User Needs and MXF Options

Preservation Planning and the AS-07 Specification

Compiled by Carl Fleischhauer for the
Federal Agencies Digitization Initiative Audio-Visual Working Group (FADGI)

September 17, 2014, updated September 10, 2015

Table of Contents

Introduction

Compiling a needs assessment

In the course of drafting the MXF AS-07 application specification, the AS-07 development team assembled a set of high level user needs. This has been a semi-formal process and, as will be explained, has primarily depended upon AS-07 team members and their connections within multiple overlapping archiving and preservation communities.

The AS-07 effort is led by the Federal Agencies Digitization Guidelines Initiative (FADGI),[1] and the development team includes representatives from two FADGI member agencies, the Library of Congress (hereafter *the Library*) and the National Archives and Records Administration (*NARA*). Both of these agencies are also users. Other AS-07 team members represent private sector companies, e.g., the Philadelphia service bureau George Blood Audio and Video and the Belgian manufacturer EVS-OpenCube. These companies provide services or conversion systems to users and are thereby familiar with their expressed needs.

AS-07 team members are in touch with other archiving and preservation organizations and intermittently discuss needs with them. For example, several on the AS-07 team are also members of the Technical Committee of the International Association of Audiovisual and Sound Archives (IASA), which launched its TC-06 video guideline project in 2009. The national libraries of Australia and Norway are active participants in the IASA TC and in the TC-06 project; others prominent in IASA represent a variety of non-commercial media archives, mostly in Europe.

During 2013-15, the development of AS-07 proceeded under the auspices of the Advanced Media Workflow Association (AMWA). This period provided a number of opportunities for the team to communicate with broadcast professionals, notably with representatives of the BBC archiving unit. Meanwhile, James Snyder, the Senior Systems Administrator at the Library's Packard Campus in Culpeper, Virginia, and a key player in AS-07 development, receives a steady stream of professional visitors from other organizations and participates in a range of organizations, including the standards group of the Society of Motion Picture and Television Engineers (SMPTE).

The most prominent user organizations in AS-07's development thus far have been *memory institutions*, a catchall term for public archives and libraries. As the communications with the BBC indicate, however, there has also been information sharing with broadcasters, principally governmental or (in U.S. terms) *public broadcasters*. In addition, throughout the development period, Oliver Morgan, the Library's expert consultant, carried out a steady stream of consultations with colleagues in the SMPTE standards group, many of whom work in the field of broadcasting.

The Library and NARA joined the AS-07 development process with some clear ideas about the needs of their respective institutions. The Library's ideas had been sharpened by its

---

[1] FADGI is an interagency activity led by the Library of Congress with membership from nineteen U.S. federal government agencies; see http://www.digitizationguidelines.gov/. FADGI includes two working groups. The Still Image Working Group develops guidelines and tools that pertain to the scanning of materials that can be reproduced as still images, e.g., books, photos, manuscripts, and maps. The Audio-Visual Working Group develops guidelines and tools that support the digitization of sound recordings, video, and motion picture film.

experience in the production of archival master MXF files on a production-line basis beginning in 2007, an experience that was echoed by work at the national libraries of Australia and Norway and, more recently, by Libraries and Archives Canada via a service contract with George Blood Audio and Video.

The needs described in this document were the most significant ones that came to the attention of the AS-07 team. Other less significant needs have also been met in the specification.

Scope and shims: today and tomorrow

AS-07 has been defined broadly, as a file format for the preservation and long-term archiving of audio-visual content, ranging from theatrical motion picture content to broadcast and non-broadcast video content to sound recordings as separate entities. Fully realized specifications for the whole range, however, await future versions or extensions of AS-07.

The matter of breadth in MXF Application Specifications is addressed by *shims*. Shims are, in effect, subtypes of an Application Specification that provide additional constraints.[2] In the case of AS-07, several additional shims, tailored to specific use cases, have been identified as appropriate for development. These use cases include the reformatting of existing videotapes, wrapping born digital essences-as-acquired, wrapping scanned film and other content types that feature RGB- and XYZ-based picture essences, and creating audio-only files. Interest has also been expressed in additional content types like telemetry data, HDR imagery, and multi- and hyper-spectral imagery.

At this writing, one shim has been prepared for release with the initial publication of the specification: a shim to support the carriage of baseband video. In this context, *baseband video* is understood to encompass both analog and uncompressed digital forms. An MXF file structure for baseband video represents the priority use case for FADGI: the reformatting of older analog and digital videotapes and, at a few agencies, the encoding and packaging of "live" content streams. (For example, the Library will be receiving, processing, and archiving high definition digital streams from congressional venues.) In all of these instances, a baseband or uncompressed digital video signal is input to an MXF-file production system.

This document: state the needs and explain how they can be met by MXF

This document is structured as pairs of descriptions. Each pairing leads off with a description of one or more important user needs--roughly speaking, user requirements. This is followed by a description of how the requirement can be implemented in an MXF file. The implementation statements cite section numbers in the draft AS-07 specification. Readers working with non-MXF formats may wish to consider whether the stated needs fit their circumstances and, if so, how they are met in the format they prefer.

At this writing, the review draft version of AS-07 is available at the AMWA Web site (http://www.amwa.tv/projects/AS-07.shtml)..

---

[2] The relationship between AMWA Application Specification and shims is discussed in Al Kovalick's 2012 paper *The AMWA Family of Application Specifications for MXF* (http://www.amwa.tv/downloads/whitepapers/AMWA_ASfamilyKovalick4-2012.pdf).

Needs and MXF Implementations

Timecode

Need: provide coherent master timecode

Systems that play back files will benefit from, if not require, the presence of a high integrity, continuous timecode. This may exist in the video recordings that are to be preserved, referred to as *source materials* or *source videos*, but it is often a good practice to record a fresh, continuous, high integrity *master timecode*. In some contexts, this is called synthetic timecode.

Need: retain legacy timecode

The source materials that are to be reformatted may carry multiple timecodes: vertical interval timecode (VITC), linear timecode (LTC), and more. Some are present on purpose, others by accident, some may have good integrity and continuity, while others may be discontinuous. The legacy timecodes in videotapes and other sources may themselves be layered in ways that an archive wishes to track, e.g., a videotape may carry LTC and may additionally carry an earlier generation of timecode recorded, say, as audio track 3. Any or all of these timecodes may provide forensic help for future researchers. A legacy or historical timecode may be keyed to old documents like tape logs, may provide clues about the older source tapes that were assembled to create the video program you are now preserving, and may (as with footage of NASA space vehicle launches) represent elapsed time that can be correlated to other data streams. In many cases, this is data you do not want to lose.

