Abstract Metadata in Public Broadcasting

Part 1: Introduction

Version 2.1
May 2017

Table of Contents

1 Introduction to This Document

Public Broadcasting Service (PBS) acts as the caretaker of the Interconnection System (IXS) and the main distributor of content. For next evolution of the IXS, PBS has implemented a hierarchical abstract metadata system using the Entertainment Identifier Registry (EIDR). This system will issue unique identifiers to newly registered content based on basic metadata, like titles, genre, etc., making new content more organized and easily discoverable by stations.

This is the first part of a series of documents concerning the metadata system. The other documents in this series are as follows:

- **Part 2: Core Models**
  This part describes the core abstract metadata models that compose the content library.
- **Part 3: Time-based Descriptive Metadata**
  This part explains the metadata that provides detailed information about media content.

This section introduces the next evolution of the IXS and the need for an hierarchical abstract metadata system.

Document Organization

This document is organized as follows:

1. "Broadcasting Content on the Interconnection System" provides the historical context behind the abstract metadata system implemented for the Interconnection System.
2. "The Metadata System" explains abstract metadata hierarchies in general, applies the general idea to broadcasting, and lists some use cases where an abstract metadata hierarchy in broadcasting would benefit station workers.
3. "Metadata in the Broadcasting Content Lifecycle" unites the content lifecycle described in Section 1 with some of the metadata concepts described in Section 2.

External Resources

The following resources provide extra information about the broadcasting lifecycle and IXS:

- [PBS-TOP1] PBS Technical Operating Procedures, Part 1, Program Submission, November 2016. http://mypbs.org/workarea/DownloadAsset.aspx?id=60130968219

- [PBS-TOP2] PBS Technical Operating Procedures, Part 2, Program Distribution from PBS, 2015. http://mypbs.org/workarea/DownloadAsset.aspx?id=60130968217

The following resources discuss metadata in the movie industry and EIDR standards:

- [CM] TR-META-CM MovieLabs Common Metadata, v2.4, http://www.movielabs.com/md/md
- [EIDR-TO] *EIDR Technical Overview*, November 2010. http://eidr.org/technology/#docs
- [PBS-EIDR] *Recommendation for adoption of EIDR for v6 Interconnection System,* January 8. Internal Document.

For metadata in general, see the following:

- [NISO-UM] *Understanding Metadata,* 2004. http://www.niso.org/publications/press/UnderstandingMetadata.pdf
- [NISO-MD] *Metadata Demystified*, 2003. http://www.niso.org/standards/resources/Metadata_Demystified.pdf

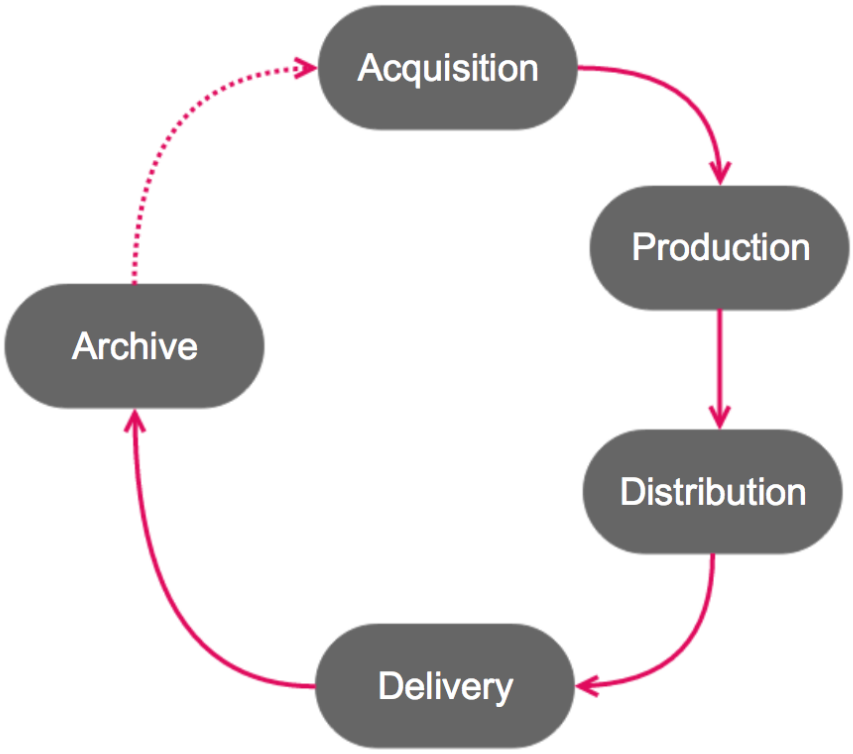
2 Broadcasting Content on the Interconnection System

The public broadcasting content undergoes a typical lifecycle, and that content is described by the metadata system explained in this document. This section explains the importance and reasoning behind the metadata system by providing contextual information about the lifecycle of broadcasting content and the IXS that delivers content to broadcasting stations.

The Broadcasting Lifecycle Phases

While there are many exceptions that can happen for individual pieces of content, the standard content lifecycle can be represented in Figure 1.

Figure 1: Content Lifecycle



The typical first-time flow for most content will follow this sequence:

1. Acquisition
2. Production
3. Distribution
4. Delivery
5. Transition
6. Archive

Acquisition

This stage represents the business and legal tasks that are required to establish a relationship or environment that facilitates the content creation.

Production

This stage represents the creation of content essence, including footage creation (Source Media) and editing (Stories, Releases).

Distribution

This stage represents the preparation required to deliver content to a wider audience, such as when PBS sends information over the Interconnection System to PBS stations.

Delivery

This stage represents content delivery from stations to viewers, such as television broadcast or digital or online availability.

Archive

If Archived, content is no longer actively produced, distributed, or broadcast. Archiving is temporary storage for content until additional lifecycles -- such as re-runs or syndication -- are triggered.

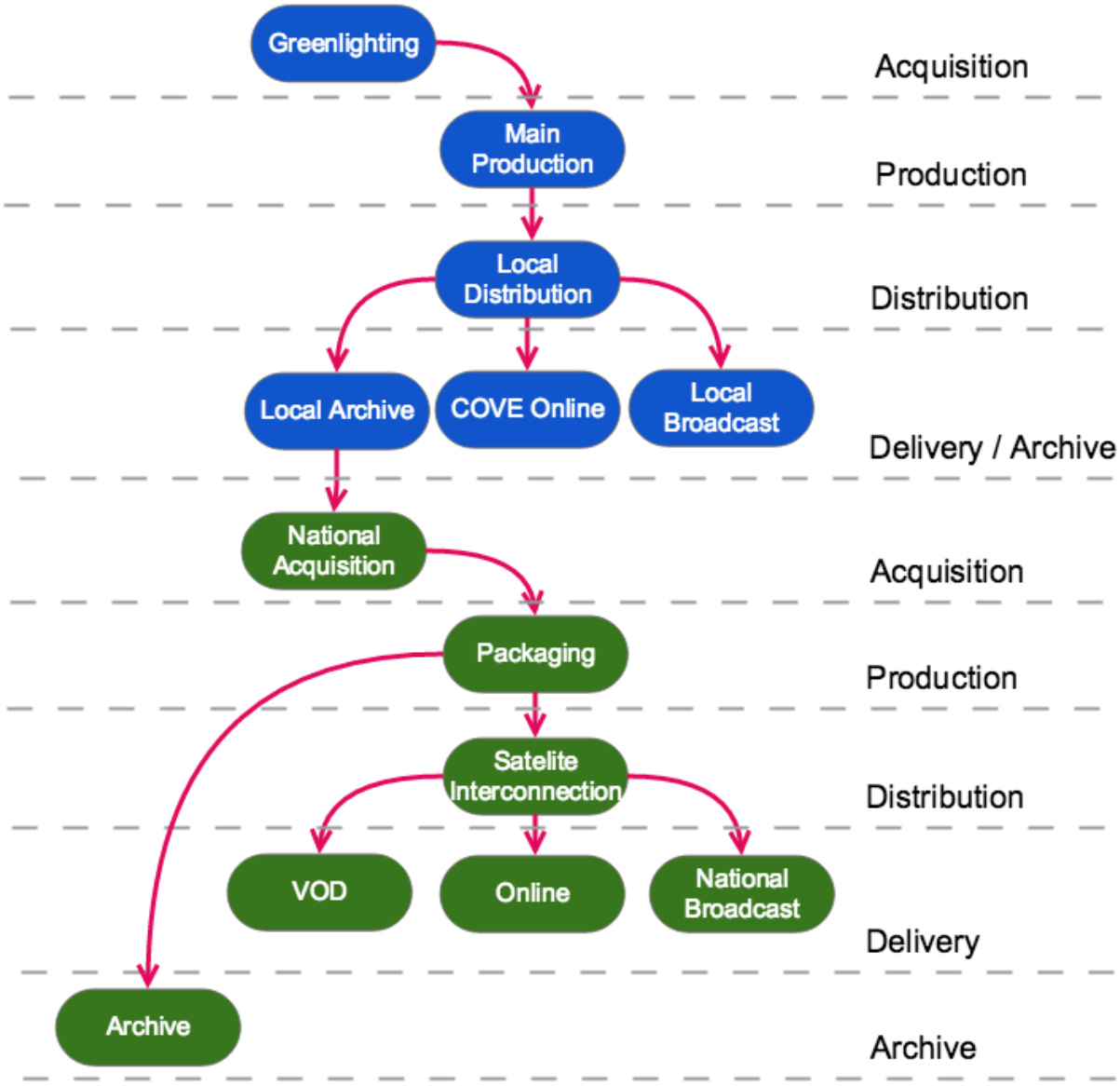
Transition

After content has completed the entire cycle at least once, the following transitions could occur:

- Content could be included in online archives and available for streaming (rights permitting), like the American Archive of Public Broadcasting or the Louisiana Digital Media Archive.
- Content could be pulled from Archive to run through the cycle again from the beginning.
- Content could be jumped to Acquisition and then to Distribution and Broadcast.

Figure 2 show how content could repeat the content lifecycle.

Figure 2: Content Moving Through Lifecycle Stages



The Interconnection System

At the Distribution phase during the content lifecycle phase, PBS distributes content to the Interconnection System.

The Interconnection System is the telecommunications system facilitating content exchange between a network of Public Media entities, including PBS, APT, NETA, and various PBS and non-PBS public broadcasting stations. Since its inception, the Interconnection System has gone through five iterations (v1 - v5) and ongoing enhancements are continually being applied. This metadata specification is intended to be implemented in a future enhancement.

This section provides an overview of past versions of the IXS. This is not intended to provide technical operating specifications. See [PBS-TOP1] and [PBS-TOP2 for detailed specifications.

v5 Interconnection System Overview

The v5 Interconnection System delivered content solely through digital satellite transmissions. Satellite distribution depends on a small number of uplink sites across the country and on participating stations with specialized equipment that had been distributed when v5 was deployed. Uplink sites are few because of prohibitive setup costs, and both the uplink sites and station receivers were prone to experience "rain fades" or snow outages caused by local weather phenomenons degrading satellite signals to or from the satellite transponders in space.

Occasional quality degradation or technological glitches aside, content delivery in v5 has been reliable due to redundant content feeds.

- There are 7 linear feeds that provided content in various formats and time zone variants. See [PBS-TOP1] for the format specifications.
- The v5 system also provided a Non-Real Time (NRT) feed, which delivered pre-recorded files sequentially and unidirectionally from PBS to recipient stations using IP-over-satellite technology.

Scheduled content is obtained from various sources: some is sent to PBS on video tape, some is pre-recorded in a digital video file, and some is delivered via another linear feed over satellite or direct fiber link. Once the content is all aggregated at PBS, PBS distributes the content on schedule over the IXS, and the individual stations chose which feed to use for downloading or streaming.

PBS sets the national schedule for the linear feeds. PBS, APT, and NETA set the NRT feed schedule. Occasionally, inconsistent metadata input from all the various content sources may erroneously indicate that a feed is present in a schedule when it is not.

The sequential nature of both the linear and NRT feed means that a station may miss the time window to access or download a particular file or stream because of technical issues or weather problems. A station may choose to cache the entire daily output of an NRT stream locally to sort through later.

Enhancement Overview

Metadata is information that describes an asset, formatted in a way to be recognized by both computers and humans. A significant enhancement to the IXS will be to overlay a hierarchical system that uses metadata within the Interconnection System to enable non-linear file delivery. This means that the metadata system overlay will allow station traffic systems to request specific video asset files from the centralized storage on the IXS at any time as long as the rights window for that content is still valid.

The ability to demand specific video content whenever schedulers want — as opposed to waiting for content to be delivered during a certain time window via linear or NRT stream — means that searching for video content on the Cloud should be streamlined and optimized.

The new metadata system is designed as a centralized database with near-real-time updates and an exposed API that station traffic systems can use to consume data directly or create interfaces as is needed. The API allows any system to query for specific metadata to find content.

Because of metadata's key role in content discovery, standards about what metadata is tagged and how it is tagged need to be established rigorously. This is especially important given the various organizations and system processes that deliver video files to the IXS. A set of standards for all new video content, whether station-submitted or distributor-submitted, will reduce duplicate submissions and improve the ability to uniquely identify content.

The metadata system borrows and relies heavily on precedents set by the Entertainment Identifier Registry (EIDR), which is a global registry that provides a unique ID for media content based on select metadata entries. This model's strength is its capability to group tangible assets under abstract concepts and tag both the abstract concepts and the tangible assets with searchable and editable metadata.

The IXS metadata system is detailed further in the following section.

3 The Metadata System

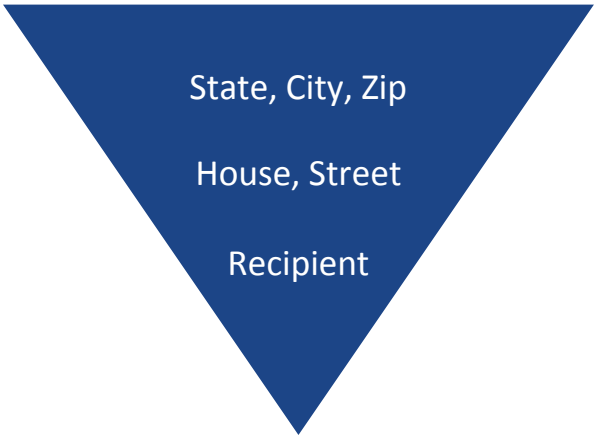
The IXS metadata system depends on a hierarchy of abstract concepts.

A hierarchical abstract metadata model is a system that provides searchable order to particular items. Items can be defined by specific metadata descriptors, and a hierarchy can be built out of logical relationships between the items. A user can then use the hierarchy to pinpoint an exact item they want.

A Non-Broadcasting Abstract Metadata System

The U.S. Postal System wants to find Fred Rogers in order to give him some letters. They could begin their search with his name, but that starting point could turn up a bevy of Fred Rogers. Then the U.S. Postal System would have to turn to other information that they'll have on hand, like City, State, or Street Address, to narrow down the right Fred Rogers.

A more efficient way to find Fred Rogers is by using the following abstract hierarchy:



Model	Precision	Example
City, State, Zip Code	General	Pittsburgh, PA 15650
House, Street	Less General	1 Mister Rogers Street
Recipient	Exact	Fred Rogers

In this hierarchy, City, State, and Zip Code are abstract models. The order in which they are defined helps to streamline the search. So for the abstract model "City" to be of optimal use, the State should be specified beforehand. Following the hierarchical order reduces confusion, like trying to find the Neighborhood of Makebelieve in Charleston, WV instead of Charleston, SC.

Each abstract model can be loaded with metadata descriptors, which are identifying details understandable by both humans and machines.

Depending on the search's purpose, some metadata descriptors may only be helpful in certain situations. For example, the "Recipient" model could contain metadata descriptors like Fred Rogers' `Height` and `Weight`. If the mailman had to identify Fred Rogers by physical appearance, these metadata descriptors would be invaluable. However, for the U.S. Postal Service's purpose of dropping mail off at the right house, the most useful metadata descriptor attached to the Recipient is his `Name`.

In the case of broadcasting, the goal is to make content easily discoverable. This includes both content distributed on the Interconnection System and content stored within a single broadcasting entity.

For an abstract metadata hierarchy to work in Public Broadcasting, the first step is identifying the most useful broadcasting abstract models, their relationship with each other, and the most useful metadata descriptors needed to accomplish common search goals. The second step is to build a system to ensure that the metadata descriptors for each broadcasting abstract model are specified in a timely way.

