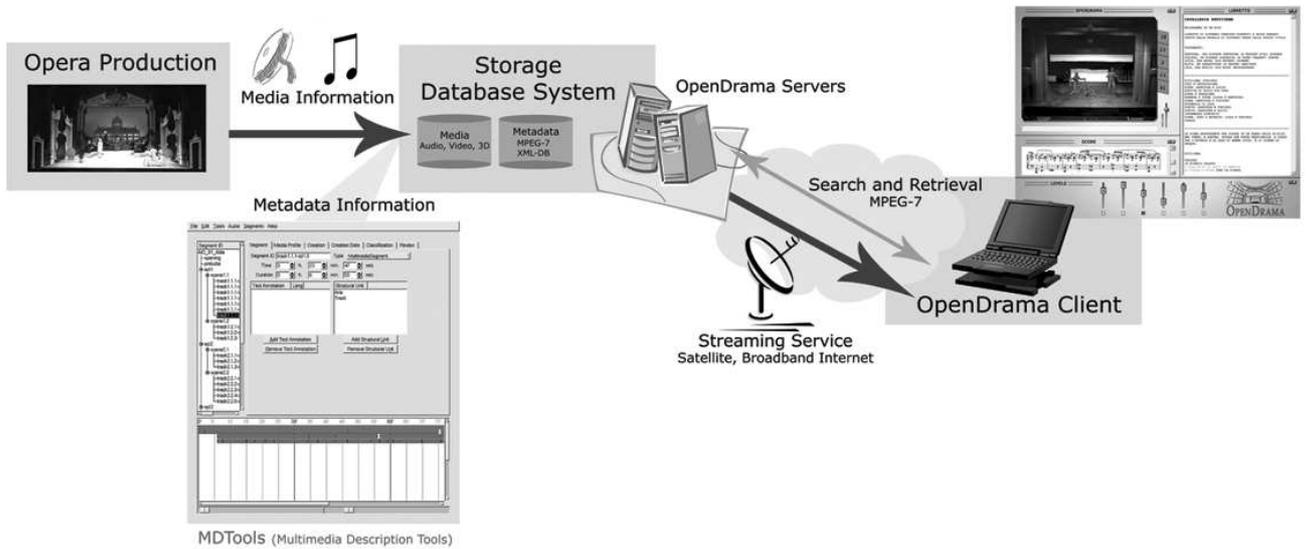


Fig. 1: OpenDrama production workflow.

and media information of the audiovisual content (storage media, coding format and compression) to adjust to different network environments. *Media Information* description scheme allows defining different media profiles. This approach is practical as it allows to attach diverse audiovisual media associated to a segment. In OpenDrama, each opera production has different profiles and, according to the client bandwidth, streaming service can send the appropriate media instance.

Example of *MediaInformationDS* with two *MediaProfileDS*:

```
<MediaInformation>
  <MediaProfile id="AID_track1.1.6-mp3">
    <MediaFormat>
      <FileFormat
        href="ODFileFormatCS:mpeg:mp3" />
      <!-- more data -->
    </MediaFormat>
    <MediaInstance
      id="AID_track1.1.6-mp3-source">
      <MediaLocator>
        <MediaUri>...</MediaUri>
      </MediaLocator>
    </MediaInstance>
  </MediaProfile>
  <MediaProfile id="AID_track1.1.6-wav">
    <MediaFormat>
      <FileFormat
        href="ODFileFormatCS:wav" />
      <!-- more data -->
    </MediaFormat>
    <MediaInstance
      id="AID_track1.1.6-wav-source">
      <MediaLocator>
        <MediaUri>...</MediaUri>
      </MediaLocator>
    </MediaInstance>
  </MediaProfile>
</MediaInformation>
```

```
</MediaLocator>
</MediaInstance>
</MediaProfile>
<MediaProfile id="AID_track1.1.6-wav">
  <MediaFormat>
    <FileFormat
      href="ODFileFormatCS:wav" />
    <!-- more data -->
  </MediaFormat>
  <MediaInstance
    id="AID_track1.1.6-wav-source">
    <MediaLocator>
      <MediaUri>...</MediaUri>
    </MediaLocator>
  </MediaInstance>
</MediaProfile>
</MediaInformation>
```

- **Content Description:** describes the structure, segmentation of the content and semantics (entities, events, relationships) of the audiovisual content. Thus, it allows to attach audio, video, annotation and content management to the multimedia segments, to depict them in detail.

