CMRT

Technical Design

Mark Reedman (Rittman Mead)

CIO Delivery

11-Mar-2016

# Document Management

## Version history

|  |  |  |  |
| --- | --- | --- | --- |
| Version | Date | Author | Description |
| V0.1 | 11-Mar-2015 | Mark Reedman (mreedman@libertyglobal.com) | Initial document |
| V0.2 | 23-Mar-2016 | Chris Redgrave (credgrave@libertyglobal.com) | Added sections on TVA, Python and OBIEE Design |
| [Click here ...] | [Click here ...] | [Click here ...] | [Click here ...] |
| [Click here ...] | [Click here ...] | [Click here ...] | [Click here ...] |
| [Click here ...] | [Click here ...] | [Click here ...] | [Click here ...] |

## Distribution list

|  |  |
| --- | --- |
| Company | Company |
| [Click here ...] | [Click here ...] |
| [Click here ...] | [Click here ...] |
| [Click here ...] | [Click here ...] |
| [Click here ...] | [Click here ...] |
| [Click here ...] | [Click here ...] |

## Approval

|  |  |  |  |
| --- | --- | --- | --- |
| Name | Job description | Date | Signature |
|  |  | [Click here ...] | [Click here ...] |
| Muhammed Shoaib (mshoaib@libertyglobal.com) | Reporting & Data Management Specialist | [Click here ...] | [Click here ...] |
| Gokduman, Halim <hgokduman@libertyglobal.com> | Business intelligence Analyst, Operational Effectiveness | [Click here ...] | [Click here ...] |
| [Click here ...] | [Click here ...] | [Click here ...] | [Click here ...] |
| [Click here ...] | [Click here ...] | [Click here ...] | [Click here ...] |

Table of contents

[Document Management 2](#_Toc446489230)

[Version history 2](#_Toc446489231)

[Distribution list 2](#_Toc446489232)

[Approval 2](#_Toc446489233)

[Table of contents 3](#_Toc446489234)

[Overview 4](#_Toc446489235)

[Background 4](#_Toc446489236)

[Business Requirements 5](#_Toc446489237)

[Functional Requirements 5](#_Toc446489238)

[Technical Design 6](#_Toc446489239)

[Architecture and Data Flow 6](#_Toc446489240)

[Parsing TVA and Ditto Files 9](#_Toc446489241)

[Overview 9](#_Toc446489242)

[Parsing the Ditto and TVA File 9](#_Toc446489243)

[ETL Design 14](#_Toc446489244)

[Data Model Design 17](#_Toc446489245)

[CMRT OBIEE Design 18](#_Toc446489246)

[Repository Design 18](#_Toc446489247)

[Reporting Layer 27](#_Toc446489248)

[Reports 28](#_Toc446489249)

[Prompts 40](#_Toc446489250)

[Issues 42](#_Toc446489251)

[APPENDIX A 43](#_Toc446489252)

# Overview

## Background

**C**ontent **M**anagement **R**eporting **T**ool or CMRT is the term for this application to provide data quality reporting on the media content metadata for VOD and Linear TV. Each asset has information about the content, such as Title, Credits, Synopsis, and many more. This application is designed to provide analysis on that content.

Data is held in an XML format known as the TVA file. The TV-Anytime contains descriptive information for each audio visual asset that is current on VOD and Linear TV. It is an international standard format to be used by content and service providers containing a set of specifications for the controlled delivery of multimedia content to a user's personal device (Personal Video Recorder (PVR)).

The TVA file is split into a number of distinct sections or tables: Program Information, Group Information, Classification, Broadcast Event, On-Demand Program, Service Information, Schedule Event. Each asset is identified with a Content reference Identifier, better known as a CRID. This allows links between each of the sections/tables to be able to gather full details for the content.

A CRID can also have links to parent CRIDs to attach an episode to a series to a show. For example, an episode of CSI, “Room Service”, is part of CSI: Series 6, and is part of the CSI Brand/Show.

In addition to the TVA file, we also are provided with a Ditto file which is in a JSON format that contains each CRID and an indication of its availability on the LGI systems (CMDC, MSS, RENG, SMARTREC, TRAXIS).

The reporting model has been developed to monitor the quality of the program information.

This document attempts to describe the technical components and the overall algorithm used to establish the reporting model.

# Business Requirements

## Functional Requirements

### Reporting Requirements

The reporting requirements concentrate on providing high level data quality. For example, the number of assets that have the same text in the long, medium and short synopsis; or the number assets that have a long synopsis that is shorter than the medium synopsis; highlight the number of assets that have credits and keywords.

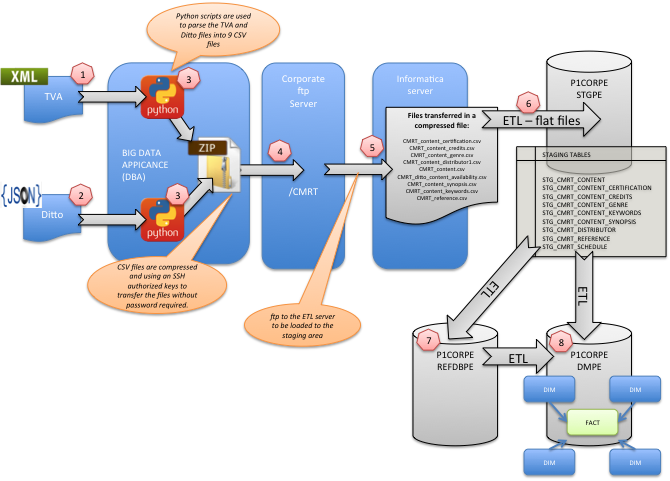
The following are the current requirements for CMRT:

|  |  |  |
| --- | --- | --- |
|  | **Data Requirement** | **Definition** |
| 1 | parental guidance values | total number of assets per parental guidance value |
| 2 | age rating coverage | total coverage of age rating |
| 3 | genre values | total number of assets per individual genres |
| 4 | genre coverage | total coverage of genres |
| 5 | sub-genre values | total number of assets per individual sub-genres |
| 6 | sub-genre coverage | total coverage of sub-genres |
| 7 | title coverage | total coverage of titles |
| 8 | title for episode coverage | total coverage of episode titles for all shows (brand) |
| 9 | title for series coverage | total coverage of series titles for all shows (brand) |
| 10 | title for show coverage | total coverage of show (brand) titles for all shows (brand) |
| 11 | episode number coverage | total coverage of episode numbers or all shows (brand) |
| 12 | series number coverage | total coverage of series (season) numbers for all shows (brand) |
| 13 | short synopsis coverage | total coverage of short synopsis (where number of words should be configurable, i.e. synopsis = < [X] | > [X] and < [X] | > [X]) |
| 14 | medium synopsis coverage | total coverage of medium synopsis (where number of words should be configurable, i.e. synopsis = < [X] | > [X] and < [X] | > [X]) |
| 15 | long synopsis coverage | total coverage of long synopsis (where number of words should be configurable, i.e. synopsis = < [X] | > [X] and < [X] | > [X]) |
| 16 | same text for synopsis S/M/L and S/L when M is not available | total number of assets where S/M/L synopsis contain same text |
| 17 | length of medium synopsis is shorter than short synopsis | total number of assets where length of medium synopsis is shorter than short synopsis |
| 18 | length of long synopsis is shorter than medium synopsis | total number of assets where length of long synopsis is shorter than medium synopsis |
| 19 | amount of characters in S/M/L synopsis coverage | amount of characters in synopsis (where the number of characters should be configurable, i.e. characters = < [X] | > [X] and < [X] | > [X])) |
| 20 | keywords coverage | total coverage of keywords (where the number of keywords should be configurable, i.e. keywords = < [X] | > [X] and < [X] | > [X])) |
| 21 | credits coverage | total coverage of credits (where the number of credits should be configurable, i.e. credits = < [X] | > [X] and < [X] | > [X]) |

# Technical Design

## Architecture and Data Flow

CMRT data file sources are transferred to the Oracle Big Data Appliance (BDA). Using the scripting language Python, the BDA parses the TVA (XML) and Ditto (Ditto) to produce 9 CSV files. These files are transferred the Corporate ftp server, and then transferred to the ETL server and loaded and loaded into the Pan European Staging area. Once in a relational database the reference data is pushed to our reference schema (REFDBPE) to allow the data to be conformed (if required). Finally the data pushed to the target reporting schema (DMPE).



