



Video Descriptor

Best Practices

Version 1.3

Updated: 2022-04-27



Copyright

Copyright © 2022 The Nielsen Company (US), LLC. All rights reserved.
Confidential and Proprietary.

Nielsen and the Nielsen Logo are trademarks or registered trademarks of
The Nielsen Company (US), LLC.

Other company names, products and services may be trademarks or
registered trademarks of their respective companies.

This documentation contains the intellectual property and proprietary
information of The Nielsen Company (US), LLC. Publication, disclosure,
copying, or distribution of this document or any of its contents is prohibited.

About Nielsen

Nielsen shapes the world's media and content as a global leader in audience
measurement, data and analytics. Through our understanding of people and
their behaviors across all channels and platforms, we empower our clients
with independent and actionable intelligence so they can connect and
engage with their audiences—now and into the future.

An S&P 500 company, Nielsen (NYSE: NLSN) operates around the world in
more than 55 countries. Learn more at www.nielsen.com or
www.nielsen.com/investors and connect with us on social media.

Gracenote

A Nielsen Company

<http://www.gracenote.com>

Contents

Overview1

Video Descriptors Dataset.....1

Video Descriptors Taxonomy1

Video Descriptors Types.....1

Video Descriptor Weights..... 3

Accessing Video Descriptors 3

 Program Annotations..... 4

 Video Descriptors Taxonomy 7

 Flattening the Hierarchy 8

 Video Descriptor Definitions and Synonyms 10

 Using Video Descriptors to Power Search & Discovery 10

 Use Case 1: Video Descriptors in recommendation engines 10

 Use Case 2: Video Descriptors to build carousels of related programs
 12

Appendix..... 13

Tables

Table 1: Video Descriptor Types 1

Table 2: Video Descriptor Weights3

Overview

Video Descriptors are a comprehensive set of descriptors bound by a unified hierarchical Video Descriptors Taxonomy. The descriptors are used to comprehensively describe a program across the various Video Descriptors Types like Theme, Character etc. The data set is created by specially-trained Gracenote editors using a rigorous tagging process to ensure a high-quality and consistent metadata product. Each video descriptor assigned to a program has a Video Descriptor Weight associated with it. The weights give the degree of importance of the descriptor for the program.

Video Descriptors Dataset

Video Descriptors are a comprehensive set of descriptors built with discovery and recommendations in mind. They are bound by a unified hierarchical Video Descriptors Taxonomy. The descriptors are used to comprehensively describe a program across the various Video Descriptors Types. The data set is created by specially-trained Gracenote editors using a rigorous tagging process to ensure a high-quality and consistent metadata product. The same taxonomy is used to annotate various program types (Movies and Series only for now), to allow forming nuanced correlations between programs and powering discovery and recommendations on platforms.

Each video descriptor assigned to a program has a Video Descriptor Weight associated with it. The weights give the degree of importance of the descriptor for the program.

Video Descriptors Taxonomy

The Video Descriptor Taxonomy is an integral part of the Video Descriptors dataset and provides the editorial team a controlled list to work with to describe a program. It is hierarchical structure which helps in forming broad as well as more nuanced correlations between programs. Each descriptor is unique in the taxonomy and is part of one hierarchy only - i.e. a descriptor will not have two different parents.

The taxonomy currently has ~16000 descriptors across 16 Video Descriptor Types. Each program is associated with descriptors which comprehensively describe the program across the various types present in the Video Descriptors Taxonomy

Video Descriptors Types

Video Descriptors are classified into the following types of descriptors:

