VIRTUAL MEDIA IN AN OAIS-ENABLED ENVIRONMENT

Sebastian Gabler, NOA Audio Solutions, Austria

1. User expectations in time

The general public more and more is accustomed to media-centric access. The user interface of about any technical equipment aims to detach the user from the actual technology. Instead, complex functionality is presented as a "service" — easy to consume. For sound and audiovisual media the struggle of content with transmission format constraints is however all but new: Radio musicians of the 1940s report that they had to adapt tempo to the remaining broadcast time, playing live on air[138], and the duration of the CD Audio was anecdotally optimized to hold Karajan's performance of Beethoven's Ninth.[139] No matter if the anecdote holds true or not, the otherwise contended playing time of the CDDA and many other engineering efforts historically prove the wish to reproduce performances coherently. More recently the buyer of the latest BD release of David Lean's *Lawrence of Arabia*[140] may enjoy not only the full 227 minunte, shiny 2012 re-master from 8 K scans without leaving the chair to change media, but will as well be able to dwell across the Interval to Maurice Jarre's epic Intermission music often left out in previous releases and screenings.

Self-service facilities are another key driver. A member of the Academy Awards jury, handed an access copy of a movie on a BD, may or may not be aware that for his or her convenience not so long ago several reels would have had to be concatenated, incurring manual work. A customer of iTunes can listen to all 22 songs of Led Zeppelin's Remasters Album in one go, while a user of a self-made CD rip will have to live with switching folders at the end of disc 1, following the carrier-oriented CDDB data model.

For digital access, offering appropriate coherence is no longer a matter of physical constraints; it ultimately depends on a suitable data model.

2. Terminology

Archival Information Package (AIP): "An Information Package, consisting of the Content Information and the associated Preservation Description Information (PDI), which is preserved within an OAIS."[141]

Archive: "An organization that intends to preserve information for access and use by a Designated Community.[142]" Remark: The terms OAIS and Archive are equivalent in the CCDM reference model; they are used synonymously in this article.

Content Information, short Content: "A set of information that is the original target of preservation or that includes part or all of that information. It is an Information Object composed of its Content Data Object and its Representation Information.[143]"

Context Information: "The information that documents the relationships of the Content Information to its environment. This includes why the Content Information was created and how it relates to other Content Information objects.[144]" Remark: at first glance, the TC-03 definitions of Primary and Secondary Information may seem as specialisations of

138 "Ich habe die Frühlingssonate noch nie so langsam gespielt", Bruno Lenz (1911-2006), vice concertmaster of Bavarian Radio Symphony Orchestra, in a personal conversation with the author, 1996.
139 The final truth of this anecdote is not proven, see: http://app.handelsblatt.com/unternehmen/management/norio-ohga-vom-kopilot-zum-kommodore/2957156.html.
140 http://www.nytimes.com/2012/09/30/movies/lawrence-of-arabia-mended-returns-to-screen-and-blu-ray.html.
141 CCSDS 650.0-M-2, 1.7.2.
142 idem.
143 Idem.
144 idem

Content- and Context information. There are however subtle differences between the two, and one can locate a dilemma in the TC-03 notion specifically during migration to digital: Whereas carrier parameters are certainly secondary once digitised, for the replay of the physical carrier they are a part of the primary Representation Information, meaning the content information cannot be retrieved at that stage without them. While going deeper on such aspects is not in the scope of this paper, the example still shows that information management has some specific challenges on the backdrop of Long Term Preservation, and that the OAIS terms are very helpful in that regard.

Reasoner: A Semantic Reasoner is a rules engine that is able to apply decisions, based on a set of axioms. Often, Reasoners use first-order quantifications, such as "For every person, if this person is a Philosopher, the person is a Scholar."

3. The 4th dimension

Time is the core criterion that differentiates sound and audiovisual from any other Content Information. Time is manifest as a continuum once we commence the (re)production of music, drama, film, or a news broadcast, and entails for the programme's duration.

There are many possibilities to store and present information aligned to a timeline. Among others, these may be marker lists (text), optical disk sub code, or XML documents. The alignment is given by a range, i.e., the distance of a "start" to the corresponding "stop" event, an event class, and sequence information.

Different event classes may require different time measures depending on their context, i.e., samples, frames, or bars/beats, to name but a few. One could try to establish different time regimes depending on the context, and relate them as required. That would be a daunting task. To the rescue, digital media provide an obvious solution to provide time information that is both generic and precise enough to support automated business procedures.