In this document and in the AS-07 specification, retained legacy timecodes are called *Historical Source Timecode*, a term borrowed from the European Broadcasting Union (EBU).[3]

Need: label multiple timecodes

Given the need to carry a master timecode and (often) one or more historical source timecodes, there is also a need to tag the timecode so that users of the file can identify each one.

Implementation: provide Master Timecode

- From section 6.4.4.2, Master Timecode in Header Metadata File Package
    - Encoders shall place uninterrupted, ascending AS-07 Master Timecode as a Timecode Track and shall identify it by setting the track number property to 1. There shall be only one timecode track with a track number property value of 1 in a Package. The Master Timecode EditRate, RoundedTimecodeBase, and DropFrame Properties shall match the frame rate and count mode of the Picture Essence in the file.
- From section 6.4.4.3, Master Timecode in Header Metadata Material Package
    - AS-07 encoders shall generate a timecode track for each Material Package. This is in addition to the Master Timecode encoded in the Top Level Source Package.

[3] The term is used in EBU R 122, *Material Exchange Format Timecode Implementation*: http://tech.ebu.ch/docs/r/r122.pdf.

For AS-07 files, the default start timecode of the material package timecode track should be equal to the timecode time address of the source package position that is referenced by the start of the first material package source clip. . . . Timecode frame rate and mode (drop-frame or non-drop frame) are required properties of a TimecodeSegment.

Implementation: retain Historical Source Timecode

- From section 6.4.5.3, Historical Source Timecode in Essence Container System Items
  - When supplied to the encoder, Historical Source Timecode shall be encoded in the second and subsequent elements of the ST 405 TimecodeArray of the ST 394 System Element. . . . Historical Source Timecode in Essence Containers shall be stored with each frame and not as a start and duration. Encoders shall accommodate discontinuities in incoming Historical Source Timecode in Essence Containers and shall record matching discontinuities within the ST405 TimecodeArray.
- From section 6.4.5.4, Historical Source Timecode Tracks in Header Metadata for TLSP
  - AS-07 encoders should generate a timecode track for each instance of Historical Source Timecode, numbered as indicated in section 6.4.3.2, referenced by the DateTimeEssenceTrackID within the DateTimeSubdescriptor.
- From section 6.4.5.5, Historical Source Timecodes in Essence Container Essence Container Picture, Sound, and Data Items
  - Additional Historical Source Timecodes may also be represented:
    - as SMPTE ST 12-2 data in ANC packages in one or more Data Items in the Essence Container.
    - as LTC in Sound Items in the Essence Container.
    - as VITC in the Picture Items in the Essence Container (such as a VBI line on the picture in D10 video essence, timecode GOP Header in MPEG-2 essence, and so on).
  - Encoders should encode a DateTimeDescriptor as specified . . . . Additional Historical Source Timecodes may also be represented as SMPTE ST 12-2 data in ANC packages in one or more Data Items in the Essence Container.
- From informative section 6.4.5.6.1, Historical Source Timecode in Lower Level Source Packages
  - EBU R 122 (Material Exchange Format Timecode Implementation) foresaw the need to identify and characterize MXF files that contain multiple expressions of Timecode. In section 3 (*Recommendations*) of this EBU standard, recommendation 2.e specifies an approach that places Historical Source Timecode(s) in timecode tracks of the Lower Level Source Package (LLSP).
- From section 6.4.5.6.2, Historical Source Timecode in Lower Level Source Packages
  - Each AS-07 shim will specify its requirements for the carriage of AS-07 Historical Source Timecode tracks in Lower Level Source Packages (LLSP) as [mandated, forbidden, encouraged, or permitted] . . . .

- From section 6.4.5.6.3, Historical Source Timecode in Lower Level Source Packages . . .
    - When Historical Source Timecode tracks are to be placed in Lower Level Source Packages, AS-07 encoders shall accommodate discontinuities in incoming Historical Source Timecode. Discontinuous timecode shall be represented as a Sequence of TimecodeComponents (ST 377-1 annex B.16). Continuous timecode shall be represented as a TimecodeComponent with Start Time and Length (ST 377-1 annex B.17). Segments with no timecode or undecodable timecode shall be represented as Filler (ST 377-1 annex B.10).

Implementation: label multiple timecodes

- *Comment*:
    - The AS-07 method for tagging multiple timecodes in an MXF file makes novel use of two SMPTE standards (ST 385 and ST 405). This tagging may appeal to those developing practices in other contexts, thereby increasing adoption.
- From informative section 6.4.3.1, Labeling Timecode in Header Metadata
    - . . . AS-07 uses elements from two SMPTE specifications: ST 405 specifies a method to construct timecode arrays in essence container System Items, while ST 385 provides a scheme for descriptors and subdescriptors. These descriptors and subdescriptors are associated with Timecode Tracks. In the case of Timecodes (all types) in essence container System Items, the tracks and descriptors are to be carried in the Top Level Source Package. When Historical Source Timecode Tracks are carried in a Lower Level Source Package, the descriptors will be carried in that location as well. The subdescriptors provide additional properties to identify the essence tracks from which the timecode data was acquired.
- From section 6.4.3.2, Labeling Timecode in Header Metadata
    - AS-07 encoders shall create Timecode Tracks in conformance with the rules of ST 377-1 B.7 for Track IDs and B.15 for Track Numbers.
    - Essence Descriptors of Source Packages should include a DateTimeDescriptor for each Timecode and should comply with [more specifications regarding Descriptors and Subdescriptors].
- From informative section 6.4.3.2.1.1, Timecode Header Label Descriptor
    - The DateTimeDescriptor for AS-07 is derived from the one specified by ST 385 table 3. The list of properties of the DateTimeDescriptor, which is derived from ST 385 table 3 and updated to match ST 377-1:2011 is provided in appendix C.3. Note that a single DateTimeDescriptor can simultaneously describe a Timecode Track, an Essence Timecode, and a SystemItem Timecode, with one DateTimeSubdescriptor for each. The LinkedTrackID property specifies the ID of the Timecode Track that is described; the DateTimeEmbedded flag indicates if the timecode data is also embedded in the essence, at the DateTimeEssenceTrackID and DateTimeChannelID given in that subdescriptor; and the distinguished value 0 together with the DateTimeChannelID describe the instance within the SystemItem. [Detailed Descriptor requirements are provided in section 6.4.3.2.1.2, not summarized here.]