Table 1: Video Descriptor Types

Type	Definition
Video Mood	A 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 being/group relevant to a work, encompassing occupation, personality, relationship, nationality, ethnicity, religion, cultural affinities, physical / mental condition, life stage, etc.
Setting - Time Period	A time period in which all, or a portion, of the plot is set
Setting - Place	A type of physical environment in which all, or a portion of, the plot is set - realm, area, specific outdoor or indoor environment
Setting - Occasion	A global cultural, religious or national holiday or festival, or transitory personal event addressed in the work
Subject – Specific Location	A continent, region, country, state, province, celestial object, planet, solar system, fantasy worlds or supernatural realm noted as a subject of the work
Subject - Issue	A mental state, emotion, personal quality or mental or physical condition noted as a subject of the work
Subject – Personal Issue	A mental state, emotion, personal quality or mental or physical condition that impact one's being noted as a subject of the work
Subject – Specific Being/Group	An actual, fictional or spiritual person, group or organization noted as a subject the work
Subject – Specific Event	An actual, fictional or spiritual event of the past, present or future noted as a subject of the work
Subject - Milieu	A well known combination of place and time, usually with additional explicit or implicit historical or cultural context that is noted as a subject of the work
Subject - Practice	A professional, vocational, activist, scholarly, religious or artistic pursuit, or other committed endeavor noted as a subject of the work
Subject -	A pursuit done for enjoyment, interest, or lifestyle outside of

Activity	a professional context noted as a subject of the work
----------	---

Video Descriptor Weights

Each Video Descriptor associated to a program is also accompanied by a weighting score. This adds a second dimension of relevancy to Video Descriptors that can be used to further refine content analysis and recommendations.

The weights are measured on a 10-point scale, but only the values 5, 7, and 9 are currently allowed to be applied. This is done to ensure the most relevant aspects of a work are tagged as well as to give the editors an ability to discern between the importance of two descriptors. The following table further explains what is meant by each weight.

The editor-assigned weights can be used as a categorization instead of a precise numerical value. I.e. you can choose to "re-weight the weights"

Table 2: Video Descriptor Weights

Weight	Description	Notes
5	Significant program aspect	
7	Major program aspect	
9	Primary program aspect	Reserved for the most important descriptors per type

At least one 9 is required in each type for it to be considered complete. 7s and 5s are applied when a Video Descriptor is applicable but is not as important for the program. This means some types in a given work will just have 9s while others will have 9s and a combination of 7s and 5s.

For example, for the movie Deadpool - Video Mood has Visceral, Irreverent, Romantic, Thrilling, and Violent as the descriptors associated. Irreverent and Thrilling are the primary descriptors. This film also showcases violence, but that is not the primary focus. So, visceral and violent are both 7s. Finally, although this film does focus on romance, it is not a major aspect relative to the other parts in the film. Therefore, Romantic is weighted as 5. Overall, this weighting allows Gracenote to show which Video Moods are the most important when discussing Deadpool and how it is similar to other programs with Thrilling and/ or Irreverent as descriptors, while also allowing for a different look for content similarity when using the lower weighted descriptors.

Accessing Video Descriptors

You must integrate with two endpoints to fully access and use the Video Descriptors data set in your application.

Endpoint	Description	More Information	xsd
Program Annotations	Contains Video Descriptors for programs build to optimize discovery use cases.	Program Annotations	http://developer.tmsapi.com/files/on_update_programAnnotations_3.5.xsd
Video Descriptors Taxonomy	A structured hierarchical relationship (parent/child tree) between descriptors.	Video Descriptor Taxonomy	http://developer.tmsapi.com/files/on_update_videoDescriptorsTaxonomy_3.5.xsd
Controlled Vocabulary	Defines terms managed (controlled) to simplify data indexing and searching		

Program Annotations

Each program is associated with descriptors which comprehensively describe the program across the various types present in the Video Descriptors Taxonomy. The weights and ordinal (which use weights in the background) give you the importance of the descriptor for the program.

Schema:

http://developer.tmsapi.com/files/on_update_programAnnotations_3.5.xsd

Sample Program Annotation XML

```
<programAnnotations>
  <programAnnotation TMSId="MV006482250000" rootId="23626"
updateId="2130645778" updateDate="2020-05-11T18:15:13Z">
    <videoDescriptors>
      <videoDescriptor ord="01">
        <videoDescriptorId>GN8C0MY9NMYHW5T</videoDescriptorId>
        <weight>9</weight>
      </videoDescriptor>
      <videoDescriptor ord="01">
        <videoDescriptorId>GN2TM7N5G3EJQGM</videoDescriptorId>
        <weight>9</weight>
      </videoDescriptor>
      <videoDescriptor ord="02">
        <videoDescriptorId>GNC4MSA58CG0CPZ</videoDescriptorId>
        <weight>9</weight>
      </videoDescriptor>
      <videoDescriptor ord="03">
        <videoDescriptorId>GNEXX2RXRAK9M9E</videoDescriptorId>
        <weight>9</weight>
      </videoDescriptor>
      <videoDescriptor ord="04">
        <videoDescriptorId>GNA5WKAKKHW091R</videoDescriptorId>
        <weight>7</weight>
    </videoDescriptors>
  </programAnnotation>
</programAnnotations>
```

```

</videoDescriptor>
<videoDescriptor ord="05">
  <videoDescriptorId>GNDE05JM9RX9EBH</videoDescriptorId>
  <weight>7</weight>
</videoDescriptor>
<videoDescriptor ord="06">
  <videoDescriptorId>GNDRNKN5CNDEKYP</videoDescriptorId>
  <weight>7</weight>
</videoDescriptor>
<videoDescriptor ord="01">
  <videoDescriptorId>GNACDCXHD8QV55S</videoDescriptorId>
  <weight>9</weight>
</videoDescriptor>
<videoDescriptor ord="01">
  <videoDescriptorId>GN8G0FMBVYF1M0V</videoDescriptorId>
  <weight>9</weight>
</videoDescriptor>
<videoDescriptor ord="02">
  <videoDescriptorId>GN4TYQ7KHS0F19P</videoDescriptorId>
  <weight>7</weight>
</videoDescriptor>
<videoDescriptor ord="01">
  <videoDescriptorId>GNB6JS41A0Y2ZC0</videoDescriptorId>
  <weight>9</weight>
</videoDescriptor>
<videoDescriptor ord="02">
  <videoDescriptorId>GNC93YGBJ8MYA1Z</videoDescriptorId>
  <weight>9</weight>
</videoDescriptor>
<videoDescriptor ord="03">
  <videoDescriptorId>GNCWFFVXK4RQ42H</videoDescriptorId>
  <weight>9</weight>
</videoDescriptor>
<videoDescriptor ord="04">
  <videoDescriptorId>GN06MC6CPG6M1RF</videoDescriptorId>
  <weight>7</weight>
</videoDescriptor>
<videoDescriptor ord="05">
  <videoDescriptorId>GN0JT1S0P3DRSBE</videoDescriptorId>
  <weight>7</weight>
</videoDescriptor>
<videoDescriptor ord="06">
  <videoDescriptorId>GN0NC6VV91PAJT0</videoDescriptorId>
  <weight>7</weight>
</videoDescriptor>
<videoDescriptor ord="07">
  <videoDescriptorId>GNCNADSSE606GYB</videoDescriptorId>
  <weight>7</weight>
</videoDescriptor>
<videoDescriptor ord="08">
  <videoDescriptorId>GNCQAMGJ3YX2GK5</videoDescriptorId>
  <weight>7</weight>
</videoDescriptor>
<videoDescriptor ord="01">