All digital media is time-discrete, and the common denominator is the sampling frequency of audio (8 kHz-192 kHz, or higher), which aligns to common moving image frame rates. A timeline resolving resolution to 1 μs is sufficient to define rounded positions of any PCM sample — or picture frame. Using an "oversampled" timeline allows the aggregation of the aforementioned event classes on the same timeline. This has the useful side effect of accommodating Content editions of different time bases, such as the 48 kHz proxy of a CD recording (44.1 kHz), or the 12 fps proxy of a movie.

4. Carrier-orientated preservation

IASA TC04 presents a mainly carrier-centric approach to preservation.

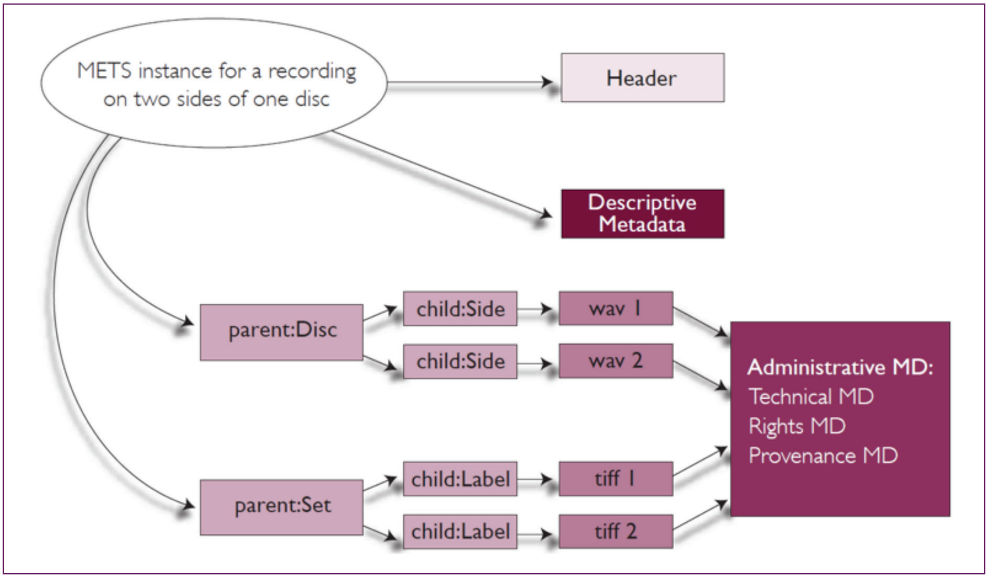

Figure 1. TC04 METS object.[145]

In the above example, the Content Information is the Carrier itself in both the digital form and the parent physical item. "wav n" stores the stream extracted from the discs' grooves, whereas "tiff n" results from scanning the respective labels. Following TC-03.3, the wave files are Primary Information, and the label scans are in the Secondary Information realm. The intrinsic access options for this AIP are as well very much like the physical counterpart: To listen to the recording, we would have to load the audio files, one by one, into a suitable player.

With this approach, apart from creating an "Eternal Media File" that may be migrated with no future information loss, the user will be able to understand the genealogy of the digital files back to their preceding physical carrier. The model could be expanded by information on how the original performance was recorded, if such information was available by the previous acquisition process or current research.

However, we should be aware that carrier migration alone may be insufficient when it comes to building of a Digital Media Archive. Pure copying will not satisfy current and future use cases, as it limits access to the features of the parent medium. Nevertheless, the safe-harbouring of the extracted signal is a required transition step that a Digital Media Archive cannot go without.

5. Virtual media

In the following, some parameters are presented to characterise the idea of Virtual Media.

5.1 Media content centric

The proposed approach assumes that the consumer is interested in the reproduction of certain content, not in the act of replaying a number of reels, disc faces, or tapes. To this end, the media modelling aims to be independent of the physical properties of underlying carriers.

145 IASA-TC04 Second Edition, 20.

5.2 Carrier mapping

Physical boundaries of media carriers are moving targets. After a less-standardized period of 35 years of silent film, the standard 2000 ft 35mm film reel, colloquially referred to as "Two-Reeler" featuring a playing time of 22 minutes, was established with the introduction of sound film around 1930.[146] It is still in use today. However, the birth of digital cinema goes together with the obsolescence of an access scenario, i.e., involving a projectionist who syncs up the reels manually during projection. Similar scenarios exist in radio production where several ATRs have to be synced up for the reproduction of programmes exceeding the typical 2400 ft (app. 30 minutes at 15 IPS) reel. Specifically here, it may even be that carrier limits do not align with content structure, e.g., a chapter border like a musical movement. Instead, carriers may overlap recorded signal to be queued up with virtually no interruption.