- From section 6.4.3.2.2.2, Timecode Header Label Subdescriptor
    - AS-07 encoders shall provide values for the Subdescriptors property that strongly references a TimecodeLabelSubdescriptor derived from the ST 377-1 annex B.3, and described in detail in appendix C.4 of this document. For ATC (described in SMPTE ST 12-2 and SMPTE ST 12-3), the value shall be DBB1.

Implementation: Illustrative diagrams for an implemented example of AS-07 timecode

- See the appendix to this document, pp. 22ff.

Embed Text-Based and Binary Data

Need: provide carriage of supplementary metadata (text-based data)

The archivists' concern addressed in this section is not about metadata in general, i.e., the issues that every archive faces in terms of overall collection management and the provision of access to researchers. Nor does this concerned pertain to basic information about file characteristics, the data that an application requires to play a file correctly, often referred to as *parametrics* metadata. MXF standards require a reasonable level of parametrics metadata and the AS-07 specification simply embraces them, with some elaborations spelled out to guide the implementation of SMPTE standards for Active Format Description (AFD), a set of codes used by broadcasters (and others) to ensure that aspect ratios, letter-boxing, pillar-boxing, and pan-and-scan are handled correctly.

The main need addressed in this document concerns what AS-07 calls *Supplementary Metadata*, i.e., metadata that is supplementary in terms of MXF technical requirements. Examples include additional technical metadata about the production or reformatting activity, sometimes called *process metadata*, information about the source item, about quality review outcomes, and *preservation metadata*, e.g., PREMIS. One example of process metadata is provided by the SAMMA reformatting system in use at the Library and elsewhere: an XML-encoded, frame-by-frame record of the metrics associated with each tape transfer. Many organizations wish to maintain process metadata and some see value in embedding such data in files, as it the case with *process history* metadata in the EBU Broadcast WAVE audio format. And beyond this technical and administrative realm, some archives (there are definite schools of thought here!) also plan to embed relatively complete *descriptive* (aka cataloging) metadata.

Need: provide carriage of captions and subtitles in the form of Timed Text (text-based data)

See the section below devoted to captions and subtitles.

Need: provide carriage of a manifest (text-based data)

A manifest for a file, especially one that may contain multiple elements, supports preservation and good housekeeping by offering an inventory of the file's parts and expresses the relationships between them. Through a mix of required and optional elements, a manifest provides a high level inventory of the parts including their identifiers, data description, MIME type, size and location. This information can help the user to better understand the composition of the file and it will also provide machine-interpretable information for content processing in

later phases of the life cycle. Manifests of one sort or another are included in several formats ranging from the digital library community's BagIt specification to the Interoperable Master Format (IMF) developed by the entertainment industry in Hollywood.

Need: provide carriage of still images, documents, EBU STL, etc. (binary data)

The AS-07 team used the term *associated materials* to name a class of binary representations of materials closely associated with the file's primary essences, e.g., scanned images and documents, video trailers, etc. Associated materials contribute to the completeness, comprehensibility, or usability of the information object represented by the AS-07 file, and often take the form of files in formats such as TIFF, JPEG, MP4, PDF, and the like. The need is to be able to embed and carry such data in the archival master file. Archives will also often also wish to retain and carry instances of the European Broadcast Union (EBU) binary subtitling format, called *EBU STL* and standardized in EBU Tech 3264 (1991).

Implementation: provide carriage of text-based and binary data

- The following bullets refer to the MXF *Generic Stream Partition*. For background information on this entity, see the sidebar at the end of this section.
- From informative section 6.2.4.1, Generic Stream Partitions and Embedding Data
    - Generic Stream Partitions (SMPTE ST 410:2008) are containers for generic data streams that could be continuous or tied to the timeline, including classes of metadata that cannot be referenced from MXF Header Metadata. Depending on the entity type, a Generic Stream Partition could be associated with an instance of a Descriptive Metadata Scheme (DMS), as specified below.
    - Data streams in AS-07 Generic Stream Partitions that consist of Timed Text or EBU STL (both specified in section 6.2.12) will be considered to be essences, will be referenced in tracks in the file's Material Package and Top Level Source Package, and will influence the determination of the file's Operational Pattern. Other text-based and binary data in AS-07 files will generally not be considered to be essences and will not influence the determination of an AS-07 file's Operational Pattern.
    - Table listing the four types of entities to embed and references to specification sections:

Category	Entity type	How described in file metadata?	Main informative and normative sections
AS-07 Essence Binary Data Objects			
	EBU STL	Caption data Descriptors, no DMS Standard: SMPTE ST 2075:2013	6.2.12.7
	Other Essence Data	*Deferred*	
AS-07 Non-Essence Binary Data Objects			
	Binary Associated Materials	AS_07_BD_GSP_DMS	6.2.4.1.2
AS-07 Essence Textual Data			
	SMPTE and EBU Timed Text	Caption data Descriptors, no DMS	6.2.12.6 6.2.12.7

AS-07 Non-Essence Textual Data			
	Supplementary Metadata and Manifest	AS_07_TD_GSP_DMS	6.2.4.1.4

- From section 6.7.1.5, XML Schema, Naming, and Carriage of Manifest
  - When a manifest is required by a shim, AS-07 encoders shall encode the AS-07 Manifest as an XML document (W3C XML 1.0), conforming to the XML schema defined in appendix H. The manifest file shall be named *manifest.xml*. Encoders shall treat the manifest file as a form of text-based data and embed it in a Generic Stream Partition in the AS-07 file . . . .
- From section 6.2.4.2, Generic Stream Partition Encoder Requirements
  - Encoders shall be capable of producing AS-07 files that contain Generic Stream Partitions (SMPTE ST 410-2008) within MXF Body Partitions and included in the Random Index Pack. Encoders shall be able to receive a Generic Steam Payload and write it to a valid Generic Stream Partition. Encoders shall accommodate any of the data stream types defined in Annex A of SMPTE ST 410-2008. . . . Encoders shall treat data streams in AS-07 Generic Stream Partitions that consist of Timed Text or EBU STL (both specified in section 6.2.12) as essences, and shall reference them in tracks in the file's Material Package, and use them to determine the file's Operational Pattern (OP1b when Timed Text is present). . . .
  - For each instance of a Generic Stream Partition containing non-essence binary or textual data . . . encoders shall create an instance of AS_07_BD_GSP_DMS or AS_07_TD_GSP_DMS [Descriptive Metadata] as appropriate.
  - When required by a shim, encoders shall wrap the Manifest according to SMPTE RP 2057:2012 and carry it as a form of non-essence textual data in a Generic Stream Partition . . . The Manifest shall conform to the formal element definition in the XML schema declaration as specified in appendix H. The Manifest shall require an instance of AS_07_TD_GSP_DMS [Descriptive Metadata] . . . .

