



# Metadata Regroup: CTAM/Gracenote

December 13, 2021

# Agenda

## → Overview of Metadata Types/Key Fields

---

### → Automated Delivery

- ☐ Process
  - ☐ Guidelines
  - ☐ Methods
- 

### → Updates/Changes

- ☐ What to expect?
  - ☐ Best Practices
-

# What Does Gracenote Need?

# Industry - Leading Foundational Video Metadata

- **Robust Data:** TV schedule and TV Program, Movie, VOD and OTT information
- **Stunning Imagery:** Movie and TV series, program and episodic level imagery and artwork, plus cast & celebrity photos
- **Search & Discovery:** Structured data and Gracenote IDs across data sets are foundational for guide, search & discovery
- **Data Schema & Delivery Options**



# Data Requirements - Bound Data



Schedule Information



Program Metadata



Program Imagery



# Data Requirements - UnBound Data



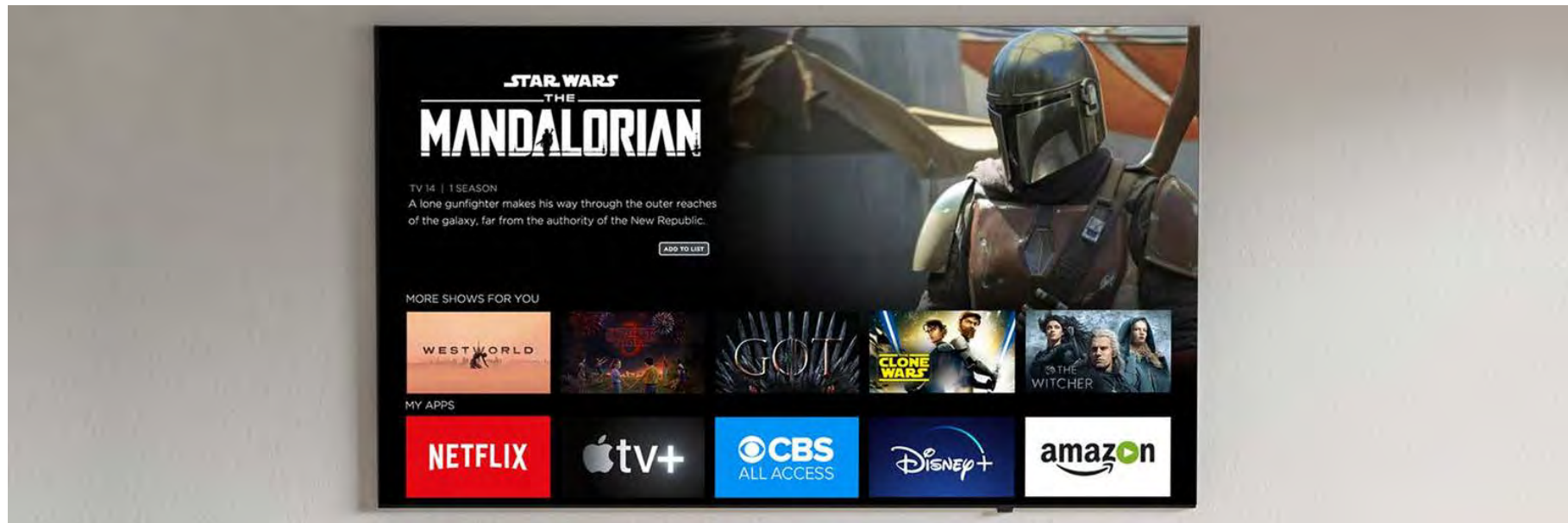
Availability Information



Program Metadata



Program Imagery



# Top Bound Data Fields



**Program Title** : The title of the program that will reside in the Gracenote entertainment database of record for perpetuity. Official on-screen title; avoid sponsors/brands. As with all max character counts, the number of allowed characters includes spaces (120 characters including spaces).



**NEW/PREMIERE Tag**: Notes when a program is airing for the first time per broadcaster. Allows viewers to identify new episodes of their favorite series. Updated programs with new/ additional footage added to a previously aired asset (or alternate versions that are explicit, extended, contain bonus content, etc.) qualify for a new tag. Note: NEW is on programs PREMIERE is on movies



**Season/Series Premiere or Finale**: Notes when an episode of a program is the first/last of the season or the series. This is for full seasons only, not for mid-season premieres or finales.