Carriers containing heterogeneous Content present a challenge for archivists as well. Also this scenario is often the result of pragmatic decisions in the past that had little to do with content-related considerations, but rather with the price of reel spools. Addressing these scenarios requires a many-to-many mapping of Carriers to Content.

5.3 Virtualized content

Getting serious about long term content preservation means both accepting the facts from the past and changing them, if required, using the means of digital media. The archive taking the role of the projectionist during access therefore has to be able to handle coherent sets of media for any Content object. That includes seamless playback across volume (carrier) borders. As physical rendering of a high-resolution copy comes at a prohibitive overhead of close to 100% storage payload, it is preferable to re-use the concatenated extractions of the carriers.

5.4 Alignment of context information

As previously defined, Context Information refers to the environment of Content, but as well to the relations of different Content objects. When providing context for A/V content pertinent during access, the archive is required to be aware of the actual time range. That is in the first place obvious for any kind of segmentation context. As inter alia information, the showdown scene of a Western movie is usually towards the end, but the archive will have to indicate the exact sequence and position of the scene to provide tangible results. For an entire news broadcast as well, there may be dozens of subjects, contributors, and rights holders. However, which one is relevant for a certain segment makes a significant difference for the information to be eventually manageable.

5.5 Support of different encodings

For the foreseeable future, archives will have to provide different encodings of content for various access scenarios. A typical scenario is the low-resolution access copy, which may differ in time or spatial resolution from the AIP. Virtual Media aim to unify the information among these versions, thus improving efficiency for content creation and access. Section 3 of this paper provided in-depth information on which consequences different encodings have on time handing.

5.6 Long term preservation context

The creation of Virtual Media is done against the backdrop of Long Term Preservation. At the same time, it aims to support present and future access scenarios, as much as it aims at supporting future migrations. This comes with a strong bias towards mathematically lossless

146 http://en.wikipedia.org/wiki/Reel.

encoding of the content, as only linear formats have the property to support automated signal-integrity quality checks. In this respect, Virtual Media are independent of chronology.

5.7 Access virtualisation

From the long-term perspective, it is safe to predict that user access scenarios will change. This is certainly true for access media formats (codecs and wrappers), but the Designated Community, too, may change, or there may be different communities to serve to simultaneously. One example the author has been working with in the recent past is Fonoteca Nacional, Mexico. The institution offers in-house access for researchers or aficionados[147], but has also started a public web service accessible over the Internet.[148] While the latter does not have online access to the actual Archive, Content and Context information still come from the same source, as provided in the AIP production in a dissemination process that is fully automated.

6. A suitable class model

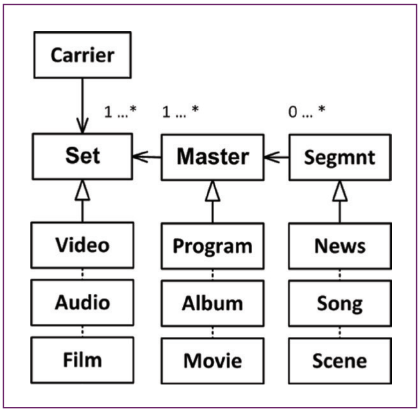

Figure 2. Virtual Media Class Model.

Figure 2 shows a UML schema of a possible approach to the matter. While traditionally flat implementation would be possible, the model should be understood as using semantically linked classes. The classes Carrier, Set, Master, and Segment have specialisations that make them usable for various Content and Media types. Each specialisation features different fields to describe the pertinent parameters.

6.1 Carrier

Carrier refers to the dedicated container where the complete Content, or part of it is stored. It may be a physical media carrier, or a media file or file set. Specialisations would be, for example, a film reel, an audio disc, a video cartridge, an open reel audiotape, or any kind of media file (e.g., DPX sequence, wav, or mov file). Structurally, it is equivalent to Volume.

6.2 Set

The Set class comprises the carrier items required for the reproduction of content, and their correct sequence.

147 http://www.fonotecanacional.gob.mx/index.php/fonoteca-nacional/espacios/audioteca-octavio-paz.
148 http://www.fonotecanacional.gob.mx/index.php/servicios/red-de-fonotecas-virtuales.