Abstract Models in Broadcasting

In the non-broadcasting example, the U.S. Postal Service relied on abstract models like Area Code, State, Zip Code, Street Address, Recipient Name.

In public broadcasting, main abstract models are as follows:

- Franchise (optional)
- Series
- Season
- Sub Series (optional)
- Sub Season (optional)
- Episode

This list also reflects the expected hierarchy. In this hierarchy, everything ultimately drills down to the identity of an exact Episode. From there, a station may want to find specific tangible assets like Versions and Manifestations:

- Episode
    - Versions
    - Manifestations

Each of these abstract models and tangible assets need to be further defined by metadata descriptors.

Types of Broadcasting Content Metadata

PBS is a large entity with many moving parts, and one department's search goal is different from another department's search goal. As a result, the broadcasting industry needs a lot of metadata descriptors in order to enable all types of searches.

At the same time, lumping all the possible metadata into a single category would be terribly unhelpful. Logical categories to better organize metadata are as follows:

1. **Technical/Structural Metadata** describes the technical aspects or renderings of a piece of content.
   Examples: HD Level, Aspect Ratio, Audio Type
2. **Relational Metadata** describes the relationships between metadata concepts.
   Example: Franchise contains a Series; Episodes are associated with One Time Onlys.
3. **Descriptive Metadata** is metadata that describes the content of the show. These fields become the basic identifying features of a show.
   Examples: Title, Primary Genre, Primary Language, Primary Genre, etc.
4. **Business Metadata** describes how and when content can be used, as well as who can use the content and what organization and people should be credited for roles or funding.
   Examples: Deals describing intellectual property rights and Usage Window information.
5. **Schedule Metadata** covers data generated by schedules and actual broadcast times.
   Examples: Channels, Feeds, and Schedule information.
   This metadata type will be addressed in a later phase.
6. **Preservation Metadata** contains information needed to archive and preserve content.
   This metadata type will be addressed in a later phase.

Each of the Abstract Models in Broadcasting can have any number of metadata fields that various searches can use to pinpoint a particular concept's unique identity. For example, a Series could have the following metadata values:

- **Title**="Downton Abbey"

- **Synopsis**="Life in the Edwardian country house of Downton Abbey has its ups and downs for the Crawley family and their cadre of servants."
- **Genre**="Historical Drama"

Having many metadata descriptors is crucial to truly identify a piece of content. But imagine if a person insisted that he be addressed as "Male, 5 feet and 9¾ inches tall, 29 years old and 4 months, with a predilection for mince pies and lemonade..." Finding a particular person using all those details can definitely work, but so would a simple identifier generated specifically for a unique set of metadata descriptors, like a full name or social security number.

Identifiers

While one could assume that using the Series Title "Downton Abbey" is a unique enough to track all associated Seasons and Episodes associated with this Series, relying on an identifier formatted like *SeriesTitle*+*SeasonNumber*+*EpisodeName* can get rather cumbersome. One station may then want to truncate the Title to something shorter, while another station may choose to represent the Season as S1 or S01, thus creating multiple, unnecessary identifiers for the same asset.

In short: there are many possible schemas, formats, and policies for assigning an ID, which can cause duplicate entries under different IDs on different systems or can lose content altogether due to miscategorization or missing metadata.

Each piece of content created for and or entered into the IXS needs to be registered in a standardized way to prevent the quirks of individual station from inhibiting content discovery.

For the metadata system, PBS uses the following four identifiers:

1. EIDR
2. UID
3. NOLA identifiers
4. Package Numbers

EIDR IDs are new to the Interconnection System, and they are the key in making a large portion of content metadata intercompatible with external platforms and entities. The Entertainment Identifier Registry provides this global identifier for media content upon registration, and this ID is recognized by any organization or platform that uses EIDR — including Neilsen, Rovi, SMPTE, Netflix, CableLabs, Comcast, and others. Registering all PBS content in EIDR reduces misidentification, duplicated entries, or poor search visibility outside of PBS.

The strength of EIDR's uniqueness and interoperability allows for increased automation; for example, correcting any metadata in EIDR disseminates this correction automatically to all platforms recognizing the EIDR ID.

The EIDR ID is still not the only ID used in PBS because EIDR is a public registry, so any piece of content will become visible to all third-party competitors once registration is complete -- even if production is not finished.

UIDs are IXS identifiers assigned to a piece of content before EIDR IDs are obtained. This ID is used until the content is ready for distribution and should never be exposed externally. Upon publication, the EIDR ID should be used instead.

NOLA code identifiers are PBS in-house identifiers assigned to specific pieces of content scheduled for the linear feeds. It has two parts: the Root code and the Episode code. The Root code is created for a Series, and the Episode code is appended to the Root code to create a Series-Episode relationship. NOLA codes get exposed externally upon publication. NOLA codes are not necessarily unique for content, but rather codes for association (i.e. an Episode may have multiple NOLA codes). Because PBS issues NOLA codes, content from other contributors to the IXS will not have a NOLA code assigned.

Package Numbers are PBS in-house identifiers assigned to Releases and Manifestations. These IDs are used to track specific versions being distributed on the IXS. This ID gets exposed externally upon distribution.

All of these IDs work to track content (and all its metadata) as it moves through the broadcasting lifecycle. Ultimately, all content should have an EIDR ID upon distribution to continue tracking versions, uses, and corrections.

The Impact of an Abstract Metadata Hierarchy in Broadcasting

Once metadata descriptors are specified for a piece of content, the values are visible and accessible by all workers at all stages in the content lifecycle. The metadata system can then use those metadata descriptor values to funnel search queries down to a precise data point, like the recipient's name in the mailing example.

Thus, required metadata fields need to have values specified as soon as possible so that decision-makers at crucial points in the content lifecycle can be informed.

Some specific examples include the following:

- A programmer would like to create a programming schedule. If expected description, genre, usage rights, and duration are documented in the Acquisition stage and properly updated throughout the Production stage, then searching for possible fits will be faster.

- A network system analyzes bandwidth capacity and assesses the impact of traffic on content options. Expected technical metadata should be made available during the Acquisition phase and updated throughout the Production phase.
- A traffic coordinator at the Distribution and Delivery stages can adjust the schedule using Episode and Release metadata to determine if the content is appropriate and up-to-date.
- A content producer can specify Deal metadata at the Acquisition phase to authorize access and enforce access levels throughout the Production stage.
- Content producers can use metadata from the Delivery phase to determine how many organizations are consuming their content to prioritize future projects.

4 Metadata in the Broadcasting Content Lifecycle

The following table gives an overview of what Metadata Types should be defined at what Lifecycle Phase. The third column lists examples of required metadata fields.

The specific abstract metadata fields that are required to describe broadcasting content are listed and detailed in Part 2: Specific Fields and Values.

Lifecycle Phase	Metadata Types	Example
Acquisition	Business/Legal Descriptive	• Organization (if new) • Usage Windows • Deals • Series • Season • Some Episode information (e. • Franchise and Sub-Series, if a
Production	Descriptive Technical Business/Legal	• Episode (episode length) • Credits (cast/crew) • Story • Source Material and Segment
Distribution	Descriptive Technical Business/Legal	• Episode • Story

Delivery	Technical Schedule	• Releases (with appropriate In • Manifestations
Temporary/Archive	[TBD]	[TBD]

Abstract Metadata in Public Broadcasting

Part 2: Core Models

Release 2.1
May 2017

Table of Contents

OrganizationStartDate
Original Release Year
Parental Rating
Primary Genre
Primary Language
Program Format
Program Identifier
Published
Release Description
Release Domain
Release Identifier
Release Type
Revision Description
Revision Identifier
Repair Identifier
Repair Description
Sample Rate
Scan Type
Screen Format
Season Name
Secondary Closed Subtitle Language(s)
Secondary Genre(s)
Secondary Language(s)
Secondary Open Subtitle Language(s)
Sensitive Language
Sensitive Material
Size
Slug
Subtitle Format Type
Subtitle Type
Synopsis100
Synopsis400
Synopsis4000
Title
TitleSortable
Title60
Title256
Track Description
Track Reference
Typical Audio Type
Typical Duration
Typical HD Level
Typical Screen Format
Variable Bit Rate
Vertical Resolution
Video Bitrate Average
Video Bitrate Max
Video Codec

Version Control

Revision ID	Date	Editor	Action
RC-1	4/5/2016	Edgar Roman	
RC-2	4/11/2016	Edgar Roman	
RC-3	4/29/2016	Edgar Roman	Parts 1 & 2 were published in draft
1.0	5/16/2016	Edgar Roman	Parts 1 & 2 were published
RC-4	7/13/2016	Edgar Roman	
2.1	5/31/2017	Interconnection Working Group	Part 2. Core Content Metadata: <u>Series</u> • Synopsis90 now Synopsis100 • NOLA Root now Program Identifier • NOLA Episode now Episode Identifier <u>Release</u> • Release Identifier added • Repair Description added Core Business Metadata • Prefix Codes added <u>Reference for Field Values</u> • Audio Type Dolby replaced with Surround • Release Type o Stacked/Unedited replaced with Stacked o Evergreen added

1 Introduction

Public Broadcasting Service (PBS) acts as the caretaker of the Interconnection System (IXS) and the main distributor of content. For the next evolution of the IXS, PBS has implemented a metadata model based on the Entertainment Identifier Registry (EIDR). This metadata system will issue unique identifiers recognized across the entertainment industry to newly registered content based on basic metadata, like titles, genre, etc.

Using an industry-standardized ID will make new content more organized and easily discoverable by stations.

This is the second part of a series of documents concerning the metadata system. The other documents in this series are as follows:

- **Part 1: Introduction to Abstract Metadata in Public Broadcasting**
  This part introduces the Interconnection System and the need for an abstract hierarchical metadata system.
- **Part 3: Time-based Descriptive Metadata**
  This part explains the metadata that provides detailed information about media content.

This document describes the core abstract metadata models that compose the content library.

Document Organization

This document is organized as follows:

1. **Core Content Metadata**: This section lists the relational, library, and technical metadata that need to be inputted about core content — like a Series, Episode, etc. — at particular times in their lifecycles.
2. **Core Business Metadata**: This section covers what metadata should be documented about Associated Organizations and other specific people who deserve Credits.
3. **Metadata Field Reference**: This section defines each metadata field mentioned in the previous sections with examples, detailed formats, and external resources.

Status

The IXS metadata system is designed as a centralized database with near-real-time updates, with an exposed API that station traffic systems can use to consume data directly or create interfaces as is needed. The API allows any system to query metadata to find content.

Standards and Conventions

Standards referenced in this document include the following:

- **EIDR**, Technical Documentation. http://eidr.org/technology/.
  - Data Fields Reference, December 2015. http://eidr.org/documents/EIDR_2.0_Data_Fields.pdf
  - EIDR ID FORMAT, v1.3, July 2015. http://eidr.org/documents/EIDR_ID_Format_v1.3.pdf
- **TR-META**, MovieLabs Specifications and Standards. http://www.movielabs.com/md/md/.
  - **TR-META-CM**, Common Metadata, v2.4, Movie Labs, October 2015. http://www.movielabs.com/md/md/v2.4/Common_Metadata_v2.4.pdf
- **RFC5646**, Tags for Identifying Languages, IETF, September 2009. https://tools.ietf.org/html/rfc5646
- **RFC3339**, Date and Time on the Internet: Timestamps, IETF, July 2002. https://tools.ietf.org/html/rfc3339
- **ISO8601**, Date and Time Format, International Organization for Standardization. http://www.iso.org/iso/home/standards/iso8601.htm
  - ISO8601-2004, Data elements and interchange formats -- Information interchange -- Representation of dates and times. http://www.iso.org/iso/catalogue_detail?csnumber=40874
- **ISO3166-1**, Codes for the representation of names of countries and their subdivisions -- Part 1: Country codes, 2007. http://www.iso.org/iso/home/store/catalogue_tc/catalogue_detail.htm?csnumber=63545
- **ISO3166-2**, Codes for the representation of names of countries and their subdivisions -- Part 2: Country subdivision code. http://www.iso.org/iso/home/store/catalogue_tc/catalogue_detail.htm?csnumber=63546

2 Core Content Metadata

This section defines the core metadata concepts of broadcasted content. The main metadata concepts include the following:

Abstract Concepts:

- Franchise (optional)
- Series
- Season
- Sub Series (optional)
- Sub Season (optional)
- Episode
- Story

Curated Compilations

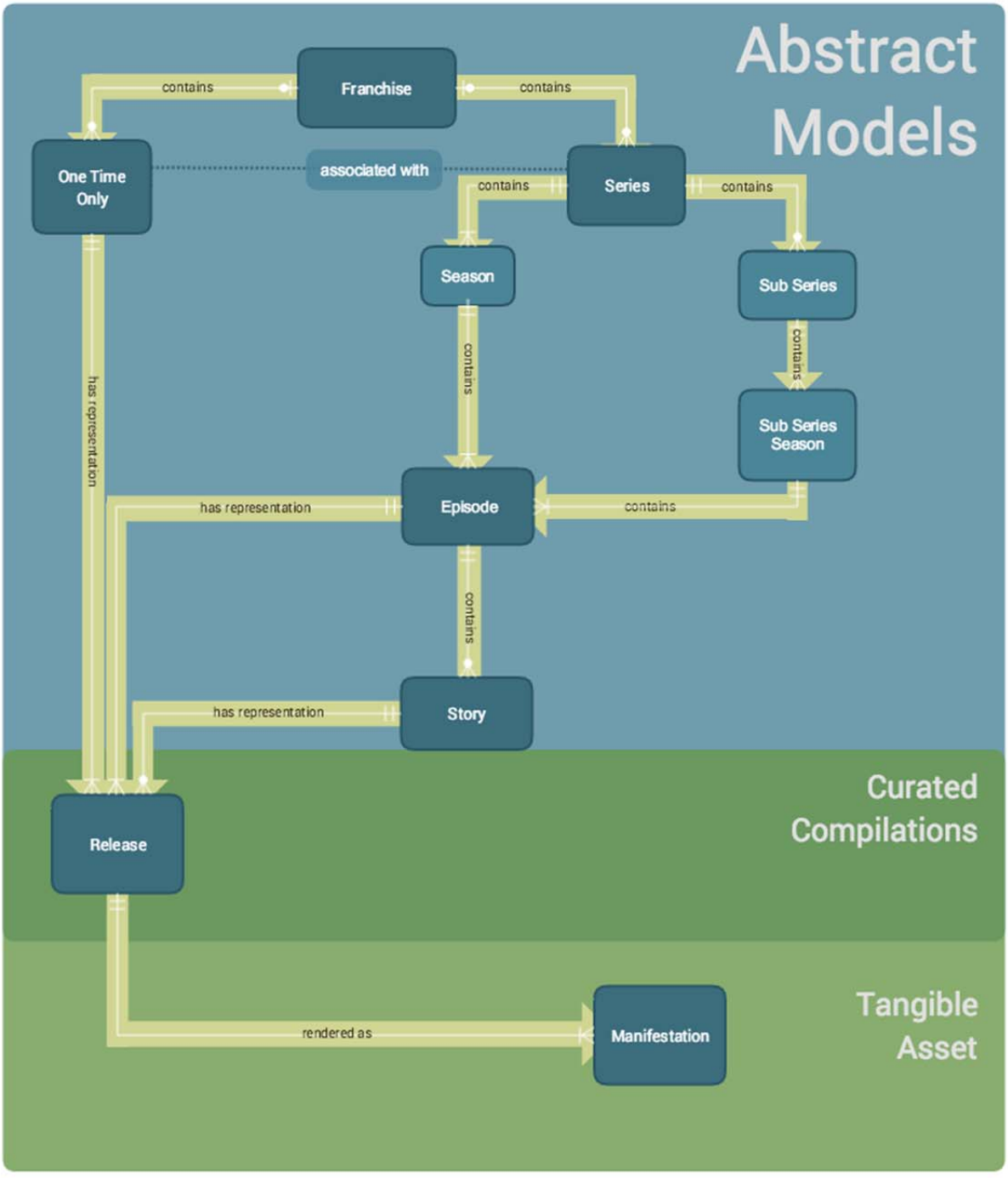
- Release

Tangible Assets

- Manifestation

Figure 1 shows an overview of the core abstract broadcasting concepts and their relationships to each other.