**TV Ratings/Advisories** : Identifies programs with objectionable content, including the levels of violence, adult language, nudity and adult situations. This allows the viewer to make more informed choices, and for set-top boxes, DVRs and other technology to block some content based on ratings & advisories. Shows, episodes, specials, TV movies and feature films on television should utilize the following ratings: TVY, TV7, TVG, TVPG, TV14, or TVMA. Feature films should follow MPAA ratings. Advisories are: D (dialogue), V (violence), S (sexual content), and L (language).

# Top Unbound Data Fields



**Deep Link URL** : Web access to content, iOS, Android Mobile, Set Top Box (STB), Android TV OS, Apple TV OS



**Content ID**: Unique identifier for event; partners use this ID for search and discovery of content so it is important to be consistent with IDs delivered to all content partners



**Business Model & Price**: T-VOD, S-VOD, free, TV Anywhere



**Availability Window** : Valid start/end for program



# Top Metadata Fields



**Program Title** : The title of the program that will reside in the Gracenote entertainment database of record for perpetuity. Official on-screen title; avoid sponsors/brands. As with all max character counts, the number of allowed characters includes spaces (120 characters including spaces).



**Program Type**: Show, Episode, Special, Sport, Movie, Compilation, Highlights, Miniseries, Music Video, Paid Program, Preview. This is required for accurate creation in the Gracenote database and allows set-top devices to group similar content together, which is helpful for end users.



**Program Description**: Descriptions should be a synopsis of what the asset is about: detail the plot/focus on the content. The description should be written in an active voice and present tense, be factual, avoid hype and promotional information, be evergreen, and meet journalistic standards. Do not repeat the asset title in the description, nor include details on when and where the program airs (avoid dates, times and networks – these are all stored in other fields).



**Genre** : A category that describes the overall artistic content of a program. Provides viewers with content information about a program; allows the set-top device to group programs together, allowing viewers to find programs like others they already enjoy. Maximum genres per show/episode/special/movies is six, in order of most broad to most specific. At least one genre should be provided.

# Additional Data Fields

- Audio Description
- Cast/Crew
- Character Names
- Duration
- Episode Title
- Episode Description
- Genre
- Language
- Movie Description
- Original Air Date
- Original Title
- Season/Episode #
- Series Level Description
- TV Ratings/Advisories
- Series Level Generic Description
- Year of Release (MV)
- Game Date/Time (Sports)



## Station 19

Drama, Action - TV14 - English - 60 mins - 2018

- ★ Practically raised at Seattle's Station 19, Andy Herrera is a confident firefighter who is also the daughter of Pruitt Herrera, the formidable head of the firehouse. Capt. Pruitt was Andy's primary inspiration to become a firefighter and is a mentor to both her and Jack Gibson, the lieutenant at the station. Jack is as fearless as Andy is by-the-books, but when the two are together, sparks fly and opposites attract. The two are joined by Maya Bishop, Dean Miller, Travis Montgomery, Victoria Hughes, and the station's newest addition, rookie Ben Warren. The team also works closely with Seattle PD, often running into Andy's former flame Ryan Tanner.
- ★ The personal and professional lives of a group of firefighters at the Seattle Fire Department, from the captain down the ranks to the newest recruit.



**Jaina Lee Ortiz**  
Andy Herrera  
Actor



**Barrett Doss**  
Victoria "Vic" Hughes  
Actor

# Image Requirements



**Resolution:** Prefer the highest-resolution images available; layered text is preferred



**Size:** Gracenote collects vertical, horizontal and square versions of the image (2x3, 3x4, 4x3, 16x9, 1x1)



**Movies:** Movie Poster, VOD Art, Key Art (no text) Backdrop (no text), Scene Still (no text)



**Series :** Series Banner, Series Iconic (no text), Backdrop(no text), Episode Still (no text), Cast in Character(no text), Cast Ensemble (no text), Channel branded AND Unbranded must be delivered

# Imagery



- Vast library of images for TV shows, movies, celebrities and channel logos
- Standard sizes & aspect ratios



# Advanced Discovery Data Set

Video Descriptors: deep content  
classifications across mood, theme, scenario  
& more





# Gracenote Video Descriptors



## In-Depth, Weighted Program Descriptors

Extensive, precise, weighted descriptors capturing the distinctive storyline and context of individual TV shows and movies



## Hierarchical, Controlled Taxonomy

Relate programs with one another with progressively granular terms across 16 classification categories



## Editorial Expertise & Coverage

Skilled, experienced in-house editorial staff curate programs using granular keywords to power nuanced discovery



# Video Descriptors - Classifications



## Mood

Tone of the work, as expressed through the combination of story, characters, setting, dialog, art direction, cinematography, music, effects, etc.



