# Multi-Channel Digital Video Data Capture, Accumulation and Processing Guidelines

Media Rating Council, Inc.

NCC media

**June 1, 2012**

**(Published on the Media Rating Council Website)**

## Table of Contents


**1 Background ..... 4**
**1.1 Objective ..... 4**
1.1.1 Scope of Data ..... 4
1.1.2 Uses of Data ..... 5
1.1.3 Applicability of Guidelines ..... 5
**1.2 Revision Process ..... 5**
**2 Definitions ..... 5**
**3 Guidelines for Linear Data ..... 8**
**3.1 General Description of Linear Data ..... 8**
3.1.1 In-House vs. Agent-Based Processing ..... 9
3.1.2 Household Environment ..... 10
3.1.3 Data Collection & Processing Environment ..... 12
**3.2 Linear Data Capture ..... 14**
3.2.1 Initial Capture Controls ..... 14
**3.3 Linear Data Accumulation ..... 14**
3.3.1 Initial Data Collection ..... 14
3.3.2 Initial Data Storage ..... 17
3.3.3 Consolidation & Front End Processing ..... 17
**3.4 Linear Data Processing ..... 18**
3.4.1 Error Processing ..... 18
3.4.2 Data Correlation and Supplementation ..... 18
3.4.3 Data Editing and Smoothing ..... 21
3.4.4 Storing Processed Linear Data ..... 24
3.4.5 Linear Data Inspection ..... 25
3.4.6 Data Release ..... 26
**3.5 Guidelines Requirements for Linear Data Set Processing ..... 29**
3.5.1 General ..... 29
3.5.2 Data Capture ..... 30
3.5.3 Accumulation ..... 30
3.5.4 Processing ..... 32
**4 Guidelines for Digital Video Recorder Time-Shifted Data ..... 39**
**4.1 General Description of Digital Video Recorder Time-Shifted Data ..... 39**
4.1.1 Impact of the Household Environment on Time-Shifted Viewership ..... 41
4.1.2 Other Digital Video Programming Delivery Streams ..... 42
**4.2 Time-Shifted Data Accumulation ..... 42**
4.2.1 Initial Data Collection ..... 42
4.2.2 Initial Data Store ..... 43

4.2.3 Consolidation & Front End Processing ..... 43
4.2.4 Storing Consolidated Data ..... 44
**4.3 Time-Shifted Data Processing ..... 44**
4.3.1 Data Correlation and Supplementation ..... 45
4.3.2 Data Editing and Smoothing ..... 45
4.3.3 Storing Processed Time-Shifted Data ..... 45
4.3.4 Time-Shifted Data Inspection ..... 45
4.3.5 Final Time-Shifted Data Processing and Data Set Release ..... 46
**4.4 Guidelines Requirements for Time Shifted Data Set Processing (Additive to Requirements in Section 3.5) ..... 46**
4.4.1 Accumulation ..... 46
4.4.2 Processing ..... 47
**5 Guidelines for Video on Demand Data ..... 49**
**5.1 General Description of Video on Demand Data ..... 49**
5.1.1 Impact of the Household Environment on VOD Viewership ..... 51
5.1.2 Other Digital Video Programming Delivery Streams ..... 52
**5.2 Supplying VOD Base Data and Video on Demand Accumulation ..... 52**
5.2.1 Initial Data Collection ..... 53
5.2.2 Initial Data Store ..... 55
5.2.3 Consolidation & Front End Processing ..... 55
5.2.4 Storing Consolidated Data ..... 55
**5.3 Video on Demand Data Processing ..... 55**
5.3.1 Data Correlation and Supplementation ..... 56
5.3.2 Data Editing and Smoothing ..... 56
5.3.3 Storing Processed Video on Demand Data ..... 56
5.3.4 Time-Shifted Data Inspection ..... 56
5.3.5 Final Video on Demand Data Processing and Reporting ..... 56
**5.4 Guidelines Requirements for Video On Demand Data Set Processing (Additive to Requirements in Section 3.5 and Section 4.4) ..... 56**
5.4.1 Accumulation ..... 57
5.4.2 Processing ..... 58
**6 Privacy Considerations ..... 59**
6.1 Consumer Notice and Interactions ..... 59
6.2 Privacy Policy Requirements ..... 60
6.3 Privacy Policy Links and Contacts ..... 60
**7 Auditing/Validation Guidance ..... 60**
**8 Conclusions and Contacts ..... 61**

# 1 Background

Consistent and fully understood digital video tuning data is critical to establish greater usage of this valuable data. Beginning in August 2010 the Media Rating Council (MRC), NCC Media and a large group of Multi-channel Video Programming Distributors (MVPDs) initiated a project to establish greater consistency in capturing, accumulating and processing multi-channel digital video tuning data, and the associated activity, which led to these guidelines.

The project had two important phases:

1. Identification and categorization of practices used by the MVPD project participants. This information was gathered through MRC staff inquiries and observations and the ongoing cooperation of a team of MVPD personnel referred to as the "working group."
2. Discussion and vetting of guidelines originally prepared by the MRC staff by a wide group of organizations including the MVPDs, NCC Media and many other organizations who were interested participants. Additionally, drafts of the guidelines were circulated to other organizations such as the AAAA and the ANA at the appropriate time as well as a public comment period.

## 1.1 Objective

This document was prepared to establish an Industry guideline for the capture, accumulation and processing of multi-channel digital video tuning data, and associated activity records, to help foster understanding and the knowledgeable, responsible use of this data by interested parties.

The guidelines identify and describe the business/operational cycles and practices used in developing this foundational data. Definitions, processing and internal control procedures and disclosures are also described – all of which are intended to develop greater understanding of these procedures and encourage the adoption of consistent methods.

The existence of these guidelines is not intended to imply that all MVPDs are monitoring, collecting or processing STB tuning data.

### 1.1.1 Scope of Data

The guidelines are intended to cover three types of digital viewership data, regardless of MVPD source:

1. Linear Viewership
2. Time-Shifted Viewership
3. Video on Demand Viewership (Non Linear Demand Viewership)
4. Reference Data

The stages of capturing, accumulating and processing are described for each type of data separately.

It should be noted that references to tuning and viewing made in this document are not intended to imply that the distributed content is displayed on the screen and available to viewers in all instances. While there is evidence that content is delivered to the STB, there are situations where the television set may be off, equipment malfunctions or equipment is configured improperly which would prevent that content from being displayed on the screen.

### 1.1.2 Uses of Data

Potential applications of digital viewership data include: advertising measurement, marketing, identifying consumer preferences and optimizing content distribution based on these preferences.

Some applications require analysis of captured tuning data from specific devices and specific accounts; for instance, consumer preferences driving content recommendations for a viewing household. Some applications may only require a representative sampling of a population of devices and accounts; for instance, to calculate a rating for a program. Note that sampling in general must be performed in specific ways in order to be representative of a wider population.

### 1.1.3 Applicability of Guidelines

The guideline is principally applicable to Multi-Channel Video Programming Distributors, herein referred to as "MVPDs "(Cable MSOs, Satellite Providers, Telephone Companies), and is intended as a guide to accepted practice, with certain best practices identified, as developed by the MRC and other project participants.

Additionally, data users such as measurement organizations, buyers and sellers of advertising, as well as the MVPDs themselves can use this document to assist in determining the quality attributes of captured, accumulated and processed digital video tuning data.

## 1.2 Revision Process

The MRC, NCC Media and MVPDs, as well as other organizations that express interest, intend to establish an ongoing process to refine and maintain these guidelines in the future.

# 2 Definitions

Definitions for key terms used throughout this document are listed below. A more complete set of definitions associated with STB data is available through the Coalition for Innovative Media Measurement (CIMM) and can be found at http://www.cimm-us.org/lexicon.htm

Anonymization/De-Identification - A series of systems, processes and controls that remove or replace personally identifiable data components from a data set thereby eliminating or reducing the likelihood that certain characteristics or attributes, whether by themselves or in

combination with other information, can be mapped to specific identifiable individuals. (Source: CIMM)

Data Supplementation/Enrichment – A process by which additional information is appended to tuning event records. It is paramount that this process be executed in a manner that protects the identity of the underlying household data.

DVR (Digital Video Recorder) - A device that allows a user to record video content to a hard drive or other storage medium within the device that can be viewed at a later time and with trick play functionality (e.g. fast forward, rewind, pause, etc)

Editing Rules – A pre-determined set of rules governing how raw data collected is modified to ensure it is complete, accurate and consistent. Editing rules may be used to check that essential fields are populated, verify values fall within pre-established limits, or identify the existence of duplicate records, among other things.

Encryption - A protection procedure where an algorithm is inserted into a process or dataset that will allow access to the content only to those who have the security key. (Source: CIMM)

FTP (File Transfer Protocol) – A network protocol that facilitates the exchange of data files from one host to another.

GUID (Globally Unique Identifier) - A unique and persistent number that may be associated with a device, household, or other entity.

Headend - The control center of a cable television system, where incoming signals are amplified, converted, processed and combined into a common cable along with any original cablecasting, for transmission to subscribers. The system usually includes antennas, preamplifiers, frequency converters, demodulators, modulators, processors and other related equipment. Or the central location on the cable network that is responsible for injecting broadcast video and other signals in the downstream direction. (Source: CableLabs)

Linear Television - Programming content delivered in real-time, as scheduled by the distributor

Linear Viewing - Viewing of programming or content in real-time, not delayed or played back. (Source: CIMM)

Metadata - Provides information about other data. In the media ecosystem, Metadata generally refers to information about content, such as title, sources, actors and other attributes. Measurement data may be associated with content metadata to provide context to the measurements

MAC Address (Media Access Control Address) - A unique identifying code assigned to a Networked device (e.g. Set-Top Box) by the manufacturer. The Set-Top Box ID.

MRC - The MRC is a non-profit Industry association established in 1964 composed of leading television, radio, print and internet companies, as well as advertisers, advertising agencies and trade associations whose goal is to ensure measurement services that are valid, reliable and effective.

MVPD - A Multichannel Video Programming Distributor delivering video programming services, usually for a subscription fee. These operators include cable television (CATV) systems, direct-broadcast satellite (DBS) providers, and wireline video providers, including telcos and competitive local exchange carriers (CLECs) using IPTV.

MVPD Data Stream – Represents the set of continuous data collected from STBs and associated ancillary information under the control of the MVPD as part of its processing. Data streams can represent linear or time-shifted data.

NCC Media - NCC Media is the ad sales, marketing and technology organization that represents MVPDs in every U.S. market. NCC Media meets marketers' needs to target consumer prospects efficiently and effectively in the communities where they live.

Non-Linear Programming - Programming that is made available to a viewer upon request. Non-linear programming implementations include network-based VOD and DVR, local DVR, Home Networked file streaming, Web Video, and others.

NTP (abbr Network Time Protocol): A protocol that uses algorithms to synchronize the time of a computer clock or other electronic devices with an internal clock with other devices within a network. Uses Coordinated Universal Time Code (UTC) to help in the synchronization of time often to a fraction of a millisecond. (Source: CIMM)

Reference Data – Information that may be associated with tuning event records during the enrichment process. An example would be program schedule data.

Return Path (also referred to as Back Channel): A communication channel that can be used by a Set-Top Box to communicate with the cable headend or a service provider. Some homes and certain types of devices (e.g. non digital Set-Top-Box) do not have return path capability. Return path communication in Satellite homes is facilitated through landline phone lines or an independent broadband connection.

Set Top Box (STB) – A physical device attached externally to a television, or resident internally, which when connected to an external source of digital or analog signal enables the display of video content on the television screen.