Sidebar: Generic Stream Partitions

SMPTE's MXF standards establish a file structure with several types of partitions. As the name implies, *Generic Stream Partitions* (SMPTE ST 410:2008) are designed to carry generic "streams" of various types. In effect, they are boxes into which you can put things. The standard requires that each Generic Stream Partition be accompanied by a block of metadata that offers a minimal amount of information about what the partition carries.

AS-07 specifies the use of Generic Stream Partitions in connection with the carriage of five types of entities: (1) supplementary metadata (*text-based data*), (2) captions and subtitles in the form of SMPTE Timed Text (*text-based data*), (3) a manifest for the file (*text-based data*), (4) "associated materials," e.g., still images of pictorial or textual items considered by an archive to be "part of the object" (*binary data*), and (5) EBU subtitles (STL; *binary data*).

The metadata that describes the content of AS-07 Generic Stream Partitions is provided by two variants on the SMPTE-standardized Descriptive Metadata Scheme (DMS), as specified in AS-07. One variant defines metadata elements for text-based streams (supplementary metadata, Timed Text, or the manifest), while the second DMS defines elements for binary data.

The AS-07 approach for Timed Text, including carriage in Generic Stream Partitions, follows the requirements of relevant SMPTE standards. Data streams in AS-07 Generic Stream Partitions that consist of Timed Text or EBU STL will be considered to be essences, will be referenced in tracks in the file's Material Package, and will influence the determination of the file's Operational Pattern. Other text-based and binary data in AS-07 files will generally not be considered to be essences and will not influence the determination of an AS-07 file's Operational Pattern.

Captions and Subtitles

Need: retain and provide carriage for captions and subtitles

Caption and subtitle texts, once extracted and indexed, have clear value for archives, supporting search and retrieval as well as other outcomes. If part of a source video recording, their retention is also required when producing an authentic copy.

Need: translate binary-format captions and subtitles to Timed Text

What form of captions or subtitles ought to be carried in a preservation file? The binary forms will be awkward for future extraction, since their use will depend on the continued availability of decoding tools and may require real-time playback. (Nevertheless, many archivists want to retain them in their original forms in order to have an *authentic* copy.) Meanwhile, easy-to-extract XML Timed Text is very desirable for archives.

Sidebar: background information on captions and subtitles

This document uses the terms *captions* and *subtitles* more or less interchangeably, to mean non-XML text intended for display over a timeline, in synchronization with image and sound essence. The term *Timed Text* carries the same meaning with the added constraint that such text is structured to comply with either the SMPTE or EBU Timed Text XML schema.

Captions and video subtitles are important features of broadcast collections although they are less frequently encountered in other classes of content. Looking back, US broadcast standards have required various flavors of binary-coded closed captioning (CC), first in line 21, right at the boundary between the picture raster and the vertical interval, and then in other places in the stream. For example, there is a specification for packet-carried captions in MPEG digital broadcast streams. Other digital video streams carry CC in packets that meet a different specification. In Europe, there are some parallels, beginning with Teletext and later, EBU subtitles (STL), another binary system.

Today, broadcast authorities on both sides of the Atlantic want to set these binary approaches aside and move toward XML-based Timed Text, derived from the W3C Timed Text standard and very applicable to the Web dissemination of video. Recent Federal Communication Commission (FCC) regulations in the United States name SMPTE ST 2052-1:2010 (Timed Text Format) as a preferred option for Internet presentations. The costs and level of effort required to convert existing systems to full XML capability, however, have led to a certain level of pushback from broadcasters.

Implementation: retain and provide carriage for captions and subtitles

- From section 6.2.12.5.1, CEA-608 and CEA-708 Data Carriage
    - If CEA-608 (CC and XDS) data or CEA-708 DTV captioning data is present, AS-07 encoders shall carry such data in a SMPTE ST 334-1/-2:2007 compliant ANC packet within a frame-wrapped Data Element in the Generic Container as described in SMPTE ST 436:2006; using 8 bit encoding. See section 6.2.13 for more information on ANC packet carriage in AS-07.
    - In addition to mandatory carriage in the ANC packet:
        - If CEA-608 (CC and XDS) data or CEA-708 DTV captioning data is present in the source material, AS-07 encoders shall preserve CEA-608 and CEA-708 in their native binary format, and
        - If CEA-608 (CC and XDS) signals are present in the source material, either as an analog signal or as a digital representation thereof, AS-07 encoders shall preserve these signals.

Implementation: translate binary-format captions and subtitles to Timed Text

- From section 6.2.12.5.2, Translation of CEA-608, and -708 to SMPTE Timed Text
    - AS-07 encoders should translate CEA-608 and -708 data to SMPTE ST 2052-1 Timed Text in the Preserve Translation Mode (ST 2052-1, section 5.1.2.1) . . . . In order to avoid confusion with the binary data as delivered, AS-07 encoders shall not translate to provide Carriage of Binary Data "tunneling," as described in ST 2052-1, section 5.4, and in ST 2052-0:2013 (now in final draft). . . .

Implementation: carriage of SMPTE Timede Text

- From section 6.2.12.6.1, AS-07 Encoder Requirements for SMPTE Timed Text
    - As described in SMPTE 429-5:2009, each Ancillary Resource in a Timed Text Track File shall be entirely contained within an MXF Generic Stream Partition defined by SMPTE ST 410:2008. . . .
    - When SMPTE Timed Text is present in an AS-07 file, encoders shall reference the Timed Text in tracks in the file's Material Package and Top Level Source Package as described in SMPTE 429-5: 2009. The Top Level Source Package shall contain one Data Essence Track with a single Data Source Clip. A single Material Package shall be present which shall contain one Data Essence Track with a single Data Source Clip referencing the Top Level Source Package. The operational pattern is designated as OP1b.

- The Top Level Source Package shall include a strong reference to a TimedTextDescriptor, which shall describe the Timed Text resource according to SMPTE 429-5:2009, Appendix A2. . . .