| Step | Description |
| --- | --- |
|  | The TVA file is requested using curl commands and transferred to the BDA. This file is in an XML format. |
|  | The Ditto file is requested using wget commands. This file is in a JSON file format. |
|  | Python scripts have been developed on the BDA to parse both the TVA and Ditto files into 9 flat CSV files. These files are then compressed and ftp’d to Corporate ftp server. Each file is named based on the country and the date it was generated:  **CMRT\_<Country>\_<date in format YYYY-MM-DD>.zip**  For example:  **CMRT\_NL\_2016-02-03.zip** |
|  | A new CMRT user and root directory has been created on the Corporate ftp server. ssh keys have been setup so that the BDA is authorized to connect without a password. |
|  | The ETL workflow ftp’s the compressed file from the Corporate ftp server to the ETL server. It then uncompresses the zip file into the following files:  CMRT\_content\_certification.csv  CMRT\_content\_credits.csv  CMRT\_content\_genre.csv  CMRT\_content\_distributor1.csv  CMRT\_content.csv  CMRT\_ditto\_content\_availability.csv  CMRT\_content\_synopsis.csv  CMRT\_content\_keywords.csv  CMRT\_reference.csv  Into:  /transfer/data/informat/SrcFiles/KPI\_PE/CMRT/<Country Code>  ie: /transfer/data/informat/SrcFiles/KPI\_PE/CMRT/NL |
|  | All the files are then loaded into the following staging tables:  STGPE.STG\_CMRT\_CONTENT  STGPE.STG\_CMRT\_CONTENT\_CERTIFICATION  STGPE.STG\_CMRT\_CONTENT\_CREDITS  STGPE.STG\_CMRT\_CONTENT\_GENRE  STGPE.STG\_CMRT\_CONTENT\_KEYWORDS  STGPE.STG\_CMRT\_CONTENT\_SYNOPSIS  STGPE.STG\_CMRT\_DISTRIBUTOR  STGPE.STG\_CMRT\_REFERENCE  STGPE.STG\_CMRT\_SCHEDULE |
|  | The following data is loading into our reference schema. An APEX application sits on top of this data and allows the business to conform, group or describe data for the following:  AGE\_RATING  BROADCASTER  CHANNEL  CONTENT\_TYPE  GENRE  LANGUAGE  STUDIO  SUBGENRE  SYSTEM |
|  | The final ETL then populates following dimensions and fact:  DIM\_CMRT\_AGE\_RATING  DIM\_CMRT\_CHANNEL  DIM\_CMRT\_GENRE  DIM\_CMRT\_ASSET  DIM\_CMRT\_LANGUAGE  DIM\_CMRT\_AVAILABLE\_ON  DIM\_CMRT\_SYSTEM  DIM\_CMRT\_STUDIO  DIM\_CMRT\_BROADCASTER  DIM\_ORGANIZATION  FACT\_CMRT\_AVAILABILITY |

# Parsing TVA and Ditto Files

## Overview

CMRT data file sources are transferred to the Oracle Big Data Appliance (BDA). Using the scripting language Python, the BDA parses the TVA (XML) and Ditto (JSON) to produce 10 CSV files. These files are then compressed and transferred the Corporate FTP server along with a flag file, ready to be processed by Informatica.

The whole process is controlled using a Unix Shell Script which will perform the following high level steps:

1. Download the TVA file from the BDA
2. Download the Ditto file using a wget command
3. Parse the TVA File
4. Parse the Ditto File
5. Upload the resulting csv files into the BDA
6. Zip the files up
7. FTP the files to the Corporate FTP Server

This is scheduled to run at 8am every day using a Unix Cron Tab record.

## Parsing the Ditto and TVA File

The diagram below details which output files are generated from which script and from which data source. The python scripts use built in libraries to traverse the XML or JSON files are export the data that is required.

C:\Users\chris\Downloads\Untitled Diagram (2).png

### CMRT\_reference\_CC\_YYYY-MM-DD.csv

The reference data file contains information about the individual age classification schemes by language and country. This is used as a lookup table in the ETL to find valid age classifications.

The file output is in the following format:

|  |  |
| --- | --- |
| Column Name | Data Format |
| classification | UTF8 String |
| language | UTF8 String |
| key | UTF8 String |
| value | UTF8 String |

### CMRT\_content\_CC\_YYYY-MM-DD.csv

The content file contains a list of all the shows, series and episodes that are in the TVA file on a particular day and how they relate to each other. Where an episode may or may not be part of a series and a series may or may not be part of a show. This file is split by crid, title, type (whether its an episode, show or series), language and parent.

The TVA file contains up to two different title values, ‘episodeTitle’ and ‘Main’ title. The titles are chosen using a priority mechanism, where the ‘episodeTitle’ value takes priority if it is present. We do not create additional records for a ‘Main’ title if an ‘episodeTitle’ exists.

This file contains a parent-child hierarchy as follows:

* Show
  + Series
    - Episode

The collection type of ‘OtherCollection’ is currently being filtered out of the file.

The file output is in the following format:

|  |  |
| --- | --- |
| Column Name | Data Format |
| crid | UTF8 String |
| title | UTF8 String |
| type | UTF8 String |
| episodeNumber | UTF8 String |
| parent\_crid | UTF8 String |
| language | UTF8 String |

### CMRT\_content\_genre\_CC\_YYYY-MM-DD.csv

The Content Genre file contains the details of all the associated genres and sub genres for a piece of content. The genres are taken from the <Genre> XML tag, whereas the sub genres are taken from the <Keyword> tag where the text contains the word ‘SubGenre::’

The file output is in the following format:

|  |  |
| --- | --- |
| Column Name | Data Format |
| crid | UTF8 String |
| language | UTF8 String |
| genre\_type | UTF8 String |
| genre\_classification | UTF8 String |
| genre\_code | UTF8 String |
| genre\_value | UTF8 String |

### CMRT\_content\_keywords\_CC\_YYYY-MM-DD.csv

The Content Keyword file contains the details of all the associated keywords for a piece of content. The keywords are taken from the <Keyword> XML tag.

The file output is in the following format:

|  |  |
| --- | --- |
| Column Name | Data Format |
| crid | UTF8 String |
| language | UTF8 String |
| keyword | UTF8 String |

### CMRT\_content\_credits\_CC\_YYYY-MM-DD.csv

The Content Credit file contains the details of all the associated credits for a piece of content. The credits are taken from the <CreditList> XML tag. This data is filtered as the credit list also contains information about the Studio that made the piece of content, to do this a filter on the attribute ‘role’ where it ends with the text ‘ACTOR’.

The file output is in the following format:

|  |  |
| --- | --- |
| Column Name | Data Format |
| crid | UTF8 String |
| language | UTF8 String |
| given\_name | UTF8 String |
| family\_name | UTF8 String |
| presentation\_role | UTF8 String |

### CMRT\_content\_synopsis\_CC\_YYYY-MM-DD.csv

The Content Synopsis file contains the details of all the associated synopsis for a piece of content. The synopsis are taken from the <Synopsis> XML tag. The synopsis have three different types, short, medium or long. This is identified using the length attribute.

The file output is in the following format:

|  |  |
| --- | --- |
| Column Name | Data Format |
| crid | UTF8 String |
| language | UTF8 String |
| given\_name | UTF8 String |
| family\_name | UTF8 String |
| presentation\_role | UTF8 String |

### CMRT\_content\_certification\_CC\_YYYY-MM-DD.csv

The Content Certification file contains the details of the associated content certification or parental guidance for a given piece of content. This is done by using the ParentalGuidance and MinimumAge tags in the content record.

The file output is in the following format:

|  |  |
| --- | --- |
| Column Name | Data Format |
| crid | UTF8 String |
| classification | UTF8 String |
| id | UTF8 String |
| lang | UTF8 String |
| crid | UTF8 String |
| classification | UTF8 String |
| id | UTF8 String |
| lang | UTF8 String |

### CMRT\_content\_distributor\_CC\_YYYY-MM-DD.csv

The Content Distributor file contains the details of all the associated studio, broadcaster and channel for a piece of content. They are taken from the following tags:

* Studio – this is taken from the <CreditList> tag where the attribute ‘role’ ends with ‘STUDIO’
* Channel – this is taken from the <CreditList> tag where the attribute ‘role’ ends with ‘CONTENT-PROVIDER’
* Broadcaster – this is taken from the <CreditList> tag where the attribute ‘href’ ends with ‘Broadcaster’

The file output is in the following format:

|  |  |
| --- | --- |
| Column Name | Data Format |
| crid | UTF8 String |
| studio | UTF8 String |
| broadcaster | UTF8 String |
| channel\_name | UTF8 String |
| lang | UTF8 String |

### CMRT\_content\_schedule\_CC\_YYYY-MM-DD.csv

The Content Schedule file contains the details of all the associated start time and end time for a piece of content. The file output is in the following format:

|  |  |
| --- | --- |
| Column Name | Data Format |
| crid | UTF8 String |
| startTime | UTF8 String |
| endTime | UTF8 String |

### CMRT\_content\_availability\_CC\_YYYY-MM-DD.csv

The Content Availability file contains the details of whether or a not a given piece of content was available on a given system at any time. The following systems are catered for:

* MSS
* TRAXIS
* CMDC
* RENG
* SMARTREC

The file output is in the following format:

|  |  |
| --- | --- |
| Column Name | Data Format |
| crid | UTF8 String |
| request | UTF8 String |
| requestTime | UTF8 String |
| exist | UTF8 String |
| exists\_ | UTF8 String |
| AvailabilityTime | UTF8 String |
| state | UTF8 String |

## ETL Design

The Extraction, Transformation and Loading (ETL) has been performed using Informatica PowerCenter (Version 9.1.0).