Smoothing – A technique that tries to compensate for the differences in successive backchannel pings from the Set-Top Box for measurement and service purposes. (Source: CIMM).

Subscriber - A contracted consumer of a particular product or service. (Source: Nielsen)

Subscription On-Demand (SVOD) – Enables subscribers to a particular tuning source to view content from that source on demand.

Switched Digital Video - A method of distributing digital video, utilizing bandwidth more efficiently by only broadcasting channels that have been requested by at least one household within a cable service area. (Source: CableLabs)

TCP/IP (Transition Control Protocol/Internet Protocol) – Part of series of protocols designed to connect different computer networks enabling functions such as e-mail, and data file transfers

Time Shifted Television - Content that has been recorded with the intent to view at another time in the future. Also, the playing back of such content. (Source: CIMM)

Transactional On-Demand (Definition Needed) – The user pays a fee to view particular content, and that content can generally be viewed multiple times within a specified timeframe (e.g. 24 hrs).

Tuning - Tuning, as opposed to viewing, is an indicator that the device (e.g. Television, Set-Top Box) is turned on to a programming source or other content and it may be unclear whether anyone is viewing at the time. Viewing is an indictor of persons level activity.

Tuning Duration - The length of time that a device (e.g. Television, Set-Top Box) is tuned into a programming source or other content. Duration is impacted by latency and can differ from viewing duration which is a measure of persons activity.

Tuning Event (Tune In, Tune Out) – Point in time events that identify any access to a single programming source or content for a duration of time. Tuning event generally have a start time, end time, and a duration

UTC (abbr Universal Time Code, also referred to as Coordinated Universal Time [with same abbreviation]) Coordinated Universal Time is a time measurement standard. Some internal Set-Top Box clocks are on UTC which replaced Greenwich Mean Time (GMT) as the main reference time scale (Source: CIMM)

VOD (Video On-Demand – non-linear demand viewership) - A service that may be provided by an operator that enables digital subscribers to access programming assets on channels that can be viewed at the viewers convenience. Video On-Demand content may be delivered to the end user at the time of request via multiple platforms such as cable, satellite, IPTV, Internet, Mobile, etc., Sometimes these channels are Free On-Demand and sometimes they are Subscription on Demand (SVOD) or Transactional On-Demand. Content can be streamed either through a Set-Top Box (allowing viewing in real-time) or downloaded to a device such as a computer or DVR (for future viewing). Internet television, accessed online, is an increasingly popular form of VOD.

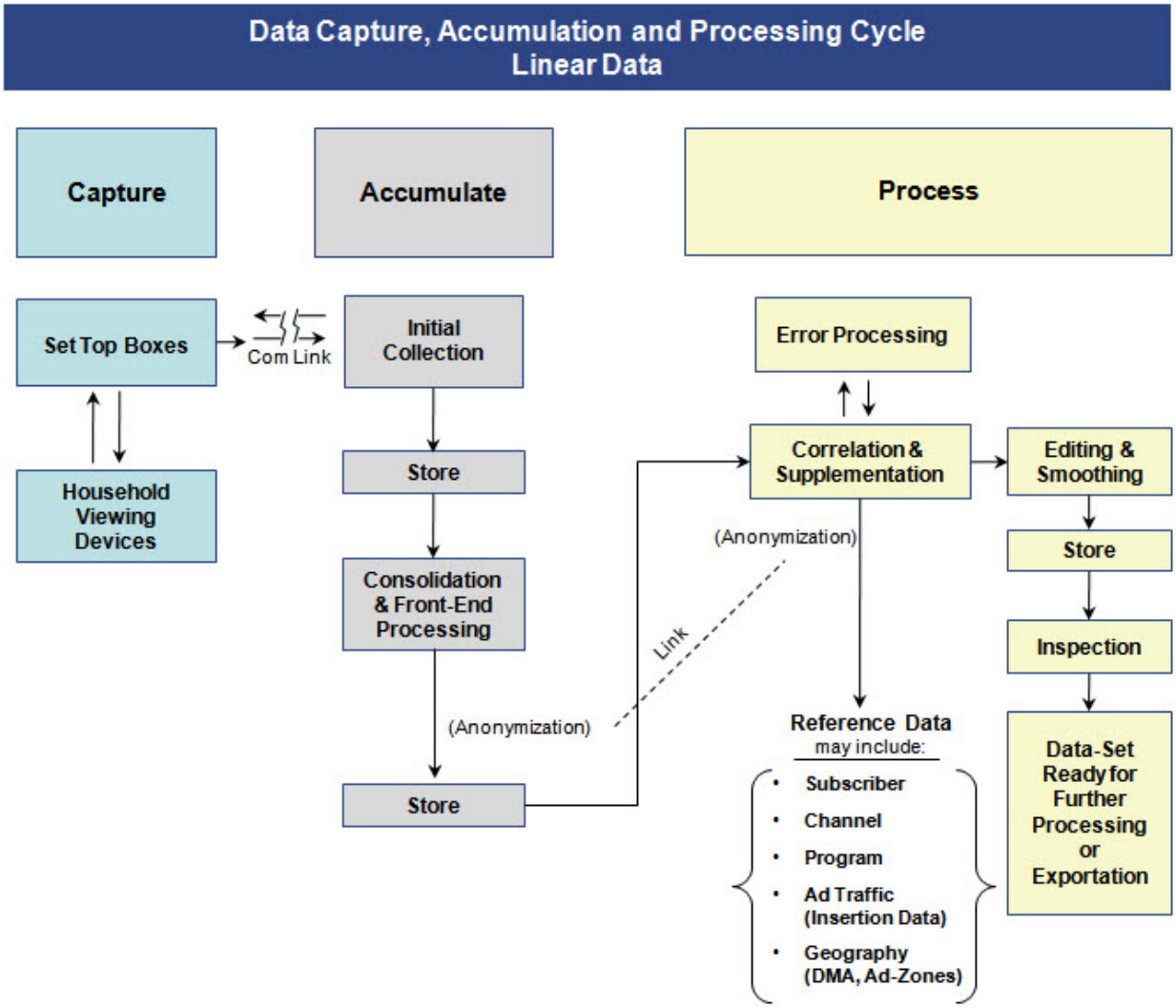
# 3 Guidelines for Linear Data

This section focuses on the capture, accumulation and processing of linear viewership data. In addition to a general description and overview, specific details will address the steps within each of these stages that will ensure high-quality, consistent data for delivery.

## 3.1 General Description of Linear Data

The following chart illustrates the data-flow for the capture, accumulation and processing of linear data, and associated activity.

**Figure 1. Linear Data Flow (For Illustrative Purposes as some processes may vary)**




As illustrated in the figure above, the process of preparing digital tuning data for delivery from an MVPD or its agent to an affiliate is accomplished in several stages. The remainder of Section 3 addresses the distinct steps associated with each stage of the process

### 3.1.1 In-House vs. Agent-Based Processing

MVPDs can choose to process multi-channel digital video tuning data, and associated activity, themselves – referred to as In-House Processing – or they can assign/contract all or part of the processing tasks to a third-party – referred to as Agent-Based Processing.

In either processing environment, it is important to note that the responsibility for accurate capture, accumulation and editing remains with the data owner (generally the MVPD, or sometimes referred to as the seller). Adequate internal controls, error investigation, processing and disclosures should be maintained regardless of processing environment.

#### *3.1.1.1 Agent-Based Processing*

Adequate communication and reporting should be maintained at various stages of processing in agent-based systems so that the MVPD can monitor the work of the agent. The frequency of this communication should align with the frequency of the processing tasks of the agent, so for example daily processing should necessitate daily monitoring and reporting from the agent to the MVPD. This type of communication facilitates the timely investigation of errors or missing data by the MVPD, helping to ensure data-capture completeness.

Agent-based processors can begin to interact within the MVPD data stream as early as on the set top box, where specialized agent software, can be loaded to facilitate data capture and transfer. In some cases, agent-based processing can occur after data capture by the MVPD.

#### *3.1.1.2 In-House Processing*

In in-house applications, the tasks within the data stream are maintained by the MVPD, including monitoring functions for data-capture completeness. Data capture and associated data consolidation functions can be performed real-time (or near real-time through back channels) through facilitations such as TCP/IP or can be based on batch information transfer through facilitations such as FTP. Batching will require storage capacity on MVPD equipment, in which case adequate storage provisions should be made to help ensure communication failures do not cause data loss.

Real time applications should have provisions to protect against data loss due to temporary loss of communication/contact with collection assets of the MVPD.

MVPDs should maintain back-up and retrieval procedures to ensure that collected data in processed form is available post processing based on the retention requirements specified herein.

### 3.1.2 Household Environment

MVPDs typically place electronic equipment into subscribing households (generally referred to as set-top-boxes, or set-top-boxes with DVRs), which facilitate connectivity/addressability, content delivery, recording, and return-path transfer of subscriber interaction data and associated transactions. The nature of these devices can vary greatly depending on manufacturer, MVPD, other household viewing equipment (TVs, VCRs, DVDs, Externally Supplied DVRs, Game Consoles, etc.) and often dictate MVPD choices around subscription attributes (pay TV, etc.). This paradigm is rapidly eroding as MVPD services proliferate to all manner of consumer owned devices, such as PCs, tablets, Smartphones, and others

The following describes existing (generalized) household viewing equipment that is supplied by MVPDs and the general functions of each. Additionally, data capture and accumulation functions are described as well as control implications for each device.

#### *3.1.2.1 Digital Video Recorder (DVR; Internal or External to Device)*

DVR is an optional hardware device and service offered by most MVPDs to their digital service subscribers. The DVR has a peripheral role in the capture of linear digital tuning data however, in some cases, devices will simultaneously record programming that is being viewed. It is the responsibility of the MVPD to ensure that usage records accurately reflect live, or linear tuning – regardless of DVR activity

DVRs can contain multiple tuners, so an MVPD should have processes to differentiate activity by tuner and to accurately represent linear tuning data being accessed by the subscriber. Certain MVPDs also have multi-room DVR capabilities so accurately representing linear tuning across all tuners and eligible devices in the household is important. For example, If a primary set is linked with a STB and DVR tuner in the same room and a DVR is linked with a different set in a different room, activity on both these sets should be additive to linear tuning, regardless of whether they are on the same program. In certain cases, multi-room DVR activity should not be additive if the same viewer is implied.

#### *3.1.2.2 Remote Control*

Many actions of the remote control are captured at the same time as linear tuning data. These actions can be used to infer certain types of behavior or engagement but are not necessary for crediting tuning events since tuning data is directly associated with activity on the STB which can be independent of remote control activity. Households may also rely on various remote controls so there may be situations where certain remote control activity is not detectable. For the purposes of digital video tuning data, several of these actions are specifically relevant:

- Power (On/Off)
- Channel Change (Up/Down or Direct Tune)
- Launch/Close PIP (picture in picture); Mosaic (multi-PIP)
- Guide (overlay of programming grid with scaled video)
- Mute
- Record
- Pause/Rewind/Fast Forward
- “On Demand” or Select A, B or C Buttons
- VOD Menu Initiation
- Stop

The level of data capture for these remote control actions varies by MVPD and device technology and the nature of the captured tuning data can change through the use of the remote.

In multi-frame environments (e.g. PIP) the video component that is audible should be identified and the relative size of each video component should be captured., noting any obstructions and the extent to which the obstruction exists.

It is the responsibility of the MVPD to ensure that the tuning records accurately reflect live, or linear, tuning – regardless of remote control interaction during the tune.