6.3 Master item

The Master Item, or Master points to Set, providing the coherent Content so that it can be reproduced directly. The Master may have specialisations of, for example, Opera, Concert, Movie, TV Programme, Radio Programme, or Compilation, to name but a few. The Master class would be a suitable target to build the AIP for a content-centric archive, however there are alternatives, such as building the Master Item as an Archive Information Collection, consisting of Parts.[149] (This is the approach followed by iTunes and many Computer Aided Media Production Systems.)

6.4 Segment

Segment points to Master, and describes any kind of segment information. It is the structural equivalent of Chapter. Specialisations of Segment may be "Song" (-> Album), Shot (-> Movie), Work and Movement (-> Concert), or Scene (-> Movie or Opera).

7. Automation aspects

A major pain point for every archive is its annotation workload. Numbers occasionally proposed by A/V archivists gravitate between 6–10 times the man hours required for physical preservation of the related carrier item for the completion of the metadata required for an AIP. This concurs with personal experience. For born-digital material the scale factor may be substantially higher, as ingesting of this kind of content information is sped-up and automated to a high degree. Linking metadata to media content is a demanding and exhausting task as it is not limited to text work, but requires working with visual and auditory events simultaneously.

7.1 Segmentation

Providing chapter information with the media is an established method in A/V production. Therefore, mature tools and methods exist that will be handled by even semi-skilled operators. Either cue sheets are typed into a text editor, or markers are set in a Wave Editor or NLE. This method is applicable to the vast majority of linear A/V carriers. The required result is a map of the chapter information, e.g., in an XML file, providing sequence, time, and hierarchy information of each event.

For some physical carriers, such as RDAT, CDDA, but as well for born-digital Content in BWF or MXF files segmentation, information may be retrievable during extraction. In some cases, auxiliary information provided by an editor workstation can be parsed as well. Even OCR-ed cue sheets containing time code lists may be considered.

7.2 Rule-based processing

Context information generated and validated in the production process of the AIP, such as access rights, provenance, and technical parameters can be post-processed and packaged on rule-basis. These rules will select the data to be included with the AIP, e.g., parsing them from process logs. The processing will be triggered automatically in an automated workflow engine.

The automated creation of Content segments and linking of the media requires advanced logics, e.g., in the form of a sematic Reasoner. NOA's MediaLinkingRulesXML is an example of such a Reasoner.

149 See CCSDS 650.0-M-2, 4-46.

8. MediaLinkingRules

MediaLinkingRules is an example of how to support the building of accessible Content, Content segmentation, and Context by employing automated rules. These rules include ontology and logical elements that, for example, allow building and maintaining a Part object of certain duration for every instance of a marker type, and aligning the Context information with it. The reverse process is supported as well, which is helpful for any advanced transformations, see section 10.1 in this article.

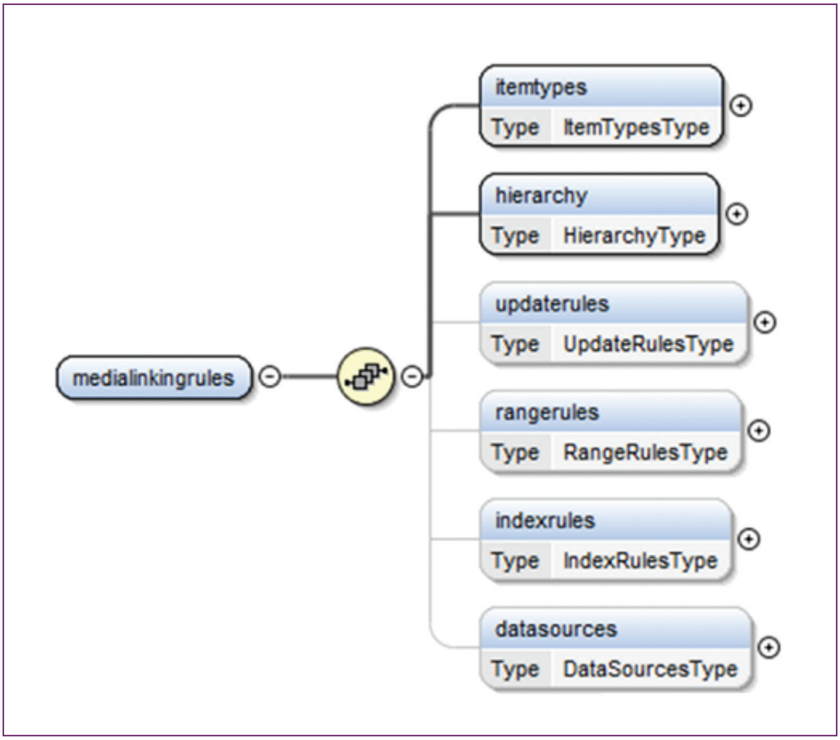

Figure 3. MediaLinkingRules schema.