### Dependencies

### Mappings

**Repository: P1ETLPE\_S1**

**Folder: KPI\_PE**

**Mapping:**

### Workflow

**Repository: P1ETLPE\_S1**

**Folder: KPI\_PE**

**Workflow:**



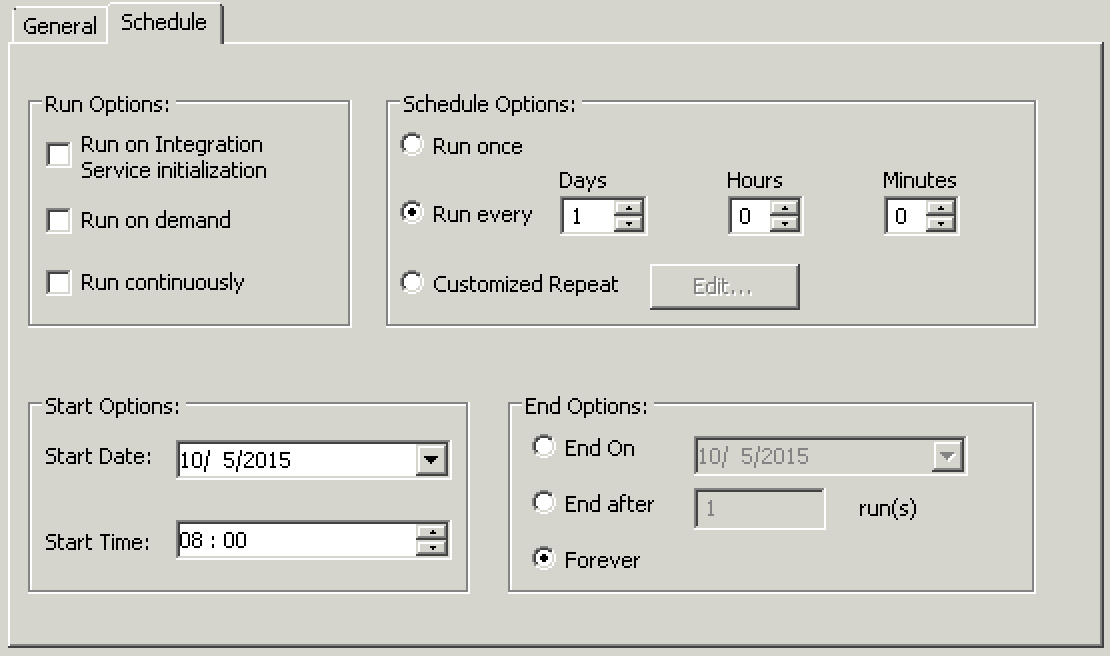
#### Parameter file

Location: /transfer/data/informat/param/KPI\_PE/

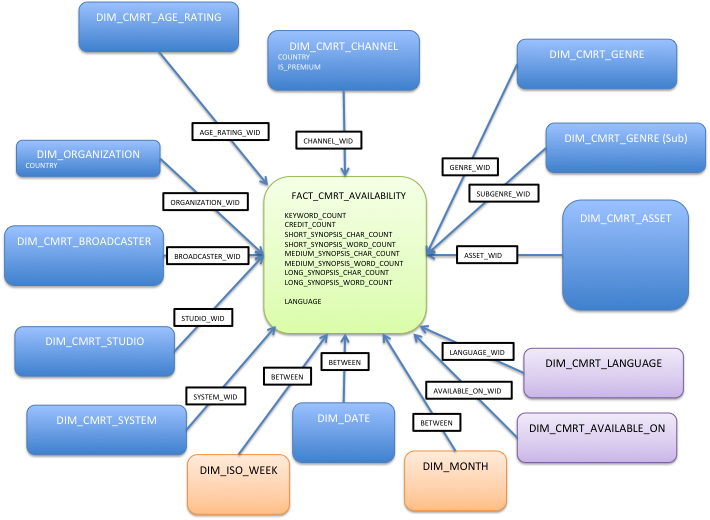
Cache location: /transfer/data/informat/Cache/KPI\_PE/

#### Schedule

Service Availability is scheduled to run each morning at 8:00:



## Data Model Design



# CMRT OBIEE Design

## Repository Design

The OBIEE repository (RPD) is split up into three sections:

* Physical
  + This is a representation of the physical tables on the data mart, along with any relationships between those tables (e.g. joins)
* Business Model and Mapping (BMM)
  + Referencing the physical layer, here is where logical structures are built and aggregation rules are defined.
* Presentation
  + Referencing the BMM, this layer presents the tables and columns to end users. For example, remove unwanted columns or rename awkwardly named columns

### Physical

The physical layer contains the following tables:

* *DIM\_DATE*
* *DIM\_MONTH*
* *DIM\_ISO\_WEEK*
* *DIM\_ORGANIZATION*
* DIM\_CMRT\_LANGUAGE
* DIM\_CMRT\_CHANNEL
* DIM\_CMRT\_GENRE
* DIM\_CMRT\_AGE\_RATING
* DIM\_CMRT\_AVAILABLE\_ON
* DIM\_CMRT\_SYSTEM
* DIM\_CMRT\_BROADCASTER
* DIM\_CMRT\_STUDIO
* DIM\_CMRT\_ASSET
* FACT\_CMRT\_AVAILABILITY

Each table has been aliased (which creates a logical representation of the table) and joined to its relevant aliases. The tables above in italic are common tables and are used as a standard data set amongst multiple projects, this ensures that we have a common set of data points to compare data. For example, the ISO week standard is used across all OBIEE projects.

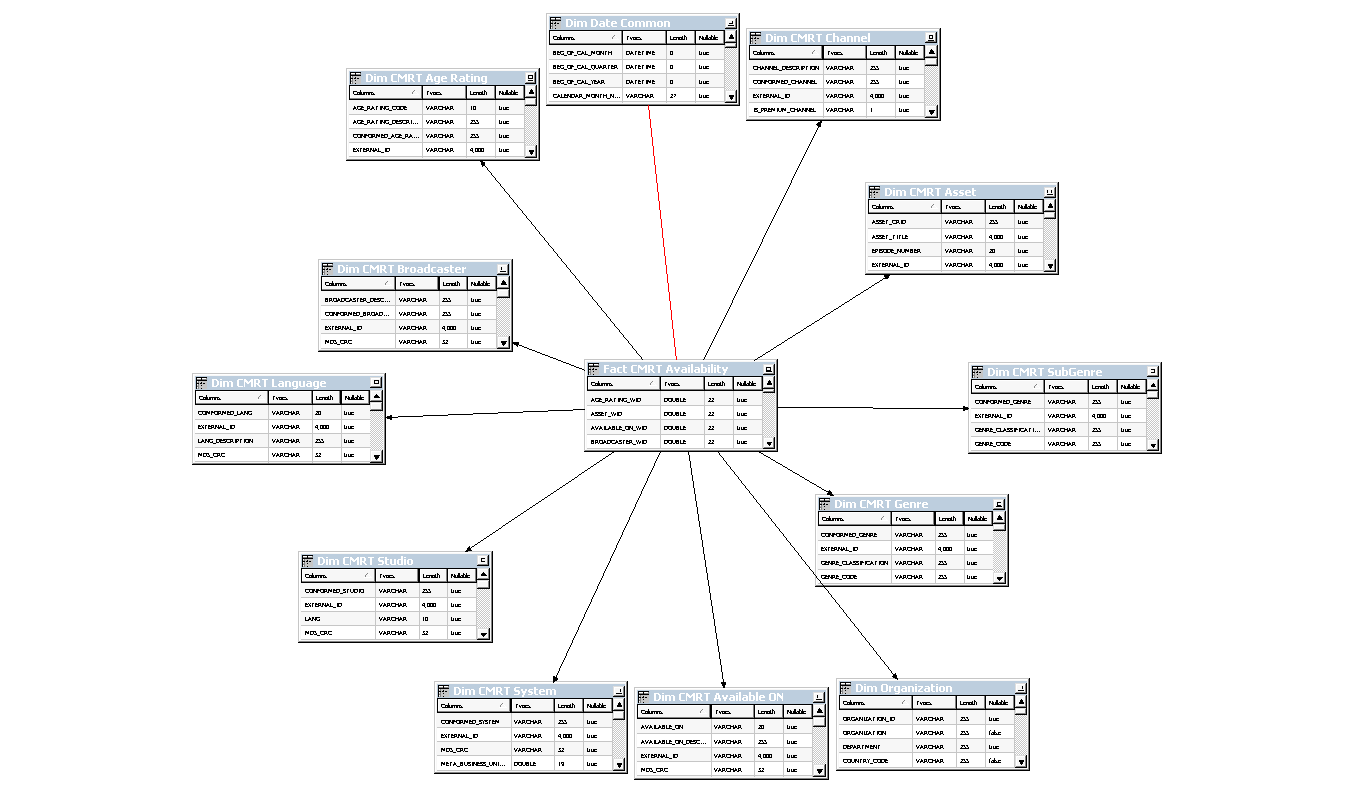
#### Joins

Each of the three star schemas shares a set of common joins, detailed below:

|  |  |  |
| --- | --- | --- |
| Driving Table | Slave Table | Join Criteria |
| "DMPE"."Dim CMRT Channel"."OBJECT\_WID" | "DMPE"."Fact CMRT Availability"."CHANNEL\_WID" | = |
| "DMPE"."Dim CMRT Asset"."OBJECT\_WID" | "DMPE"."Fact CMRT Availability"."ASSET\_WID" | = |
| "DMPE"."Dim CMRT SubGenre"."OBJECT\_WID" | "DMPE"."Fact CMRT Availability"."SUBGENRE\_WID" | = |
| "DMPE"."Dim CMRT Genre"."OBJECT\_WID" | "DMPE"."Fact CMRT Availability"."GENRE\_WID" | = |
| "DMPE"."Dim Organization"."OBJECT\_WID" | "DMPE"."Fact CMRT Availability"."ORGANIZATION\_WID" | = |
| "DMPE"."Dim CMRT Available ON"."OBJECT\_WID" | "DMPE"."Fact CMRT Availability"."AVAILABLE\_ON\_WID" | = |
| "DMPE"."Dim CMRT System"."OBJECT\_WID" | "DMPE"."Fact CMRT Availability"."SYSTEM\_WID" | = |
| "DMPE"."Dim CMRT Studio"."OBJECT\_WID" | "DMPE"."Fact CMRT Availability"."LANGUAGE\_WID" | = |
| "DMPE"."Dim CMRT Language"."OBJECT\_WID" | "DMPE"."Fact CMRT Availability"."LANGUAGE\_WID" | = |
| "DMPE"."Dim CMRT Broadcaster"."OBJECT\_WID" | "DMPE"."Fact CMRT Availability"."BROADCASTER\_WID" | = |
| "DMPE"."Dim CMRT Age Rating"."OBJECT\_WID" | "DMPE"."Fact CMRT Availability"."AGE\_RATING\_WID" | = |

The table FACT\_CMRT\_AVAILABILITY has been aliased three times to cater for three separate time scales. Each one of the FACT\_CMRT\_AVAILABILITY aliases is used to create a star schema, as shown below. This is complimented by a set of three different ‘Date’ dimensions. This enables us to generate more efficient queries at runtime and help sustain performance of the system over a longer period of time. The details of each of these star schemas is shown below.

#### Fact CMRT Availability



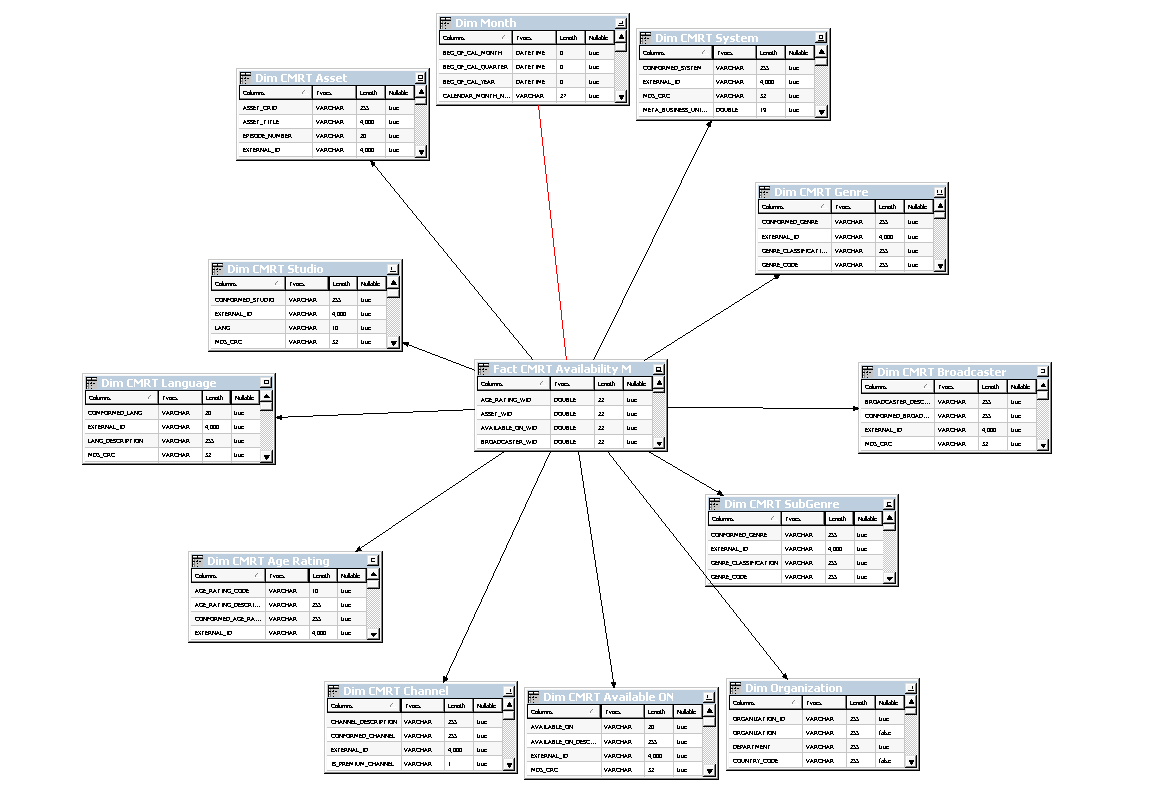
The ‘Fact CMRT Availability’ table is joined to the ‘Dim Date Common’ alias table using the following criteria:

"DMPE"."Fact CMRT Availability"."valid\_start\_date\_wid" <= "DMPE"."Dim Date Common"."OBJECT\_WID" AND

"DMPE"."Fact CMRT Availability"."valid\_end\_date\_wid" >= "."DMPE"."Dim Date Common"."OBJECT\_WID"

This join enables us to set a given fact record to be applicable for a certain date range. In the case of CMRT a given piece of content can be valid on multiple days before it is cycled out of the TVA file.

#### Fact CMRT Availability M



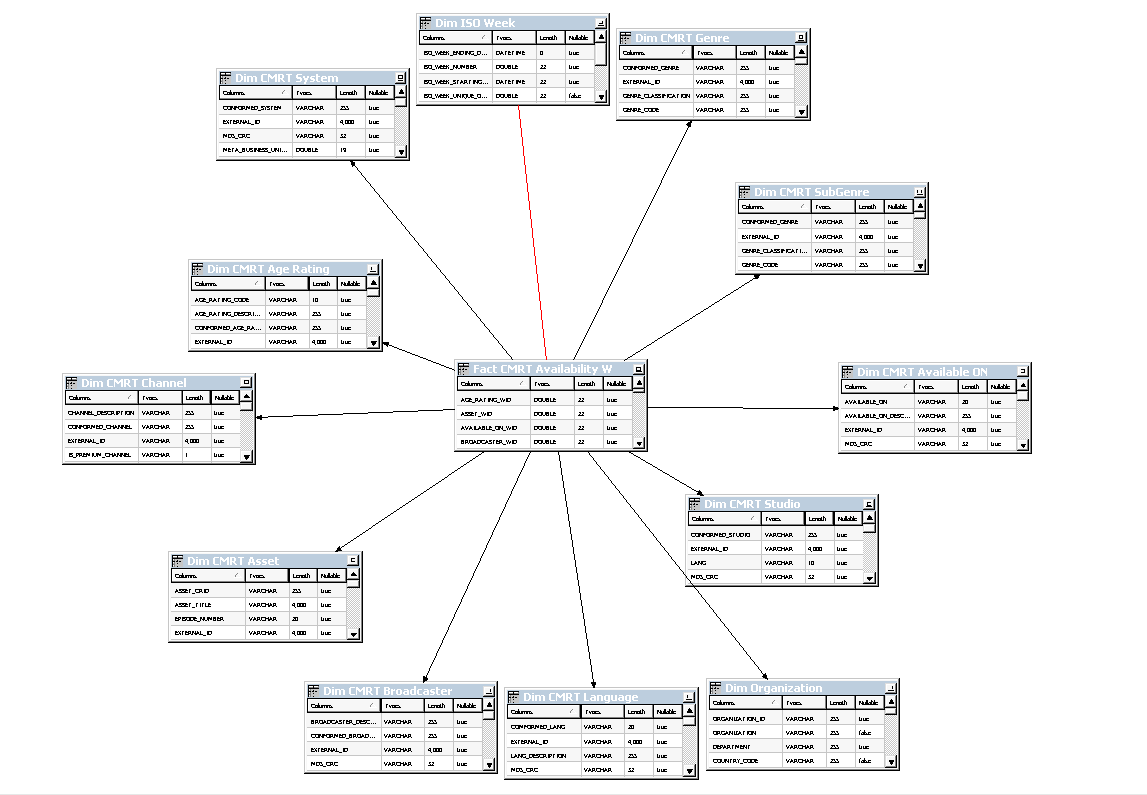
The ‘Fact CMRT Availability M’ table is joined to the ‘Dim Month’ alias table using the following criteria:

"DMPE"."Fact CMRT Availability M"."VALID\_START\_MONTH\_WID" <= "DMPE"."Dim Month"."OBJECT\_WID" AND

"DMPE"."Fact CMRT Availability M"."VALID\_END\_MONTH\_WID" >= "DMPE"."Dim Month"."OBJECT\_WID"

This join enables us to set a given fact record to be applicable for a certain date range. In the case of CMRT a given piece of content can be valid on multiple days before it is cycled out of the TVA file. It also means that we are not going to be doing ~30 joins to this table each time we need to report on a months worth of data.