#### *3.1.2.3 Agent Software (Can be Present within Device)*

Certain MVPDs use vendor created software that is loaded on STBs to facilitate the capture, accumulation, and transmission of tuning data and other relevant ancillary data. This vendor software is created and maintained by a third party. Although ultimately the responsibility for the functions performed resides with the MVPD. Agent software generally facilitates agent-based processing.

### 3.1.3 Data Collection & Processing Environment

Regardless of which processing environment is used by the MVPDs, the data must first be collected. The following describes existing (generalized) processing equipment used by MVPDs and the general functions of each.

#### *3.1.3.1 Collection Servers*

The collection server is the device where tuning data first resides before being accumulated (e.g., across broader time periods, geographies, etc.). There are two primary collection server types based on where the data collection takes place: Client Side and Server-Side:

##### *3.1.3.1.1 Client-side collection: At The SetTop Box*

With client-side collection, the data collection is done on the settop box, which collects or logs tuning data over the course of a defined timeframe. From here, the data is sent to a more centralized server that further accumulates the data.

The amount of data stored on the client-side, and how frequently the data is sent to a more centrally located server where it is accumulated can vary by MVPD.

##### *3.1.3.1.2 Server-side collection: At MVPD Distribution Point*

With Server-side collection, tuning information originates on a server (generally in the head-end) as a result of an action on the set-top box. In most cases where switch-digital video (SDV) is used, the SDV server's primary function is to optimize bandwidth for the delivery of digital video to the set-top box, and, secondarily is used as an intermediary collection and storage point for tuning data.

###### *3.1.3.1.2.1 Head End*

For MSOs, the head end architecture is the final decoding and distribution point for video streams before delivery to subscriber households. The equipment located at the head end monitors and manages the configuration of other components and optimizes the video transmission system. As part of that process, it also collects and stores change requests that originate from set-top boxes or the SDV server.

#### *3.1.3.1.2.2 Satellite Downlinks/Uplinks*

For Satellite operators, the digital video feed enters the system through the Broadcast Center and is transmitted by satellite to the Dish and Receiver in the digital household. Subscribers may also connect to the satellite operator through a telephone land line, which transmits the tune records from the Receiver back to the collection facilities.

#### *3.1.3.1.2.3 Centralized General Processing Environment/Storage/Retention*

The centralized processing environment refers to the transmission of regionally or locally stored data into the MVPD's data warehouse prior to the beginning of general editing and processing procedures. MVPD's are strongly encouraged to have redundant delivery methods throughout the process, from original collection point to point of processing to minimize the potential for data loss.

As described in section 3.1.3.1.1, the data is transmitted from the household to the collection server on a regular basis. The frequency of data transmission depends on the communication technology.

- Broadband: Data is communicated over broadband on a perpetual basis. IPTV/Telco uses this communication technology.
- Coaxial / Fiber-optic Cable: Data is communicated from the set-top box to the server on a daily basis and in some instances near real-time. MSOs and Telcos use this communication technology.
- Telephone Dial-up: Data is communicated on a daily basis over a land line. Satellite operators generally rely on this communication technology.

In some cases, despite advance connectivity, MVPDs may choose to gather data less frequently due to bandwidth efficiency. In these cases, data collection methods and associated frequencies should be disclosed.

MVPDs should maintain a retention policy over significant captured data and key interim data files from processing; this retention policy should recommend storing raw data for sufficient time to ensure accurate data loading and completeness checks and thereafter the retention of key downstream processed records for a minimum of thirteen (13) months, to ensure an MVPD's ability to investigate issues that may arise as well as the ability to restore data for reference purposes.

### *3.1.3.2 Cable, Satellite or Telco Device and Systems Software*

Refers to specifications and technologies used by MVPDs to determine subscriber's level of service as well as prevent theft of service, unauthorized device tampering and other non-permitted activities.

These are generally unrelated to the collection and processing of linear tuning data, although there may be some overlap within ties to reference data for subscriber households. Much of this data is highly proprietary.

## 3.2 Linear Data Capture

Data capture includes processes to acquire tuning and associated activity records from subscriber households and their viewing equipment. MVPDs also capture and retain information (non-viewing data) from other sources such as subscriber information, channel information, program line-ups, geographic and ad-traffic data (collectively called reference data).

### 3.2.1 Initial Capture Controls

Adequate internal controls should exist to provide reasonable assurance that data is derived solely from subscriber initiated tuning activity and that this captured data is not lost or altered in unintended manners. Relevant data may also be generated that is not directly subscriber initiated, such as a scheduled recording These controls should include:

- Adequate storage and timely transmission of data by set-top box or other presentation device to ensure collection of all events on digital data capture device(s)
- Limitations on access to captured data, while in the household or at a head-end of other collection point, by unauthorized individuals
- Secure storage and transmission of sensitive data or tuning events
- Time stamp controls (accuracy of time stamps and consistency with other MVPD systems, if resident on device); later linked to consolidation processes
- Warning and monitoring systems to alert the MVPD as to loss or mis-attribution of household system functions
- Coding of captured event records to ensure completeness of capture, for example, using source identifiers and segmentation indicators such as household ID, device ID, header records and trailer records to allow data collectors to properly aggregate packetized data.

Controls required for maintaining reference data by MVPDs are covered in section 3.4.2 of these guidelines.

## 3.3 Linear Data Accumulation

The data accumulation stage includes processes to: (1) gather data from the household environment at initial stages, representing the entry of household and device data (tuning and other associated transactions) into the MVPD-housed server infrastructure, and (2) later accumulation of this initial data into time-period based grouped data (larger geographic groupings than head-end) at the sub-market or market level.

### 3.3.1 Initial Data Collection

There are five mandatory fields identified in the initial data collection that are necessary for accumulation and processing. These fields are:

1. Unique HH Identifier

2. Unique Device Identifier
3. Tuning Record Identifier
4. Program Source Identifier
5. Tuning Time Identifiers

Following collection of these base data, further data is generally appended to facilitate downstream processing and analyses. These fields are typically as follows:

1. Geographic Information
2. MVPD Technical Platform (through device ID) – The specific set of devices, and therefore functionality located within the household
3. Interactive or 2-way Capability – In addition to receiving content, the household equipment is capable of communicating back to the MVPD through a backchannel or other communication mechanism
4. Link between Household and Device

Each of these nine elements is described below.

#### *3.3.1.1 Unique HH Identifier*

A unique, anonymous household identifier must be assigned to each tuning record that is consistent over time for the same household. Ideally, this unique HH ID should be consistent across all data elements that can be collected, while still preserving the integrity of the anonymization. These unique identifiers are for an MVPD, not across MVPDs.

#### *3.3.1.2 Unique Device Identifier*

A unique, anonymous device identifier must be assigned to each tuning record that is consistent over time for the same device. These unique identifiers are for an MVPD, not across MVPDs and are used to make reference to information about the capabilities of the device.

#### *3.3.1.3 Link Between HH and Device Identifier*

The link between HH and Device identifier is required so that both can be tracked over time. This link between Unique HH Identifier and Unique Device Identifier must be maintained and updated frequently (monthly at minimum), and agreed upon by each MVPD to ensure that the link remains valid. The valid link ensures that tuning linked to a HH is valid.

However, in some cases the link between Unique HH Identifier and Unique Device Identifier can be “severed” if a device is relocated without appropriate documentation in the reporting system. Causes for relocation may include:

- Subscriber churn
- Device upgrade/replacement
- Subscriber relocation
- Device malfunction

MVPDs should have procedures to periodically assess the number of false linkages or missing linkages (where household data and reference data do not align or no linkages can be made) within their subscriber base. The ranges of time these conditions exist should also be analyzed and an MVPD should strive to minimize the time needed to update data to minimize link errors.

#### *3.3.1.4 Tuning Record Identifier*

- E.g., tuning record would include Tune-in and Tune-out, although some MVPDs may collect these elements separately
- Should contain options for all interactions (clicks, guide, dvr, etc)

#### *3.3.1.5 Program Source Identifier*

The Program Source Identifier is required so that the tuning record can been aligned to both a source (e.g. network) and a program. While the initial ID may vary by market, the final source ID should try to remain consistent across markets, local stations notwithstanding. For this reason, there is typically a translation step in the data supplementation step (section 3.4.2.1.1) to convert the source IDs into standard network IDs.

Program source identifier can also contain a channel number.

A universal ID for content sources such as television networks would be a helpful structure to ensure consistent processing decisions by MVPDs. Currently there is no such ID, but this development is encouraged by the guidelines participants.

#### *3.3.1.6 Tuning Time Identifiers*

A key factor in data capture and consolidation is identifying and maintaining time-stamp information. At a minimum all tuning data should be captured at the second level and this level of granularity requires precise time-stamp information.

Adequate controls should be maintained to ensure consistent and accurate time stamps associated with multi-channel digital tuning data during processing, regardless of stage or location of processing. Time stamp data should initially be maintained in accurate local time (for the household or head-end) and that time can be equalized or converted in later processing stages based on known and widely accepted business rules.

#### *3.3.1.7 Device Identifier*

A device identifier is required to allow the MVPD or its agent to determine the technical equipment contained in the household. In most instances specific data capture rules or techniques as well as data collection methods are applied to specific devices based on operational requirements.

#### *3.3.1.8 Geographic identifier*

- Zip code at a minimum, if not more granular (zip+4, block group, etc). As these geographic identifiers become more discrete, potential privacy concerns are magnified – sub-zip code level identifiers necessitate involvement of additional privacy measures to ensure non-reversal of personal information.
- Local time zone of the subscriber household.

### 3.3.2 Initial Data Storage

Certain MVPDs collect data from subscribers in batches or other subset which can necessitate storage processes. Data is then grouped until time, volume and/or completeness requirements are met and further processing is executed (e.g. data consolidation and front-end processing)

### 3.3.3 Consolidation & Front End Processing

Consolidation techniques vary based on the operating environment of the MVPD, for example whether switched digital video server environments are employed, or agent-based processing or other linear data collection techniques are used. Many of these techniques evolved with the primary goal of optimizing bandwidth usage, but this document focuses on these aspects from a tuning data collection perspective.

In certain instances, the first stage of data consolidation would be the original collection of data into the head-end structure and comingling of this data with other data from other subscriber households. In other situations, these stages may be combined due to direct transmission of household environment level data directly to allow centralized accumulation at a large scale.

Adequate internal controls should be present to ensure complete data capture from subscriber households and, accordingly, gaps in data collection from applicable subscriber households should be identified and investigated on a timely basis. These gaps should be flagged and trended over time to aid in assessing and improving data gathering and transmission success rates. Additionally, adequate internal controls should be present to ensure data is not lost or altered in unintended manners post initial collection from subscriber households in later accumulation stages. These controls should include:

- Header and trailer information on consolidation records
- Run-to-run control totals (for batched environments)
- Data count and sequence checks
- Data Limit or reasonableness tests
- Alerts for missing files or consolidation records
- Data transmission monitoring and alerts
- Accounting for device-based time delays

## 3.4 Linear Data Processing

The linear data processing stage includes several critical steps needed to ensure that the data is complete, accurate and consistent. These steps are:

- Error processing
- Correlation and Supplementation
- Editing and Smoothing
- Inspection
- Data Release & Disclosures

### 3.4.1 Error Processing

Error processing describes the process and steps for investigating data anomalies and message flags that occur in data capture, accumulation and processing steps. Specifics within each step will be described within each area respectively.

The role of the MVPD is to serve as the final clearing house for investigation processes, regardless of whether the data is processed in-house or by a third party. The initial review and investigation of data errors can be conducted by the primary data processor, but the ultimate responsibility for addressing data errors lies solely with the MVPD.