Figure 1: Overview of Abstract Content Models



This diagram shows the hierarchy of abstract content models. Child content models may inherit some metadata values from a parent content model. For example, a Series and an One Time Only from the same Franchise may share metadata with each other.

Franchise

This element is an optional super container for special circumstances, such as the following:

- Grouping a large body of work (e.g. Ken Burns Collection)

- Aggregating content for an anthology Series (e.g. Masterpiece)

How a Franchise Relates to Other Concepts

As represented in Figure 1, a Franchise has the following relationships with other concepts:

Relationship	Description
contains	A Franchise contains a Series and sometimes a One Time Only show. This means a Series and its descendants will inherit metadata from the Franchise. This also means that a Series and a One Time Only may share metadata values.

Franchise Metadata Fields

The metadata fields associated with a Franchise include the fields detailed in the following table:

Lifecycle Phase	Metadata Fields
During Acquisition	UID Title TitleSortable Title60 Title256 Slug Synopsis100 Synopsis400 Synopsis4000 Organizations
During Distribution	Aliases

Series

For most programs, the Series is the name of the program and contains multiple Seasons that run for several Episodes. Examples of a Series include NOVA, Nature, PBS Newshour.

Series Relational Metadata

As represented in Figure 1, a Series has the following relationships with other concepts:

Relationship	Description
partOf	A Series can be part of a Franchise, in association with another Series or an One Time Only show. This means a Series and its descendants may inherit metadata from the Franchise. A Series and its descendants may share metadata with another Series or an One Time Only of the same Franchise.
contains	A Series typically contains one or more Seasons. This means that Seasons and their descendants may inherit metadata from the Series container.

Series Metadata in the Content Lifecycle

This section lists the core library, technical, and business metadata associated with a Series and what part of the Lifecycle Phase by which their values should be documented.

Lifecycle Phase	**Metadata Fields**
During Acquisition	UID Title TitleSortable Title60 Title256 Alternate Titles Slug Synopsis100 Synopsis400 Synopsis4000 e/i Categories Parental Rating Primary Language Secondary Language(s) Country of Origin Primary Genre Secondary Genre(s)

	Program Format Typical HD Level Runtime Organization
During Production	EIDR ID Typical Audio Type Typical Screen Format Credits
During Distribution	Program Identifier Original Release Year Aliases

Season

A Season is a container for Episodes. Seasons are created by the Producer or Distributor.

Season Relational Metadata

As represented in Figure 1, a Season has the following relationships with other concepts:

Relationship	Description
isPart	A Season is part of a Series. This means that the Season may inherit metadata from the Series.
contains	A Season contains Episodes. This means that a Season may pass metadata onto Episodes.

Season Metadata in the Content Lifecycle

A Season is mainly a wrapper used to delineate a group of Episodes by production year or order.

Lifecycle Phase	Metadata Fields
During Acquisition	UID Season Name
During Production	EIDR ID

During Distribution	Aliases

Sub Series

A Sub Series is an optional metadata wrapper to group episodic content within a Series.

A Sub Series can be used to arrange Episodes in a particular order inside a regular Season, such as NOVA's "Making Stuff" sub series inside NOVA Season 2011.

A Sub Series can also contain Episodes inside a Series but outside of a regular Season. An example is the Antiques Roadshow Tucson Hours 1, 2, 3, which is a Series of Episodes that are part of the Antiques Roadshow Series but not within a typical Antiques Roadshow Season.

Lifecycle Phase	Metadata Fields
During Acquisition	UID Synopsis100 Synopsis400 Synopsis4000 Title TitleSortable Title60 Title256 Organizations
During Production	EIDR ID
During Distribution	Aliases

Sub Series Season

A Sub Series Season is an optional metadata wrapper to manage Sub Series Episodes that span Seasons. For Example, the Antique Roadshow: Austin, TX (2015) is a Sub Series Season separate from Antiques Roadshow: Austin, TX (2011).

Lifecycle Phase	Metadata Fields
During Acquisition	UID Season Name
During Production	EIDR ID
During Distribution	Aliases

Episode

An Episode represents `the essence of the program`.

Relational Metadata for an Episode

As represented in Figure 1, an Episode has the following relationships with other concepts:

Relationship	Description
partOf	An Episode is contained by a Season or Sub Series Season, which are in turn contained by a Sub Series or Series. This means that an Episode may inherit metadata from the Series, Sub Series, Season, or Sub Series Season.
contains	Episodes `optionally` contain Stories. This means that Stories may inherit Episode metadata.

Episode Lifecycle Metadata

Each Episode of a Season or Sub Series Season has its own unique metadata, as well as metadata inherited from the Series container. All inherited information can be overridden at this level.

Lifecycle Phase	Descriptive Metadata
Entering Acquisition	Primary Language Secondary Language(s) Country of Origin Primary Genre

	Secondary Genre(s) Program Format Categories Parental Rating Content Descriptors Sensitive Language Sensitive Material Runtime HD Level Organizations
During Acquisition	UID
Entering Production	Title TitleSortable Title60 Title256 AlternateTitles Slug Synopsis100 Synopsis400 Synopsis4000
During Production	EIDR ID Broadcast Premiere Date Open Subtitles Open Subtitles Language Closed Subtitles Closed Subtitles Language Credits
Distribution	Episode Identifier Aliases

One Time Only (OTO)

An One Time Only (OTO) is a collection of essence formatted into a standalone show that is not part of a Series.

Relational Metadata for an OTO

As represented in Figure 1, an OTO has the following relationships:

Relationship	Description
partOf	An OTO may be contained by a Franchise. This means that an OTO inherits metadata from a Franchise.
IsAssociatedWith	An OTO may be related to a Series as a special standalone "Episode" outside of a regular Season. This means an OTO may share metadata with a Series.
contains	An OTO contains Stories. This means that the Stories contained within an OTO may inherit some metadata.
	An OTO may also be completely independent from a Series or Franchise. This means that all of its metadata is unique.

OTO Lifecycle Metadata

This section describes the core data fields associated with an OTO.

An OTO contains metadata fields very similar to an Episode. However, because an OTO is not part of a Series, it will not have any inherited values.

Lifecycle Phase	Descriptive Metadata
During Acquisition	UID Title TitleSortable Title60 Title256 AlternateTitles Slug Synopsis100 Synopsis400 Synopsis4000 e/i

	Categories Parental Rating Content Descriptors Primary Language Secondary Language(s) Country of Origin Primary Genre Secondary Genre(s) Program Format Sensitive Language Sensitive Material Runtime HD Level Organizations
During Production	EIDR ID Broadcast Premiere Date Audio Type Screen Format Open Subtitles Language Closed Subtitles Language Credits
During Distribution	Program Identifier Episode Identifier Aliases

Story

A Story is a standalone piece of content within an Episode or OTO. It could be published by itself to represent a stand-alone piece of content but would not represent full long-form content.

For example, an Episode of Clifford typically includes two 12-minute Stories with an interstitial between the two Stories.

Relational Metadata for a Story

As represented in Figure 1, a story has the following relationships:

Relationship	Description
partOf	A Story may be contained by an Episode or OTO.

	This means that a Story may inherit metadata from its containing Episode or OTO.

Story Metadata

This section describes the core data fields associated with a Story.

Stories do not undergo the entire content lifecycle, like a Series or Episode. Rather, it is created during Production.

Lifecycle Phase	Descriptive Metadata
Entering Production	Primary Language Secondary Language(s) Country of Origin Primary Genre Secondary Genre(s) Categories Content Descriptors Runtime Sensitive Language Sensitive Material Duration HD Level
During Production	UID Title TitleSortable Title60 Title256 Slug Synopsis100 Synopsis400 Synopsis4000 EIDR ID Credits
Distribution	Aliases

Release

A Release is a version of an Episode or OTO created for a specific purpose or time. It is created during Production and does not undergo the entire Content Lifecycle.

Relational Metadata

Releases have the following relationship:

Relationship	Description
rendered as	Releases can be rendered as multiple Manifestations.

Release Metadata

The following metadata fields need to be defined when a Release is created at the end of Production:

- Aliases
- Duration
- EIDR ID
- Published
- Release Identifier
- Release Description
- Release Domain
- Release Type
- Revision Identifier
- Revision Description
- Synopsis100
- Synopsis400
- Synopsis4000
- UID

Manifestation

Manifestations are the actual video renderings of a Release. Each Manifestation derived from the same Release has identical content but different technical details.

The metadata for a Manifestation should be defined during the Production stage and before the beginning of the Distribution stage.

Core Manifestation Metadata

The following metadata fields help identify the Manifestation container as a whole:

- EIDR ID

- EIDR Manifestation Class
- HD Level
- Duration
- Size
- Container Type
- Container Subtype
- Repair Identifier
- Repair Description
- MD5
- Aliases

Video Track Metadata

The following technical metadata fields describe the video tracks contained in the manifestation.

- Screen Format (inherited from Series)
- Video Track Type
- Video Codec
- Video Codec Type
- MPEGProfile
- MPEGLevel
- Video Bitrate Max
- Video Bitrate Average
- Vertical Resolution
- Horizontal Resolution
- Scan Type
- Frame Rate
- Field Dominance
- Chroma Subsampling
- Color Space
- Closed Subtitles
- Closed Subtitles Language
- Start of Message
- End of Message

Audio Track Metadata

The following technical metadata fields describe the audio tracks contained in the Manifestation.

- Primary Language (inherited from Series)
- Track Reference
- Track Description
- Audio Track Type
- Audio Dubbed
- Audio Codec

- Audio Codec Type
- Audio Bitrate Max
- Audio Bitrate Average
- Audio Bit Depth
- Variable Bit Rate
- Sample Rate
- Channels
- Channel #
- Loudness
- Dial Norm
- Audio Track Duration

Subtitle Metadata

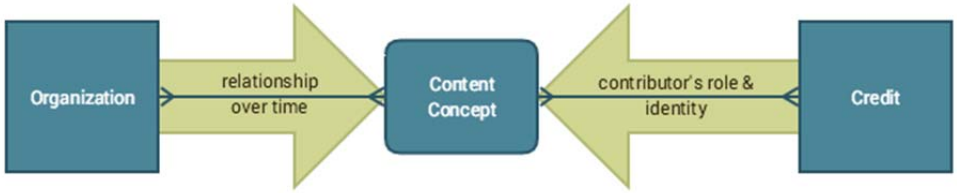
- Track Description
- Subtitle Type
- Subtitle Format Type
- Subtitle Language
- Track Reference

4 Core Business Metadata

Organizations and Credits are business concepts that can be associated with each of the content concepts described in Section 2: Core Content Metadata.

The basic relationships between Organizations, Credits, and the abstract content models are described in Figure 2.

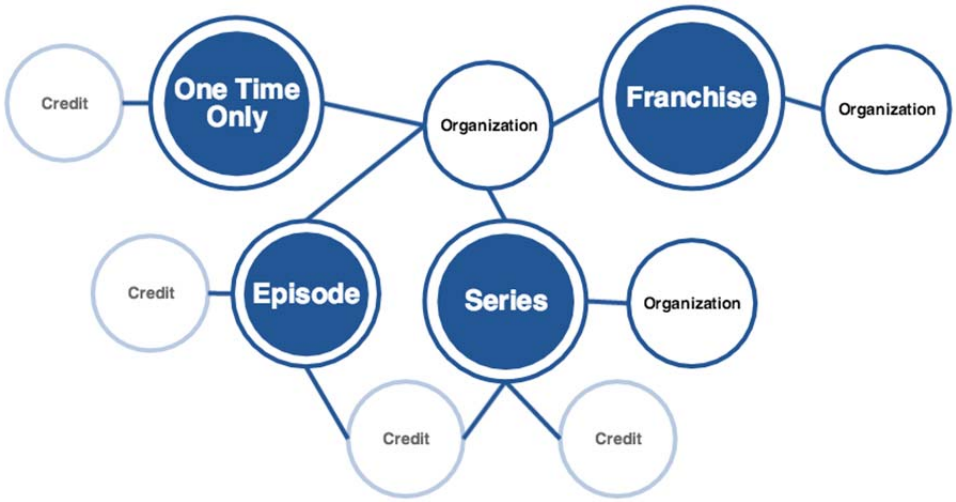
Figure 2: Metadata Relationships Between Organizations, Content, and Credits



Once registered, an Organization contains metadata describing its relationship to a content concept over time. Similarly, a Credit contains metadata that identifies a contributor and describes the contributor's specific role in relation to a content.

Many Organizations and Credits can be associated with one content concept, and Organizations can be associated with various pieces of content. Figure 3 illustrates these possibilities.

Figure 3: Example of Multiple Content Concepts, Organizations, and Credits



This web illustrates how Credits and Organizations can be associated with multiple content concepts.

Organizations

An Organization is some business entity with a role in content distribution, such as a Distributor, a Station, or a Producer. These Organizations may have some information like contact information and mailing addresses that would be useful to associate with content for legal reasons and planning.

Relationship	Description
isAssociatedWith	Organization metadata should be associated with a Franchise, Series, Sub Series, Episode, or OTO.

Setup for New Organization

During the Acquisition phase, a PBS system administrator should register new Organizations in the metadata system with the following information:

Field	Description	Format
Name	Name of the Organization. **Example:** PBS	String (100)
Description	A description of the business entity. **Example**: The Public Broadcasting Service is an American public broadcaster and television program distributor.	String (500)
Address	The physical address of the Organization's headquarters.	
Phone	The best phone number to reach the Organization.	
Prefix Codes	A list of short (up to 4 characters) codes that represent this entity. Used for compact display purposes and/or other string representations such as filenames, etc. Allowed characters are A-Z and case insensitive.	String (4)

After registration, the Organization will receive an EIDR ID, which should be added to their profile.

Adding Organization Metadata During Acquisition

Once an Organization has a profile with an EIDR ID, it can be associated to content with the following metadata fields:

Field	Description	Accepted Values
Role	This captures the Organization's role.	Distributor Primary Producer Secondary Producer Primary Presenter Secondary Presenter Copyright Holder Post-Production Other
Start Date	The beginning of the Organization's Role in the associated project.	Date and timestamp in the following format: `YYYY-MM-DD HH:MM:SS`
End Date	The end of the Organization's Role in the associated project. A null end timestamp represents a current relationship.	Date and timestamp in the following format: `YYYY-MM-DD HH:MM:SS`
Published	This field captures the status of a curated compilation or tangible asset that does not undergo the typical content lifecycle.	• Unpublished Default value. Content is only visible to the creator. • Published Content is visible to other entities on the network. • Invalid A blocking issue prevented the content from reaching publication.

Credits

Credits tell the identity and role of someone who contributed to a Series, Episode, OTO, or Story. For example: Directed by John Smith.

Some Credit metadata may be known at the Acquisition stage, such as Producer and Director. However, a lot of the Credit metadata cannot be filled in for sure until the end of Production.

Multiple Credits can be added to a piece of content.

Relational Metadata for Credits

As seen in the Overview of Core Business Metadata Concepts, Credits have the following relationships with other concepts:

Relationship	Description
ContributedTo	Series, Episode, OTO, or Story

Setup for New Credit

Field	Description	Accepted Values
UID	Unique ID for internal tracking purposes during the Acquisition phase.	See UID
DisplayName	The accredited person's name for display purposes.	String (500)
FirstGivenName	The person's first name.	String (500)
SecondGivenName	The person's second name.	String (500)
FamilyName	The person's family name.	String (500)
Suffix	e.g. Jr., J.D., etc.	String (500)
Prefix	Miss, Mrs. Mr.	String (500)

Adding Credit Metadata During Acquisition

Once the individual's profile is set up, it can be associated to content with the following metadata fields:

Field	Description	Accepted Values
Role	This field captures the type of contribution that the person made.	Producer Director Cast Crew Writer
Aliases	A Credit may have have different identifiers	See Aliases

	from external repositories, such as IMDB.	

5 Reference for Field Values

This section contains descriptions, examples, and accepted values or formats for the metadata fields in each model.

Aliases

This field captures a list or array of system identifiers used by external entities. These are identifiers for a Franchise, Series, OTO, Season, etc. that are not an EIDR ID or internal PBS identifiers like the UID or Package Number.

External companies that may use different identifiers include the following:

- iTunes
- Netflix
- Amazon
- Station Call Sign (e.g. WNET, WXXI, KOPB)
- PBS
- APT
- NETA

Applicable to: Series, Season, Episode, OTO, Story, Release, Manifestation, Franchise, Sub Series, Sub Series Season, Credits

Alternate Titles

If a content's Title changes, deprecated titles are captured in this field. Deprecated titles include working titles used during the Acquisition phase.