```

```

    <videoDescriptorId>GN68ZBT1HDZZYWA</videoDescriptorId>
    <weight>9</weight>
  </videoDescriptor>
  <videoDescriptor ord="02">
    <videoDescriptorId>GN3DJYSJFMSFEC8</videoDescriptorId>
    <weight>7</weight>
  </videoDescriptor>
  <videoDescriptor ord="03">
    <videoDescriptorId>GN6CV9XK1GATCV5</videoDescriptorId>
    <weight>7</weight>
  </videoDescriptor>
  <videoDescriptor ord="04">
    <videoDescriptorId>GN73Z028FDD10HA</videoDescriptorId>
    <weight>7</weight>
  </videoDescriptor>
  <videoDescriptor ord="01">
    <videoDescriptorId>GN6SGF1791W3GNX</videoDescriptorId>
    <weight>9</weight>
  </videoDescriptor>
  <videoDescriptor ord="01">
    <videoDescriptorId>GN1QADGG707NQ50</videoDescriptorId>
    <weight>9</weight>
  </videoDescriptor>
  <videoDescriptor ord="02">
    <videoDescriptorId>GN46C8JJJHC7466</videoDescriptorId>
    <weight>9</weight>
  </videoDescriptor>
  <videoDescriptor ord="03">
    <videoDescriptorId>GN2Z11HHPB3AX1Q</videoDescriptorId>
    <weight>7</weight>
  </videoDescriptor>
  <videoDescriptor ord="01">
    <videoDescriptorId>GN69MTZCWC3Q19</videoDescriptorId>
    <weight>9</weight>
  </videoDescriptor>
  <videoDescriptor ord="02">
    <videoDescriptorId>GN1QRM8JKQJWS7P</videoDescriptorId>
    <weight>7</weight>
  </videoDescriptor>
  <videoDescriptor ord="01">
    <videoDescriptorId>GNE11CBQ7J9F2N8</videoDescriptorId>
    <weight>9</weight>
  </videoDescriptor>
  <videoDescriptor ord="01">
    <videoDescriptorId>GN2JS4K4QA0A9W2</videoDescriptorId>
    <weight>9</weight>
  </videoDescriptor>
  <videoDescriptor ord="02">
    <videoDescriptorId>GNEAK15ZSR9MA9Q</videoDescriptorId>
    <weight>7</weight>
  </videoDescriptor>
  <videoDescriptor ord="03">
    <videoDescriptorId>GNEKBJF83NPPTK1</videoDescriptorId>
    <weight>7</weight>

```

```

    </videoDescriptor>
  </videoDescriptors>
</programAnnotation>
</programAnnotations>

```

Video Descriptors Taxonomy

This endpoint provides access to the video descriptor taxonomy which is a structured hierarchical relationship (parent/child tree) between descriptors. You can get the full list of Video Descriptors, grouped by Type, in this endpoint. In addition to the Video Descriptor ID and Name, this file also provides the following additional information:

Immediate parent of a Video Descriptor, using which the hierarchical taxonomy can be built

Type (ID and Name) that the video Descriptor belongs to (For example: Video Mood, Theme, Character, ...)

Video descriptors are grouped into a specified number types required to completely and accurately describe the video program for discovery use cases and allow forming correlations between various programs. A given Video Descriptor will only belong to a single Type and have a single parent. Programs will only be annotated with Video Descriptors that belong to the Video Descriptor Taxonomy.

Sample Video Descriptors Taxonomy XML

```

<on xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:noNamespaceSchemaLocation="http://developer.tmsapi.com/files/on_update_videoDescriptorsTaxonomy_3.5.xsd" schemaVersion="3.0">
  <header>
    <content>On - Updates: Video Descriptors Taxonomy</content>
    <created>2019-02-05T00:43:11Z</created>
    <copyright>Copyright 2019 Gracenote. All rights reserved.</copyright>
  </header>
  <videoDescriptorsTaxonomy>
    <taxonomyItem updateId="769286025" updateDate="2018-08-15T04:56:16Z">
      <videoDescriptors>
        <videoDescriptor>
          <videoDescriptorId>GN71R1EZQRZ9T26</videoDescriptorId>
          <videoDescriptorName>Dreamlike</videoDescriptorName>

          <videoDescriptorParentId>GN3289YNZFRPDVC</videoDescriptorParentId>
          <typeId>GN22FN81QZSVXN0</typeId>
          <typeName>Video Mood</typeName>
        </videoDescriptor>
        <videoDescriptor>
          <videoDescriptorId>GNB9BTX3Z1SJ8CF</videoDescriptorId>
          <videoDescriptorName>Romantic</videoDescriptorName>

          <videoDescriptorParentId>GNERHBFMG67PKRZ</videoDescriptorParentId>
          <typeId>GN22FN81QZSVXN0</typeId>
          <typeName>Video Mood</typeName>
        </videoDescriptor>
      </videoDescriptors>
    </taxonomyItem>
  </videoDescriptorsTaxonomy>
</on>