### 3.4.2 Data Correlation and Supplementation

Data correlation and supplementation includes processes to enrich consolidated household and device data (tuning and other associated transactions) with reference data. Reference data may include network information, program information, geographic information such as ad-zones, ad traffic information if MVPD-controlled ad insertion is employed and subscriber information. These reference data sets can originate from inside the MVPD, such as with subscriber information or ad-zones, or can be sourced from outside organizations such as with program names information.

Reference data accuracy should be considered in context of reporting applications associated with tuning data. For example, if tuning data is summarized and reported on a monthly basis, then reference data sets should be updated and with database structure changes controlled and identified on a basis aligned with monthly reporting frequency. Many of these processes require coordination with field staff of the MVPD where certain forms of data collection are originally conducted and where periodic updating occurs on an as-needed basis.

MVPDs should seek to establish data dictionaries associated with their output data, as well as for incoming third party data to be used during data integrity checks.

#### *3.4.2.1 Reference Data Types*

Further information about significant types of reference data is presented below:

#### 3.4.2.1.1 Program source identifier

MVPDs must have reliable and up to date reference data that can translate internal system codes for video content to the more common industry names for content originators. This information should be sufficiently granular to distinguish local distributors, where necessary and/or digital sub-channels. The MVPD should maintain separate codes and translations for each copy of a video feed that is distributed to subscribers with the potential to be viewed.

Separate codes should be maintained for breakouts of the same feed on different channel numbers, distinguish between HD and SD feeds for the same programming, and to distinguish between identical video running with different audio feeds.

MVPDs are required to update this data on a regular (at minimum monthly) basis in order to identify changes in channel lineups, channel names and feed patterns on a timely basis.

The industry should seek to maintain a universal database of identifiers for all services to reference.

#### 3.4.2.1.2 Program Names Data

The MVPDs may also choose to supplement the tuning data with program information. For example, schedules might be linked to the tuning records by using the following fields as a key:

- Station/Channel/Network ID (as defined above)
- Start Time (DayDateTime)
- Duration
- Genre(s)
- Description

MVPD's should take measures to account for anomalies that may be encountered when attempting to correlate tuning data to broadcast schedules, for instance, breaking news and sports overtimes, or line-ups which can vary locally

MVPDs should make note of start and end time offsets to account for time zone of the market and network feed to ensure proper assignment of program names to tuning events.

#### 3.4.2.1.3 Geographic Data (e.g., DMA, Ad-Zone, and/or Zip Code)

The MVPDs may supplement the tuning records with geographic information, including DMA, Ad Zone and zip code. This information which may be based on internal subscriber billing information that would be linked by the anonymous HH ID. The levels of geography for data delivery must be agreed-upon by the MVPDs.

#### 3.4.2.1.4 Subscriber Information

In addition to geographic assignments, subscriber level information may also include the following at the household or device level:

- HH level: Count of STBs per HH by Feature
  - Capable of supporting Return Path Data

  - Not Capable of supporting return path data
  - DVR Capable
  - SD/HD
- Headend Identifier
- Collection Platform Hardware/Software Information

Other optional HH and device level data can be inserted at different processing stages at the discretion of the MVPD's based on need, availability (internal or external) and compliance with company privacy standards. Examples of these types of data include:

- HH Level: VOD or Interactive capability
- Device Level: HD Set (Y/N)
- Device Level: DVR Sets

#### *3.4.2.1.5 Ad Traffic Data (present if ad-insertion is performed by the MVPD)*

In some circumstances, MVPDs may replace or insert advertising into content, Such ad traffic and log data may be inserted into a data set at the discretion of the MVPD. This element is not in scope for this project.

#### *3.4.2.1.6 Time Zone Data*

When data is collected the time stamp will be based on the local time, or the time zone associated with the regional programming feed. MVPDs must insure that all time based data for associated reference data (e.g., programming info, VOD release dates, etc) be aligned for each market, based on that market's time zone information. If a market itself crosses time zones, the time stamp should match the feed or programming that is delivered to subscribers.

MVPDs should not adjust timestamps for the raw tuning data, however data may be adjusted in the reporting or processing phase if combining multiple markets that cross multiple time zones. The time adjustment will have to take into account several factors:

- The time zone of the programming feed
- The time zone of the HH_ID
- The time zone that will be used as the standard, to apply to all tune records

### *3.4.2.2 Preserving Anonymization in Data Supplementation Process*

The supplementation process is often reliant on a HH Identifier as a primary key in merging additional information to the tune records. While the HH Identifier itself is anonymous, the data supplementation must preserve the controls around this identifier to ensure that the subscriber cannot be identified.

Association of supplemental data after anonymization may be done at the MVPD's discretion, either by the MVPD itself or through licensed and authorized other parties.

#### *3.4.2.3 Potential Issues with Data Supplementation*

A key process associated with data correlation is accounting for household device based time-delays, which can sometimes impact appropriate pairing of household data with tuning event reference data (channel, program). Time delays can occur because of DVR delays, I-Frame related delays, MVPD plant latency, data transmission latency, uplink/downlink latency, etc. A MVPD should study, document and account for these delays properly, as applicable.

Reference data sets should be maintained on a current (updated) basis to ensure the accuracy of correlated data and inferences made from this data in downstream processing of the MVPD. Adequate internal controls should exist to identify, investigate and resolve errors or mismatches that occur in the data correlation and supplementation process. Errors or mismatches should be followed-up on by MVPDs or agents of MVPDs in a timely manner, with necessary corrections made to the reference data when discovered.

Error rates encountered during data correlation and supplementation should be maintained and trended over time to use in evaluating data sources, internal controls and other accuracy metrics.

Material errors or mismatches that impact the accuracy of the accumulated multi-channel digital tuning data for the processed time-period should be corrected and these errors should be disclosed to down-stream users of the MVPD data.

### 3.4.3 Data Editing and Smoothing

Data Editing and Smoothing is a complex process that involves completing or appending data records or otherwise making inferences about data activity to enhance the usability and understanding of the data. Data editing and smoothing is intended to remove anomalies from the data, reduce the size of the data set by removing unnecessary data elements, and prepare the data for processing.

In some cases these data edits and smoothing functions can bridge missing or interim data or pad data prior-to or post logged events. Data editing and smoothing is an accepted technique for improving the quality and completeness of data when applied in a consistent and reasonable manner. All rules used should be documented and available to users of the processed data.

Data editing and smoothing rules should be supported by empirical evidence and rationale demonstrating the accuracy-benefits and validity (as related to the underlying activity of the subscriber household on the device) of choices made. Empirical evidence should be established and maintained for each type of equipment the MVPD delivers to its customers. Additionally, periodic updates and validations to those rules must be performed to ensure that any changes made by the MVPD, whether to lineups, physical plant or transmission methods, do not impact data collection and stability.

To ensure the objectivity, appropriateness and replicability of these data adjustment procedures, the MVPD should have appropriate controls over approval and monitoring of the various types of these adjustments. Additionally, significant data adjustment types should be supported by appropriate empirical evidence as to their appropriateness.

Significant Types of Data Editing and Smoothing Rules, in the order they generally are applied, Include:

#### *3.4.3.1 Data Cleaning Rules*

- Inter-Communication Data Gap Analyses and Closure
- Removal of Invalid Records
- Removal of Non-Essential Records
  - Generally non-tuning records, or non-significant other transactions
- Filtration Rules
  - Removal of Commercial Sites
  - Handling of Anonymous Sites/Devices
  - Systematic Outages or Known Data Problems
  - Below Tolerance Tuning Thresholds

MVPDs must have controls to identify and segregate from tuning data any records that reflect known non-user initiated device activity such as:

- Events resulting from MVPD diagnostic activity
- MVPD channel changes to save bandwidth
- Actions to disable PIP

#### *3.4.3.2 Data Anomaly Removal*

Generally related to issues associated with the household environment equipment, or other technical issues including Television Status Determination (i.e., On-Off). An On-Off situation exists when a television connected to a presentation device such as a STB is off while the device remains on. If the connected device cannot detect the power status of the TV then tuning events could be recorded and used in further processing when, in fact, the TV was off. On-Off determination edits, in this case, are critical to the accurate interpretation of the recorded tuning data.

On-Off determination rules can:

- Vary by device type or household configuration
- Be impacted by household behavior in how the householders power down devices and which devices are powered down
- Vary significantly based on the presence of a DVR or HDMI devices
- Be impacted by very long-form events such as the Superbowl or lengthy News contents

The On-Off issue should decrease with:

- Technology that turns off the device when the TV is turned off
- An increase in viewers who turn off the device to save energy
- Availability of clickstream data that will report device activity beyond tuning and on/off. This clickstream data will provide more variables on which to base capture and edit rules. Clickstream data for these guidelines purposes means application 'instrumentation' in which an Electronic Program Guide or other consumer facing application

#### *3.4.3.3 Linear Data Sampling Rule/Representativeness*

Data is collected only from Digital video subscribers, which typically do not include all MVPD subscribers. Additionally, not all sets associated with a Digital video subscriber may allow for return path data. The resulting digital video viewership is not captured at a true census level. Additionally some MVPDs may not store the tune records for all of their digital video subscribers. Instead, they accumulate and process data from a sample of their digital subscriber households.

Other MVPDs store tune records from all of their digital video subscribers, but from day to day, data from some digital households are unavailable. These "out-of-tab" households are missing for any of the following reasons:

- New device is not yet synched with HH ID
- Power outage or equipment failure
- Technical Failures
- Communication Issues

MVPDs should track "out-of-tab" situations for each household, by type of reason, and flag those households that exhibit persistent conditions for remediation.

It is critical that data sampling and data capture success information is retained and disclosed by the MVPD so further use of the data downstream can be informed as to coverage and representation issues

MVPDs must disclose the capture rate and how it may impact the representation of digital video tuning data. When samples are used, characteristics of the sampled sites/households should be available as well as the comparative percentages for the total MVPD footprint (and the market when available).

#### *3.4.3.4 Tuning Event Linkage Rules*

The tuning event linkage rules address how to link common-channel records separated by short timing gaps indicating they are part of the same tuning event and these linkages are generally made using bridging rules. Bridging rules are designed to enhance the transaction stream coming from the household environment such that all likely tuning seconds are accounted for and ensuring they are attachable to a valid source. Occasionally, hints from remote IR activity

captured simultaneous with the tuning data can be used to help inform these linkages in complex and developed processing environments.

#### *3.4.3.5 Minimum Tuning Event Length Rules*

For raw data, all events of one second or longer should be captured. Later in processing, MVPDs generally drop certain tuning events simply for the purpose of bandwidth and storage efficiency purposes. Typically events less than 15 seconds with no adjacent similar events can be dropped.

Consecutive remote keystrokes of very short durations (sub-second) are generally not retained.

Consideration should be given to the fact that in some instances there is a multi-second delay (up to three seconds) for video to display subsequent to a channel change and steps should be taken to account for this latency.

#### *3.4.3.6 Maximum Tuning Event Length (Cap) Rules*

Capping rules are designed to identify situations of clear unlikely viewing activity and to identify situations where the television status is not likely to be "on." These rules can legitimately vary by content, daypart or the tuning history of the subscriber household.