Applicable to: Series, Episode, OTO

Audio Codec

The field captures the name of the supported codec.

For a list of the supported types, see the most recent TR-META-CM.

Example: AAC, AC-3, ALAC, DOLBY-TRUEHD, DST, MPEG1, MPEG4-ALS, MP3, WAV, WMA

Applicable to: Manifestation

Audio Codec Type

This field captures the formal reference identification of the Audio Codec.

For a list of the supported types, see the most recent TR-META-CM.

Format: <namespace>:codectype

Example: mpegra:ac-3

Applicable to: Manifestation

Audio Bit Depth

This field captures the number of bits per audio sample.

Format: Positive integer

Example: If an audio sample has 16 bits, then Audio Bit Depth=16.

Applicable to: Manifestation

Audio Bitrate Average

Bitrate averaged over the entire audio track in kilobits/second.

Format: Positive integer

Example: If the bitrate average is 36 kilobits/second, the Audio Bitrate Average=36.

Applicable to: Manifestation

Audio Bitrate Max

Peak bitrate (kilobits/second) averaged over a short period.

Format: Positive integer

Example: If the bitrate max is 96 kilobits/second, the Audio Bitrate Max=96.

Applicable to: Manifestation

Audio Dubbed

This boolean field indicates if the spoken language in an audio track is dubbed.

Format: True/False

Applicable to: Manifestation

Audio Track Duration

Length of an audio track according to the ISO 8601:2004 standard format for time duration.

See IETF RFC 3339 for a full description of this format.

Format: HH:MM:SS

Applicable to: Manifestation

Audio Track Type

This field describes the purpose of the track in a controlled vocabulary.

Accepted values are as follows:

- Primary
- Narration
- Video Descriptive Service

Applicable to: Manifestation

Audio Type

The actual Audio Type for an Episode, OTO, or Manifestation. The default value is inherited from the Typical Audio Type at the Series level.

The accepted values are as follows:

- Stereo
- Surround 5.1
- Surround 7.1

- AC-3
- TrueHD
- Mono
- *none*

Applicable to: Episode, OTO

Broadcast Premiere Date

The date that an Episode is aired live. This may need to be adjusted from the expected value to the actual value upon distribution.

See the IETF RFC 3339 for information on the full-date format.

Format: *YYYY-MM-DDTHH:MM:SSZ*

Example: The Broadcast Premiere Date for an Episode aired at 4PM on December 24, 1999 is 1999-12-24T16:00:00Z.

Applicable to: Episode, OTO

Categories

A comma-delineated list of topics that can be associated with the content.

Format: 5000 characters

Example: Women, Youth, Social Services

Applicable to: Series, Episode, OTO, Story

Channels

This field captures the total number of channels in an audio track.

Format: Integer

Example: If an audio track has 16 channels, then Channels=16.

For each channel, there should be a sequentially numbered field to describe the audio channel's content. For two examples of these channel descriptions, see Channel #.

Applicable to: Manifestation

Channel

This field captures the content on a particular channel in the audio track, identified by a number (#).

Accepted values include the following:

- Left
- Right
- Center
- LFE
- Left Surround
- Right Surround
- Descriptive Video Service
- Alt Language: <language code>
- Unused
- DVI

This is a field that can be duplicated for however many Channels exist in an audio track. The # in the field label would change to identify a new Channel.

Example: Channel 1=Right, Channel 2=Left, Channel 3=Center

Applicable to: Manifestation

Chroma Subsampling

The luma and color channel sampling rate in an encoded video should be specified using the $Y'C_BC_R$ model.

Accepted values are as follows:

- 4:1:1
- 4:2:0
- 4:2:2
- 4:4:4

Applicable to: Manifestation

Closed Subtitles

Closed subtitles can be toggled on or off by the display device. This field indicates whether or not closed captions exist in the video.

Format: Yes/No

Applicable to: Episode, OTO, Manifestation

Closed Subtitles Language

This specifies the language of the closed subtitles. See Primary Language for accepted language subtags.

A Description flag can be added to the language code to indicate that the subtitle includes descriptions of sound effects for the deaf or hard of hearing. If the flag is not there, then the subtitles only transcribe dialogue lines.

Format: *Language Code, Language Code* - Description

Example: *Downton Abbey* includes Open Subtitles for English, English with Descriptions, Spanish, and French. The accepted values for this include en, sp, fr, and en - description.

Applicable to: Episode, OTO, Manifestation

Color Space

This field indicates the Colorimetry Encoding for a video track. Accepted values include the following standards:

- 601
  ITU Recommendation BT.601, Studio encoding parameters of digital television for standard 4:3 and widescreen 16:9 aspect ratios
- 709
  ITU Recommendation BT.709, Parameter values for the HDTV standards for production and international programme exchange.
- 2020
  ITU Recommendation BT.2020, Parameter values for ultra-high definition television systems for production and international programme exchange.

Format: 601, 709, 2020

Applicable to: Manifestation

Container Subtype

This field captures subtypes of a specified Container Type. See the most recent version of TR-META-CM for values.

Example: Operational Pattern 1a (OP1a) is a subtype of MXF.

Applicable to: Manifestation

Container Type

The container is a file comprised of a metadata wrapper and a subcontainer holding audio, video, subtitle, or image tracks.

Some acceptable values include the following:

- 3GP
- AC
- AV
- DIVX
- DTS
- FLV
- ISO
- JPEG
- M4V
- MP4
- MPEG
- OGG
- MOV
- other

For a complete list of values, see the most recent version of TR-META-CM for Container Type encoding.

Some Container Types may have a Subtype that can be specified further.

Applicable to: Manifestation

Content Descriptors

Content descriptors warn about specific thematic elements associated with the content. This field can be used to designate up to four content descriptors.

- D Suggestive Language
- L Coarse or crude language
- S Sexual situations
- V Violence
- FV Fantasy Violence (exclusive to TV-Y7)

Each piece of rated content may have up to four content descriptors along with their Parental Rating. By default, none are selected.

For more information, see the TV Parental Guidelines.

Applicable to: Episode, OTO, Story

Country of Origin

This field specifies the country from which the Title originates. The country code format should be from the ISO 3166-1 alpha 2 codes (Wikipedia has an easy-to-access list here).

Example: The country of origin for CRTC Recording productions is Canada, so the ISO 3166-1 alpha 2 code is CA.

Applicable to: Series, Episode, OTO, Release, Story

Dialnorm

Dialnorm stands for dialog normalization. It is an integer from 1-31 corresponding with a playback gain of -30 to 0 dB, where higher values allow for audio levels to have high peaks.

Format: dBFS

Example: -24dBFS

Applicable to: Manifestation

Distributor

The code of the Entity offering the content. Values are expressed as one to four alphabetic characters.

Examples: "PBS", "BBC", "APT", "WDSE"

Applicable to: Series, OTO

Duration

The frame-accurate time length of a piece of content expressed as a SMPTE timecode.

Format: Timecode

Example: 01:27:59;05

Applicable to: Release, Manifestation

e/i

This boolean field flags content as educational and informational (e/i) or not.

If True, then content is e/i.

If False, then content is not e/i.

Applicable to: Series**,** OTO

EIDR ID

Unique ID generated when the required metadata has been submitted into the MDM during the Acquisition phase. Every concept (Series, Season, Episode, etc.) has its own unique EIDR ID.

Format: 10.*EIDR*/*XXXX-XXXX-XXXX-XXXX-XXXX-C*

Where:

- 10 indicates that the ID string is a Digital Object Identifier (DOI)
- *EIDR* is a 4-integer string assigned by EIDR
- *XXXX-XXXX-XXXX-XXXX-XXXX-C* is the unique DOI suffix for the item

See the EIDR ID FORMAT documentation for more details on the format.

Example:

- The EIDR ID for *Downton Abbey* the Series is 10.5240/5BC6-2FA3-4F64-B17E-0B7D-H
- The EIDR ID for *Downton Abbey: Season 1* is 10.5240/FAEB-23C2-EB99-E52C-C356-T
- The EIDR ID for *Downton Abbey: Season 1: Episode 1* is 10.5240/12DC-C92E-76A3-8706-BC6D-I
- The EIDR ID of an Release of *Downton Abbey: Season 1: Episode 1* is 10.5240/B268-3C0B-107A-2E60-1271-R
- The EIDR ID for *Downton Abbey Rediscovered* (an OTO) is 10.5240/8AF4-86FD-E4D7-2A8B-7079-S

Applicable to: Series, Season, Episode, OTO, Story, Release, Manifestation, Franchise, Sub Series, Sub Series Season, Organizations

EIDR Manifestation Class

This field captures the purpose of a rendered Manifestation. The rendering may have been made for a specific platform (e.g. game machine, mobile, etc.), a file type (e.g. master, mezzanine), a definition setting (HD, SD), or a new dub (version language).

The acceptable values are as follows:

- Version Language
- Video on Demand (VOD)
- Electronic Sell-Through (EST)
- Mobile
- Web
- Master
- Mezzanine
- Proxy
- Screener
- DVD
- Blu-ray
- HD
- SD
- UHD
- Other

See the "Manifestation Class Details" section in the most recent EIDR Data Fields Reference for more information.

Applicable to: Manifestation

End of Message

The media timecode for the end of the media.

Format: Timecode

Applicable to: Manifestation: Video

Episode Identifier

The episode identifier string used to identify Episodes. For BroadView distributed content, this represents the NOLA Episode string.

Format: String (6) of integers

Example: 101

Applicable to: Episode, OTO

Field Dominance

This boolean field only needs to be filled if the Scan Type=Interlaced. This value specifies whether the first frame of picture is Field 1 or Field 2.

Format: 1 / 2

Applicable to: Manifestation

Frame Rate

This field only needs to be filled if the Scan Type=Interlaced. This value specifies the frame rate of the video scanning system.

Applicable to: Manifestation

HD Level

This is the actual HD Level for an Episode, Release, or Manifestation. The default value is inherited from the Typical HD Level at the Series Level.

Acceptable values include the following:

- 1080i

- 1080p
- 720i
- 720p
- 480i
- 480p
- Unknown
- None

Applicable to: Series, Episode, OTO, Manifestation, Story

Horizontal Resolution

This specifies the total number of pixel columns in the active portion of a frame in the video pixel matrix.

Accepted values include the following:

- 3840
- 1920
- 1280
- 720
- 640

Applicable to: Manifestation

Loudness

The loudness of an audio track in decibels.

Standard: LKFS

Applicable to: Manifestation

MD5

The MD5 message-digest algorithm is a cryptographic hash function used to verify the data integrity of a Manifestation. It is expressed as a 32-digit hexadecimal number.

Format: String (32)

Example: 94864ec28716b3be9c3bc1bc7d7edf90

Applicable to: Manifestation

MPEGLevel

The MPEG level if supported by the Video Codec.

See the "Video Encoding Details" section in the most recent EIDR Data Fields Reference for some common values.

Format: String (20) based on ISO/IEC 14496-10 or ISO/IEC 13818-2

Example: For MPEG-4: 4.1

Applicable to: Manifestation

MPEGProfile

The MPEG profile if supported by the Video Codec.

See the "Video Encoding Details" section in the most recent EIDR Data Fields Reference for some common values.

Format: String (20) based on ISO/IEC 14496-10 or ISO/IEC 13818-2

Example: For MPEG-2: Spatial

Applicable to: Manifestation

Open Subtitles

Open subtitles are burned into the video and cannot be removed. This field (Yes/No) indicates whether or not open captions exist in the video. If they do exist, the Open Subtitle Language needs to be specified.

Format: Yes/No

Applicable to: Episode, OTO, Manifestation

Open Subtitles Language

This indicates the language of the open subtitles. See Primary Language for accepted values.

A Description flag can be added to the language code to indicate that the subtitle includes descriptions of sound effects for the deaf or hard of hearing. If the flag is not there, then the subtitles only transcribe dialogue lines.

Format: *Language Code, Language Code* - Description

Example: *Downton Abbey* includes Open Subtitles for English, English with Descriptions, Spanish, and French. The accepted values for this include en, sp, fr, and en - description.

Applicable to: Episode, OTO, Manifestation

OrganizationAddress

The physical address of the Organization's headquarters.

Format: U.S. Mailing system standards

Example: 1111 Mouse Way, Turducken MD 21087

Applicable to: Series, OTO

OrganizationEndDate

The end date of the Organization's role. See ISO 8601 for the standard formatting.

Format: YYYY-MM-DD

OrganizationName

The name of an associated Organization. There may be as many duplicates of this field as is necessary to add all associated Organizations.

Format: String (100)

Applicable to: Series, OTO

OrganizationPhone

The best phone number to reach the Organization.

Format: (country code) area code - etc.

Example: (1) 932-854-1442

Applicable to: Series, Episode, OTO

OrganizationRole

This field describes the role of an associated Organization.

Accepted values include the following:

- Producer
- Broadcaster
- Distributor
- Editor
- Encoding
- Post-production
- Licensor
- Other

Applicable to: Series, OTO

OrganizationStartDate

The beginning of the Organization's role in a project. See ISO 8601 for the standard formatting.

Format: YYYY-MM-DD

Applicable to: Series, OTO

Original Release Year

The original release year for a Series.

Format: Year timestamp (YYYY)

Applicable to: Series

Parental Rating

Parental Rating codes indicate the recommended age group for this content. The following values

are acceptable:

- TV-Y
- TV-Y7
- TV-G
- TV-PG
- TV-14
- TV-M

For more information, see the TV Parental Guidelines.

Applicable to: Series, Episode, OTO, Release

Primary Genre

The primary genre of a piece of content. Genre refers to the kind of content a show will have. Only one option should be chosen from the following acceptable values:

- Arts
- Children's
- Cultural
- Drama
- Educational
- How-To
- History
- Outreach
- Public Affairs
- Science/Nature
- Self-Help
- Sports

Applicable to: Series, Episode, OTO, Story

Primary Language

A subtag indicating the primary language spoken in the original production.

An acceptable value for this field includes the following language code standards:

- a two-letter code from ISO 639-1 (2002)
- a three-letter code from ISO 639-2 (1998), ISO 639-3 (2007) or ISO 639-5 (2008)

Some common values are as follows:

- en (English)
- sp (Spanish)
- de (German)
- fr (French)
- ja (Japanese)

See the IETF RF 5646 for more information on using language subtags.

Applicable to: Series, Episode, OTO, Release, Story, Manifestation: Audio, Manifestation: Subtitle

Program Format

This field describes what formatting conventions will be used to present the show's content. The following values are acceptable:

- Children's Live/Animation
- Demonstration/ Instructional
- Documentary
- Event Coverage
- Feature Film/ Video Drama
- Interview/ Discussion/Review
- Magazine
- News
- Other
- Performance

Applicable to: Series, Episode, OTO

Program Identifier

The root code for the program identifier. For BroadView distributed content, this represents the NOLA Root

Format: Four character string of capital characters.

Example: MAST for Masterpiece, ANRO for Antiques Roadshow, NOVA for NOVA

Applicable to: Series, OTO

Published

This field captures the status of a curated compilation or tangible asset that does not undergo the typical content lifecycle.

The accepted values are as follows:

- Unpublished
  Default value. Content is only visible to the creator.
- Published
  Content is visible to other entities on the network.
- Invalid
  A blocking issue prevented the content from reaching publication.

Applicable to: Release, Manifestation, Organizations

Applicable to: Release

Release Description

Detailed description of a Release's content.

Applicable to: Release

Release Domain

The Entity Code of the agency assigning the Release Identifier. Value is limited to the four character alphanumeric code assigned by the sIX office to eligible metadata submission entities.

Examples: "P", "KFME", "NETA"

Applicable to: Release

Release Identifier

An identifier that represents the unique number for a release. In the general sense, this should be a string that is unique to a given Entity's Release Domain code. For purposes of content exchange, it is recommended this is represented as a six digit integer.

Example: "465967"

Release Type

This field identifies the type of Release from a set list.

The accepted values are as follows:

- Base
- Unedited
- Stacked
- Embedded Promo
- Reinvented Breaks
- Pledge
- Pledge Event
- Promo
- Short
- Spot
- Evergreen

Applicable to: Release

Repair Description

Detailed description of the repair of any technical impairment that was corrected.