Following (partial) example shows a temporal decomposition of an opera:

```
<AudioVisual id="DGI_01_Don_Giovanni">
  <MediaTime>
    <MediaRelTimePoint>PT0H0M0S
    </MediaRelTimePoint>
    <MediaDuration>PT2H41M36S
    </MediaDuration>
  </MediaTime>
  <TemporalDecomposition gap="true"
  overlap="false" criteria="temporal">
    <AudioVisualSegment id="DGI_01_act1">
      <MediaTime>
        <MediaRelTimePoint>PT0H6M5S
        </MediaRelTimePoint>
        <MediaDuration>PT1H16M46S
        </MediaDuration>
      </MediaTime>
      <StructuralUnit
        href="ODOperaticFormCS:Act"/>
      <!-- more data ... -->
    <TemporalDecomposition gap="true"
    overlap="false" criteria="temporal">
      <AudioVisualSegment id="DGI_01_scene1.1">
        <StructuralUnit
          href="ODOperaticFormCS:Scene"/>
        <!-- more data ... -->
        <TemporalDecomposition gap="true"
        overlap="false" criteria="temporal">
          <AudioVisualSegment
            id="DGI_01_track1.1.1-cd1">
            <MediaTime>
              <MediaRelTimePoint>PT0H6M5S
              </MediaRelTimePoint>
              <MediaDuration>PT0H5M21S
              </MediaDuration>
            </MediaTime>
            <StructuralUnit
              href="ODOperaticFormCS:Aria"/>
            </AudioVisualSegment>
          <!-- more tracks... -->
          <AudioVisualSegment
            id="DGI_01_track1.1.6-cd1">
            <StructuralUnit
              href="ODOperaticFormCS:Duet"/>
            </AudioVisualSegment>
          </TemporalDecomposition>
        </AudioVisualSegment>
      </AudioVisualSegment>
    </TemporalDecomposition>
  </AudioVisualSegment>
</AudioVisualSegment>
```

```
</TemporalDecomposition>
</AudioVisualSegment>
<AudioVisualSegment id="DGI_01_act2">
  <!-- more data... -->
</AudioVisualSegment>
</TemporalDecomposition>
</AudioVisual>
```

- Complementary to these descriptors, MPEG-7 **Classification Schemes** define taxonomies to classify a subject area with a set of terms, organized into a hierarchy. A term represents one well-defined concept in the domain covered by the classification scheme. A term has: an identifier—that uniquely identifies it—, a name (useful for a search in a target database) and a definition that describes the meaning of the term.

Example of a Classification Scheme (extracted from OpenDrama Operatic Form CS):

```
<ClassificationScheme
  id="ODOperaticFormCS">
  <Term termID="Aria">
    <Name xmlang="it" preferred="true">
      Aria</Name>
    <Name xmlang="en" preferred="false">
      Air</Name>
    <Definition xmlang="en">A closed
      lyrical piece for solo voice...
    </Definition>
  </Term>
  <Term termID="Recitative">
    <Name xmlang="en" preferred="true">
      Recitative</Name>
    <Name xmlang="it" preferred="false">
      Recitativo</Name>
    <Definition>A type of vocal writing,
      normally for a single voice...
    </Definition>
  </Term>
</ClassificationScheme>
```

A term in a classification scheme is referenced in a description with the *TermUse* data type:

```
<AudioVisualSegment>
  ...
  <StructuralUnit
    href="ODOperaticFormCS:Aria">
```

```
...
</AudioVisualSegment>
```

### 3.2. Multimedia content description: MDTools

The *MDTools* application, developed in the framework of OpenDrama project, is an authoring tool to generate MPEG-7 MDS descriptions. In this project it has been used to create related metadata from an opera production, but it can be used to describe any type of media. *MDTools* implements most of the MPEG-7 MDS features.

The main MDS tools that are integrated are:

- Segmentation of the content: performs temporal segmentation of the content within a hierarchical segment structure.
- Segment information: start time, duration, text annotations, segment identifier, segment type (audio, video, audiovisual, multimedia) and StructuralUnit —a term in a Classification Scheme to describe structural aspect of the segment.
- Classification Schemes (CS): creation of taxonomies and term definitions.
- Media Information: Segments are linked with its multimedia files. Media profiles can be created for delivery purposes, according to client bandwidth.
- Creation and production information: includes editorial information such as creators (author, singers, characters, etc.), title, abstract (with multilanguage support).

## 4. MPEG-7-BASED RETRIEVAL SYSTEM

An MPEG-7-based retrieval system has been implemented to search and retrieve audiovisual data of multimedia productions, generated with the *MDTools* application.

Since XML has been adopted as the format for MPEG-7 standard, eXist 0.9.1 [5], a native XML database (XML:DB) [6] is used in the OpenDrama project. XML databases make use of collections as internal folders for repositories of XML documents.

So, XML documents created with *MDTools* application are introduced into the database. Each document contains information of an opera performance.

In this context, for simple pull application (a client requests data to the XML database) the *Client-Server* computing model is suitable. Hence, a web-based application queries the XML database according to MPEG-7 descriptors. In order to query the XML database, the server application transforms the MPEG-7-based query to an XPath expression. The next step is to search into the database for segments matching the criteria, to process the XML results with an XSLT transformation, and to return results in HTML to the browser (as well as the multimedia files retrieved from the streaming server). Figure 2 illustrates this process.

## 5. CONCLUSIONS

Applications based on the MPEG-7 standard are emerging in the areas of multimedia archive, digital broadcasting, digital library, etc. MPEG-7 provides description mechanisms for multimedia content; however, applications are still immature and are not really explored in concrete fields. In this article we propose the application of MPEG-7 for an opera information retrieval system.