Data Adjustment and Ascription Controls:

- For tuning events, a maximum tune event should be indentified to eliminate "stale" tunes that would reflect a powered box with no viewing. This is generally considered after 6 hours of no customer activity.
- Cap rules can vary by content, daypart or tuning history
- Non-channel change events can also influence occurrence of capping. For example, viewer interaction with the remote such as volume can be used as a hint to reset inactivity rules.
- Homes and devices with repeated excessive tuning indicating suspect behavior should be flagged for follow-up analysis and possible exclusion from the dataset
- When creating cap and edit rules it's important that they be founded on empirical evidence and consider the following:
  - Percent of devices left on all day on any given day
  - Over longer periods of time the percent of boxes never turned off
  - Samples of devices that are turned off daily (measured over time) may be used to help guide the creation of cap and edit rules as persons in these households may more reliably turn off the device when not viewing

### 3.4.4 Storing Processed Linear Data

The incidence of editing and smoothing rules and the overall impact of these rules on the tuning data volumes should be known and trended over time as a monitoring process.

Pre-edited/smoothed data should be retained in its original form for sufficient time to ensure valid downstream data loads, front end processing and completeness checks. Post edited/smoothed data should be retained for thirteen (13) months at minimum. These data should be accessible but are not required to be stored online.

### 3.4.5 Linear Data Inspection

MVPDs should have adequate procedures to ensure competent review of multi-channel digital tuning data prior to finalization. This review should focus on the completeness of data processed, for example did known subscriber households report or are they otherwise accounted for, and full execution of data consolidation, correlation and supplementation, and editing and smoothing processes.

Inspection processes are generally a combination of automated and manual processes, with manual processes being typically focused on exception investigation. Data volumes and other key data attributes should be trended on an overall, household and geographic level. Trending by source or program can also be useful, although this is typically not audience data but instead some general volumetric information such as minutes captured. Actions should be taken to:

- Validate tuning event minutes do not exceed the total number of minutes of the timeframe it is contained within (e.g. cannot exceed 60 minutes for one hour reporting)
- Assess the extent to which variations occur in the volume of data captured and investigate those situations where levels fall outside the expected mean based on historical patterns. These analyses should be performed at a granular level (e.g. head-end by MSO). Situations that might cause a change in data volume include:
  - Utility outages
  - Technical difficulties by the Service provider (e.g. Head-end)
  - Coverage area revisions
  - Subscriber turnover
- Compare data to previously defined benchmarks that were established based on historical information accounting for:
  - File Size
  - Count of Tuning Events
  - Households reporting
  - Devices reporting

Inspection checklists should be relied on to ensure completeness of the checkout process and as a control to provide management with assurance that these processes are administered. The inspection process should clearly identify and escalate points of failure with an indication of the required action, including, for example:

- Need to Reprocess
- Re-acquire MVPD data
- Gap Filling

- Remove data element
- Discard record

MVPD's will retain up to 3 months of de-identified exception data, both from initial collection phase and through the processing phase. Privacy requirements may prevent the MVPD from storing or releasing any non-anonymized subscriber data for these purposes, so that testing of specific behaviors in the home and review of the resultant pre and post data can only be performed in a controlled MVPD environment or through the use of data relating to volunteers that opt in for this purpose.

Suspect data-conditions identified during inspection processes should be investigated prior to data finalization, to minimize the release of suspect data for further processing or to outside organizations.

MVPDs should take measures to ensure that the data-set is formatted correctly in conformance with the agreed upon format, and that all necessary data elements are represented. To the extent that they might exist, supplementary data components (beyond the standard format) should be clearly identified and labeled.

MVPDs should maintain a data-reissue policy which contains objective criteria for when tuning data would necessitate re-issue as a result of errors or omissions that are discovered in the normal course of business. This policy should specify how long data is subject to potential corrective measures and the types of disclosures that would accompany data reissues.

### 3.4.6 Data Release

The Data Release step involves placing hereto processed multi-channel digital video tuning data into agreed upon format(s) which may involve an internal data translation process. The common format data can be distributed through a number of secure methods either within the MVPD or to outside organizations for further processing, such as audience measurement applications. The Data Release step can also include data on other associated transactions, which can add insight into the tuning data itself and/or enhance the value of this data though in all instances supplemental data will be anoymized to protect the identity of household subscriber information.

#### *3.4.6.1 Data Release Rules*

- Data owners and other stakeholders will contribute to proposed data use rules that determine what levels of data can be made available, including:
  - Minimum reporting sizes for daily, weekly, and monthly average basis reporting
  - Specifications for data intended for internal use only
  - Guidelines for custom projects – what is acceptable for release outside a basic product
  - Minimum geographic reporting rules.
  - Minimum rules for merging and/or fusing additional datasets onto base set top box data

- A clear description of terms for data elements included in the release.

#### *3.4.6.2 Data Factoring Rules*

- Data factoring rules may be developed to conform the measured dataset into a "projectable" tuning database which may include the following:
  - Sample balancing
  - HH or Device weighting
  - Creation of performance reports that highlight impact of cap and edit rules
  - Imputation of usage by household or device
  - Other factoring of household or device counts, events, history

#### *3.4.6.3 Header File Specifications*

**Basic Data Contents:**

| Header | Description |
|---|---|
| **Day Definition** | 12:00 midnight to 11:59PM |
| **Time Zone** | • Full disclosure of time-zone processing is required. It is preferable for compliant MVPDs to have the ability to produce accumulated data reports in a common, equalized, time-zone so down-stream users can understand activity across MVPDs.<br>• For U.S. reports it is recommended that reports be available on the basis of the Eastern Time zone. For non-U.S. reports it is recommended that reports be available on the basis of GMT. |
| **Week Definition** | Monday through Sunday |
| **Month** | Three reporting methods can be available:<br>1. TV broadcast month definitions<br>2. 4 week periods (13 per year) consistent with media planning for other media<br>3. A calendar month.<br>Month definitions selected should be disclosed. |
| **Dayparts** | Specified similar to those currently in use by syndicated audience measurement organizations – specifics to be disclosed by the MVPD. |

#### *3.4.6.4 Methods-Oriented Disclosures*

A key attribute of trusted data is the full disclosure of the basis for producing that data. In accordance with accepted industry practice, the producers of the multi-channel digital video tuning data should therefore disclose to data recipients pertinent aspects of their process on a routine basis.

Specifically, the nature of data capture, accumulation, methods of editing and processing and limitations of this data should be disclosed.

The preparation of a written Description of Methodology (DOM) to house the disclosures specified herein is strongly encouraged to facilitate distribution of this information. The DOM, or whatever method is chosen for disclosure, should be maintained with currently used procedures; a monthly or more frequent update is suggested.

The following presents the suggested contents for the DOM:

- Name of MVPD
- Accumulation Period (By major geography should it vary)
- Subscriber Households Included (ideally these statistics are reported with each data accumulation run)
  - Number of Households Reported
    - Total
    - Reporting
    - Tabulation Rate
    - Non-Tabulation Dispositions
    - Known Systematic Outages
  - Sites per Household on Average
  - Estimates of Persons Included in Households Reported
  - Overall MVPD Penetration for Accumulated Geography
  - Number of Sites Reported
    - Total
    - Reporting
    - Tabulation Rate
    - Non-Tabulation Dispositions
    - Known Systematic Outages
- Types of Data Accumulated Reported
  - Time periods included
  - Days included
  - Geographic Areas
  - Significant Sub-Groupings of data
- Formats of Reported Data (Record Layouts)
- Nature of Inspection Procedures Applied Prior to Reporting
- Any Sampling/Projection Methods Used
- Data Collection Methods Employed
  - Methods of Data Capture
  - Types of Data Captured by Method
  - Contacts with Users

  - Agents Used, If Present
- Editing or Data Adjustment Procedures Applied
- Data De-Identification Procedures Used
  - Types of Data Protected, Nature of Protection
  - Consumer Notice and Privacy Policy
- Computations Applied to Data, If Any
- Reporting
  - Frequency
- Limitations of Data Use
- Method of Disclosing Changes in Methodology

## 3.5 Guidelines Requirements for Linear Data Set Processing

The following summarizes the requirements for the processing of Linear Tuning Data that are noted within the narrative portion of Section 3 of the guidelines and is provided to ease reference.

### 3.5.1 General

- The responsibility for accurate capture, accumulation and editing remains with the data owner (generally the MVPD, or sometimes referred to as the seller). Adequate internal controls, error investigation, processing and disclosures should be maintained regardless of processing environment. (Section 3.1.1)
- Real time applications should have provisions to protect against data loss due to temporary loss of communication/contact with collection assets of the MVPD. (Section 3.1.1.2)
- MVPDs should maintain back-up and retrieval procedures to ensure that collected data in processed form is available post processing based on the retention requirements specified herein. (Section 3.1.1.2)
- It is the responsibility of the MVPD to ensure that usage records accurately reflect live, or linear tuning – regardless of DVR activity. (Section 3.1.2.1)
- MVPD's should have processes to differentiate activity by tuner and to accurately represent linear tuning data being accessed by the subscriber. (Section 3.1.2.1)
- In instances where multi-room DVR capabilities exist it is important to accurately represent linear tuning across all tuners and eligible devices in the household. (Section 3.1.2.1)
- In multi-frame environments (e.g. PIP) the video component that is audible should be identified and the relative size of each video component should be captured., noting any obstructions and the extent to which the obstruction exists. (Section 3.1.2.2)

- It is the responsibility of the MVPD to ensure that the tuning records accurately reflect live, or linear, tuning – regardless of remote control interaction during the tune. (Section 3.1.2.2)
- MVPD's are strongly encouraged to have redundant delivery methods throughout the process, from original collection point to point of processing to minimize the potential for data loss. (Section 3.1.3.1.2.3)
- MVPDs should maintain a retention policy over significant captured data and key interim data files from processing; this retention policy should recommend storing raw data for sufficient time to ensure accurate data loading and completeness checks and thereafter the retention of key downstream processed records for a minimum of thirteen (13) months, to ensure an MVPD's ability to investigate issues that may arise as well as the ability to restore data for reference purposes. (Section 3.1.3.1.2.3)

### 3.5.2 Data Capture

- Adequate internal controls should exist to provide reasonable assurance that data is derived solely from subscriber initiated viewing activity and that this captured data is not lost or altered in unintended manners. Relevant data may also be generated that is not directly subscriber initiated, such as a scheduled recording. These controls should include (Section 3.2.1):
  - Adequate storage and timely transmission of data by set-top box or other presentation device to ensure collection of all events on digital data capture device(s)
  - Limitations on access to captured data, while in the household or at a head-end of other collection point, by unauthorized individuals
  - Secure storage and transmission of sensitive data or tuning events
  - Time stamp controls (accuracy of time stamps and consistency with other MVPD systems, if resident on device); later linked to consolidation processes
  - Warning and monitoring systems to alert the MVPD as to loss or mis-attribution of household system functions
  - Coding of captured event records to ensure completeness of capture, for example, using source identifiers and segmentation indicators such as household ID, device ID, header records and trailer records to allow data collectors to properly aggregate packetized data.