#### Fact CMRT Availability W



The ‘Fact CMRT Availability W’ table is joined to the ‘Dim ISO Week’ alias table using the following criteria:

"DMPE"."Fact CMRT Availability M"."VALID\_START\_MONTH\_WID" <= "."DMPE"."Dim Month"."OBJECT\_WID" AND

"DMPE"."Fact CMRT Availability M"."VALID\_END\_MONTH\_WID" >= "DMPE"."Dim Month"."OBJECT\_WID"

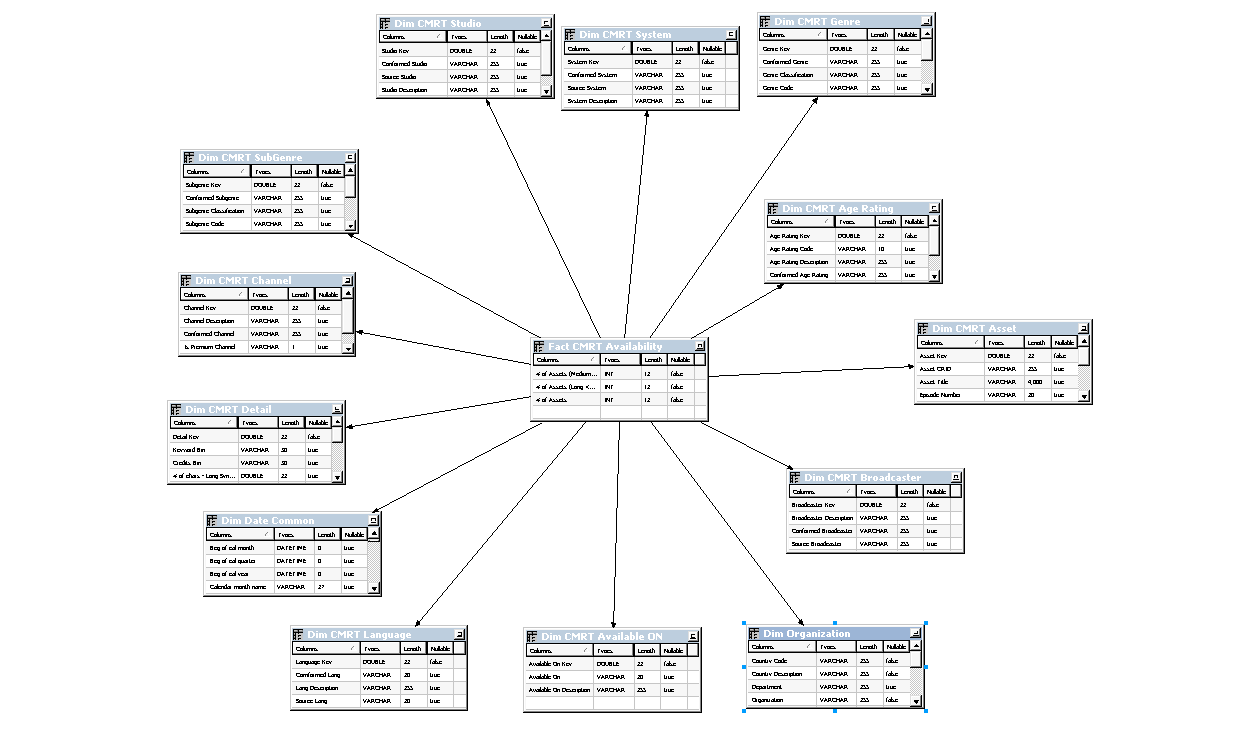
This join enables us to set a given fact record to be applicable for a certain date range. In the case of CMRT a given piece of content can be valid on multiple days before it is cycled out of the TVA file. It also means that we are not going to be doing ~30 joins to this table each time we need to report on a months worth of data.

### Business Model and Mapping Layer

#### Overview

All three star schemas described in the physical layer are consolidated into a single BMM layer. The RPD is designed to use federation to cater for the physical joins between the three facts and the three ‘Date’ dimensions described previously. This is accomplished by having three logical table sources for both the Fact CMRT Availability and Dim Date Common logical tables. OBIEE is then able to extrapolate the most efficient query from the attributes and measures chosen on a report.

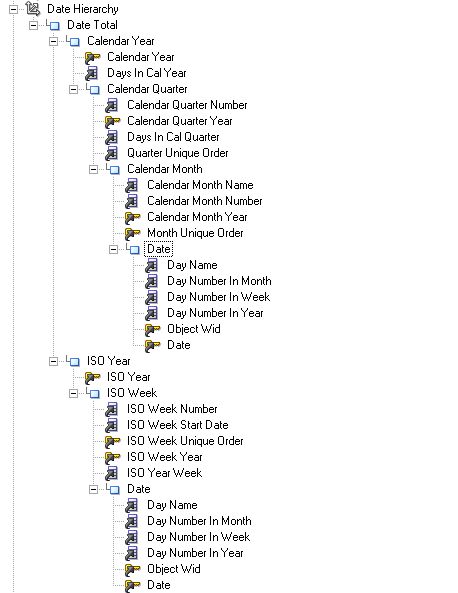
A diagram of the BMM layer is shown below showing the full BMM model, as you can see we now have a single star schema:



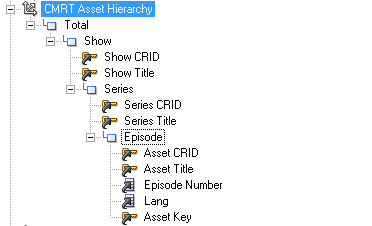
#### Dimensions

Each dimension is given an OBIEE hierarchy with two levels (a total and a detail) by default. The exceptions to this are:

1. The Date hierarchy which has the following structure:

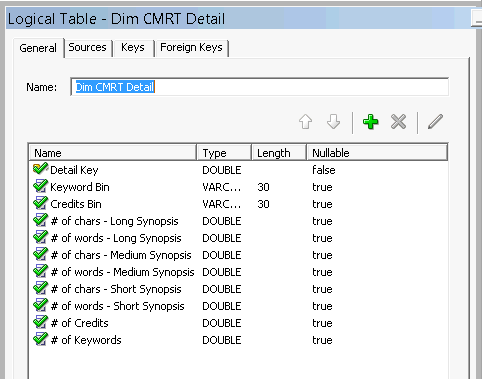


1. Content Asset, which has logic to cater for the relationship between a Show, Series and Episode:



The BMM layer also has one logical dimension, ‘**Dim CMRT Detail’**. This dimension table doesn’t exist on the database or Physical Layer in the RPD, it is based on the fact tables and contains the attributes used to group the counts of measures such as Keywords, Credits, and Synopsis etc. This is implemented as the ‘measures’ are non-aggregable and would be confusing should they be modelled with the main fact table.

As you can see from the screenshot below:



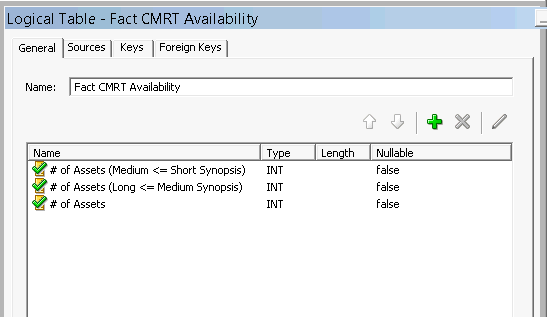
#### Facts

‘Fact CMRT Availability’ contains three logical table sources:

* Fact CMRT Availability W
* Fact CMRT Availability M
* Fact CMRT Availability

These correspond to the three CMRT fact aliases that were described in the Physical Layer.

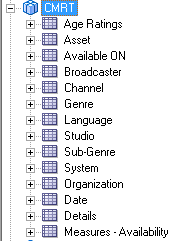
Fact CMRT Availability contains three measures:



|  |  |
| --- | --- |
| Measure Name | Description |
| # of Assets (Medium <= short Synopsis) | The number of assets where the length of the short synopsis is greater than or equal to the length of the medium synopsis. This is split by the dimensions/attributes that are in the particular OBIEE report. |
| # of Assets (Long <= Medium Synopsis) | The number of assets where the length of the medium synopsis is greater than or equal to the length of the long synopsis. This is split by the dimensions/attributes that are in the particular OBIEE report. |
| # of Assets | The number of assets split by the dimensions/attributes that are in the particular OBIEE report. This it the measure that is used in the majority of reports. |

#### Presentation Layer

The presentation layer is called CMRT. This contains all the tables and columns that are accessible to the reporting layer. This is shown below:

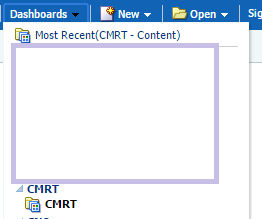


## Reporting Layer

### Overview

The reporting layer contains the physical reports, prompts and filters that are accessible in the front end to create reports and dashboards.