Applicable to: Manifestation

Repair Identifier

An identifier that represents number of repair transcoding. In the case where the original transcode has a technical impairment, then subsequent transcodes can be created to fix the technical impairment. These subsequent transcodes can then be labeled using this field using monotonically increasing integers. The default value of this field shall be zero. The maximum value shall be 9.

Example: "1"

Applicable to: Manifestation

Revision Description

This field allows for detailed documentation as to why a particular revision exists.

Format: String (5000)

Example: [*Revision Identifier*] contains a new promo different from previous Releases.

Applicable to: Release

Revision Identifier

This field is a simple numerical identifier for a Release to keep track of versions. The default value of this field shall be "1". The maximum value shall be 999.

Format: Integer

Example: "3"

Applicable to: Release

Runtime

The time length of a piece of content in minutes.

If this value is set at the Series level, it determines the default Runtime for newly produced Episodes. This value only represents the Runtime for current and future Episodes and does not need to capture the HD Level for past Episodes.

If the value of the Episode differs from this expected value, the actual value may be set at the Episode level. Otherwise, the default value is inherited from the Runtime field at the Series level.

Example: An episode of *Downton Abbey* runs 60 minutes, so Runtime=60.

Applicable to: Series, Episode, OTO, Story

Sample Rate

Samples per second in kilobits/second.

Format: Positive integer

Example: Sample Rate=44

Applicable to: Manifestation

Scan Type

This field specifies whether the Scan Type for a Video Track is Progressive or Interlaced.

Format: Progressive/Interlaced

If Interlaced, the Frame Rate and Field Dominance must be specified.

Applicable to: Manifestation

Screen Format

The actual screen format for a rendered video. The default value is inherited from Typical Screen Format at the Series level but should be adjusted accordingly.

Format: IAR,CAR PF

Where:

- IAR is the Image Aspect Ratio. It defines the pixel ratio of the width vs the height.
- CAR is the Content Aspect Ratio. This can be omitted if identical to IAR. It describes the video content (i.e. non-black padding) in a numerical ratio of the width and height.
- PF is picture format which describes any additional manipulation or padding applied to the content to ensure content fits within the image. It is a term: Full Frame, Pillarbox, Letterbox, Other.

Acceptable Screen Formats are as follows:

- 16:9 Full Frame
- 16:9, 4:3 Pillarbox
- 16:9, 14:9 Pillarbox
- 16:9 Other
- 4:3 Full Frame
- 4:3, 16:9 Letterbox
- 4:3, 14:9 Letterbox
- 4:3 Other

Applicable to: Series, Episode, OTO, Manifestation

Season Name

A Season's name is formatted as a string followed by an ordering integer.

- The string can be the unique name or it can be the word Season.
- The ordering integer is a numerical identifier for a Season. They can be ordered by cardinal numbers, typically formatted as 100 for Season 1, then 200 for Season 2, etc. A Season with daily Episodes may be formatted as 1000, 2000, etc. Seasons also can also be ordered by year (1999, 2000, 2001...).

The first Season should be numbered 1, and if it is the only one in the sequence, it is numbered 1.

Format: String+Ordering Integer

Example: *Downton Abbey: Season 2*

Applicable to: Season, Sub Series Season

Secondary Closed Subtitle Language(s)

This field captures any secondary languages for Closed Subtitles. See Primary Language for accepted values.

Applicable to: Episode, OTO, Manifestation

Secondary Genre(s)

Any secondary genres that could apply. Multiple secondary genres can be selected for a Series. See Primary Genre for accepted values.

Applicable to: Series, Episode, OTO, Story

Secondary Language(s)

This field lists the secondary languages, if any, used in the original production. See Primary Language for accepted values.

Applicable to: Series, Episode, OTO, Story, Release

Secondary Open Subtitle Language(s)

This field captures any secondary languages for Open Subtitles. See Primary Language for accepted values.

Applicable to: Episode, OTO, Manifestation

Sensitive Language

This boolean field indicates whether or not the video contains language considered sensitive by the FCC. The default value is No.

Format: Yes/No

Applicable to: Episode, OTO, Story

Sensitive Material

This boolean field indicates whether or not the video contains images considered sensitive by the FCC. The default value is No.

Format: Yes/No

Applicable to: Episode, OTO, Story

Size

This field captures the size of a Manifestation container in bytes.

Format: Integer

Applicable to: Manifestation

Slug

A short, unique term used to identify the show quickly.

Format: String (100)

Example: american-experience, antiques-roadshow

Applicable to: Series, Episode, OTO, Story, Franchise

Start of Message

The media timecode for the start of the content. Note that the timecode of the media is not required to be 00:00:00;00 -- more commonly the Source Media content will begin at 01:00:00;00 to accommodate the preamble. The preamble usually includes a slate, bars, or a countdown.

Format: Timecode

Applicable to: Manifestation: Video

Subtitle Format Type

Acceptable format types include the following values:

- 3GPP
- Blu-Ray
- DCI
- DVB
- DVD
- SMPTE 2052-1 Timed Text
- SCC
- SRT
- TTML
- WEBVTT

For a full description of this subtitle format types, see the values for FormatType Encoding in the most recent version of the TR-META-CM.

Applicable to: Manifestation: Audio, Manifestation: Subtitle

Subtitle Type

This field indicates if the subtitle track is one of the following types:

- Caption
  Transcription of the dialogue.
- SDH
  Descriptive subtitles for the deaf and hard of hearing.
- Other
  Translations or other subtitle text

Applicable to: Manifestation: Subtitle

Synopsis100

Summary of the plot limited to 100 characters for Electronic Program Guides.

Applicable to: Series, Season, Episode, OTO, Story, Release, Franchise, Sub Series

Synopsis400

Summary of the plot limited to 400 characters. If this field is not user-specified, it will inherit the value for Synopsis100.

Applicable to: Series, Episode, OTO, Story, Franchise, Release, Sub Series

Synopsis4000

Summary of the plot limited to 4000 characters. If this field is not user-specified, it will inherit the value for Synopsis400.

Applicable to: Series, Episode, OTO, Story, Franchise, Release, Sub Series

Timecode

Timecode notation is specified using the following form: HHpMMpSSqFF

Where:

- **HH** is the two digit hour representation
- **MM** is the two digit minute representation
- **SS** is the two digit second representation
- **FF** is the two digit frame representation
- **p** is the punctuation delimiter. Generally, this is the colon (:) character but sometimes is represented as a semi-colon (;) to indicate drop-frame timecodes
- **q** is the punctuation delimiter for Frames. For drop-frame timecodes, this shall be the semi-colon character (;). At present, virtually all media at 29.97 FPS and thus uses the drop-frame timecode. In the extremely rare case for non drop-frame timecodes, the colon (:) character should be used.

Examples::

"01:00:00;00" - Represents one hour
"12:38:59;28" - Represents twelve hours, 38 minutes, 59 seconds and 28 frames

Title

The full title for a piece of content.

Format: String limited to 100 characters.

Example: *A Chef's Life*

Applicable to: Series, Season, Episode, OTO, Story, Release, Franchise, Sub Series

TitleSortable

A sortable version of the Title with leading articles moved to the back.

Format: String limited to 100 characters.

Example: *Chef's Life, A*

Applicable to: Series, Season, Episode, OTO, Story, Release, Franchise, Sub Series

Title60

A version of the Title limited to 60 characters for Electronic Program Guides.

Applicable to: Series, Episode, OTO, Story, Release, Franchise, Sub Series

Title256

A version of the Title limited to 256 characters.

Applicable to: Series, Episode, OTO, Release, Franchise, Sub Series, Story

Track Description

This field captures a brief, human-friendly description that helps identify what the track contains.

Format: String (500)

Applicable to: Manifestation: Audio, Manifestation: Subtitle

Track Reference

This field captures an identifying number for an audio or subtitle track container within a Manifestation.

Format: Integer

Applicable to: Manifestation: Audio, Manifestation: Subtitle

Typical Audio Type

This value is set at the Series level and determines the default Audio Type for newly produced Episodes. This value only represents the Audio Type for current and future Episodes and does not

need to capture deprecated values for past Episodes.

The actual value may be set at the Episode level as Audio Type.

Applicable to: Series

Typical HD Level

This value is set at the Series level and determines the default HD Level for newly produced Episodes. This value only represents the HD Level for current and future Episodes and does not need to capture deprecated values for past Episodes.

The actual value may vary at the Episode level and should be adjusted appropriately. See HD Level for acceptable values.

Applicable to: Series

Typical Screen Format

This value is set at the Series level and determines the default Screen Format for newly produced Episodes. This value only represents the Screen Format for current and future Episodes and does not need to capture deprecated values for past Episodes.

The actual value may vary at the Episode, OTO, or Manifestation level as Screen Format. See Screen Format for acceptable values.

Applicable to: Series

Variable Bit Rate

This boolean field indicates if the audio bit rate is variable or constant. If set to Yes, then the audio track contains a variable bit rate.

Format: Yes / No

Applicable to: Manifestation

Vertical Resolution

This specifies the total number of pixels rows in the active portion of a frame in the video pixel matrix.

Accepted values include the following:

- 2160
- 1080
- 720
- 480
- Other

Applicable to: Manifestation

Video Bitrate Average

Expressed in megabits/second and averaged over the entire video track.

Format: Integer

Example: If the bitrate average is 100 megabits/second, the Video Bitrate Average=100.

Applicable to: Manifestation

Video Bitrate Max

Expressed in megabits/second.

Format: Integer

Example: If the bitrate max is 500 megabits/second, the Video Bitrate Max=500.

Applicable to: Manifestation

Video Codec

The name of the video codec used to encode video data.

Some common codecs include the following:

- AVI Uncompressed
- Cineform HD
- DIVX
- DV
- H.264
- JPEG2000
- MOBICLIP
- PHOTOJPEG
- PRORES
- PRORESHQ
- PRORES422
- QT Uncompressed
- REAL
- Spark
- WMV8
- WMV9
- VC1
- VP6
- VP7
- VP8
- XVID

- MPEG1
- MPEG2
- On2
- SVQ
- WMV
- WMV7
- OTHER

For a full list of codecs, see the values for DigitalAssetVideoEncoding-type in the most recent version of the TR-META-CM.

If the Video Codec supports an MPEG profile or level, the MPEGProfile and MPEGLevel need to be specified.

Applicable to: Manifestation

Video Codec Type

The formal reference identification of the codec used in the track.

For a full description of this format, see the values for DigitalAssetVideoEncoding-type in the most recent version of the TR-META-CM.

Example: IANA:video/h264

Applicable to: Manifestation

Video Track Type

This field describes the purpose of the track with controlled vocabulary.

Accepted values include the following:

- Primary This is the default value.
- Overlay This refers to Picture-in-Picture (PiP) or other overlay track, intended for use with a primary track.
- Angle This is an alternate angle track.
- Other For none of the above.

For a full list of codecs, see the values for Video Track Type encoding in the most recent version of the TR-META-CM.

Applicable to: Manifestation

UID

Unique ID for internal tracking purposes during the Acquisition phase.

Format: urn:pbs:content:<uuid>

Example: urn:pbs:content:899310eb-b7cd-4364-acf0-8da105f46966

Applicable to: Franchise, Series, Season, Episode, OTO, Sub Series, Sub Series Season, Story, Release, Credits

Abstract Metadata in Public Broadcasting

Part 3: Time-based Descriptive Metadata

Version 2.1
May 2017

Table of Contents

1 Introduction

This document explains the metadata that provides detailed time-based information about media content. Time-based descriptive metadata is important in a number of ways. It can provide:

- A list of primary elements to traffic personnel to ensure confidence in content placement and programming continuity
- A list of secondary graphical elements to brand/marketing personnel to allow additional placement of graphics or other treatments
- A list of content or chapter breaks to allow logical placement of underwriting or pledge breaks
- Information that can be passed to station playout automation for automated workflows (e.g. 'Channel in a box' and many more workflows)

This is the third part of a series of documents concerning the Interconnection System's metadata system. The other documents in this series are as follows:

- **Part 1: Introduction to Abstract Metadata in Public Broadcasting**
  This part introduces the Interconnection System and the need for an abstract hierarchical metadata system.
- **Part 2: Core Values**
  This part describes the core abstract metadata models that compose the content library.

Timecode Format

Timecode notation is specified using the following form: `HHpMMpSSqFF`

Where:

- **HH** is the two digit hour representation
- **MM** is the two digit minute representation
- **SS** is the two digit second representation
- **FF** is the two digit frame representation
- **p** is the punctuation delimiter. Generally, this is the colon (`:`) character but sometimes is represented as a semi-colon (`;`) to indicate drop-frame timecodes
- **q** is the punctuation delimiter for Frames. For drop-frame timecodes, this shall be the semi-colon character (`;`). At present, virtually all media at 29.97 FPS and thus uses the drop-frame timecode. In the extremely rare case for non drop-frame timecodes, the colon (`:`) character should be used.

Example of time codes:

> “01:00:00;00” - Represents one hour
> “12:38:59;28” - Represents twelve hours, 38 minutes, 59 seconds and 28 frames

Conceptual Overview

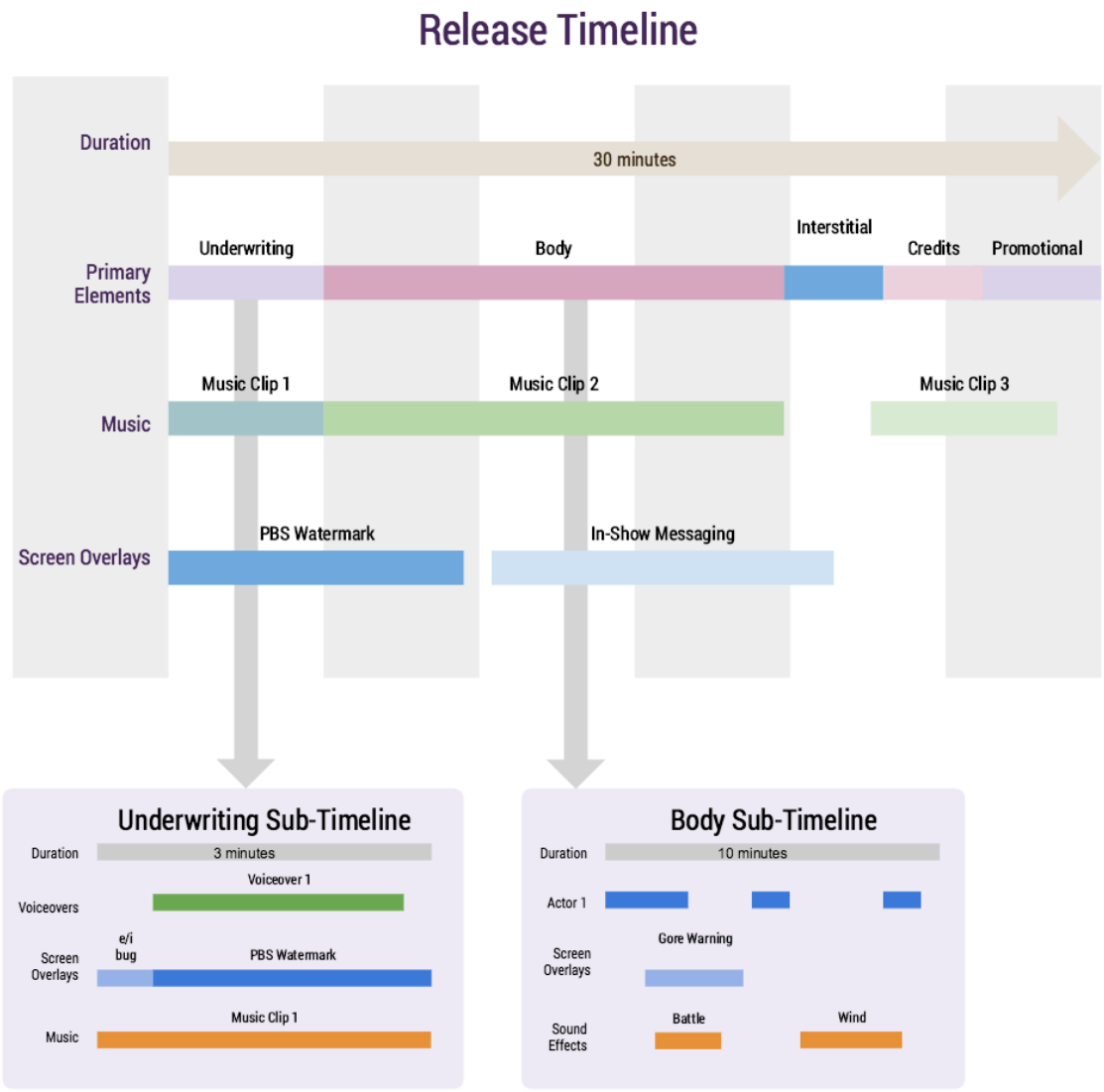
Figure 1 shows the relationships between segmentation concepts (Source Media and Timelines) and abstract content concepts (Episode and Release).