Implementation: retain and translate EBU STL, carry EBU TT

- From section 6.2.12.7.2, AS-07 Encoder Requirements for EBU STL
  - AS-07 encoders shall convert EBU STL data to EBU-TT following the mapping provisions of EBU-TT part 2 (EBU Tech 3360, v.0.9 for comment, June 2013). Additionally, AS-07 encoders may place EBU STL (EBU Tech 3264) data in Generic Stream Partitions in accordance with SMPTE ST 2075:2013. When EBU STL is present in an AS-07 file, encoders shall ensure that the EBU STL resource be described by a Top Level Source Package as described in SMPTE ST 2075:2013. . . . The Top Level Source Package shall include a strong reference to the STLEssenceDescriptor according to SMPTE 2075: 2013, including the STLSubdescriptor to describe multiple languages that are stored in a single STL file which is mapped into a single MXF file.
- From section 6.2.12.7.3, AS-07 Encoder Requirements for EBU TT
  - As described in SMPTE 429-5:2009, each Ancillary Resource in a Timed Text Track File shall be entirely contained within an MXF Generic Stream Partition defined by SMPTE ST 410:2008. . . .
  - [Other requirements that closely parallel those for SMPTE TT, not summarized here.]

Sidebar: implementing captions, subtitles, and timed text in AS-07

SMPTE MXF and EBU standards both offer multiple options for carriage of caption and subtitle original forms and for Timed Text, and the AS-07 team consulted SMPTE and EBU standards as they drafted the specification. If present in the source recording, the original binary CC remains in line 21 in the image raster. The MPEG packets stay where they were (in that stream) while the other type of digital-stream packets are carried in what are called System Items, side by side with picture and sound essences within the file. Finally, XML Timed Text is to be carried in MXF Generic Stream Partitions (see the Generic Stream Partition sidebar above). Since EBU STL, although binary, has generally been handled as a "sidecar" file, AS-07 also calls for STL to be carried in Generic Stream Partitions.

In order to meet this specification, AS07 encoding systems will have to be able to convert the binary text to XML Timed Text. The archivists' need for re-usable textual data for indexing systems that support search-and-retrieval, led to the AS-07 specification's request that the "tunneling" approach, so convenient for broadcaster handling of Timed Text, not be employed for AS-07 archive files.

Audio Track Layout and Labeling

Need: audio track layout and labeling

Archive users wish to reformat audiovisual content from a wide variety of source material with widely varying sound tracks. In terms of sound or aural field, examples range from silent research footage to monaural oral history recordings to performances with stereo, surround, or multichannel audio. In other cases, the tracks on a source item will include Descriptive Video Service (DVS), Second Audio Program (SAP), annotations (like a director's commentary for a dramatic program), as well as other types of multiple language content or other versioning elements. Sound tracks on certain videotape formats may also carry timecode data, e.g., the carriage of LTC on track three of the 1-inch type C format. Archivists seeking to produce authentic copies wish to retain this source data and require metadata that labels the tracks in a manner that will serve future users.

Source material audio tracks may or may not be labeled according to a standard or industry convention. When so labeled, the tagging may be in terms of such standards as SMPTE Multi-Channel Audio (MCA; SMPTE ST 377-4), the EBU track allocation templates specified by EBU R 48 or EBU R 123, or by an industry convention promulgated by a broadcast network, such as the PBS Audio Configuration specification cited in AS-03. Existing tagging should be retained in archive or preservation files.

Implementation: audio track layout and labeling

- From section 6.2.11.7.2, Audio Track Layout Identification in AS-07-Core-DMS
    - AS-07 encoders shall identify audio track layouts by placing the coded values listed in appendix B in the AS-07-Core-DMS-Audio_Track_Layout element . . . . Encoding devices shall provide a method to permit archiving organizations to input the coded value prior to encoding. If organizations do not provide values in advance, the encoder shall make a best effort to identify the tracks and to use codes as defined in tables 0 through 5 in appendix B.
- From section 6.2.11.7.4, Audio Track Layout Descriptors and Subdescriptors for SMPTE MCA
    - When the video content in an AS-07 file consists of SMPTE Multi-Channel Audio (MCA; SMPTE ST 377-4), and when such information is provided by the encoding organization, AS-07 encoders shall provide the Descriptors specified in SMPTE ST 377-1 and the Subdescriptors specified for MCA in SMPTE ST 377-4. Additional relevant information is provided in SMPTE ST 2035-2009 (Audio Channel Assignments for Digital Television Recorders (DTRs).

Language Tagging

Need: provide a means to tag Timed Text languages

Archives may wish to tag primary and secondary languages in Timed Texts. In addition, primary language tagging is required by SMPTE RP 2057:2011 (*Text-Based Metadata Carriage in MXF*), including Am1:2013. General practices in broadcasting and archiving often rely upon the IETF RFC 5646 and/or upon the coding approaches that underpin RFC 5646, especially ISO 639-2, which is in turn an important basis for the MARC cataloging standard and other library-oriented specifications.

Because collections are often relatively homogeneous in terms of language, many organizations will employ default language values in tags. In the U.S., for example, this will often be the code value for American English ("en-US").

Need: retain language tagging associated with binary caption or subtitle data

Language tagging may be present in caption and subtitle source material, notably examples that employ the standardized approaches established for CEA-608, and -708 (caption service descriptors), EBU STL, and SMPTE and EBU Timed Text. This language identification information should be retained in the output from a reformatting process.

Need: provide a means to tag soundtrack languages

Archives may wish to tag primary and secondary languages in soundtracks. As noted above, general practices in broadcasting and archiving often rely upon the IETF RFC 5646 and/or upon the coding approaches that underpin it.

Implementation: provide a means to tag Timed Text languages

- From section 6.2.12.6.2, AS-07 Encoder Requirements for SMPTE Timed Text
    - The Top Level Source Package shall include a strong reference to a TimedTextDescriptor, which shall describe the Timed Text resource according to SMPTE 429-5:2009, Appendix A2. In accordance with SMPTE 429-5:2009 Appendix A3, the Timed Text resource may be additionally described by the TimedTextResourceSubdescriptor set which may be strongly referenced by the TimedTextDescriptor via the MXF Generic Descriptor (as defined in SMPTE 377-1:2011).

SMPTE Descriptor and Subdescriptor current status

As the preceding bullet indicates, AS-07 depends upon two SMPTE standards for the Descriptors and Subdescriptors for Timed Text. At this writing, SMPTE has not standardized an appropriate Descriptor or Subdescriptor that includes language coding, and the AS-07 team does not feel that it is practical to invent such a specialized element within the confines of its own specification. Parties that implement AS-07 and wish to record the language of Timed Texts may optionally use the PartAnnotationText element in the AS-07 Manifest (6.7.1.3.6) to informally tag the language.