The CMRT dashboard is named ‘CMRT’, and can be accessed from the Dashboards menu, as shown below:

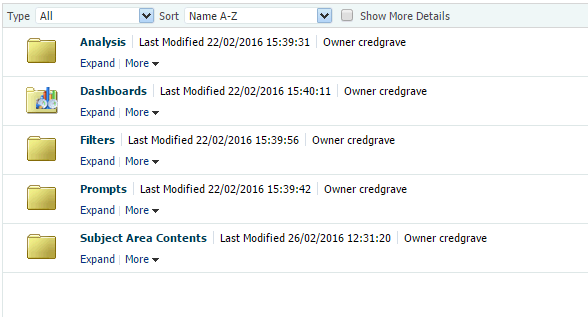


### Catalog

The catalog is called CMRT, and is located in the Shared Folders directory. It is organised into the following sub folders:

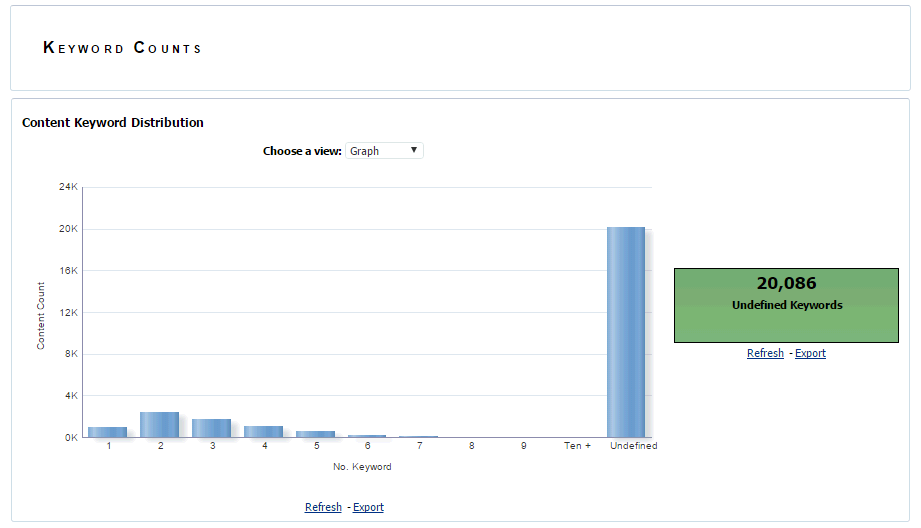
|  |  |
| --- | --- |
| Folder Name | Description |
| Analysis | This contains the report files |
| Dashboards | This contains the dashboard itself |
| Filters | This contains a set of filters that are used to determine which reports determine which reports are displayed on a given dashboard |
| Prompts | This contains prompts that are used to slice and dice the data on a dashboard |
| Subject Area Contents | This contains a set of ‘canned’ filters that can be applied to a report as opposed to applying each individual filter. |

The catalog is shown below:



## Reports

### Keyword Counts



#### Report Columns

|  |  |
| --- | --- |
| Presentation Column | Calculation |
| "Details"."Keyword Bin" |  |
| "Measures - Availability"."# of Assets" |  |

#### Filters Used

|  |  |
| --- | --- |
| Presentation Column | Filter Criteria |
| "Date"."Day" / "Date"."Calendar month year" | Is Prompted |
| "Organization"."Country Description" | Is Prompted |
| "Asset"."Lang" | Is Prompted |
| "Studio"."Conformed Studio" | Is Prompted |
| "Broadcaster"."Conformed Broadcaster" | Is Prompted |
| "Genre"."Conformed Genre" | Is Prompted |
| "Channel"."Conformed Channel" | Is Prompted |
| "Channel"."Is Premium Channel" | Is Prompted |

### Credit Counts



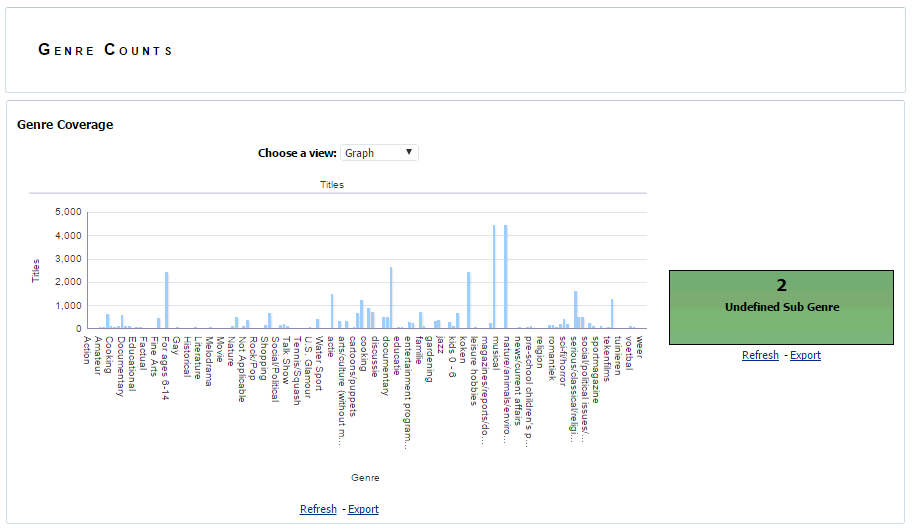
#### Report Columns

|  |  |
| --- | --- |
| Presentation Column | Calculation |
| "Details"."Credit Bin" |  |
| "Measures - Availability"."# of Assets" |  |

#### Filters Used

|  |  |
| --- | --- |
| Presentation Column | Filter Criteria |
| "Date"."Day" / "Date"."Calendar month year" | Is Prompted |
| "Organization"."Country Description" | Is Prompted |
| "Asset"."Lang" | Is Prompted |
| "Studio"."Conformed Studio" | Is Prompted |
| "Broadcaster"."Conformed Broadcaster" | Is Prompted |
| "Genre"."Conformed Genre" | Is Prompted |
| "Channel"."Conformed Channel" | Is Prompted |
| "Channel"."Is Premium Channel" | Is Prompted |

### Genre Counts



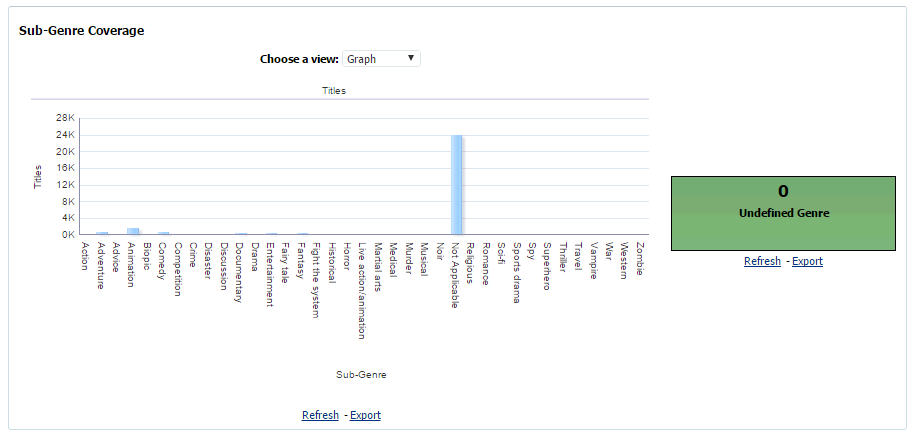
#### Report Columns

|  |  |
| --- | --- |
| Presentation Column | Calculation |
| "Genre"."Conformed Genre" |  |
| "Measures - Availability"."# of Assets" |  |

#### Filters Used

|  |  |
| --- | --- |
| Presentation Column | Filter Criteria |
| "Date"."Day" / "Date"."Calendar month year" | Is Prompted |
| "Organization"."Country Description" | Is Prompted |
| "Asset"."Lang" | Is Prompted |
| "Studio"."Conformed Studio" | Is Prompted |
| "Broadcaster"."Conformed Broadcaster" | Is Prompted |
| "Genre"."Conformed Genre" | Is Prompted |
| "Channel"."Conformed Channel" | Is Prompted |
| "Channel"."Is Premium Channel" | Is Prompted |

### Sub Genre Counts



#### Report Columns

|  |  |
| --- | --- |
| Presentation Column | Calculation |
| "Sub-Genre"."Conformed Genre" |  |
| "Measures - Availability"."# of Assets" |  |