### 3.5.3 Accumulation

- There are five mandatory fields identified in the initial data collection that are necessary for accumulation and processing. These fields are (Section 3.3.1):

  Unique HH Identifier

  Unique Device Identifier

  Tuning Record Identifier

Program Source Identifier

Tuning Time Identifiers

- Following collection of the mandatory fields noted above, further data is generally appended to facilitate downstream processing and analyses. These fields are typically as follows (Section 3.3.1):
  - Geographic Information
  - MVPD Technical Platform (through device ID) – The specific set of devices, and therefore functionality located within the household
  - Interactive or 2-way Capability – In addition to receiving content, the household equipment is capable of communicating back to the MVPD through a backchannel or other communication mechanism
  - Link between Household and Device
- The link between Unique HH Identifier and Unique Device Identifier must be maintained and updated frequently (monthly at minimum), and agreed upon by each MVPD to ensure that the link remains valid. (Section 3.3.1.3)
- MVPDs should have procedures to periodically assess the number of false linkages or missing linkages (where household data and reference data do not align or no linkages can be made) within their subscriber base. The ranges of time these conditions exist should also be analyzed and an MVPD should strive to minimize the time needed to update data to minimize link errors. (Section 3.3.1.3)
- At a minimum all tuning data should be captured at the second level and this level of granularity requires precise time-stamp information. (Section 3.3.1.6)
- Adequate controls should be maintained to ensure consistent and accurate time stamps associated with multi-channel digital tuning data during processing, regardless of stage or location of processing. Time stamp data should initially be maintained in accurate local time (for the household or head-end) and that time can be equalized or converted in later processing stages based on known and widely accepted business rules. (Section 3.3.1.6)
- Adequate internal controls should be present to ensure complete data capture from subscriber households and, accordingly, gaps in data collection from applicable subscriber households should be identified and investigated on a timely basis. These gaps should be flagged and trended over time to aid in assessing and improving data gathering and transmission success rates. Additionally, adequate internal controls should be present to ensure data is not lost or altered in unintended manners post initial collection from subscriber households in later accumulation stages. These controls should include (Section 3.3.3):
  - Header and trailer information on consolidation records
  - Run-to-run control totals (for batched environments)
  - Data count and sequence checks
  - Data Limit or reasonableness tests

- Alerts for missing files or consolidation records
- Data transmission monitoring and alerts
- Accounting for device-based time delays

### 3.5.4 Processing

- MVPDs should seek to establish data dictionaries associated with their output data, as well as for incoming third party data to be used during data integrity checks. (Section 3.4.2)
- Reference data associated with content originators should be sufficiently granular to distinguish local distributors, where necessary and/or digital sub-channels. The MVPD should maintain separate codes and translations for each copy of a video feed that is distributed to subscribers with the potential to be viewed. (Section 3.4.2.1.1)
- Separate content originator codes should be maintained for breakouts of the same feed on different channel numbers, distinguish between HD and SD feeds for the same programming, and to distinguish between identical video running with different audio feeds. (Section 3.4.2.1.1)
- MVPDs are required to update content originator reference data on a regular (at minimum monthly) basis in order to identify changes in channel lineups, channel names and feed patterns on a timely basis. (Section 3.4.2.1.1)
- MVPD's should take measures to account for anomalies that may be encountered when attempting to correlate tuning data to broadcast schedules, for instance, breaking news and sports overtimes, or line-ups which can vary locally. (Section 3.4.2.1.2)
- MVPDs should make note of start and end time offsets to account for time zone of the market and network feed to ensure proper assignment of program names to tuning events. (Section 3.4.2.1.2)
- In addition to geographic assignments, subscriber level information may also include the following at the household or device level (Section 3.4.2.1.4):
  - HH level: Count of STBs per HH by Feature
    - Capable of supporting Return Path Data
    - Not Capable of supporting return path data
    - DVR Capable
    - SD/HD
  - Headend Identifier
  - Collection Platform Hardware/Software Information
- Other optional HH and device level data can be inserted at different processing stages at the discretion of the MVPD's based on need, availability (internal or external) and compliance with company privacy standards. Examples of these types of data include (Section 3.4.2.1.4):
  - HH Level: VOD or Interactive capability

  - Device Level: HD Set (Y/N)
  - Device Level: DVR Sets
- MVPDs should not adjust timestamps for the raw tuning data, however data may be adjusted in the reporting or processing phase if combining multiple markets that cross multiple time zones. The time adjustment will have to take into account several factors (Section 3.4.2.1.6):
  - The time zone of the programming feed
  - The time zone of the HH_ID
  - The time zone that will be used as the standard, to apply to all tune records
- While the HH Identifier itself is anonymous, the data supplementation must preserve the controls around this identifier to ensure that the subscriber cannot be identified. (Section 3.4.2.2)
- A MVPD should study, document and properly account for household device based time delays. (Section 3.4.2.3)
- Reference data sets should be maintained on a current (updated) basis to ensure the accuracy of correlated data and inferences made from this data in downstream processing. (Section 3.4.2.3)
- Adequate internal controls should exist to identify, investigate and resolve errors or mismatches that occur in the data correlation and supplementation process. Errors or mismatches should be followed-up on by MVPDs or agents of MVPDs in a timely manner, with necessary corrections made to the reference data when discovered. (Section 3.4.2.3)
- Error rates encountered during data correlation and supplementation should be maintained and trended over time to use in evaluating data sources, internal controls and other accuracy metrics. (Section 3.4.2.3)
- Material errors or mismatches during data supplementation that impact the accuracy of the accumulated multi-channel digital tuning data for the processed time-period should be corrected and these errors should be disclosed to down-stream users of the MVPD data. (Section 3.4.2.3)
- Data editing and smoothing rules should be documented and available to users of the processed data. (Section 3.4.3)
- Data editing and smoothing rules should be supported by empirical evidence and rationale demonstrating the accuracy-benefits and validity (as related to the underlying activity of the subscriber household on the device) of choices made. Empirical evidence should be established and maintained for each type of equipment the MVPD delivers to its customers. Additionally, periodic updates and validations to those rules must be performed to ensure that any changes made by the MVPD, whether to lineups, physical plant or transmission methods, do not impact data collection and stability. (Section 3.4.3)

- To ensure the objectivity, appropriateness and replicability of data adjustment procedures, the MVPD should have appropriate controls over approval and monitoring of the various types of these adjustments. Additionally, significant data adjustment types should be supported by appropriate empirical evidence as to their appropriateness. (Section 3.4.3)
- Significant Types of Data Editing and Smoothing Rules, in the order they generally are applied, Include (Section 3.4.3):
  - Data Cleaning Rules
    - Inter-Communication Data Gap Analyses and Closure
    - Removal of Invalid Records
    - Removal of Non-Essential Records
      - Generally non-tuning records, or non-significant other transactions
    - Filtration Rules
      - Removal of Commercial Sites
      - Handling of Anonymous Sites/Devices
      - Systematic Outages or Known Data Problems
      - Below Tolerance Tuning Thresholds
    - MVPDs must have controls to identify and segregate from tuning data any records that reflect known non-user initiated device activity such as:
      - Events resulting from MVPD diagnostic activity
      - MVPD channel changes to save bandwidth
      - Actions to disable PIP
  - Data Anomaly Removal
  - Linear Data Sampling Rules/Representativeness
    - It is critical that data sampling and data capture success information is retained and disclosed by the MVPD so further use of the data downstream can be informed as to coverage and representation issues.
    - MVPDs should track “out-of-tab” situations for each household, by type of reason, and flag those households that exhibit persistent conditions for remediation.
    - MVPDs must disclose the capture rate and how it may impact the representation of digital video tuning data. When samples are used, characteristics of the sampled sites/households should be available as well as the comparative percentages for the total MVPD footprint (and the market when available).
  - Tuning Event Linkage Rules
  - Minimum Tuning Event Length Rules
    - For raw data, all events of one second or longer should be captured. Later in processing, MVPDs generally drop certain tuning events simply for the purpose of bandwidth and storage efficiency purposes. Typically events less than 15 seconds with no adjacent similar events can be dropped.

  - Consecutive remote keystrokes of very short durations (sub-second) are generally not retained.
  - Consideration should be given to the fact that in some instances there is a multi-second delay (up to three seconds) for video to display subsequent to a channel change and steps should be taken to account for this latency.
- Maximum Tuning Event Length (Cap) Rules
  - A maximum tune event should be indentified to eliminate "stale" tunes that would reflect a powered box with no viewing. This is generally considered after 6 hours of no customer activity.
  - Cap rules can vary by content, daypart or tuning history
  - Non-channel change events can also influence occurrence of capping. For example, viewer interaction with the remote such as volume can be used as a hint to reset inactivity rules.
  - Homes and devices with repeated excessive tuning indicating suspect behavior should be flagged for follow-up analysis and possible exclusion from the dataset
  - It is important that cap and edit rules be founded on empirical evidence and consider the following:
    - Percent of devices left on all day on any given day
    - Over longer periods of time the percent of boxes never turned off
    - Samples of devices that are turned off daily (measured over time) may be used to help guide the creation of cap and edit rules as persons in these households may more reliably turn off the device when not viewing

- The incidence of editing and smoothing rules and the overall impact of these rules on the tuning data volumes should be known and trended over time as a monitoring process. (Section 3.4.4)
- Pre-edited/smoothed data should be retained in its original form for sufficient time to ensure valid downstream data loads, front end processing and completeness checks. Post edited/smoothed data should be retained for thirteen (13) months at minimum. These data should be accessible but are not required to be stored online. (Section 3.4.4)
- MVPDs should have adequate procedures to ensure competent review of multi-channel digital tuning data prior to finalization. This review should focus on the completeness of data processed, for example did known subscriber households report or are they otherwise accounted for, and full execution of data consolidation, correlation and supplementation, and editing and smoothing processes. (Section 3.4.5)
- Actions should be taken to: (Section 3.4.5)
  - Validate tuning event minutes do not exceed the total number of minutes of the timeframe it is contained within (e.g. cannot exceed 60 minutes for one hour reporting)

- Assess the extent to which variations occur in the volume of data captured and investigate those situations where levels fall outside the expected mean based on historical patterns. These analyses should be performed at a granular level (e.g. head-end by MSO). Situations that might cause a change in data volume include:
  - Utility outages
  - Technical difficulties by the Service provider (e.g. Head-end)
  - Coverage area revisions
  - Subscriber turnover
- Compare data to previously defined benchmarks that were established based on historical information accounting for:
  - File Size
  - Count of Tuning Events
  - Households reporting
  - Devices reporting

- Inspection checklists should be relied on to ensure completeness of the checkout process and as a control to provide management with assurance that these processes are administered. The inspection process should clearly identify and escalate points of failure with an indication of the required action, including, for example: (Section 3.4.5)
  - Need to Reprocess
  - Re-acquire MVPD data
  - Gap Filling
  - Remove data element
  - Discard record
- MVPD's will retain up to 3 months of de-identified exception data, both from initial collection phase and through the processing phase. Privacy requirements may prevent the MVPD from storing or releasing any non-anonymized subscriber data for these purposes, so that testing of specific behaviors in the home and review of the resultant pre and post data can only be performed in a controlled MVPD environment or through the use of data relating to volunteers that opt in for this purpose. (Section 3.4.5)
- Suspect data-conditions identified during inspection processes should be investigated prior to data finalization, to minimize the release of suspect data for further processing or to outside organizations. (Section 3.4.5)
- MVPDs should take measures to ensure that the data-set is formatted correctly in conformance with the agreed upon format, and that all necessary data elements are represented. To the extent that they might exist, supplementary data components (beyond the standard format) should be clearly identified and labeled. (Section 3.4.5)

- MVPDs should maintain a data-reissue policy which contains objective criteria for when tuning data would necessitate re-issue as a result of errors or omissions that are discovered in the normal course of business. This policy should specify how long data is subject to potential corrective measures and the types of disclosures that would accompany data reissues. (Section 3.4.5)
- Data owners and other stakeholders will contribute to proposed data use rules that determine what levels of data can be made available, including (Section 3.4.6.1):
  - Minimum reporting sizes for daily, weekly, and monthly average basis reporting
  - Specifications for data intended for internal use only
  - Guidelines for custom projects – what is acceptable for release outside a basic product
  - Minimum geographic reporting rules.
  - Minimum rules for merging and/or fusing additional datasets onto base set top box data
  - A clear description of terms for data elements included in the release.
- Data factoring rules may be developed to conform the measured dataset into a "projectable" tuning database which may include the following (Section 3.4.6.2):
  - Sample balancing
  - HH or Device weighting
  - Creation of performance reports that highlight impact of cap and edit rules
  - Imputation of usage by household or device
  - Other factoring of household or device counts, events, history
- The header file should be structured to contain the following basic contents (Section 3.4.6.3):
  - Day Definition - 12:00 midnight to 11:59PM
  - Time Zone
    - Full disclosure of time-zone processing is required. It is preferable for compliant MVPDs to have the ability to produce accumulated data reports in a common, equalized, time-zone so down-stream users can understand activity across MVPDs.
    - For U.S. reports it is recommended that reports be available on the basis of the Eastern Time zone. For non-U.S. reports it is recommended that reports be available on the basis of GMT.
  - Week Definition
    - Monday Through Sunday
  - Month, of which three reporting methods can be available:
    1. TV broadcast month definitions
    2. 4 week periods (13 per year) consistent with media planning for other media
    3. A calendar month.
    - Month definitions selected should be disclosed.