8.1 Itemtypes

A list of the item categories maintained by the rules, e.g., Carriers, Set, Master, and Segment with their specialisations.

8.2 Hierarchy

The hierarchy definition establishes a tree hierarchy, mapping item categories to the required roles of carrier, and their relations.

```
<medialinkingrules>
  <itemtypes>
    <!-- ... -->
  </itemtypes>
  <hierarchy>
    <itemtype id="1000" nameinfo="CarrierSet" role="master">
            <itemtype  id="500"  nameinfo="Carrier"  linktype="10000"
role="ingest"/>
      <itemtype id="50010" nameinfo="Album" linktype="10010">
        <itemtype id="50020" nameinfo="Title" linktype="10020"/>
      </itemtype>
    </itemtype>
  </hierarchy>
</medialinkingrules>
```

8.3 UpdateRules

Define the data handling per update mode upon creation or update of a target item instance.

```
<medialinkingrules>
  <itemtypes>
    <!-- ... -->
  </itemtypes>
  <hierarchy>
    <!-- ... -->
  </hierarchy>
  <updaterules>
    <updaterule name="album-create-rule" mode="create update">
      <updatefields>
                 <field  name="Item.StringId"  mode="create">Metadata.
GetNextUid('Record - ID', 1, 'U')</field>
        <field name="Item.Float01">@@MediaDuration@@</field>
      </updatefields>
    </updaterule>
  </updaterules>
</medialinkingrules>
```

They include a mode rule, depending if a change should be infered only on creation or als well on update of the target, and field definitions. I.e., in the above example, the UID of the target would be generated during creation, but it would remain unchanged during an update.

8.4 RangeRules

RangeRules provide for a default format to describe the running order of Segments, and of the Carriers. Along with the Index information of the carrier members, the absolute position can be detected.

8.5 IndexRules

IndexRules provide for a default format to express the set information, e.g., amount and position of carriers. IndexRules and RangeRules both provide options to convert a customized numbering schema to the default format.

8.6 Data sources

Data sources define types of data sources to be used during the execution. These may be XML structures, such as Marker Lists that are generated during production, but they could equally be existing data in the same database, or an external source.

9. Building of virtual media

From criteria mentioned in the OAIS Reference Model, that is an AIP consists of Content, Provenance, Rights, Fixity, Reference, and Context,[150] there are no specific circumstances to be mentioned in the context of Virtual Media. In each individual Archive, decisions will have to be taken which SIPs will have to contribute, which mile stones have to be set up along the process, and which Packaging parameters will finally contribute to the final result, so that the AIP satisfies the set criteria required for access by the Designated Community.

What is specific to A/V is the required building of a coherent time relation of all parts of the AIP or AIC. For this, the following common use cases exist:

9.1 Concatenation

The content may be reproduced by playing several set members in their correct sequence. If these members contain taper areas, an Edit Decision List function is required to exclude these ranges from the Master. Automated rule models can support concatenation in most cases. Concatenation includes the processing of timeline, media, and time-related metadata information. It is necessary to establish the global context for any sequence information, for example Track 3 on Carrier 2 in the global context can be seen as Track 002.03, or maybe even as Track 77 (or any other current global counter). Figure 4 outlines the schema of a concatenated Media Object, including Context information.