## Theme

An abstract concept of human experience that the work addresses



## Scenario

A specific situation, often personal or interpersonal, that sets the plot into movement, or moves it forward



## Concept Source

The type of original source material that provided the inspiration or story for the work



## Character

Actual or fictional person/being/group relevant to the work



## Setting

When and where all, or a portion, of the plot is set:

- Time
- Place
- Occasion



## Subject

A specific topic that the work addresses:

- Issue
- Personal Issue
- Activity
- Practice
- Event
- Location
- Milieu
- Being/Group



# Video Descriptors - Classification Examples



## Mood

Upbeat



Cheerful



Carefree



Light

GNK831JY2



## Theme

Darkness



Oppression



Injustice



Betrayal

GNK174LC5



## Scenario

Challenging Situation



Conflict Situation



Duel



Man vs Machine

GNK034NH9



## Concept Source

Written Work



Book



Young Adult Book

GNK260A7



## Character

Outdoor Enthusiasts



Mountaineers



Ice Climbers

GNK450B3



## Setting: Place

Lodging



Inn



Bed & Breakfast

GNK4568Z4



## Subject: Practices

Arts



Design



Fashion



Modeling

GNK749V2







## VIDEO DESCRIPTORS

### MOOD

Thrilling, Fantastic  
Electrifying, Emotional

### THEME

Heroism, Rivalry  
Revenge,  
Transformation

### SCENARIO

Rise to Power,  
Saving the World  
Banding Together,  
Haunted by the Past,  
Pursuit, Race Against  
Time

### CONCEPT SOURCE

Comic Books  
Hollywood Films



### SETTING

TIME: 2010s

PLACE:

Kingdom, Futuristic Worlds  
Big Cities, Laboratories, Secret  
Locations  
Vehicles

OCCASION: Showdown Fights

### SUBJECT

PERSONAL ISSUE:

Courage & Bravery  
Bitterness, Obsessions with  
the Past, Rage  
Painful Memories

ISSUE:

African Society & Culture

PRACTICE:

Technology



INTRODUCING



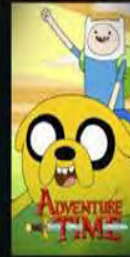
# Gracenote Personalized Imagery

Hunt for the  
Wider-people

A boy (Julian Dennison) and his foster father (Sam Neill) become the subjects of a manhunt after they get stranded in the New Zealand wilderness.

★★★★

THRILLING ADVENTURE



Hunt for the  
Wider-people

A boy (Julian Dennison) and his foster father (Sam Neill) become the subjects of a manhunt after they get stranded in the New Zealand wilderness.

★★★★

WORKING TOGETHER



## Display the Best Program Image for Every Viewer

Display captivating images portraying key aspects of a program tailored to individual viewers in order to supercharge visual content merchandising, drive engagement and increase viewership



# Search & Discovery – Use Case

Emma

WORKING TOGETHER



Ethan

THRILLING ADVENTURE







# Personalized Imagery Example



**Descriptors:**  
Game Show Contestants, Playful



**Descriptors:**  
Struggle, Personal Challenge



**Descriptors:**  
Winners, Going For The Win

# Leverage Advanced Discovery?

→ Provide Gracenote with robust data as far in advance as available

---

→ Multiple levels of description

---

→ Expanded Genres

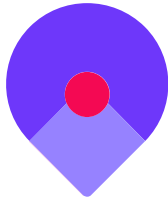
- ☐ Gracenote will list 5 (broad to generic)
  - ☐ Additional genres guide data experts on proper classification
- 

→ Multiple Images- specifically at Series/Movie level

---

# Automation!

# Automation Increases Quality & Speed



## Ingestion

Consistent files are quickly ingested into the Gracenote database



## Process

Content that has already been mapped will update channel days near instantly; new or updated content presents itself to a data expert for matching, creating or updating



