briefing paper

digital preservation europe

The PrestoSpace data model

One of the main themes of the PrestoSpace project was the study of techniques to improve the access to large audiovisual collections. In this brief paper the salient characteristics of the designed data model are outlined.

The PrestoSpace [1] Integrated Project (February 2004 – January 2008) was launched in February 2004 under the Information Society Technologies priority of the Sixth Framework Programme of the European Community (IST FP6 507336). The consortium included several European broadcasters and audiovisual archive owners, universities and research centres and industry representatives for a total of 35 active partners.

The goal of the project was to provide technical devices and systems for the digital preservation of all types of audio-visual collections with the aim to foster the birth of preservation factories to deliver affordable services to all kinds of collection owners in the field of asset digitisation, management and distribution.

The project was organised in four main areas: preservation, where specific digitisation techniques were studied and optimised for different types and conservation conditions of media, restoration, that dealt with defect analysis and correction of degraded material, storage, that performed surveys on archive management technologies and products and metadata access and delivery, concerned with content description and retrieval.

The theme of the Work Area on metadata was that of improving the accessibility of audiovisual collections via advanced indexing techniques and subsequently defining a proper optimized process for the effective and efficient processing of vast amounts of content, introducing automation wherever feasible. Traditionally, content documentation was produced through a completely manual process based on the catalogue information about the item, complemented by descriptive annotations and technical information about the format of the media. A crucial point is the level of detail provided, as often a television programme contains several items that can be the object of the search, interviews or performances, requiring a proper segmentation of the content where salient parts are isolated from the main programme. Hierarchical models have therefore been introduced. In PrestoSpace several techniques have been exploited to automate at least part of the process, including video and audio signal analysis to extract features like segmentation into scenes, clustering of scenes with visually similar content, speaker tracking, speech recognition and semantic analysis of texts to mine hidden pieces of information like named entities, to classify items into specified categories and to correlate items to external sources on the web. The envisaged documentation process was first to import existing legacy documentation bits, including identification information, then to process the digitized audiovisual content with automatic tools and finally to manually verify and complete the documentation.

The analysis of the state of the art has lead to the identification of the following four basic metadata classes:

•Identification information, e.g. titles, credits, programme publication information.

•Editorial parts information, i.e. information about the relevant editorial sub-items of a programme (e.g. news items, interviews, artistic performances).

•Content-related information, e.g. speech transcript, topics, descriptions, aural and visual low level descriptive features.

•Enrichment information, i.e. information coming from external sources generically or topically related to the programme content.

The subsequent work was that of designing a data model to represent them, together with a data format capable of implementing the data model without any loss of information.

Aside from various proprietary data models in use at several audiovisual archives, there are a number of standards models available. This was the obvious starting point for the selection of the PrestoSpace data model. The evaluated formats were:

• Dublin Core [2], defined by the American Library of Congress. It is a flat model constituted by a set of 15 main attributes that can be further specialized via modifiers. Hierarchical descriptions are not possible.

• P_META [3], defined by the European Broadcasting Union (EBU). This model has been defined on the basis of the experience of several European broadcasters. It allows hierarchical descriptions and can be considered a superset of Dublin Core.

• DMS-1 [4], defined by the Society of Motion Picture and Television Engineers (SMPTE). It has been conceived as a complement of the MXF container for audiovisual content. It is similar as descriptive power to P_META.

• MPEG-7 [5], defined by the ISO/IEC Committee. This model has a number of unique characteristics, and is suited especially in the field of automatic audiovisual analysis. It provides a wealth of low level feature descriptors but is not very suitable for high level structured descriptions.

Unfortunately, none of the above could satisfactorily fulfil all the requirements, being conceived for use cases partly different from PrestoSpace. There are however large overlaps so it has been agreed to assemble a new model using existing formats as much as possible and adding new structures only when unavoidable.
This was achieved through the creation of a single XML-based document format, taking the best from each of two metadata standards natively orientated to the description of audiovisual objects, MPEG-7 and P_META.
P_META was adopted due to its complete set of information structures for the identification, classification and publication-related features of a programme. The MPEG-7 standard is the only one to offer powerful temporal segmentation structures and a comprehensive set of standard audiovisual descriptors, core features for the automatic audiovisual analysis tools. In order to achieve clear semantics of the elements in the MPEG-7 description, the Detailed Audiovisual Profile [6] is used for the MPEG-7 part of the format.
Figure 1 schematises the resulting document format, also indicating the ad-hoc data structures introduced to realise those structures not covered by any of the two standards, for example to express the relation between the original media and their digitized versions. A common problem in audiovisual archives is the management of multiple copies of the same content, stored on different media types. Even if these copies can have different quality or slightly differ in the package, it is often advantageous to let them share the same descriptive metadata and express their mutual relation. An ad-hoc structure has been defined to map the logical timeline on which events have been annotated to the physical timelines of each identified copy of the same content in the archive. Another custom structure has been defined to reference external information sources.

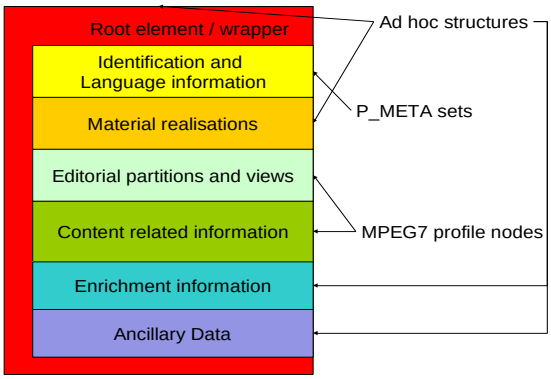

Figure 1 - document format schema.

The unique identification of the instances of audiovisual material has been achieved through the adoption of the SMPTE UMID [7] standard for globally unique identifiers generation and formatting. This includes original media, digitally re-mastered media, and all the materials created during the documentation process (e.g. key frames, low resolution videos).
The full data model description and the corresponding XML Schema are freely available at [8].
In conclusion, the PrestoSpace model for audiovisual content description has been obtained by combining parts of existing standard models, thanks to the schema import mechanism intrinsic of XML specifications, complemented with few custom structures. The model proved to be entirely satisfactory for the envisaged application. The reuse of widely known standard structures (P_META, MPEG7, UMID) simplifies the mapping operations required to export the generated information to other environments, like content publication on federated heterogeneous archives that make use of different (generally simpler) data models.

References

[1] http://prestospace.org, official PrestoSpace project web site

[2] ISO/IEC 15836, The Dublin Core Metadata Set, 2003

[3] EBU Tech3295, P_META Metadata Exchange Scheme, v 2.0, European Broadcasting Union (EBU), 2007

[4] SMPTE 380M, Material Exchange Format (MXF) - Descriptive Metadata Scheme-1, Society of Motion Pictures and Television Engineers (SMPTE), 2004

[5] ISO/IEC 15938, Multimedia Content Description Interface

[6] W. Bailer, P. Schallauer, "The Detailed Audiovisual Profile: Enabling Interoperability between MPEG-7 Based Systems", Proc. of 12th International Multi-Media Modeling Conference, Beijing, CN, Jan. 2006

[7] SMPTE 330M, Unique Material Identifier (UMID), Society of Motion Pictures and Television Engineers (SMPTE), 2000

[8] G. Dimino, L. Boch, A. Messina, W. Bailer, C. Bauer, V. Tablan, PrestoSpace Documentation Platform http://prestospace.org/projects/deliverables/D15.2.pdf, 2008

Author: Giorgio Dimino, *RAI Research Centre*
g.dimino@rai.it