- Dayparts
  - Specified similar to those currently in use by syndicated audience measurement organizations – specifics to be disclosed by the MVPD.

- The nature of data capture, accumulation, methods of editing and processing and limitations of this data should be disclosed. (Section 3.4.6.4)
- The preparation of a written Description of Methodology (DOM) to house the disclosures specified herein is strongly encouraged to facilitate distribution of this information. The DOM, or whatever method is chosen for disclosure, should be maintained with currently used procedures; a monthly or more frequent update is suggested. The following presents the suggested contents for the DOM (Section 3.4.6.4):
  - Name of MVPD
  - Accumulation Period (By major geography should it vary)
  - Subscriber Households Included (ideally these statistics are reported with each data accumulation run)
    - Number of Households Reported
      - Total
      - Reporting
      - Tabulation Rate
      - Non-Tabulation Dispositions
      - Known Systematic Outages
    - Sites per Household on Average
    - Estimates of Persons Included in Households Reported
    - Overall MVPD Penetration for Accumulated Geography
    - Number of Sites Reported
      - Total
      - Reporting
      - Tabulation Rate
      - Non-Tabulation Dispositions
      - Known Systematic Outages
  - Types of Data Accumulated Reported
    - Time periods included
    - Days included
    - Geographic Areas
    - Significant Sub-Groupings of data
  - Formats of Reported Data (Record Layouts)
  - Nature of Inspection Procedures Applied Prior to Reporting
  - Any Sampling/Projection Methods Used
  - Data Collection Methods Employed
    - Methods of Data Capture

  - Types of Data Captured by Method
  - Contacts with Users
  - Agents Used, If Present
- Editing or Data Adjustment Procedures Applied
- Data De-Identification Procedures Used
  - Types of Data Protected, Nature of Protection
  - Consumer Notice and Privacy Policy
- Computations Applied to Data, If Any
- Reporting
  - Frequency
- Limitations of Data Use
- Method of Disclosing Changes in Methodology

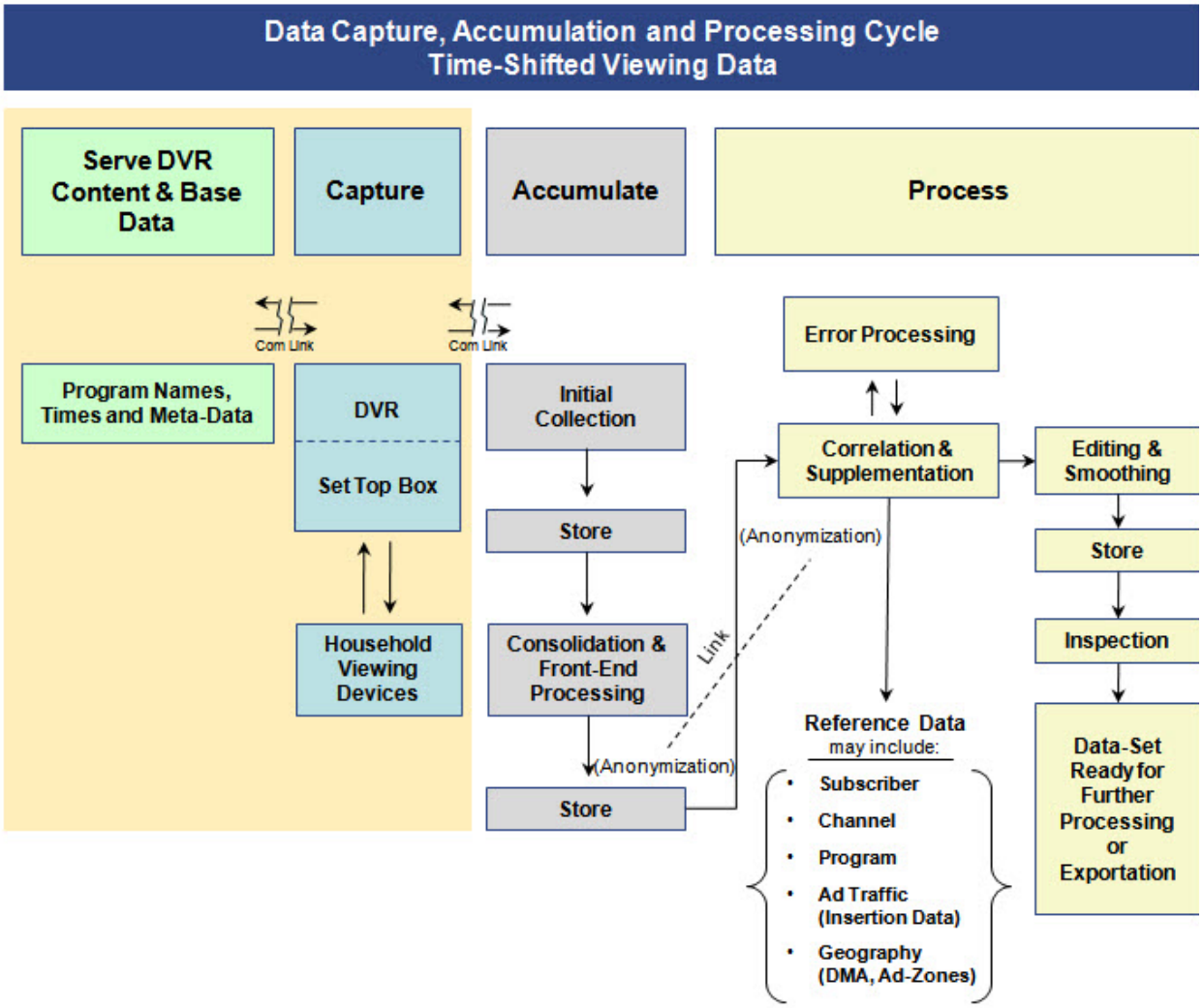
# 4 Guidelines for Digital Video Recorder Time-Shifted Data

This section of the Guidelines covers processing of Digital Video Recorder Time-Shifted Data. Information is presented to the extent it differs from the linear processing information described previously – unless otherwise noted, the Guidelines for linear programming apply to Digital Video Recorder data.

## 4.1 General Description of Digital Video Recorder Time-Shifted Data

The following chart illustrates the data-flow for the capture, accumulation and processing for Digital Video Recorder Time-Shifted Data, and associated activity.

**Figure 2. Digital Video Recorder Time-Shifted data (For Illustrative Purposes as some processes may vary)**




As illustrated in the figure above, the process of preparing Digital Video Recorder Time-Shifted data for delivery is accomplished in several stages. The remainder of Section 4 addresses the distinct steps associated with each stage of the process.

The household environment described in the linear tuning section of this Guideline is similar to that necessary for the Digital Video Recorder Time Shifted Data environment, except for the required presence of equipment that includes Digital Video Recorder (DVR) functionality and an appropriate communication connection. The DVR device and its significant internal components – which may include one or more tuners, digital storage (RAM and physical disk-based storage), software, remote, etc. – provides several capabilities, most notably the ability to time-shift the consumption of television content.

A DVR allows the consumer to delay consumption of television content for any length of time, from one or more seconds to much longer – days, weeks, etc. In fact, recorded content does not necessarily have to be viewed at all. DVRs also introduce the capability to manipulate the consumption of content in forms such as fast forward, skip content, re-view content and pause. This manipulation is generally referred to as "trick-play."

DVRs typically record all content consumed on the primary tuning source on an internal hard-disk which causes very small time delays for the consumer, even if no manipulation from "live" television has been selected. Any time shifting adds to this underlying time delay. The time delay associated with un-manipulated content due to latency and writing to a DVR hard disk is not considered time shifting.

DVR functionality generally requires some form of communication back channel (telephone, satellite or Internet) to the MVPD for supplying program guide information, which is necessary for many DVR functions, such as program selection, season-pass recording, etc.

In many instances the DVR and its associated remote control takes the place of the digital set-top-box in the household, providing essentially all the capabilities of the set-top-box and the additional time-shifting capabilities described previously. DVRs are used by all types of MVPDs, but they are not present in all MVPD subscribing households, or necessarily available on all TV sets in DVR households, so time shifting should be properly accounted for in the accumulation of set-top-box data in DVR households to ensure television tuning is accurately attributed to time-periods, channels and programs when it is used by households.

Also considered under the heading of time-shifted data are those MVPD features that allow viewers to return to the beginning of a program that is in process in real time, and may offer limited trick play functionality without the need for a DVR or other special equipment in the household. This replay functionality generally delivers the same ad content as associated with the live programming stream unless otherwise noted.

Many MVPDs have DVR arrangements whereby recorded content can be viewed on other non-DVR devices as well as other networked DVRs within the household. The time-shifting processing rules of the MVPD should reflect these complexities, if offered to subscribers. For example, multi-processing may be necessary to remove duplicate tuning if the same viewer is implied. In many instances multi-tuning instances should be counted (non-multi-processed) because the viewer status is unknown. If networked DVR-type arrangements are available within subscribing households, processing rules should be disclosed.

### 4.1.1 Impact of the Household Environment on Time-Shifted Viewership

As noted previously, the household's environment can determine what type of information can be gathered by the MVPD, for example the connection method with the DVR and back-channel communication frequency dictates processing timing and other parameters. In general, when an MVPD-supplied DVR is present at the viewing installation, no digital set-top-box is necessary.

### 4.1.2 Other Digital Video Programming Delivery Streams

Other digital video programming delivery systems such as Internet enabled programming schedulers provide the ability for consumers to time-shift programming. In general, the parameters described in section 3 of the Guidelines are applicable to these other systems, however, the collection of data from these systems generally require outside-sourced collection systems for some or all of the tuning transaction information. This outside-sourced information may include the tuning information itself.

## 4.2 Time-Shifted Data Accumulation

Capture of time-shifted data includes additional data-elements necessary to accurately interpret the tuning information. These data elements should be added to tuning data for linear datasets described in previous sections of this guideline. The additional data-elements include:

- Time elements – original air time, time of recording, time of playback linked associated trick modes, relative position in the content playback from a time perspective
- Trick-play elements – operations selected such as fast forward (x-speeds), re-view backward, skipping, return to "live," pause, etc.
- Program guide usage elements
- DVR recording events
- Buffer playback
- Recording functions, with and without tuning
- Play-back elements
- DVR hardware identifier
- Tuner identifier within DVR
- Authorization codes

Each of these data elements can be necessary to successfully identify and process time-shifted data and successfully link the data elements with program and channel identifiers previously described in the linear data section of this guideline.