#### Filters Used

|  |  |
| --- | --- |
| Presentation Column | Filter Criteria |
| "Date"."Day" / "Date"."Calendar month year" | Is Prompted |
| "Organization"."Country Description" | Is Prompted |
| "Asset"."Lang" | Is Prompted |
| "Studio"."Conformed Studio" | Is Prompted |
| "Broadcaster"."Conformed Broadcaster" | Is Prompted |
| "Genre"."Conformed Genre" | Is Prompted |
| "Channel"."Conformed Channel" | Is Prompted |
| "Channel"."Is Premium Channel" | Is Prompted |

### Age Ratings Count



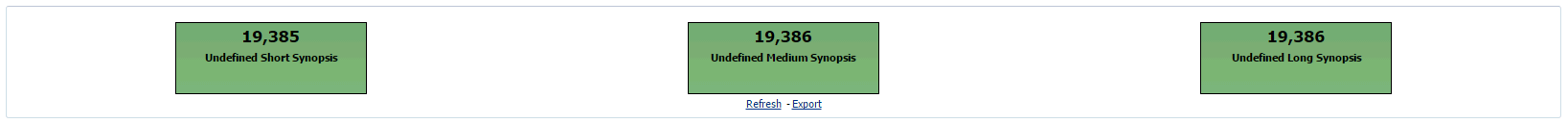
#### Report Columns

|  |  |
| --- | --- |
| Presentation Column | Calculation |
| Age Ratings"."Conformed Age Rating" | IFNULL(Measure,’Undefined’) |
| "Measures - Availability"."# of Assets" |  |

#### Filters Used

|  |  |
| --- | --- |
| Presentation Column | Filter Criteria |
| "Date"."Day" / "Date"."Calendar month year" | Is Prompted |
| "Organization"."Country Description" | Is Prompted |
| "Asset"."Lang" | Is Prompted |
| "Studio"."Conformed Studio" | Is Prompted |
| "Broadcaster"."Conformed Broadcaster" | Is Prompted |
| "Genre"."Conformed Genre" | Is Prompted |
| "Channel"."Conformed Channel" | Is Prompted |
| "Channel"."Is Premium Channel" | Is Prompted |

### Undefined Synopsis Counts



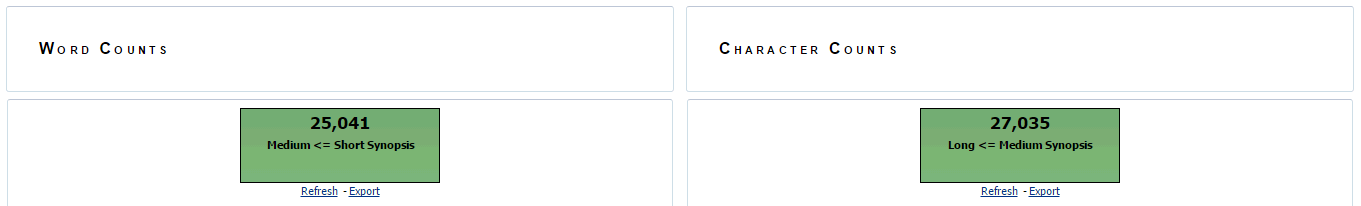
#### Report Columns

|  |  |
| --- | --- |
| Presentation Column | Calculation |
| "Measures - Availability"."# of Assets" | FILTER("Measures - Availability"."# of Assets" USING ("Details"."# of words - Short Synopsis" = 0)) |
| "Measures - Availability"."# of Assets" | FILTER("Measures - Availability"."# of Assets" USING ("Details"."# of words - Medium Synopsis" = 0)) |
| "Measures - Availability"."# of Assets" | FILTER("Measures - Availability"."# of Assets" USING ("Details"."# of words - Long Synopsis" = 0)) |

#### Filters Used

|  |  |
| --- | --- |
| Presentation Column | Filter Criteria |
| "Date"."Day" / "Date"."Calendar month year" | Is Prompted |
| "Organization"."Country Description" | Is Prompted |
| "Asset"."Lang" | Is Prompted |
| "Studio"."Conformed Studio" | Is Prompted |
| "Broadcaster"."Conformed Broadcaster" | Is Prompted |
| "Genre"."Conformed Genre" | Is Prompted |
| "Channel"."Conformed Channel" | Is Prompted |
| "Channel"."Is Premium Channel" | Is Prompted |

### Synopsis Length Counts



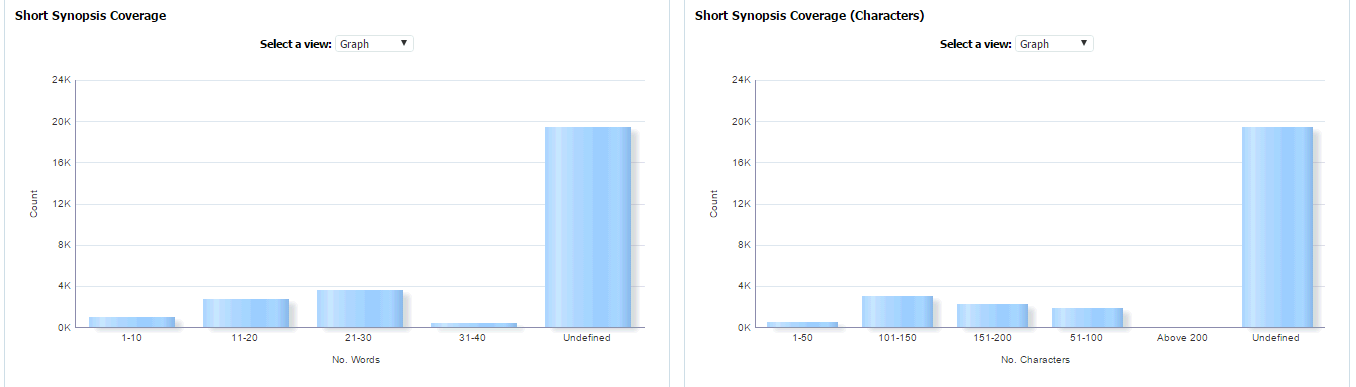
#### Report Columns

|  |  |
| --- | --- |
| Presentation Column | Calculation |
| "Measures - Availability"."# of Assets (Medium <= Short Synopsis)" |  |
| "Measures - Availability"."# of Assets (Long <= Medium Synopsis)" |  |

#### Filters Used

|  |  |
| --- | --- |
| Presentation Column | Filter Criteria |
| "Date"."Day" / "Date"."Calendar month year" | Is Prompted |
| "Organization"."Country Description" | Is Prompted |
| "Asset"."Lang" | Is Prompted |
| "Studio"."Conformed Studio" | Is Prompted |
| "Broadcaster"."Conformed Broadcaster" | Is Prompted |
| "Genre"."Conformed Genre" | Is Prompted |
| "Channel"."Conformed Channel" | Is Prompted |
| "Channel"."Is Premium Channel" | Is Prompted |

### Short Synopsis Coverage



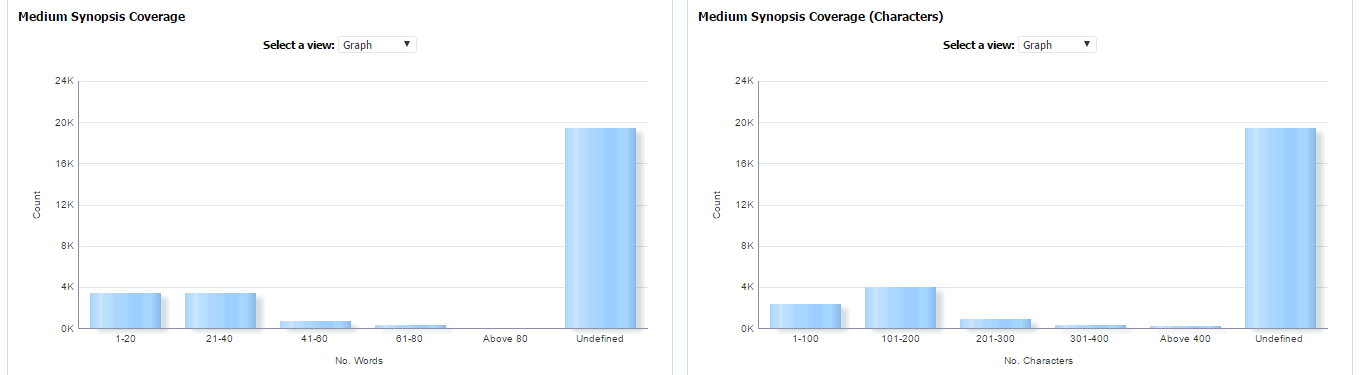
#### Report Columns

|  |  |
| --- | --- |
| Presentation Column | Calculation |
| "Details"."# of words - Short Synopsis" | CASE  WHEN "Details"."# of words - Short Synopsis" BETWEEN 1 AND 10 THEN '1-10'  WHEN "Details"."# of words - Short Synopsis" BETWEEN 11 AND 20 THEN '11-20'  WHEN "Details"."# of words - Short Synopsis" BETWEEN 21 AND 30 THEN '21-30'  WHEN "Details"."# of words - Short Synopsis" BETWEEN 31 AND 40 THEN '31-40' WHEN "Details"."# of words - Short Synopsis" > 40  THEN 'Above 40' WHEN "Details"."# of words - Short Synopsis" = 0 THEN 'Undefined'  ELSE ''  END |
| "Details"."# of chars - Short Synopsis" | CASE  WHEN "Details"."# of chars - Short Synopsis" BETWEEN 1 AND 50 THEN '1-50'  WHEN "Details"."# of chars - Short Synopsis" BETWEEN 51 AND 100 THEN '51-100'  WHEN "Details"."# of chars - Short Synopsis" BETWEEN 101 AND 150 THEN '101-150'  WHEN "Details"."# of chars - Short Synopsis" BETWEEN 151 AND 200 THEN '151-200'  WHEN "Details"."# of chars - Short Synopsis" > 200 THEN 'Above 200'  WHEN "Details"."# of chars - Short Synopsis" = 0 THEN 'Undefined'  ELSE ''  END |
| "Measures - Availability"."# of Assets" |  |