```

```

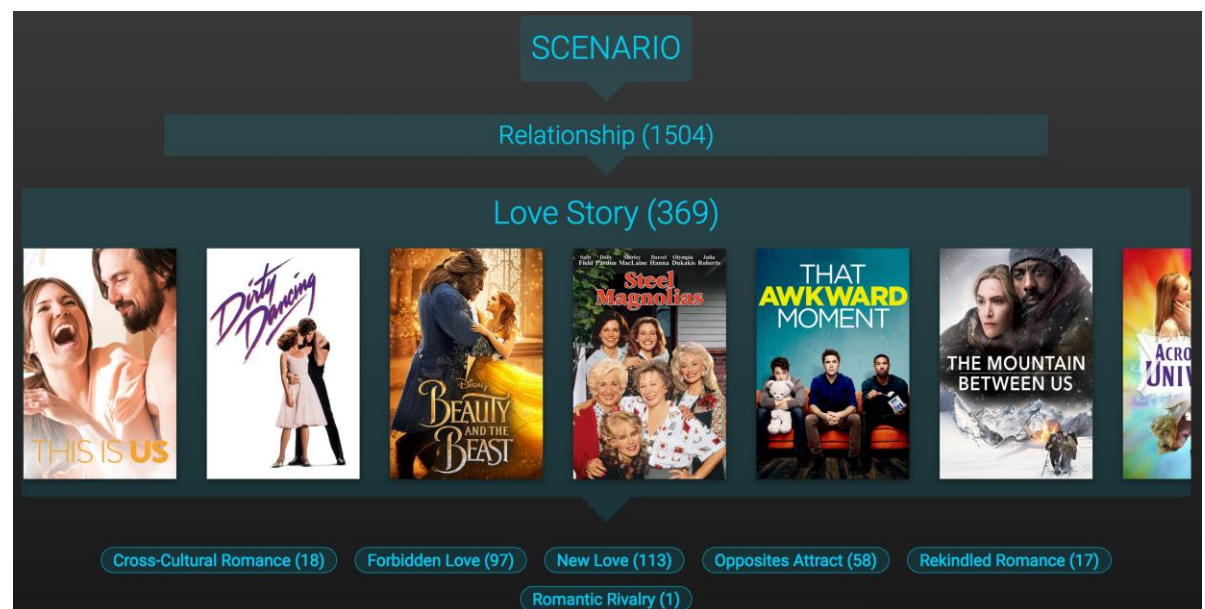
    </videoDescriptor>
    ...
  </taxonomyItem>
</videoDescriptorsTaxonomy>

```

Flattening the Hierarchy

The Video Descriptors Taxonomy as mentioned above is a tree-like structure. The editors while annotating a program assign the most granular descriptors as applicable. These explicitly assigned tags are available in programAnnotations endpoint while the ancestors of the descriptor should be assumed to be implicitly assigned to the program.

The Video Descriptors taxonomy is a tree-like structure as represented below:



Scenario | Relationship | Love Story is one branch in the Video Descriptors Taxonomy tree. Each node in the tree has a GNID associated with it.

Programs can be tagged with descriptors at any level in the tree, i.e. programs can be tagged with 'Relationship', 'Love Story' or even 'Opposites Attract' – depending on what is most relevant for the program in question

While flattening the hierarchy – if a program is tagged with 'Opposites Attract', consider that it is also tagged with 'Love Story' and 'Relationship'. If a program is tagged with 'Love Story', consider it is also tagged with 'Relationship' – i.e. all ancestors of the descriptor which is assigned to the program.

Here we describe how to traverse the tree to get all the descriptors as applicable to the program:

1. Step 1: Fetch the videoDescriptorId/s explicitly associated with the program from the programAnnotations XML file

Example: videoDescriptorId GN67NVSXKN4VXG9

```

<created>2020-04-28T17:23:57Z</created>
<copyright>Copyright 2020 Gracenote, a Nielsen Company. All rights reserved.</copyright>
<requestParameters>
  <requestParameter name="tmsId">MV001296160000</requestParameter>
</requestParameters>
</header>
<programAnnotations>
  <programAnnotation TMSId="MV001296160000" rootId="31064" updateDate="2020-03-19T13:30:20Z" updateId="1999447680">
    <videoDescriptors>
      <videoDescriptor ord="01">
        <videoDescriptorId>GN67NVSXKN4VXG9</videoDescriptorId>
        <weight>9</weight>
      </videoDescriptor>
      <videoDescriptor ord="02">
        <videoDescriptorId>GN9550CHAPAKHYC</videoDescriptorId>
        <weight>7</weight>
      </videoDescriptor>
      <videoDescriptor ord="01">
        <videoDescriptorId>GN2BP755FJH39H5</videoDescriptorId>
        <weight>9</weight>
      </videoDescriptor>
    </videoDescriptors>
  </programAnnotation>
</programAnnotations>