Implementation: retain language tagging associated with binary caption or subtitle data

- From section 6.2.12.1.4, Retain Line 21, CEA-608, and CEA-708 data as Delivered
    - If either CEA-608 line 21 (CC and XDS) data or CEA-708B DTV captioning data are present in the source material, AS-07 encoders shall preserve CEA-608 and CEA-708 in the form in which they are delivered.

Implementation: provide a means to tag soundtrack languages

- From 6.2.11.5, Language repertoire and tagging [Carl says OK]
    - AS-07 producers are encouraged to tag soundtrack languages (primary and secondary) in AS-07-Core-DMS (section 6.6.1) but this is optional unless required by a shim. . . .

Content integrity

Need: provide support for within-file content integrity data

Content destined for archival preservation files will often be managed for long term archiving and preservation management. This objective is supported by a number of actions, including the creation of fixity or hash values and the monitoring of those values for change over time. In other MXF contexts, such codes are referred as *MIC*, glossed as *Message Integrity Code* or *Media Integrity Check*. For digital library specialists, content or media integrity usually turns on whole-file fixity values, critical for a well-run asset management system. But whole-file fixity data cannot be embedded in the file itself: that action would change the file, making the hash value "next time" different, thus invalidating it for comparison and monitoring. Whole-file checksums are a critical part of storage and repository systems but have no place in a file-wrapper specification. For file wrappers, a good fit is provided by specifying a carriage location for hash values on segments of the file, e.g., on a frame or some other small unit of video.

There are several models for within-file fixity data, including examples from the BBC and the from the digital cinema community. The BBC Archive Preservation File Format specified in *BBC White Paper 233*[4] employs an approach dubbed *frame-level checksum*. The digital cinema approach is standardized in SMPTE ST 429-6:2006, *D- Cinema Packaging -- MXF Track File Essence Encryption*. In the latter example, fixity data is conjoined with data pertaining to encryption.

Frame-level hash values (often referred to as checksums or Cyclic Redundancy Checks, CRCs) are sometimes employed for use cases such as monitoring production. For example, some specialists use ffmpeg's *framecrc* and *framemd5* checksums to judge the success of lossless compression processes. Process monitoring, however, is not a goal for the AS-07 specification.

---

[4] http://downloads.bbc.co.uk/rd/pubs/whp/whp-pdf-files/WHP233.pdf

Sidebar: Background on creating and embedding AS-07 "frame-level" content integrity data

The AS-07 approach described below calls for the embedding of fixity data on the *V* or value data in the KLV triplets that represent frame-wrapped essences. Similar approaches are used in other standards and specifications and, writing informally, this is often referred to as frame-level or edit-unit-level fixity; the latter term is defined in SMPTE ST 377-1:2011. AS-07 files will generally be frame-wrapped, with the exception of files that carry long-GOP D-10 essences. For D-10, content integrity systems native to long-GOP are to be retained in AS-07 files.

Frame-wrapped picture may be progressive-scanned or interlaced. Picture data for progressive-scanned content will be represented as the *V* in a KLV triplet, and the calculation of fixity is straightforward. Picture data for interlaced video will very often be carried with the data from both fields represented as a single *V* in a KLV triplet. This is the case for uncompressed video mapped according to SMPTE ST 384 and ST 377-1 (annex G.2.25), and also for JPEG 2000 compressed video case *I2* (frame wrapping, interlaced two fields per KLV triplet) mapped according to SMPTE ST 422:2013.

The exception to the general rule outlined in the preceding paragraph is the JPEG 2000 interlaced picture wrapping identified as case *I1* in SMPTE ST 422:2013, where each field is wrapped as a separate KLV triplet. In this case, AS-07 requires that the concatenated *V* values for pairs of KLV triplets be hashed as one. AS-07 uses this approach so that the integrity data for interlaced video is always at the frame (edit unit) level. The same hash value would be calculated as from case *I2*, and this outcome supports integrity monitoring if an essence is re-wrapped from *I1* to *I2* or vice versa.

The AS-07 approach borrows from two important precedents: (1) SMPTE ST 429-6:2006 (D- Cinema Packaging -- MXF Track File Essence Encryption) and (2) the BBC Archive Preservation File Format described in section 5 in the BBC White Paper 233, cited above. From SMPTE ST 429-6:2006, AS-07 re-uses the equivalent of a DMS (Descriptive Metadata Scheme) system for fixity data. In the digital cinema context represented by this standard, fixity data is conjoined with data pertaining to the encryption of the triplet.

Although the use of encryption will be very rare in AS-07 files, in order to allow for this rare use and also to remain consistent with ST 429-6:2006, the AS-07 specification uses that standard's terminology: *Cryptographic Context Set* (like a DM Scheme), *Cryptographic Framework* (like a DM Framework), and *Cryptographic Framework DM Tracks*. The Cryptographic Context Set implemented in AS-07 includes three adaptations from the ST 429-6:2006 implementation: (1) the addition of the optional MICCarriage item, (2) specifying the permitted Null value as the default value for the CipherAlgorithm item and (3) specifying 0 (zero) as the default value for the CryptographicKeyID item.

When content integrity data is created for an AS-07 file, however, the specification does not require the *Encrypted Triplet Variable Length Pack* specified by ST 429-6:2006 to carry the hash values. Instead AS-07 employs the System Item in the Generic Container, like the BBC.

SMPTE ST 429-6:2006 specifies the SHA-1 algorithm for integrity. In contrast, the AS-07 specification calls for the more easily created Castagnoli CRC-32C. The Encrypted Triplet Variable Length Pack from ST 429-6:2006 also carries an element called *Sequence Number*, defined as "Sequence number of this Triplet within the Track File." In AS-07, the required

carriage of the Master Timecode in a System Item (see section 6.4.4.4) provides a one-up set of numbers that can be consulted to the same effect. To allow decoders to differentiate between AS-07 use of System Items and ST429-6:2006 Encrypted Triplets, AS-07 defines an optional item MICCarriage in the Cryptographic Context Set in which a SystemItem value indicates the AS- 07 usage and whose absence indicates use of Encrypted Triplets.