#### Filters Used

|  |  |
| --- | --- |
| Presentation Column | Filter Criteria |
| "Date"."Day" / "Date"."Calendar month year" | Is Prompted |
| "Organization"."Country Description" | Is Prompted |
| "Asset"."Lang" | Is Prompted |
| "Studio"."Conformed Studio" | Is Prompted |
| "Broadcaster"."Conformed Broadcaster" | Is Prompted |
| "Genre"."Conformed Genre" | Is Prompted |
| "Channel"."Conformed Channel" | Is Prompted |
| "Channel"."Is Premium Channel" | Is Prompted |

### Medium Synopsis Coverage



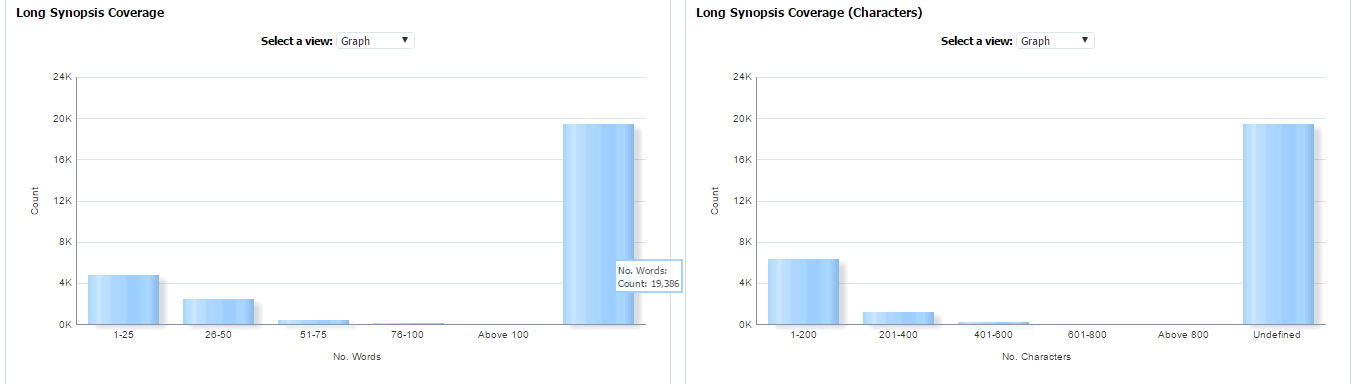
#### Report Columns

|  |  |
| --- | --- |
| Presentation Column | Calculation |
| "Details"."# of words - Medium Synopsis" | CASE  WHEN "Details"."# of words - Medium Synopsis" BETWEEN 1 AND 20 THEN '1-20'  WHEN "Details"."# of words - Medium Synopsis" BETWEEN 21 AND 40 THEN '21-40'  WHEN "Details"."# of words - Medium Synopsis" BETWEEN 41 AND 60 THEN '41-60'  WHEN "Details"."# of words - Medium Synopsis" BETWEEN 61 AND 80 THEN '61-80'  WHEN "Details"."# of words - Medium Synopsis" > 80 THEN 'Above 80'  WHEN "Details"."# of words - Medium Synopsis" = 0 THEN 'Undefined'  ELSE ''  END |
| "Details"."# of chars - Medium Synopsis" | CASE  WHEN "Details"."# of chars - Medium Synopsis" BETWEEN 1 AND 100 THEN '1-100'  WHEN "Details"."# of chars - Medium Synopsis" BETWEEN 101 AND 200 THEN '101-200'  WHEN "Details"."# of chars - Medium Synopsis" BETWEEN 201 AND 300 THEN '201-300'  WHEN "Details"."# of chars - Medium Synopsis" BETWEEN 301 AND 400 THEN '301-400'  WHEN "Details"."# of chars - Medium Synopsis" > 400 THEN 'Above 400'  WHEN "Details"."# of chars - Medium Synopsis" = 0 THEN 'Undefined'  ELSE ''  END |
| "Measures - Availability"."# of Assets" |  |

#### Filters Used

|  |  |
| --- | --- |
| Presentation Column | Filter Criteria |
| "Date"."Day" / "Date"."Calendar month year" | Is Prompted |
| "Organization"."Country Description" | Is Prompted |
| "Asset"."Lang" | Is Prompted |
| "Studio"."Conformed Studio" | Is Prompted |
| "Broadcaster"."Conformed Broadcaster" | Is Prompted |
| "Genre"."Conformed Genre" | Is Prompted |
| "Channel"."Conformed Channel" | Is Prompted |
| "Channel"."Is Premium Channel" | Is Prompted |

### Long Synopsis Coverage



#### Report Columns

|  |  |
| --- | --- |
| Presentation Column | Calculation |
| "Details"."# of words - Long Synopsis" | CASE  WHEN "Details"."# of words - Long Synopsis" BETWEEN 1 AND 25 THEN '1-25'  WHEN "Details"."# of words - Long Synopsis" BETWEEN 26 AND 50 THEN '26-50'  WHEN "Details"."# of words - Long Synopsis" BETWEEN 51 AND 75 THEN '51-75'  WHEN "Details"."# of words - Long Synopsis" BETWEEN 76 AND 100 THEN '76-100'  WHEN "Details"."# of words - Long Synopsis" > 100 THEN 'Above 100'  ELSE ''  END |
| "Details"."# of chars - Long Synopsis" | CASE  WHEN "Details"."# of chars - Long Synopsis" BETWEEN 1 AND 200 THEN '1-200'  WHEN "Details"."# of chars - Long Synopsis" BETWEEN 201 AND 400 THEN '201-400'  WHEN "Details"."# of chars - Long Synopsis" BETWEEN 401 AND 600 THEN '401-600'  WHEN "Details"."# of chars - Long Synopsis" BETWEEN 601 AND 800 THEN '601-800'  WHEN "Details"."# of chars - Long Synopsis" > 800 THEN 'Above 800'  WHEN "Details"."# of chars - Long Synopsis" = 0 THEN 'Undefined'  ELSE ''  END |
| "Measures - Availability"."# of Assets" |  |

#### Filters Used

|  |  |
| --- | --- |
| Presentation Column | Filter Criteria |
| "Date"."Day" / "Date"."Calendar month year" | Is Prompted |
| "Organization"."Country Description" | Is Prompted |
| "Asset"."Lang" | Is Prompted |
| "Studio"."Conformed Studio" | Is Prompted |
| "Broadcaster"."Conformed Broadcaster" | Is Prompted |
| "Genre"."Conformed Genre" | Is Prompted |
| "Channel"."Conformed Channel" | Is Prompted |
| "Channel"."Is Premium Channel" | Is Prompted |

## Prompts

There are currently three prompts on the dashboard:

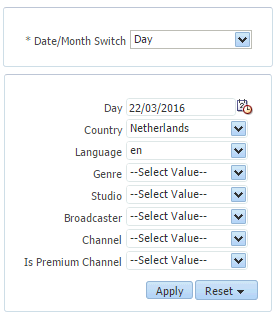
* Timescale Switch – this enables the user to switch between the month based view of the dashboard and the date based view of the dashboard
* Month Based prompt – this contains the standard set of prompts plus a prompt for the Month
* Date Based prompt - this contains the standard set of prompts plus a prompt for the Date

The standard set of prompts are

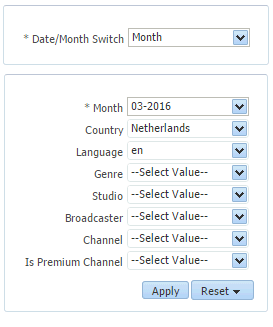
|  |  |
| --- | --- |
| Prompt | Default |
| "Organization"."Country Description" | Netherlands |
| "Asset"."Lang" | en |
| "Studio"."Conformed Studio" | All Values |
| "Broadcaster"."Conformed Broadcaster" | All Values |
| "Genre"."Conformed Genre" | All Values |
| "Channel"."Conformed Channel" | All Values |
| "Channel"."Is Premium Channel" | All Values |

The month and date prompt values are defaulted to the current date/month. As you can see from the screenshots below, the user will only see the Timescale Switch and Month OR Date prompts:

#### Date Switch



#### Month Switch



# Issues

# APPENDIX A