```

- Step 2: For the videoDescriptorId/s from step 1, find the ancestor VideoDescriptors using the Video Descriptors Taxonomy data and VideoDescriptorParentId

Looking up videoDescriptorId GN67NVSXKN4VXG9 in videoDescriptorsTaxonomy XML, finding videoDescriptorParentId GNATK82Q3CP1F3P and looking up(joining to the same table again in the datastore once ingested) again in videoDescriptorsTaxonomy XML

```

<typeId>GN0262ZZFABVGFK</typeId>
<typeName>Setting - Occasion</typeName>
</videoDescriptor>
<videoDescriptor>
  <videoDescriptorId>GN67NVSXKN4VXG9</videoDescriptorId>
  <videoDescriptorName>Police Pursuits</videoDescriptorName>
  <videoDescriptorParentId>GNATK82Q3CP1F3P</videoDescriptorParentId>
  <typeId>GN0262ZZFABVGFK</typeId>
  <typeName>Setting - Occasion</typeName>
</videoDescriptor>
<videoDescriptor>
  <videoDescriptorId>GN8YG3K3NE5TVHY</videoDescriptorId>
  <videoDescriptorName>Political Campaigns</videoDescriptorName>
  <videoDescriptorParentId>GN1ZBVQDKSAPM0D</videoDescriptorParentId>
  <typeId>GN0262ZZFABVGFK</typeId>
  <typeName>Setting - Occasion</typeName>
</videoDescriptor>

```

- Stop iterating at videoDescriptorId GNATK82Q3CP1F3P since it has no videoDescriptorParentId available. Repeat this for all the videoDescriptors assigned to the program above.

```

<videoDescriptor>
  <videoDescriptorId>GN0262ZZFABVGFK</videoDescriptorId>
  <videoDescriptorName>Setting - Occasion</videoDescriptorName>
</videoDescriptor>
<videoDescriptor>
  <videoDescriptorId>GNATK82Q3CP1F3P</videoDescriptorId>
  <videoDescriptorName>Law Enforcement Events</videoDescriptorName>
  <typeId>GN0262ZZFABVGFK</typeId>
  <typeName>Setting - Occasion</typeName>
</videoDescriptor>
<videoDescriptor>
  <videoDescriptorId>GN6HZCZD1RV7QG1</videoDescriptorId>
  <videoDescriptorName>Layoffs</videoDescriptorName>
  <videoDescriptorParentId>GNBZ7YPMVZFZQSQ</videoDescriptorParentId>
  <typeId>GN0262ZZFABVGFK</typeId>
  <typeName>Setting - Occasion</typeName>
</videoDescriptor>
<videoDescriptor>
  <videoDescriptorId>GN481XDYRV8CT9X</videoDescriptorId>
  <videoDescriptorName>Legal Battles</videoDescriptorName>

```

After this process, the program above will have two descriptors associated with it because of the videoDescriptorId GN67NVSXKN4VXG9.
Setting – Occasion | Law Enforcement Events | Police Pursuits

4. While doing this apply the same or a lower weight to the ancestors nodes.

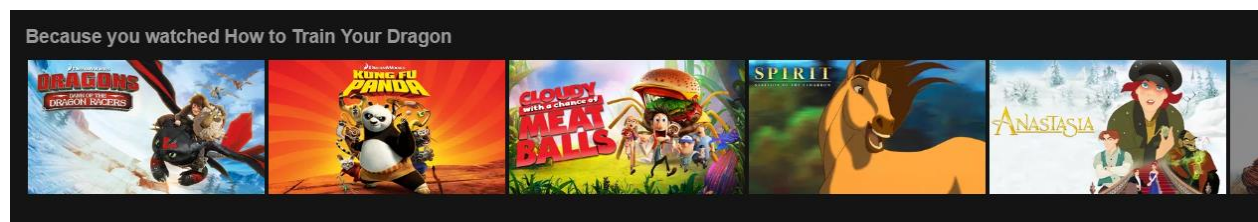
Video Descriptor Definitions and Synonyms