The BBC Archive Preservation File Format provides AS-07 with the structure that carries the fixity data itself, as specified in BBC White Paper 233, which refers to the approach as a frame- level checksum. There is one small variation: BBC calls for the use of the PNG CRC-32 Cyclic Redundancy Code algorithm; instead, AS-07 specifies the Castagnoli CRC-32C.

It is beyond the scope of a wrapper specification to specify when in an organization's workflow the initial MIC hash value should be calculated. It is worth noting, however, that many experts counsel that hash creation should occur at the moment of initial encoding, a possibility enhanced by the selection of the Castagnoli CRC-32C hash, which is easy and fast to calculate. Generating the initial hash at the time of encoding means that a sophisticated file-creation system can use this data to verify that the file has been correctly written to media the first time file-writing occurs, thereby supporting quality control at an early stage in the life cycle.

Implementation: provide support for within-file content integrity data

- From section 6.7.2.2, CRC-32C values per KLV essence triplets
    - When required by a shim, AS-07 encoders shall calculate a Castagnoli CRC-32C Cyclic Redundancy Code (IETF RFC 3385) value for every V or value data unit in the KLV triplets that represent frame-wrapped essences, with the exception of interlaced JPEG 2000 that is wrapped according the case I1 specified in SMPTE ST 422:2013, the case in which each field is wrapped as a separate KLV triplet. In the latter case, when integrity data is required by a shim, AS07 encoders shall calculate the Castagnoli CRC-32C for the concatenated values of the two Vs in the pair of KLVs. . . .
- From informative section 6.7.2.3, Content integrity values carried in arrays in Essence Container System Items
    - The structure of data arrays of the type described here, and in the section devoted to Timecode (6.4), are governed by the batch syntax for KLV values specified in ST 2003:2012. For AS-07, the TimecodeArray is a single property whose value is an array, with the first element MasterTC, and with second and subsequent elements representing other Historical Source Timecodes. The integrity data is represented in a HashArray with a single property whose value is an array, with the first element EssenceTrack Hash, and with second and subsequent Hashes for other EssenceTracks. Generally speaking the first EssenceTrack is picture and the second and subsequent elements are sound, as in the BBC illustrative example [from *White Paper 233*]. . . .

- From section 6.7.2.4, Content integrity array in Essence Container System Items
    - The CRC-32C values shall be stored in essence System Items as arrays that comply with SMPTE ST 2003:2012.
- From section 6.7.2.6 Cryptographic Context Set, Cryptographic Framework, and Cryptographic Framework DM Tracks
    - When CRC-32C hash values are created for frame-wrapped essences, AS-07 encoders shall also create and populate Cryptographic Context Set, Cryptographic Framework, and Cryptographic Framework DM Tracks as specified in SMPTE ST 429-6:2006, with the optional item MICContainer in the Cryptographic Context Set in which a SystemItem value indicates the AS- 07 usage and whose absence indicates use of Encrypted Triplets. . . .

Segmentation

Need: provide support for segmented content

Segmentation refers to the presence of regions in the program's essence data that represent parts of a larger whole (e.g., episodes in a series) or points where the program content may be broken (interrupted) in playback. Segmentation may be useful to archives, e.g., if a content asset is a complete movie, a segmentation structure would indicate where the reels start and stop. If the content is episodes of television series, a segmentation structure would indicate where the episodes start and stop. Another example is the film strip genre, where the timing and linkage to the sound track could be described in terms of segmentation.

Sidebar: note on DMS-based Segmentation

A DMS-based segmentation structure takes on added importance in an MXF context because one of the elements in the MXF standard has not been implemented by vendors and users. This is the family of operational patterns under the OP2 heading: OP2a, OP2b, and OP2c. On paper, MXF OP2 files can be comprised of separate essence units--one for each segment--with playback governed by the file's Material Package, in effect a set of sequencing instructions. However, since OP2 has not been implemented in today's marketplace, other methods are required in order to achieve a similar outcome. Thus DMS-based segmentation. The precedent for AS-07 comes from AS-11, an application specification for broadcast files in which video is segmented in order to permit the broadcast-time insertion of non-program content like advertising.

Implementation: provide support for segmented content

- From section 6.7.5.2.1.2, Segmentation Track Detailed Requirements
    - If AS_07_Segmentation_DMS is used in an AS-07 file, encoders shall represent program segmentation by creating an MXF Timeline track in the file's Material Package, referred to as the Segmentation Track. Encoders shall construct the

Segmentation Track's descriptive metadata in accordance with the recommendations of SMPTE EG 42:2004 and SMPTE ST 377:2011. . . .

- From informative section 6.7.5.3, DM_AS_07_Segmentation_Framework
    - The DM_AS_07_Segmentation_Framework extends the generic MXF DM Framework class. It contains the segment's part number and the total number of parts in the program. These metadata items represent part numbers of the form "1 of 3", "2 of 3", "3 of 3". Refer to appendix G for the complete definitions of DM_AS_07_Segmentation_Framework and DM_AS_07_Segmentation_Scheme.

Appendix. Timecode Descriptors and Subdescriptors

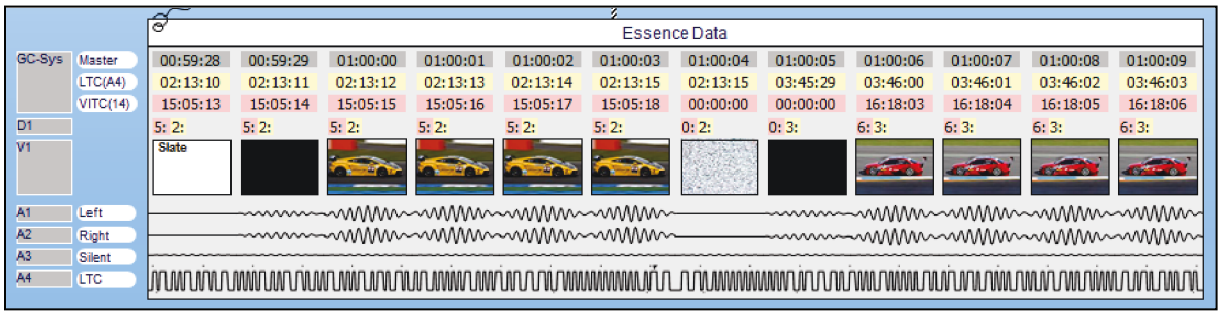
1. Explanatory Illustrative Diagram

1.1 Source videotape illustrative example

- Source: 1-inch videotape with timecode
- Picture; footage with slate and camera starts and stops.
  - *VITC* in line 14 (discontinuous, jumps to zero in gaps)
  - *Visual representation of AS-07 MXF Essence Container carriage is offered in the diagram in the next section.*
- Audio: four channels
  - stereo audio on A1 and A2
  - silence on A3
  - *LTC* on A4 (discontinuous, repeats a frame number)