150 CCSDS 650.0-M-2, 4.2.2.3.

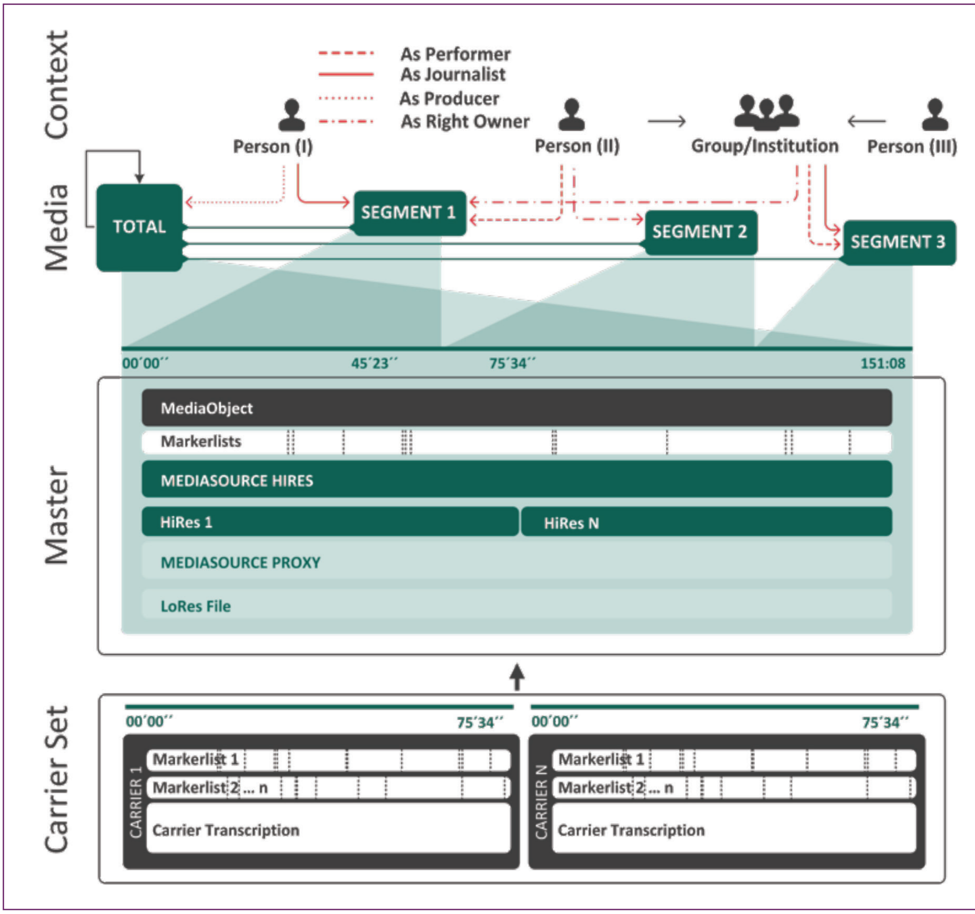

Figure 4. Media Object.

9.2 Re-use

If the content is playable coherently on the original carrier or the transcript thereof, it can be re-used. Leaders or footers may be excluded from the target Master Range. As well, segmentation of heterogeneous content on a single carrier (Mix Tapes) falls under the re-use scenario in Virtual Media. In fact, Re-use can be seen as a specialisation of Concatenation — there is only one member in the Set.

9.3 Multiple master

The case of multiple, heterogeneous content is common in Broadcast. For economic reasons, and as carrier playing time extended over time, multiple programmes have been copied to a single carrier. Here, during the building of the Master, the opposite of the concatenation case is taking place: Multiple Masters will reference a single set, and each of them will have segmented ranges for Content, Fixity, time-related metadata, and Context. The approach mentioned in 10.1 also supports this scenario.

10. Conclusions and outlook

As demonstrated, Virtual Media offers a concept to manage A/V Content with efficiency and agility. It follows from the core principle that media are (re-)producing in time and it aims to render tangible access experience for the Archive's users. Using other components of the Archive, such as Rights Management and Access Services, they offer a foundation of automated procedures and help to avoid redundant effort for access scenarios.

10.1 Ambiguous segments

Segmentation of Media typically does not lead to uniform results. Commonly, there will be multiple levels of segmentation, depending on the context. For example, an Opera may be divided into Acts, Scenes, Arias and Recitative, which may have further divisions. A BD or DVD movie may be divided into chapters that have Scenes and Shots. Therefore, segments need to be implemented in a way that they can change their hierarchy as required by the context. For MediaLinkingRules, one possible solution is to define multiple instances of the rules, each describing one of the required hierarchy levels.

10.2 Re-mastering

Re-mastering is not an actual use case for building a Master Item. Instead, it subsumes any pre- and post-processing to render a coherent content stream from bits and pieces. This includes re-splicing of physical material, re-synchronisation of specific media sources (audio or subtitle tracks), or DSP processing. This may be an iterative process. From the OAIS perspective, re-mastering is required if the SIPs cannot be used for the direct production of an AIP. In fact, it is a borderline case, as in many cases the original SIPs and the re-mastered SIP will have to be preserved, at least if the process is meant to be reversible.