A key point when building an information system is to have a good term ontology, general enough to cover as much real cases and applications as possible. In the context of the OpenDrama project, Classification Schemes have been proposed for musical concepts such as operatic form, singer, operatic style and creator role. However, that is only a departing point to create a real musicological ontology that would need a great effort of expert musicologists.

Furthermore, a web system for browsing, searching and filtering metadata information has been implemented.

Finally, we present an authoring application for MPEG-7 multimedia descriptions (*MDTools*) developed in the frame of the OpenDrama project. *MDTools* is designed to support general MPEG-7 MDS descriptors suitable for any application, not only for opera data. Next development foresees to extend the functionality to cover audio and visual descriptors and integrate automatic low-level description tools.

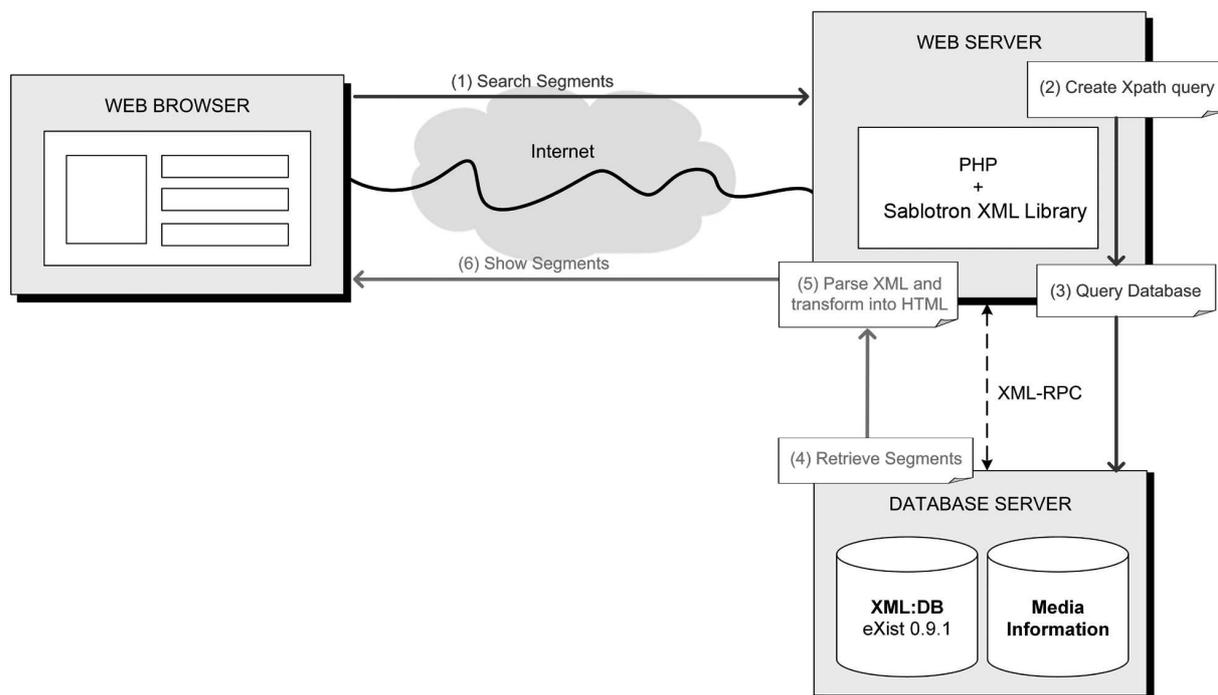


Fig. 2: Overview of the MPEG-7 Web Browser architecture.

## 6. ACKNOWLEDGMENTS

This work has been partially funded by the OpenDrama IST European project.

## 7. REFERENCES

- [1] Manjunath, B. S., Salembier, P. and Sikora, T. "Introduction to MPEG 7: Multimedia Content Description Language". Ed. Wiley, 2002.
- [2] Gómez, E., Gouyon, F., Herrera, P. and Amatriain, X. 2003. "Using and enhancing the current MPEG-7 standard for a music content processing tool". Proceedings of Audio Engineering Society, 114th Convention Amsterdam, The Netherlands, 2003 March 22–25.
- [3] Gómez, E. Gouyon, F. Herrera, P. and Amatriain, X. "MPEG-7 for Content-based Music Processing". Proceedings of 4th WIAMIS-Special session on Audio Segmentation and Digital Music. London, United Kingdom, 2003 April 9–11.
- [4] Van Beek, P., A.B., Heuer, J., Martínez, J., Salembier, P., Smith, J. and Walker T. "MPEG-7: Multimedia Description Schemes, ISO/IEC FDIS 15938–5:2001, International Standard Document, 2001.
- [5] Meier, W. "eXist An Open Source Native XML Database. Web, Web-Services, and Database Systems". NODE 2002 Web and Database-Related Workshops, Erfurt, Germany, 2003 October 7–10.
- [6] XML:DB Initiative for XML Databases. <http://www.xmldb.org>