1.2 Elements in the resulting AS-07 MXF File

1.2.1 Essence in Generic Container



Top (labeled GC-Sys): Generic Container System Items:

- Gray: Master Timecode (synthetic), GCSys Item, element 0
- Yellow: converted LTC, additional GCSys Item
- Pink: ATC (Advanced Timecode, SMPTE ST 12-2); VITC converted to ANC packets

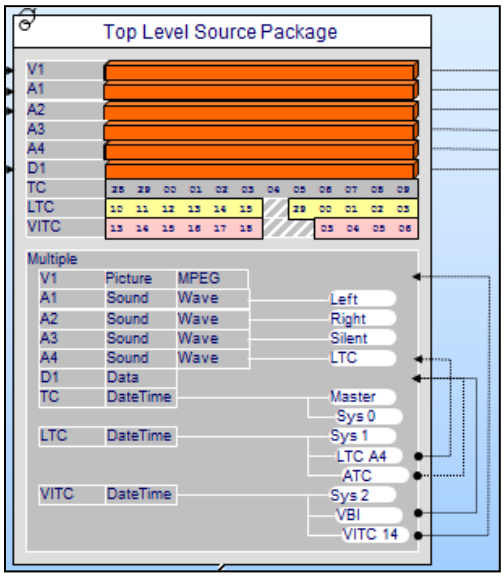
Top (labeled D1): Packetized VITC and LTC

Middle (labeled V1): picture essence (row of images, including starts/stops/snow)

Bottom:

- **Labeled A1**: left
- **Labeled A2**: right
- **Labeled A3**: silent
- **Labeled A4**: LTC (as PCM waveform)

1.2.2 Top Level Source Package, with Descriptors and Subdescriptors



MXF Top Level Source Package (TLSP)

- Six essence and data tracks (orange):
  - 1 picture
  - 4 audio
  - 1 data
- Three TC tracks:
  - *Gray*: TC (Master TC)
  - *Yellow*: LTC
  - *Pink*: VITC
- At bottom:
  - Descriptors (gray)
  - Subdescriptors (white)

Master TC in two places, thus two Subdescriptors

- Master TC track, with the symbolic label *Master*
- GCSys with Master, in element 0 (zero) of the GCSys, symbolic label *Sys 0*

LTC in two places, thus two Subdescriptors

- Second item in GCSys, symbolic label *Sys 1*
- Audio track 4, symbolic label *LTC* with the added Essence TrackID property: *A4*
  - *This Subdescriptor has a LinkedTrackID to connect it to the A4 audio track (dotted line arrow)*
- ATC version of LTC, symbolic label *ATC*
  - *This Subdescriptor has a LinkedTrackID to connect it to the D1 data track (dotted line arrow)*

VITC in three places, thus three Subdescriptors

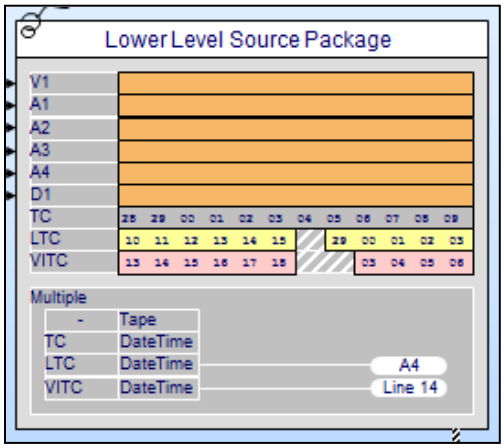
- VITC ingested into the GCSys, symbolic label *Sys 2*
- VBI as Data Item in GC, symbolic label *VBI*
  - *This Subdescriptor has a LinkedTrackID to connect it to the D1 data track (dotted line arrow)*
- VITC in video raster retained on line 14, symbolic label *VITC 14*
  - *This Subdescriptor has a LinkedTrackID to connect it to the V1 picture track (dotted line arrow)*

Note: TLSP track data is metadata. Note that in this example, there is a sequence of components in the LTC and VITC tracks, showing the first segment, filler, and then the second segment.

Regarding the audio tracks: Subdescriptors employ tags from SMPTE ST 377-04 MCA (Multichannel Audio):

left, *right*, *silent*, *LTC.*

1.2.3 Lower Level Source Package

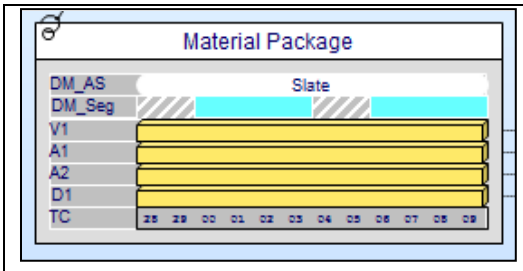


Lower Level Source Package

- Six essence and data tracks (golden):
  - 1 picture
  - 4 audio
  - 1 data
- Three TC tracks:
  - *Gray*: TC (Master TC)
  - *Yellow*: LTC
  - *Pink*: VITC
- At bottom:
  - Descriptors (gray)
  - Subdescriptors (white)

The structure of Descriptors and Subdescriptors is simpler than for the Top Level Source Package. Subdescriptors are provided only for audio track 4 and line 14 in the vertical interval. These Subdescriptors are text (not symbolic) labels.

1.2.4 Material Package



Material Package

- Governs playout of program content
- Four essence and data tracks (yellow)
  - 1 picture
  - 2 audio ("stereo")
  - 1 data
- AS-07 DMS metadata for the slate
- AS-07 DMS metadata for segmentation (if any)

1.2.5 Unified Diagram and Selected Identifiers

The next section in this appendix presents all of the preceding elements in a single, unified diagram. The diagram also shows the presence and linking for some selected identifiers, all of which are part of the normal set required by the MXF family of standards. These identifiers have limited connection to the AS-07 timecode specifications. Each of the packages--Material Package (MP), Top Level Source Package (TLSP), and Lower Level Source Package (LLSP)--has a PackageID in the form of a UMID, drawn to resemble baggage tags. In addition, the tracks inside the packages have TrackIDs that, together with other metadata, establishes the linking relationships shown as dotted arrow lines.

2. Unified Diagram for AS-07 Ingest of 1-inch Type C