During the Production phase of a content’s lifecycle, **Source Media** are placed in sequence to create Releases. **Timelines** can be created to better categorize similar elements in sequential order.

All Releases contain Source Medias with individual Timelines, and individual Source Medias can have its own Sub-Timelines. In Figure 1, the Release has a Primary Events Timeline and a Secondary Events Timeline layered on top of each other.

Multiple Timelines is discussed more in the Timelines section.

Figure 1: Release Timeline and Sub-Timelines



This is an example of a Release's master timeline, which is composed of different Source Media: Primary Elements, Music, Screen Overlays, etc.. Each Source Media can have its own Sub-Timelines, as demonstrated by the Underwriting and Body Sub-Timelines.

Simply put, each piece of media has the possibility to contain one or more discrete timelines. When combined with other media, Timelines can be combined to form a single master Timeline of all the newly combined media.

2 Segmentation Metadata Values

Source Media

Source Media includes all the raw audio, video, and flag components that can be edited together into Episodes.

Source Media is created during Production, and the metadata around it should be documented before Distribution.

- Description
- Media ID
- Source Media Duration
- Source Media Type
- Start of Message
- End of Message
- Title
- UID

Description
The identifying characteristics of the media's content.
Format: String (4000)

Media ID
The House ID of the respective asset manager, which is a unique value generated by the metadata store for each content provider. Additional identification numbers can be associated to the source media using the 'Alias' relationships.
Format: String (50)

Source Media Type
The physical/logical format of the Source Material. If the `Source Media Type=Video`, the Source Media will have associated Source Media Durations and Timelines.
Format: video, graphic, audio

Source Media Duration
If `Source Media Type=Video`, the total expected duration of a Source Media is calculated from the Timeline's Start Time and End Time.
Format: `Timecode`

Title

A descriptive title for a piece of content, such as *A Chef's Life.* More information could be included to help identify the Source Media, such as: "A Chef's Life; webtags; credits over content"
Format: String limited to 400 characters.

Start of Message

The media timecode for the start of the content. The timecode of the media is not required to be 00:00:00;00—more commonly the Source Media content will begin at 01:00:00;00 to accommodate the preamble. The preamble usually includes a slate, bars, or a countdown.

Format: Timecode

End of Message

The media timecode for the end of the media.
Format: Timecode

UID

Unique ID for internal tracking purposes during the Acquisition phase.
Format: `urn:pbs:content:<uuid>`
Example: `urn:pbs:content:899310eb-b7cd-4364-acf0-8da105f46966`

Events

Events are time-based elements that make up a Source Media Video's content. Multiple Events that have the same Type can be temporally sequenced together to form a Timeline.

Events are defined by the following metadata:

- Visual Location (optional)
- Start Time
- End Time
- Event Details
- Event Type
- Relationship Element (optional)
- Title (optional)

Visual Location

This field captures the physical location of where a digital onscreen graphic (a graphic overlay, ticker, watermark, or bug) will appear on the screen. For example, the value for this field could be "lower right."
Format: String

Start Time
The timecode at the beginning of an Event. This value is used with an End Time to calculate the Source Media Duration. Alternatively, the Source Media Duration could be specified to calculate the End Time.
Format: `Timecode`

End Time
The timecode at the end of an Event. This value is used with a Start Time to calculate the Source Media Duration.
Format: Timecode

Event Details
Specific information relevant to an Event.
Format: String limited to 1000 characters.

Event Type
Describes the category of an Event. Possible values are as follows:

- Primary Element: Essence
- Primary Element: Supplemental Media
- Secondary Element: Supplemental Media
- Other, based on Timeline

This list should remain flexible for Timeline needs, as each Timeline will dictate the types of Events that occur during the Timeline. For example, a 'Rundown' Timeline will consist solely of contiguous Primary Elements. But a Timeline that is capturing a series of events for playback on a digital tablet device could trigger interactive features and games, and thus may have an Event Type of 'Trigger: Interactive Game.'

Relationship Element
A reference to the Source Media or the Supplemental Media.
Format: Reference to Source ID or Supplemental Media EIDR ID.

Title
A human-friendly identifier for the Event.
Format: String limited to 100 characters.

Timelines

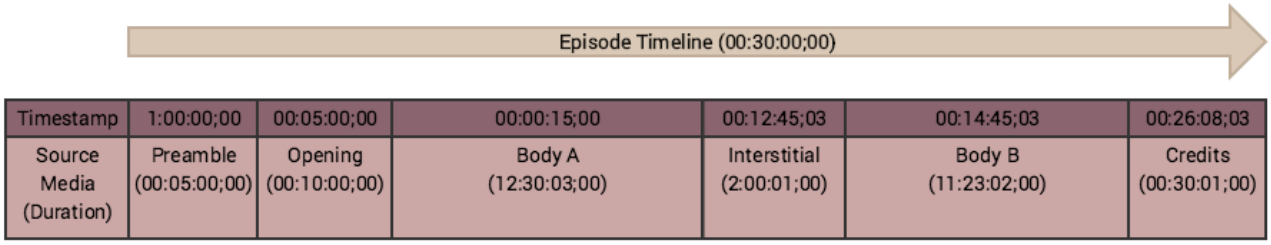
Timelines are time-based constructs that describe the content of video media. Each Timeline represents a specific type of Events that occur within the video media.

Since Timelines are generic constructs, they may be associated with any abstract metadata model that represents video. In the current model, the following models[1] may have associated Timelines:

- Source Media (when Source Media Type is video)
- Release
- Manifestation
- Supplemental Media

Timelines can be created as a general catch-all for similar events. For example, the Episode in Figure 2 has one Timeline known as the 'Rundown' that represents all the Source Medias that appear in the Episode, in order.

Figure 2: Single Timeline



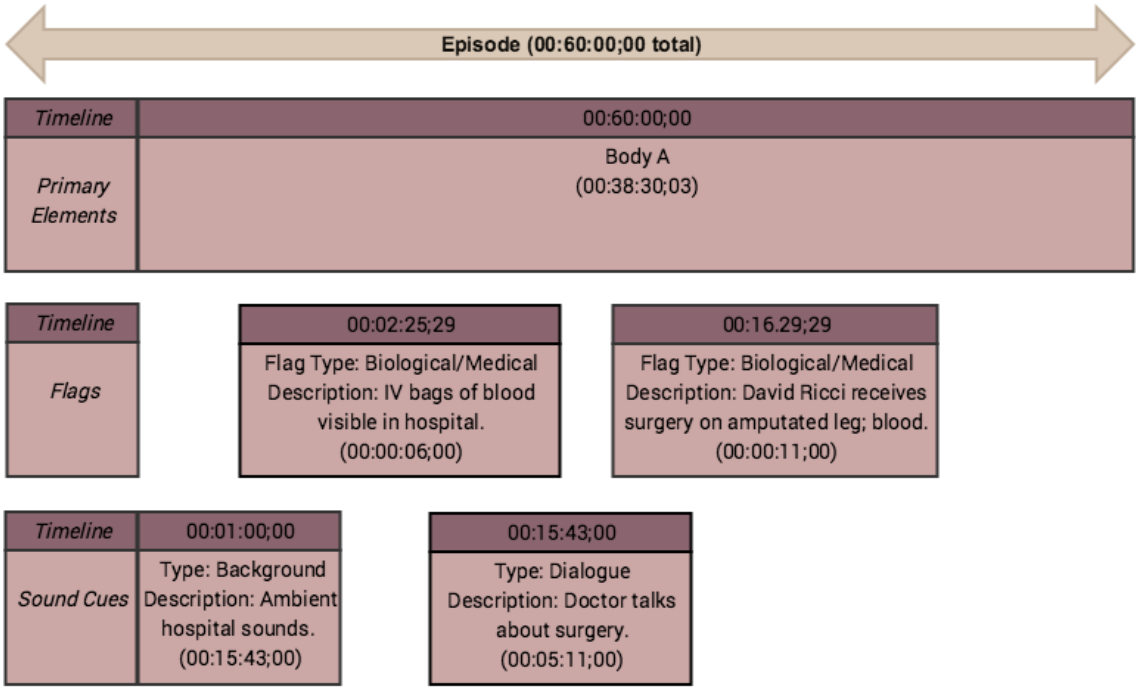
This figure shows a single Timeline containing all the Primary Elements of an Episode. Each piece of Source Media (teal) contains a Duration. Its place on the Timeline is indicated by the timestamp above the Source Media box. The long blue arrow above the timestamp line indicates the Episode's total duration.

Timelines of different types can be created in a single video, as shown in Figure 3. Timelines of different types in the same video media exist in parallel and do not interact in any way. For example, an Episode or OTO can contain a Timeline for Graphic Events only, a separate Timeline for Audio Events, and even a separate Timeline for when actors appear and disappear on-screen.

The duration of the Timeline, while expressed as a timecode, is always relative to the Start of Message (SOM) of the associated model. Thus while the timecode of the SOM for Source Media is often `01:00:00;00`, the associated Timelines will always start at `00:00:00;00`. This means that consumers of the Timeline who wish to know absolute timecodes of Timeline events must add the SOM timecode to the Timeline timecodes.

[1] Only Source Media is discussed in this Part. The other models (Episode, Story, etc.) are discussed and defined in Part 2 of this document series.

Figure 3: Multiple Timelines



This figure shows three different Timelines running concurrently in the same Episode. The first Timeline contains the Timestamps, Name, and Duration for Source Media categorized as Primary Elements. The second Timeline contains the Timestamps, Name, and Duration for Source Media categorized as Flags. The third Timeline contains the Timestamps, Name, and Duration for Source Media categorized as Sound Cues.

Events do not need to form a contiguous block of events to be a complete Timeline, as illustrated in Figure 3.

During the production and distribution phases of content, many pieces of media (Source Media, Supplemental Media, etc.) will be combined to form the final distributed content. During this process, the Timelines of these pieces of media will be combined, ultimately producing a finished piece of media that not only combines the visual and audio but also is associated with a combined set of Timelines that represent all the Timelines of the component pieces of media.

During the combination of component Timelines, Timelines of the same type can simply be added together, assuming they have compatible frame rates. The finished piece will have a distinct number of Timelines that represents the total number of distinct Timeline types. Each combined Timeline will be the union of all Timelines from the source components.

Here is an example:

- Source Media A represents a 12-minute kids' show with an associated Timeline that records educational curriculum opportunities at periodic points in the show. The events in this Timeline are calculated from the start of Source Media A, starting with a Start Time of `00:00:00;00`.

- Source Media B represents an 11-minute kids' show with a different associated Timeline that also records educational curriculum opportunities. This Timeline also uses a Start Time of `00:00:00;00`.
- When both Source Medias A and B are combined into an Episode, the Episode now has a single Timeline for educational curriculum opportunities with all events from both Source Media Timelines. However, the Events from Timeline B have been temporally shifted to start after A's Timeline.

This process is repeated for all Timelines in each Source Media. Consequently, when a Release is created from Episodes, One-Time-Onlys, or traffic/promotional Source Media, the Timelines of all components are combined by Timeline type.

Timelines are defined by the following metadata:

- Timeline Type
- Contiguous
- Related Content
- Frame Rate
- Events

Timeline Type

The Events contained by the Timelines can have specific meaning to systems that interpret the data. Common Timeline Types include the following:

- **Rundown**: This type of Timeline is a contiguous set of events that defines the primary elements included in the media.
- **Visual Overlays**: This type of Timeline identifies various graphics and overlays applied on top of the video essence.
- **Content Flags**: This type of Timeline identifies various events in the video media that must be identified and communicated to consumers for viewing discretion and or broadcast decisions.
- **Chapters**: This type of Timeline identifies logical breaks in a story according to the content creator.
- **In-Show Messaging**: This type of Timeline identifies the temporal location, duration, placement, and associated content
- **In-Show Messaging Opportunities**: This type of Timeline identifies the temporal location, duration, and placement of opportunities for broadcasters to insert additional overlays or secondary elements.
- **Interactive Events:** This type of Timeline allows markers that may instruct playback devices to trigger an external interactive feature.

Additional types of timelines may be added as needed.

Format: String

Contiguous

This field describes if the Events in the Timeline are contiguous. If True, then the Timeline must have no undefined gaps between events. And the events must cover the entire time period of the Timeline. In the case of the **Rundown** Timeline, all events are contiguous.
Format: True/False

Related Content

This field captures the UID of the abstract metadata model to which this Timeline is associated.
Format: `urn:pbs:content:<uuid>`

Frame Rate

`The expected Frame Rate of all Events in the Timeline. This is used to ensure Timeline compatibility and nomenclature.` The frame rate is derived from the Source Media itself and tells us if it is drop/non-drop, interlaced/non-interlaced. SMPTE timecode.
`Format:` String in Frames Per Second (FPS). Example: 29.97

`Events`

This field captures the associated Events in the Timeline.
Format: Collection of relationships to Events.

Abstract Metadata in Public Broadcasting

Part 4: Deals and Rights

Version 2.1
May 2017

Table of Contents

Introduction

This part discusses how the metadata associated with broadcasting deals, rights, permissions, and content visibility is represented. This is the fourth part of a series of documents concerning the Interconnection System's metadata system. The other documents in this series are as follows:

- **Part 1: Introduction to Abstract Metadata in Public Broadcasting**
  This part introduces the Interconnection System and the need for an abstract hierarchical metadata system.
- **Part 2: Specific Fields and Values**
  This part describes the core abstract metadata models that compose the content library.
- **Part 3: Segmentation Metadata**
  This part explains the metadata that provides detailed information about media content.

Goal

In previous iterations of the Interconnection System, Deals and Rights were dealt with by systems outside of the Interconnection System.

The newest iteration of the Interconnection System acts as a commonly accessible repository that passively captures and distributes information about Deals and Rights in a standardized format. The enforcement of the captured Deals and Rights still relies on human interpretation and external Rights Management Systems. However, by giving the metadata system parseable data to calculate usage windows, PBS can establish an Interconnection System rights model that can drive future automation.

Terms and Definitions

This document specifically discusses deals and rights in the Interconnection System. The following terms are used throughout and should be interpreted with the definitions below.

Actors:

- **Content providers** (or distributors) are the organizations that hold the creative license and rights for content created or obtained to be broadcasted.
- **Content consumers** are the organizations that wish to use content, such as broadcasting stations within PBS, PDP (program differentiation plan) stations, and independent broadcasting stations.

Systems:

- The **Interconnection System** (IXS) is responsible for calculating and enforcing the Delivery Window of content from a content provider to a content consumer. It also sends outbound notifications when the content provider makes changes to a Usage Window.

- The **Station Traffic System** is responsible for scheduling the delivery of content based on the Delivery Window and for enforcing the Rights Usage Window.
- **Rights Management Systems** are software used by the content provider to establish the Rights Usage Window. These systems hold the contract and legal information for a piece of content.
- The **Content Packaging System** is software used by the content provider to establish the Context Usage Windows.
- **Feed schedule systems** are software that control the feed schedule for content providers.

Terms:

- **Broadcasting rights** are granted by **content providers** to a broadcasting station (either directly or via a distributor) to air a specific program on a station's channel or channels. These rights describe how and how often the rights holder can use the content.