A singular concept is represented once within a type and has a specific definition in which editors utilize the Video Descriptors to describe programs. To understand the context or a similar descriptor to the one explicitly associated with the program, you can access definitions and synonyms information for a subset of video descriptors using this file [VideoDescriptors_ControlledVocabulary.xml](#)

Using Video Descriptors to Power Search & Discovery

Use Case 1: Video Descriptors in recommendation engines

Video Descriptors dataset has various features which could be tuned to generate recommendations for last-watched program by the user or a particular program from the user's watch history.



Hierarchical structure of the Taxonomy: The hierarchical structure of the taxonomy allows forming correlations between more programs. The editors while tagging a program, choose the most granular descriptor available. We recommend using the parent descriptors or ancestors of each descriptor explicitly tagged at the program or in other words flatten the hierarchy. It is worthwhile to also highlight here that a descriptor will only exist in 1 Video Descriptor type and can only have 1 parent.

Note *We highly recommend using the hierarchy completely to form correlations between more programs, calculating and finding the most similar programs. For computational ease however, if required, you can choose to ignore only the level 1(L1) parent descriptor for all types.*

For example:

Character | Educators(L1) | Education Administrators(L2) | Admissions Officers(L3)

If a program is tagged with Admissions Officers explicitly (data you get from Program Annotations), assume it is also tagged with Education Administrators and Educators implicitly (because of the hierarchical nature; this data you get from the Video Descriptors taxonomy – you can choose to ignore Educators(L1) as being tagged) while preparing the data for generating similar programs

Weights associated to each descriptor: Weights are another tuning parameter which can be used to further optimize your algorithm. Using weights allows you to take advantage of the comprehensive tags associated with a program while appropriately taking into account their importance for the program. We recommend using all the tags associated with a program – weighting the descriptors with weight 9s the highest and 5s the least. The editor-assigned weights can be used as a categorization instead of a precise numerical value and therefore 9, 7 or 5 s are not the exact weight to be used in your algorithms.

Note *If you need to reduce the computational burden [for a downstream task], you can ignore annotations with weight 5s.*

Types: Out of the 16 types available in the Video Descriptors Taxonomy - Video Mood, Theme, Scenario and Character provide the most lift when calculating content-to-content similarity. These four types therefore should be weighted higher than the other types and also at least the same as genres. These types Setting – Time, Setting – Place, Setting – Occasion and Subject Types (Subject – Specific Location, Subject – Personal Issue, Subject Practice, Subject – Activity) should be weighted medium.

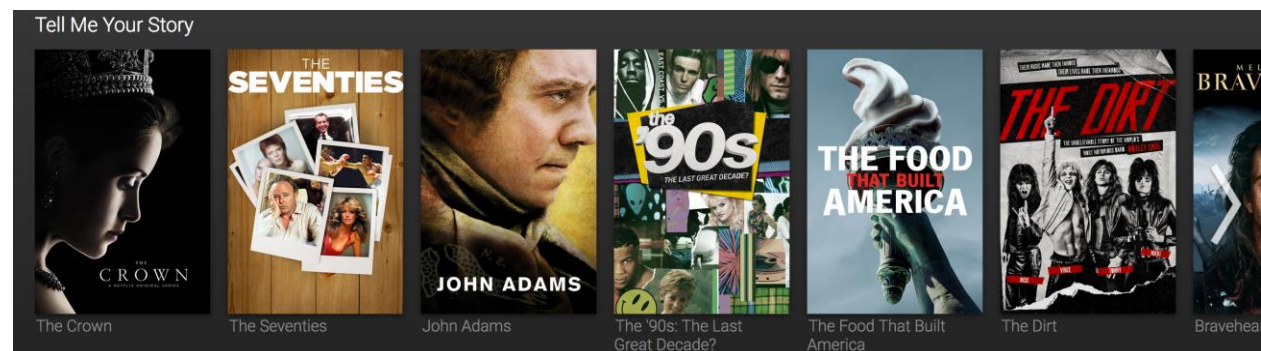
Note *For computation ease, Concept Source and subject types (Subject - Specific Being / Group, Subject – Milieu, Subject - Specific Event) can be weighted least or ignored completely.*