## Delivery

Speedy completion of data allows for availability to Gracenote customers

# Required Data Fields - Bound Data



**Provider ID & Name** : Unique identifier assigned by the provider to identify a channel (*se.tv3, se.kanal5*)(*TV3, Kanal 5*)



**Unique and Persistent Asset ID**: Unique identifier for asset (*INPK010506000*)



**Air Date and Time**: The date and time a show will air (*2009-10-28 19:30:00 UTC+1*)



**Duration** : The length of time a program is scheduled to air. (*120 Minutes, 1 hour*)

\*\*\*\*Bound Data requires a full 24-hour schedule be provided; gaps can cause the process to fail and/or inaccurate schedule listings



# Required Data Fields - UnBound Data



**Provider ID & Name** : Unique identifier assigned by the provider to identify a channel  
(*se.tv3, se.kanal5*)(*TV3, Kanal 5*)



**Unique and Persistent Asset ID**: Unique identifier for asset (*INPK010506000*)



**License/air window start** : First and last date and time the show will be available on the service.(2009-10-28 00:00:00 UTC+1) (2009-12-28 00:00:00 UTC+1)



**Program type** : Series, special, movie, sport, episode, short form



**Program Title** : The full title of each program as it is displayed on screen



**Original Air Date/Year of Release** : Date the asset first premiered globally



**Duration/Run Time** : Length of the asset without commercials

# Delivery

# Delivery Methods



**Preferred Data Format** : XML, XMLTV, JSON (can also automate: HTML, XLS, XLSX, CSV or TSV). Deep links to URL can only be provided via XML, JSON or HTML. SCTE-224 (or other real-time event files) can be implemented which is useful for very timely updates – Sports and News!



**Delivery Method:** Gracenote prefers that exported or imported data files for automatic ingestions are made available at online sources such as APIs, FTPs, S3 bucket, RSS and MRSS feeds. Gracenote can pull from a partner's source or a partner can push to a Gracenote FTP. Gracenote prefers to the partner to provide the name for the S3 bucket first, rather than depend on a Gracenote naming convention



**Consistency!** The key to a great format is consistency. No matter which format you choose, consistency is crucial (no swapping fields/columns around). If changes must be made (e.g., you wish to include additional data points), please alert Gracenote so we can adjust our parser settings.



**Tips** : Use XML or JSON Validator to ensure the feed is well-formed. A text editor can be helpful for spotting illegal characters. Use binary (rather than ASCII) when delivering via FTP

\*\*\*\*Bound Data requires a full 24-hour schedule be provided; gaps can cause the process to fail and/or inaccurate schedule listings

# Timing

# Data “IN”



**Initial Set -up:** Initial set-up of a feed can take 4-6 weeks; robust data and high quality structure of the feed will speed up integration. Data dictionary is extremely helpful



**Consistency is key:** The format, the data and delivery are all dependent on consistency!



**Processing:** Data that has been mapped will flow through Gracenote database within minutes of delivery and available to clients via Gracenote products. There is no need for manual intervention of established mapped content. New content to be mapped will be processed within 24 business hours – more data points allow for a quicker map. New content that can't be mapped and either needs deeper review or creation can take up to 72 business hours for completion. Schedules will get a placeholder until correct program is identified or created. Images are applied 24 hours after creation of new asset.



**What else?** : Gracenote normal business hours are M-F 9a-5p est. We do have staff working around the clock and “on-call” procedures for needs beyond traditional working hours so data “in” will be completed beyond 9a-5p but a partner can't assume they will be and should follow on-call procedures.

# Data “OUT”



**Availability** : Once data has been processed in the Gracenote system it is available to Gracenote clients.



**Client Timing:** Data availability on the Gracenote side is only part of the equation; each client “pulls” data from Gracenote, processes and pushes to consumers at their own cadence. This can range from very quickly (a few hours) to slowly (days)



**Delayed Changes?:** Each client has their own process once they receive Gracenote data and the volume of the files as well as engineering nuances on the client side may delay changes to be viewable. Generally - schedule and/or availability changes are viewed more quickly than metadata and images.



**What can be done to speed this up?** : Full, robust data provided to Gracenote in a consistent manner WELL IN ADVANCE of premiere or schedule data is key to ensuring information is accurately viewed by viewers. Identify where you are seeing delays, work directly with Gracenote and the place you are seeing delays to see if gaps can be closed. Work with Gracenote to understand “off-hours” procedures and how to get late changes processed.