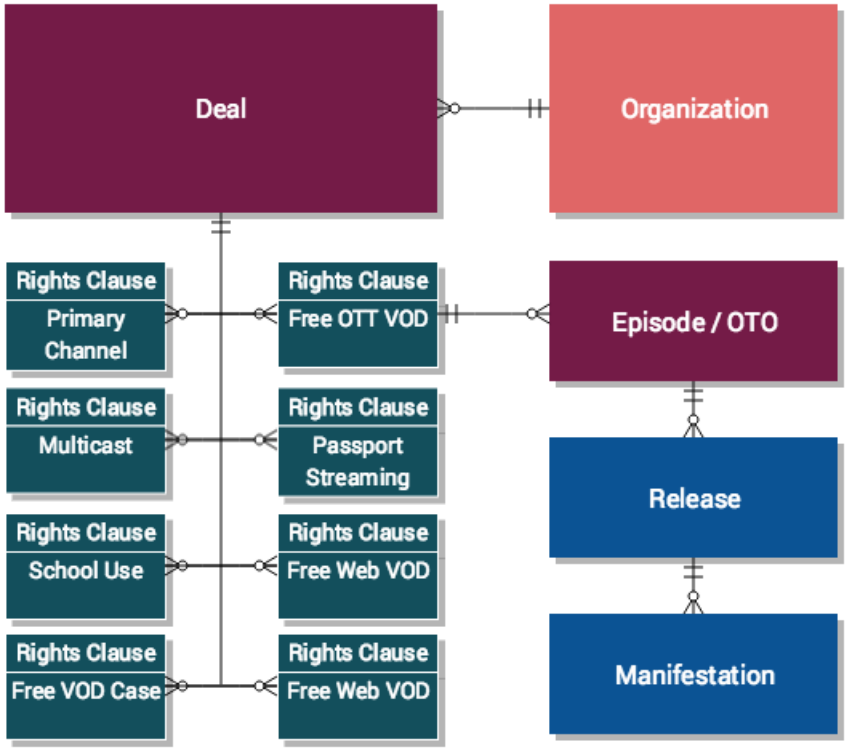
The abstract models that will be defined in following sections are Rights, Deals, and Usage Windows.

Core Metadata Concepts

Structural Metadata Overview

The abstract models discussed in this section include Deals, Rights, and the Effective Usage Window.

Their relationship to abstract content metadata concepts (Episode/OTO, Release, Manifestation) and Organizations are shown below:



Deals

Deals are constructed during the Acquisition phase of content development cycle. They represent specific usage agreements between the content providers and content consumers about a specific piece of content, such as an Episode, OTO, Release, or Manifestation.

Before a Deal can be created, the metadata for the content in question and the Organizations involved must already be registered to the metadata system.

Deals act as containers for the following relationships:

- Relevant Organizations (Content providers and consumers)

    - A Deal can support multiple content providers. As long as multiple providers do not try to distribute the same piece of content simultaneously to the same station on the IXS, there should be no conflict.
    - Multiple content providers can still distribute the same piece of content at the same time to the same station as long as they are not all using the IXS.
- Relevant Abstract Content Model (Specific IDs for Episode, OTO, Release, etc.)

Deal UID

The universal Identification number for the Deal. This allows the IXS to associate Episodes, Usage Windows, and Underwriting with this particular Deal. The format for a UID is an URN string, such as fe393d2a-0d5b-4366-b0ea-7aa8ecf715f2.

Deal Reference ID

This field captures the local ID given to a Deal by the content provider's rights management system. The format for this ID may vary from system to system.

Deal Name

The name of the deal. For example, the Deal name for the show History of Bikes could be History of Bikes.

Rights Holder Organization

This field links to the content provider's profile.

Rights Clauses

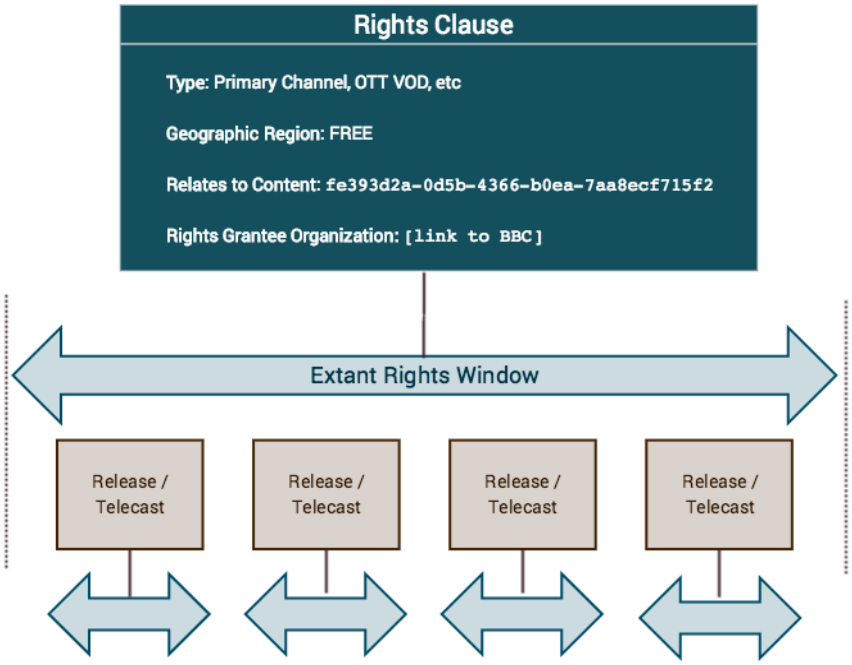
Rights are represented as one or more clauses. A Rights Clause defines how content can be used by licensee organizations.

A Rights Clause is defined by the following metadata fields:

- Relates to Content
- Rights Grantee Organization
- Rights Type
- Extant Rights Window
- Usage Model
- Geographic Region

The relationship between these metadata fields is represented in Figure 1.

Figure 1: Rights Clause Metadata



Extant Rights Window

This is the Start Date/Time and End Date/Time for a Rights Clause in a format consistent with ISO 8601 (YYYY:MM:DD:HH:MM). The time zone should be set to the licensee's local time. The Extant Rights Window should include any blackout windows.

In cases where there is no End Date, the Extant Rights Window can be set to Perpetuity.

Geographic Region

This field represents the geographic region that is permitted or excluded by the Rights Clause. In the case where there are no geographic restrictions, this clause is omitted.

Attributes of the Geographic Region are as follows:

- **List of Geographic Regions**: This attribute captures a list of relevant geographic regions.
- **Inclusive / Exclusive**: This boolean flag determines if the list of geographic regions are allowed or not allowed. Often, the Exclusive flag is used when the range of "allowed" geographic regions are much larger than the "not allowed" regions.
- **Type of Geographic Regions**: This attribute defines the resolution of the geographic regions. Acceptable values are:
    - Country
        - Based on ISO 3166-1 alpha 2
    - State
        - Based on ISO 3166-2
    - Television Market Area (TMA)

        - Based on FCC list

Some examples:

1. For U.S. based broadcast:
    a. Inclusive = True
    b. Type = Country
    c. List = US
2. For content only in a single TMA
    a. Inclusive = True
    b. Type = Television Market Area
    c. List = "Washington (Hagerstown, MD), DC"
3. For content only in a single State
    a. Inclusive = True
    b. Type = State
    c. List = NE

Relates to Content

A Deal concerns a specific abstract content model, such as an OTO, Episode, or Release. This field captures the UIDs of content models under that Deal.

Rights Grantee Organization

This field links to the content consumer's profile. There may be multiple organizations.

Rights Type

Media can be exploited by the content consumer in different ways: broadcast, published, etc. The exact exploitation rights for a piece of media are generally defined in a legal agreement between the Content Provider (or Distributor) and the Content Consumer. The process of how that legal agreement is created is beyond the scope of this document; however, once that legal agreement is established and the exploitation rights are defined, the IXS metadata system will facilitate the exchange of that information in a precise, machine-readable format.

However, for the metadata system to work, the Content Provider and the Content Consumer must agree on unique terms that represent each exploitable content rights. These unique terms are represented by a **Rights Type** and consist of a simple string (1000 characters) that is considered opaque by the metadata system and merely passed along.

As of the writing of this document, a common set of exploitation rights for use in the Interconnection System is still being defined and will be documented independently from this series of documents concerning the Metadata model.

Some examples of **Rights Type** could be:

- Linear

    - Primary Channel
    - Multicast
    - Live Streaming
- Non-Linear (Video on Demand [VOD])
    - Free Cable
    - Free OTT
- Education
    - School Off-Air Recording
    - PBSLM

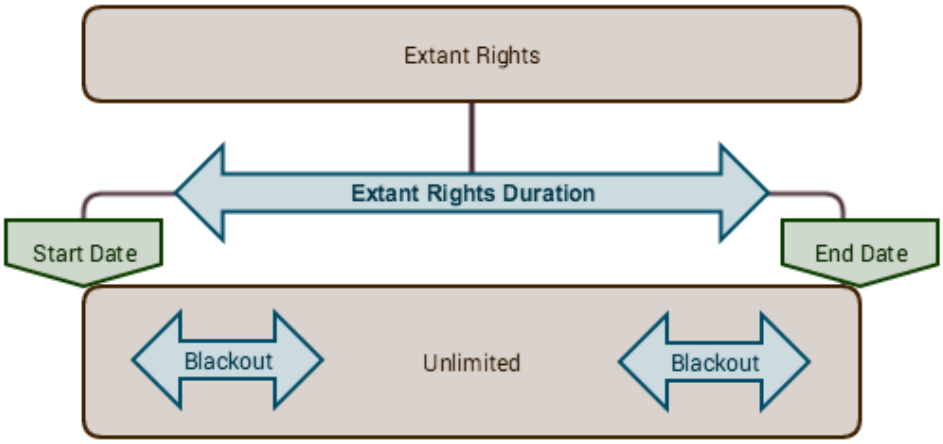
Once again, these Rights Types are included here merely as context for understanding the Metadata models.

Usage Model

Content providers must designate the usage of their content as one of the following three options: unlimited, release, or telecast. The usage model will define the **Rights Usage Window**.

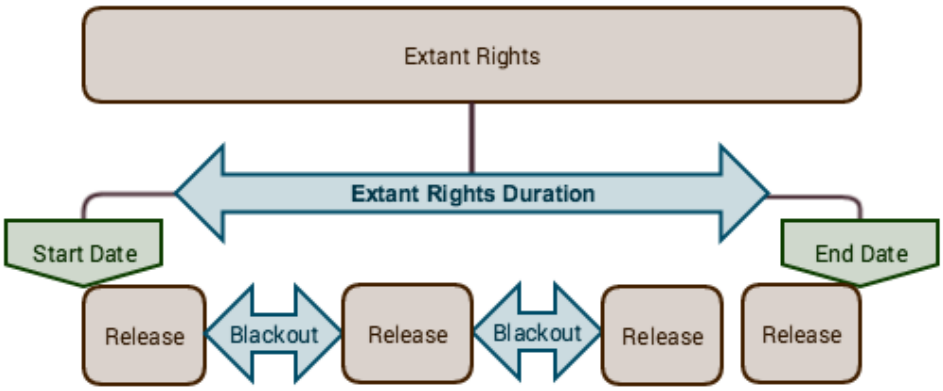
An **Unlimited** model (Figure 2) means that the content consumer may use the content as many times as it wants within the Extant Rights Window. There may be blackout dates defined in the Extant Rights Window. The **Rights Usage Window** is the same as the **Extant Rights Window**.

Figure 2: Illustration of Unlimited Usage Model



A **Release** model (Figure 3) means that the content consumer can trigger a predefined duration window of air time for a specified number of times. The **Rights Usage Window** must be defined with a Start/End Date and Time consistent with ISO 8601 (YYYY:MM:DD:HH:MM); alternatively, the **Rights Usage Window** may be defined as a Start Date and Time and a Duration.

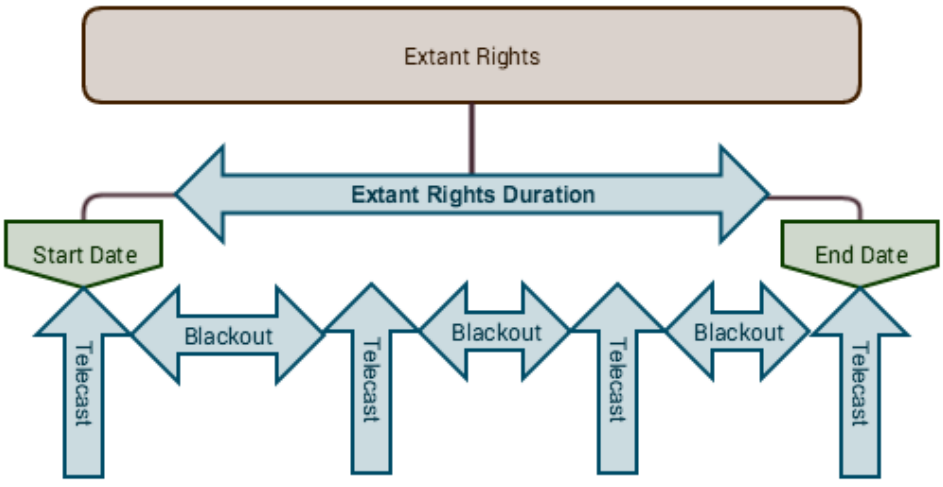
Figure 3: Illustration of Release Usage Model



For example, a Rights Clause for History of Bikes may stipulate that the content consumer may have 3 Releases, and each Release is defined as 7-day window. The consumer may choose to trigger a Release window on January 17. History of Bike thus airs on January 17, and for the seven days after January 17, the consumer may run History of Bikes as many times as desired. Then, after January 24, the consumer has to use one more of its two remaining Releases in order to run History of Bikes again.

A **Telecast** model (Figure 4) means that a consumer can broadcast the specified content for a specified number of times. Unlike a Release usage model, a Telecast usage model does not include a time window for unlimited runs. Telecasts are one-time-only uses. The **Rights Usage Window** is the same as the runtime of the program.

Figure 4: Illustration in Telecast Usage Model



For example, the Rights Clause may stipulate that a consumer may telecast *History of Bikes* 4 times within the Extant Rights Window of January 2014 and January 2016.
Thus, a valid usage scenario be:

- *The consumer telecasts* History of Bikes *on February 2 2014, June 13 2014, February 4 2015, and May 9 2015.*

Invalid usage scenarios include the following:

- *The consumer telecasts* History of Bikes *on February 2 2014, June 13 2014, February 4 2015, and May 9 2015. The content consumer then attempts to telecast* History of Bikes *one more time on December 13 2015. This scenario violates the number of uses listed in the Rights Clause.*
- *The consumer telecasts* History of Bikes *three times between January 2014 and January 2016. It then attempts to telecast* History of Bikes *for the fourth time on February 13 2016. Because February 13 2016 is outside of the Extant Rights Window, this is an invalid use.*

Effective Usage Window

Up to this point, the Rights discussed in this document have centered on the abstract core models. However, the metadata system also includes metadata for Releases (which include Supplemental Media) and Manifestations, the curated compilations and tangible assets that are actually exploited by the Content Consumer. The Effective Usage Window describes when a Release and its Manifestations can be used, depending on the Supplemental Media packaged with the Release and according to the associated Rights.

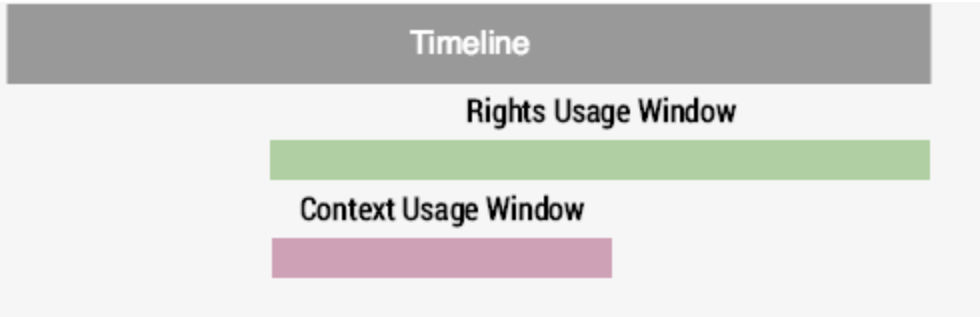
Upon creation, a Release is a combination of Primary Media (or essence) and Supplemental Media (underwriting, system cues, brand spots, teasers; these are defined further in Part 5 of this document series). Often, the included Supplemental Media reference other programs that occur on a linear feed. These programs could have occurred in the past or are scheduled to occur in the future, but the Supplemental Media must be in the proper temporal location to make sense. For example, Past Time Teasers ("Previously on Downton Abbey…") and Next Time Teasers ("Next time on Downton Abbey….") are Supplemental Media that are only valid and relevant for the end consumer when put in a linear context.

The consumable time window in a linear feed is known as the **Context Usage Window**. This means that if the Release (and associated Manifestation files) are scheduled in the linear feed outside of the **Context Usage Window**, the content will not make sense to the end consumer if he/she were consuming the media in a linear fashion.

The Context Usage Window is combined with the Rights Usage Window of the Release's parent abstract model to calculate the start and end date and time that constitutes allowed usage. The result of this calculation is the **Effective Usage Window**.