Other notes on Video Descriptors usage:

- Using Inverse Document Frequency (IDF) techniques after applying the hierarchy helped further to improve the results. This means weighting the descriptors occurring less frequently higher than the ones occurring more frequently.
- Layering with other metadata like Genre, program Cast information, Parental ratings also helped in getting improved results
- We use the same Video Descriptors taxonomy to describe all kinds of content to allow forming correlations between a variety of programs. We do this to ensure you get good correlations between programs and we keep away from defining the business rules on what to show to a user on the client. Once the correlated programs are generated, you could use the various genres and parental ratings data to filter, sort and decide on what to actually surface to the user for discovery. This is an important callout to consider for Children or Kids content for example and for other programs with different Parental Ratings.
- Using Gracenote popularity data also helped in creating smaller subsets of programs to calculate similarity.

Use Case 2: Video Descriptors to build carousels of related programs

Video Descriptors dataset has various features which could be tuned to generate related clusters of programs



Hierarchical structure of the Taxonomy: The hierarchical structure of the taxonomy allows forming correlations between more programs. The editors while tagging a program, choose the most granular descriptor available. It is worthwhile to also highlight here that a descriptor will only exist in 1 Video Descriptor type and can only have 1 parent.

Note *We highly recommend using the hierarchy to form correlations between more programs, calculating and finding the most similar programs. For creating correlated clusters, use the descriptor explicitly tagged at the program level plus one parent*

For example:

- Character | Educators(L1) | Education Administrators(L2) | Admissions Officers(L3)

If a program is tagged with Admissions Officers explicitly (data you get from Program Annotations), assume it is also tagged with Education Administrators implicitly (because of the hierarchical nature; this data you get from the Video Descriptors taxonomy)

Weights associated to each descriptor: Weights are another tuning parameter which can be used to further optimize your algorithm. Using weights allow you to take advantage of the comprehensive tags associated to a program while appropriately taking into account their importance for the program. For creating correlated clusters, we recommend using descriptors with weights 7s and 9s

Types: Out of the 16 types available in the Video Descriptors Taxonomy – using Video Mood, Theme, Scenario provided clusters with minimum

Appendix

Using the features of Video Descriptors and configurations (different weights across types, using all descriptors vs only 7,9s) as described in the sections above, here are a few samples how Video Descriptors can help surface different programs for different user profiles (Different “More Like Sleepless in Seattle” program lists) or different program lists that could be surfaced if for example Sleepless in Seattle was not available for streaming on your platform

More like Sleepless in Seattle

List 1	List 2	List 3	List 4
Hav Plenty	Just Like Heaven	Letters to Juliet	Hav Plenty
You've Got Mail	Serendipity	Why Him?	Serendipity
Just Like Heaven	Hav Plenty	When in Rome	Manhattan Love Story
When in Rome	Roman Holiday	My Big Fat Greek Wedding 2	The Mirror Has Two Faces
One Spring Night	One Fine Day	Serendipity	Letters to Juliet
The Giant	Can We Get Married	Because of Winn-Dixie	Permission
Mechanical Man	When in Rome	He's just not that into you	Something Borrowed
Manhattan Love Story	When Harry Met Sally	Her	One Fine Day
Happy-Go-Lucky	You've Got Mail	Sweet Home Alabama	Addicted to Love
Serendipity	Moonstruck	Anastasia	Frankie and Johnny
The Mirror Has Two Faces	See You in the Morning		Can We Get Married
	Enough Said		Roman Holiday
	Permission		Just One Night
	Three to Tango		I Hate Valentine's Day
	Addicted to Love		The Giant Mechanical Man

--	--	--	--

Note

These are just samples from test runs using an off-the-shelf algorithm to understand how various features are influencing the output.