The Rights Usage and Context usage windows are illustrated in Figure 5. Their overlap is the Effective Usage Window.

Figure 5: Time Windows



Workflow

As of the publication of this document, the workflow in this section represents a possible workflow based on plans for how the Interconnection System (IXS) could work in the future. The IXS may not operate in this manner, but due to the nature of Rights and Usage, there are numerous systems outside of the MOS that are relevant to the end-to-end workflow.

1. Content consumer negotiates agreement.
    - Content provider and Content consumer agree on terms and conditions.
    - Provider records legal agreement in the Provider's Rights Management System.
2. Content provider enables content delivery.
    - Creates a Deal in MOS system of record (TBD) and associates it with content consumers.
    - Creates one or more Rights Clauses
3. Content Provider updates Rights based on events.
    - Some Rights Usage Window are based on Events in time (e.g. NPS Feed).
    - Content Provider / Distributor uses data in their own rights management system to calculate start/end dates.
    - Updates Rights Usage Window for event-based Rights Clause.
4. Content consumer schedules content to be broadcasted.
5. Content provider prepares content.
    - Creates Release which includes Supplemental Media.
    - Calculates Context Usage Window
    - Calculates Effective Usage Window
6. Provider uploads content.
    - Uploads content to IXS.
    - Establishes Effective Usage Window start/end timestamp.
7. IXS calculates Delivery Window (outside scope of this document).
8. IXS delivers content as long as current time is in Delivery Window.

Abstract Metadata in Public Broadcasting

Part 5: Supplemental Material

Version 2.1
May 2017

Table of Contents

Introduction

This document is part of the Abstract Metadata in Public Broadcasting Specification series. This part describes the relationships between supplemental media and primary media.

The other parts in this series are as follows:

- **Part 1: Introduction to Abstract Metadata in Public Broadcasting**
  This part introduces the Interconnection System and the need for an abstract hierarchical metadata system.
- **Part 2: Core Models**
  This part describes the core abstract metadata models that compose the content library.
- **Part 3: Time-based Descriptive Metadata**
  This part explains the metadata that provides detailed information about media content.

Terms and Definitions

The following terms are used throughout and should be interpreted with the definitions below.

Systems:

- The **Interconnection System** (IXS) is responsible for calculating and enforcing the Delivery Window of content from a content provider to a content consumer.
- The **Entertainment Identifier Registry** (EIDR) a global registry that provides a unique ID for media content based on select metadata entries. The IXS abstract metadata system borrows from and relies on EIDR precedents, and the IXS will include EIDR IDs to better sync metadata with external entities.

Metadata:

- The metadata system discussed in the document series so far is an **abstract hierarchy** that borrows from EIDR, MovieLabs, and established broadcasting standards.
- The abstract hierarchy system is composed of **abstract models** and **packaged content/digital files**.
  - Abstract models act as containers for **metadata descriptors**, which are information structured for both humans and machines and abstract models or packaged content/digital files easier to find, retrieve, and use. Examples from the abstract metadata hierarchy system for content, which is detailed in Part 2, include Series, Episodes, Seasons, and One-time-onlys (OTO).
  - Packaged content/digital files include **Releases**, which are final editorial sequences, and **Manifestations**, which are concrete renderings of Releases. They contain the Primary Essence.

    - **Primary Media** includes the main storyline of a Release. The main storyline may be represented as one or more bodies.
- Types of metadata include but are not limited to **Relational, Descriptive,** and **Structural.**
    - **Relational Metadata** describes the relationships between metadata concepts.
    - **Descriptive Metadata** is metadata that describes the content of the show. These fields become the basic identifying features of a show.
    - **Structural Metadata** describes the technical aspects or renderings of a piece of content.

Conceptual Overview

The abstract metadata hierarchy system presented thus far has been focused on the identification and use of Primary Media: the video and audio material containing the defining plots and characters of a show. Any other media that contributes to or supports the Primary Media is called Supplemental Media. Examples include promotional materials, short clips, and underwriting.

Supplemental Media are not standalone programs and must:

1. Promote or support one or multiple Primary Medias
2. Augment a primary work with additional content
3. Promote or support an organization, abstract model, channel, or other construct relevant to the purpose of the media

Supplemental Media metadata shall be relevant in two ways:

1. **Cataloging all Supplemental Media embedded within a primary work's 'Release' and merely referenced and/or described with metadata.** In this way, the Release is associated with deep-linked metadata, allowing stations to fully understand the included component media elements.
2. **Tracking and linking individual media files that are distributed separately from the Primary Media, such as a promotional tease or sizzle**. This is especially relevant when individual Supplemental Media files are distributed in the IXC system. By making Supplemental Media metadata standardized and searchable, stations and programmers will be able to identify, track, find, and obtain the appropriate files to suit their needs.

Types of Supplemental Media

The following list of possible Supplemental Media is not complete but is meant to define the scope of what can be classified as Supplemental Media.

The list is divided into types: Promotional, Brand, Program, and Sponsorship. Subtypes of each fall underneath.

Promotional

- Series Sell (SS)
- Long Lead Tease (LLT)
- Episodic
- Next-On
- Next-Time-Teaser

- Previous-Time-Teaser
- Theme
- Sizzle
- Press Tour
- Clip
- Clip Reel
- Interstitials
- Evergreen
- Crossover
- Other

Brand

- Seasonal
- Testimonial(s)
- Donate
- Fundraising
- Awards
- Support
- Brand Package
- Genre
- Channel (linear feeds)
- System Cue

Program

- Preamble
- Clip
- Bumper
- Featurette
- Other

Sponsorship

- Ad
- Underwriting

Other

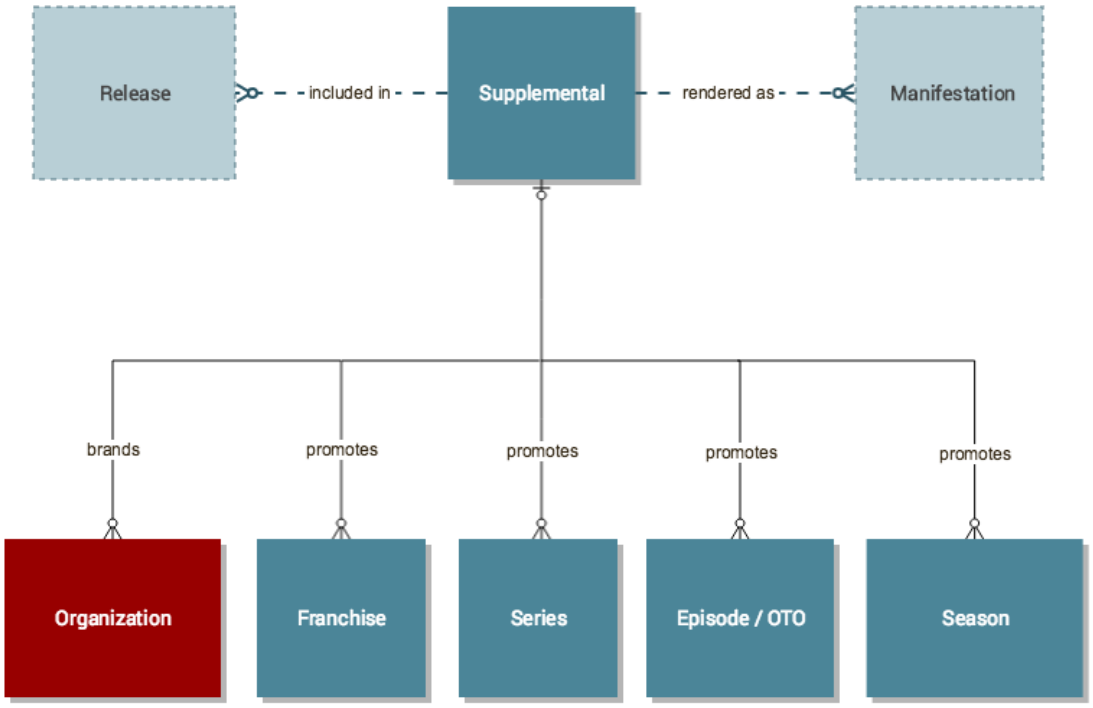
- External Event
- Organization

Supplemental Media Metadata

As represented in Figure 1, Supplemental Media can:

1. be included in any Release (package) that is created and distributed
2. have multiple Manifestations to support multiple versions in different video formats
3. relate to any abstract model, organization, or external construct (e.g. promotion of a local event). If the Supplemental Media refers to an external construct, it is beyond the metadata model to capture full metadata information.

Figure 1



In the current abstract metadata hierarchy system for public media, Supplemental Media can be seen as an abstract model separate from but connected to the primary abstract content hierarchy.

Relational Metadata

Supplemental Media has the following relationships:

Relationship	Description
promote	Supplemental Media, regardless of the type, must support or promote some sort of Primary Media: Franchise, Series, Episode/OTO, Season, or

	Organization. For example, a standalone promo can have any one of the following relationships: • The media promotes a Series • The media promotes a Franchise • The media promotes a Season of a Series • The media promotes a single Episode • The media promotes two Episodes from different Series
brand	Supplemental Media may exhibit or showcase an organization or branding concept.
supplement	Supplemental Media may be additional secondary content to primary media.
support	Supplemental Media may showcase organizational sponsorship or otherwise show some supporting relationship.
included in	Supplemental Media can be included in any Release.

Descriptive Metadata

Supplementary Media inherit many metadata field values from associated Primary Media, and the number and types of metadata fields vary depending on the type of Supplementary Media. Metadata fields common to all Supplementary Media include the following:

- Title
- Description
- Keywords
- EIDR ID

Supplemental Media has unique characteristics in descriptive metadata. Descriptive metadata germaine to Supplemental Media include the following:

- Asset Type
- Allowable Business Use
- Allowable Business Use Time Window

Asset Type
See the Types of Supplemental Media.

Allowable Business Use
Offers guidance to consumers of standalone Supplemental Media as to the allowed usage. For example, the Business Use could be 'Broadcast' or 'Online.'

Allowable Business Use Time Window
Indicates the start / stop time window for the aforementioned business use.

Structural Metadata Fields
Most structural metadata is inherited from the Primary Media's Manifestation values, which can be seen in Part 2 of this document series. Examples of structural metadata germaine to Supplemental Media include the following:

- Video Codec
- Bit Rate
- Audio Mappings
- And more...

Example of Embedded Supplemental Media
The example that follows is taken from the show *Arthur*, package number P339442-010.

Figure 2 shows the exact timing of elements in the package. This package contains two pieces of Primary Media: Cue #4 BOD and Cue #11 CWT. All the other cues are Supplemental Media.

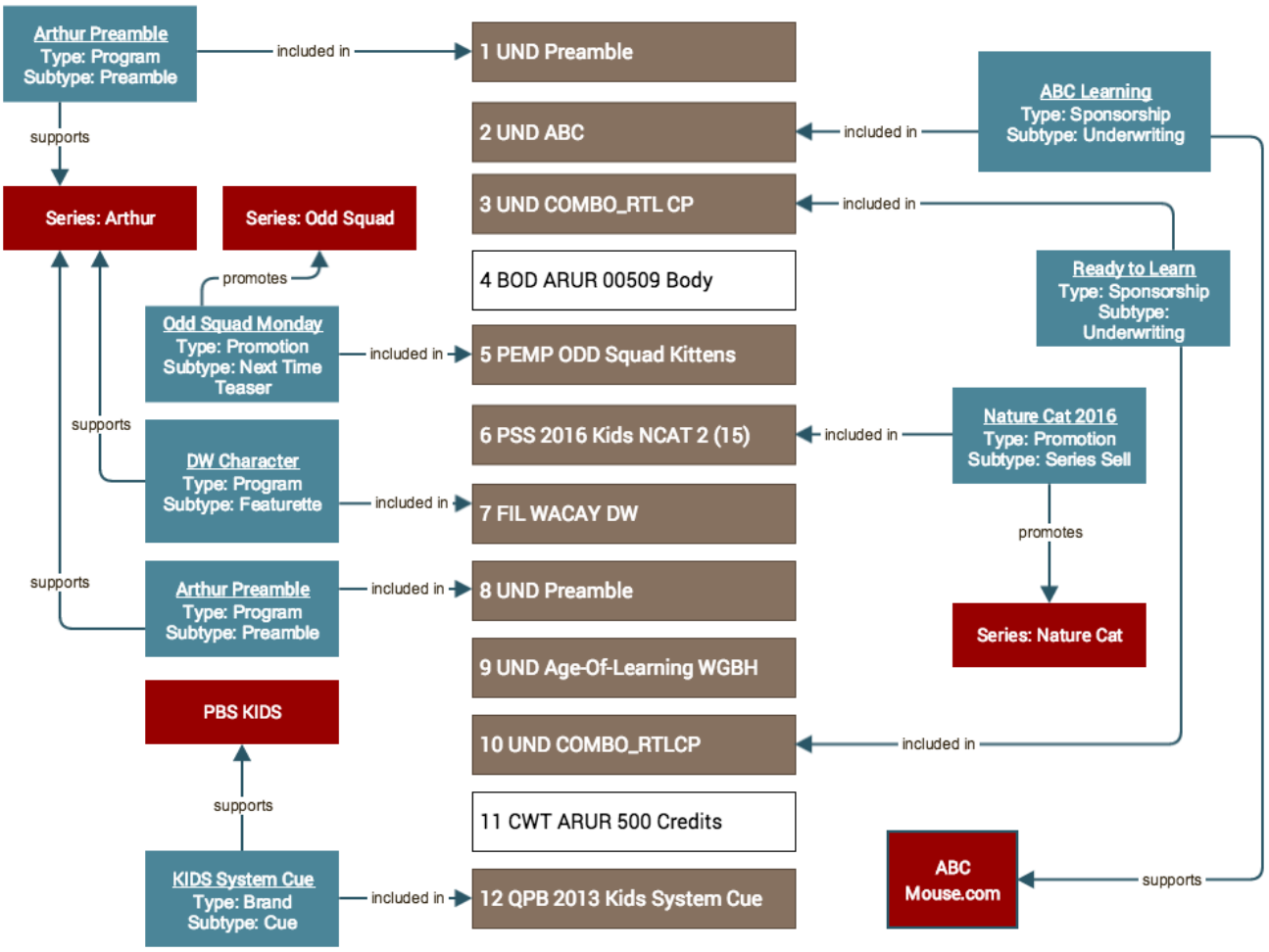
Figure 2

Cue Sheet

#	Type	Description	Audio	Ratio	Req MS	Start Time	Duration	Variable
1	UND	PREAMBLE 01 08132012 ARUR WGBH	Stereo	16:9	☑	:00.00	:05.00	☐
2	UND	AGE-OF-LEARNING Newintro 12032015 ARUR WGBH	Stereo	16:9	☑	:05.00	:15.00	☐
3	UND	COMBO_RTLCP V4-S5-And-S7 09302014 ARUR WGBH	Stereo	16:9	☑	:20.00	:10.00	☐
4	BOD	ARUR 000509 Body 09302014	Stereo	4:3	☑	:30.00	25:40.00	☐
5	PEMP	ODDS AoC "Kittens" :30 Monday	Stereo	16:9	☑	26:10.02	:30.00	☐
6	PSS	2016 Kids Crossover NCAT 2 (15)	Stereo	16:9	☑	26:40.02	:15.00	☐
7	FIL	WACAY DW	Stereo	16:9	☑	26:55.02	:30.02	☐
8	UND	PREAMBLE 01 08132012 ARUR WGBH	Stereo	16:9	☑	27:25.06	:05.00	☐
9	UND	AGE-OF-LEARNING New-Outro 12032015 ARUR WGBH	Stereo	16:9	☑	27:30.06	:15.00	☐
10	UND	COMBO_RTLCP V4-S5-And-S7 09302014 ARUR WGBH	Stereo	16:9	☑	27:45.06	:10.00	☐
11	CWT	ARUR 500 Credits 07282014	Stereo	16:9	☑	27:55.06	:45.00	☐
12	QPB	2013 KIDS System Cue - Trampoline (Ster)	Stereo	16:9	☑	28:40.08	:05.22	☑

The breakdown of the Supplemental Media in the package shown in Figure 2 is represented in Figure 3:

Figure 3



Cues from the package shown in Figure 2 are displayed in the central column as brown and white boxes. The blue boxes represent the Supplemental Medias included in those cues. The rex boxes represent brands or organizations promoted or supported by Supplemental Media.