### 4.2.1 Initial Data Collection

It is common for the data records from DVR activity to include second-by-second data (or change-oriented records that are convertible to second-by-second data), which means the applicable data sets can be very large. Frequently, the data accumulated from DVR activity (including Tuning data, trick play or remote key-strokes, etc.) has to be compressed or filtered to reduce the file sizes for data collection by the MVPD. This compression/filtration can result in data being converted to different file structures or file formats. Appropriate controls should be maintained over the data compression and translation process (mapping, etc.) to ensure continued understanding of data fields.

Back-channel communication methods can cause varying frequency of data collection, from real-time (especially, as it relates to interactivity), to daily to weekly collection (less frequent data capture is often used for telephone connections versus IP connections). Appropriate data transmission controls should exist to ensure data collection from expected households and appropriate investigation of collection delays and/or failures.

### 4.2.2 Initial Data Store

Data should be stored by the MVPD or their processing agent after collection. This may include accumulating data over varying time periods depending upon collection method from each household. Appropriate internal controls should exist to ensure alignment of resulting data sets accurately by date and time.

Strong consideration should be given to de-identifying sensitive data early in the data-capture and storage process. Specific time-shifting tuning data is generally considered sensitive, so therefore specific household identifying information should be linked with this data only on an anonymous basis.

Because replay of DVR time-shifted data can occur over lengthy periods, data storage techniques should provide for this potential time lag between record and replay the necessary translation and mapping of data back to program, tuning source and episode specifics for accurate tuning assignment.

### 4.2.3 Consolidation & Front End Processing

Front end processing can include aligning data received to ensure data is captured in its entirety for a telecast date (program episode), even though the time-shifted playback and data collection can occur at different dates, sometimes considerably delayed. Data can also be aligned by DVR type and model (for the application of sometimes different editing rules), data communication method, geography, head-end, which can facilitate further processing, editing or special analyses.

Adequate controls should exist to ensure identification of households where DVR data is expected and recognition of DVR capabilities in data sets from these households. Data from DVR households is sometimes sampled, rather than gathered as a census. Adequate controls should exist to identify collection failures and recognize situations that may impact sample processes. If failures occur by geographic region because of certain natural disasters or problems (weather problems, power outages, etc.), these impacts should be recognized and impacts on data gathering and sample methods should be considered/disclosed. In general, the recognition of these failures should occur early in the data collection, storage and front-end processing phases through data completeness checks.

The MVPD should have automated alarms to notify proper operations personnel of significant collection failures. These controls can include periodic communication with the DVR regardless of activity, to ensure the box remains operational – this enables differentiating no device activity from data capture/collection failure.

### 4.2.4 Storing Consolidated Data

Consolidated data should be stored in a de-identified manner, but in a manner able to be cross-referenced to household information during the data enrichment process, described later in this Guideline.

Often, the data captured from DVRs includes several orientations, for example, tuning data, trick-play or other data manipulations, time stamps, tuning source, logistical or performance data on the DVR itself (storage capacity, connection quality, etc.) and hardware and software identifiers. Some data is useful for customer service, quality assessment and trouble-shooting perspectives, not necessarily for external reporting purposes.

These data can be stored in more than one file, often with a common de-identified linking record.

## 4.3 Time-Shifted Data Processing

The processing of time-shifted data is more complex than linear data processing because of the linkages that are necessary to translate time-shifted data back to the exact nature of the content and time of original telecast (episode). Because of this added complexity, generally, time-shifted processing occurs using different job streams than linear or VOD data even for activity occurring in the same household. It is also common that DVRs are not connected to all televisions in the household so individual tuning locations within a household can be processed using different job streams.

Upon completion of processing, adequate controls should exist to realign the activity and attribute complete activity to the tuning device and household.

Time-shifted processing should account for the following major functions and/or attributes:

- Power On/Off
- Identification of time-shifting instances and content impacted by time-shifting
- Identifying trick play associated with time-shifted data, or moving in and out of "live" tuning situations.
- Pause events, pause buffer information
- Recording content independent of television tuning
- Identifying play-back events, from previously stored content
- Device functionality indicator
- Consumer alerts or action initiators
- PIP Activity and PIP switches
- On/off status of various ports
- Obstruction of screen
- External to DVR Tuning Source Indicators – OTA, TV Anywhere, Slingbox, Google TV, Apple TV, etc.

- Delete
- Game Consoles

In addition to tuning data events, the following data should also be used in processing:

- Guide data, including future program lineups for future recording appointments

A key attribute of accuracy is to account for time (shifting, where applicable) and speed of display on the television screen, which are important to align the content with the original telecast (episode) of the programming.

### 4.3.1 Data Correlation and Supplementation

Similar to linear tuning, time-shifted tuning records need to be correlated and supplemented to include household, subscriber and other reference data. Generally, this reference data is appended to tuning data records after they have been collected, front-end processed and initially stored.

Data correlation and supplementation may be performed for household records that contain both (or either) linear and time-shifted data at the same time.

As in linear programming, free previews of channels should be identified and distinguishable for analysis purposes.

### 4.3.2 Data Editing and Smoothing

Data editing of time-shifted data will contain the necessary steps to align content based on program, channel, episode and tuning attributes (speed, pause, etc.) which were otherwise not covered in front-end processing and other earlier processing steps.

An important aspect of data editing is accounting for time delays introduced by DVRs, over and above the linear digital tuning environment insertion delays. An MVPD should have methods for identifying and maintaining information about time delays introduced by its equipment and communications methods, including DVRs, and this information should be accounted for in the processing stream and disclosed to users.

### 4.3.3 Storing Processed Time-Shifted Data

The storage of de-identified time-shifted data is similar to that of linear data, except that it should be inclusive of edited and consolidated time-shifted tuning events and associated activity indicators (such as trick-play). Time-shifted data can build over time (re-play) so storage can/should be appended to reflect follow-on tuning events.

### 4.3.4 Time-Shifted Data Inspection

As with linear data, the MVPD should have adequate internal controls and processes to periodically inspect time-shifted tuning data for accuracy and to ensure complete processing. Part of this includes updating of known time-delays, channel maps, program lineups, equipment and household records, populations and geographic data and sample selection methods (where applicable).

Inspection should also include quality checking of guide data supplied to households with DVRs. This guide data is important to certain elements of DVR functionality, most notably identifying and establishing pre-recording appointments for programming content.

### 4.3.5 Final Time-Shifted Data Processing and Data Set Release

Time-shifted tuning activity can be reported for further processing on a consolidated or separate basis for analysis purposes. Reporting attributes are similar to linear tuning, except that time-shifted tuning should be separated from "live" tuning as it may be assessed differently by the advertising community. The length of the time-shift (perhaps using various time intervals) should be indicated within the reported tuning records so that at minimum, the following time shifted intervals are distinguishable:

- Near live (within 15 minutes)
- Within one hour of telecast
- Same Day
- Within 1, 2, 3 days reported separately
- Within week
- Total

"Live" tuning data, not impacted by any trick-play or other time manipulation (other than standard insertion time delays), should be indicated. It is sufficient for "live" records to be indicated because they have no time-shifting or trick-play indicators present.

## 4.4 Guidelines Requirements for Time Shifted Data Set Processing (Additive to Requirements in Section 3.5)

The following is a summary of the requirements for Time-Shifted Data Sets that should be considered in conjunction with the Linear Data Set requirements outlined in Section 3.5. These requirements are contained within the narrative portion of Section 4 of the guidelines and are summarized below for ease of reference,

### 4.4.1 Accumulation

- Time shifted data-elements that should be added to tuning data for linear datasets include (Section 4.2):
  - Time elements – original air time, time of recording, time of playback linked associated trick modes, relative position in the content playback from a time perspective
  - Trick-play elements – operations selected such as fast forward (x-speeds), re-view backward, skipping, return to "live," pause, etc.
  - Program guide usage elements
  - DVR recording events
  - Buffer playback
  - Recording functions, with and without tuning

  - Play-back elements
  - DVR hardware identifier
  - Tuner identifier within DVR
  - Authorization codes
- Appropriate controls should be maintained over the data compression and translation process (mapping, etc.) to ensure continued understanding of data fields. (Section 4.2.1)
- Appropriate data transmission controls should exist to ensure data collection from expected households and appropriate investigation of collection delays and/or failures. (Section 4.2.1)
- Strong consideration should be given to de-identifying sensitive data early in the data-capture and storage process. Specific time-shifting tuning data is generally considered sensitive, so therefore specific household identifying information should be linked with this data only on an anonymous basis. (Section 4.2.2)
- Because replay of DVR time-shifted data can occur over lengthy periods, data storage techniques should provide for this potential time lag between record and replay the necessary translation and mapping of data back to program, tuning source and episode specifics for accurate tuning assignment. (Section 4.2.2)
- Adequate controls should exist to ensure identification of households where DVR data is expected and recognition of DVR capabilities in data sets from these households. Data from DVR households is sometimes sampled, rather than gathered as a census. Adequate controls should exist to identify collection failures and recognize situations that may impact sample processes. If failures occur by geographic region because of certain natural disasters or problems (weather problems, power outages, etc.), these impacts should be recognized and impacts on data gathering and sample methods should be considered/disclosed. In general, the recognition of these failures should occur early in the data collection, storage and front-end processing phases through data completeness checks. (Section 4.2.3)
- The MVPD should have automated alarms to notify proper operations personnel of significant collection failures. These controls can include periodic communication with the DVR regardless of activity, to ensure the box remains operational – this enables differentiating no device activity from data capture/collection failure. (Section 4.2.3)
- Consolidated data should be stored in a de-identified manner, but in a manner able to be cross-referenced to household information. These data can be stored in more than one file, often with a common de-identified linking record. (Section 4.2.4)

### 4.4.2 Processing

- Upon completion of processing time-shifted data, adequate controls should exist to realign the activity and attribute complete activity to the tuning device and household. (Section 4.3)

- Time-shifted processing should account for the following major functions and/or attributes (Section 4.3):
  - Power On/Off
  - Identification of time-shifting instances and content impacted by time-shifting
  - Identifying trick play associated with time-shifted data, or moving in and out of "live" tuning situations.
  - Pause events, pause buffer information
  - Recording content independent of television tuning
  - Identifying play-back events, from previously stored content
  - Device functionality indicator
  - Consumer alerts or action initiators
  - PIP Activity and PIP switches
  - On/off status of various ports
  - Obstruction of screen
  - External to DVR Tuning Source Indicators – OTA, TV Anywhere, Slingbox, Google TV, Apple TV, etc.
  - Delete
  - Game Consoles
- In addition to tuning data events, the following data should also be used in processing (Section 4.3):
  - Guide data, including future program lineups for future recording appointments
- A key attribute of accuracy is to account for time (shifting, where applicable) and speed of display on the television screen, which are important to align the content with the original telecast (episode) of the programming. (Section 4.3)
- Time-shifted tuning records need to be correlated and supplemented to include household, subscriber and other reference data. Generally, this reference data is appended to tuning data records after they have been collected, front-end processed and initially stored. (Section 4.3.1)
- Data correlation and supplementation may be performed for household records that contain both (or either) linear and time-shifted data at the same time. (Section 4.3.1)
- Free previews of channels should be identified and distinguishable for analysis purposes. (Section 4.3.1)
- An important aspect of data editing is accounting for time delays introduced by DVRs, over and above the linear digital tuning environment insertion delays. An MVPD should have methods for identifying and maintaining information about time delays introduced by its equipment and communications methods, including DVRs, and this information should be accounted for in the processing stream and disclosed to users. (Section 4.3.2)

- The MVPD should have adequate internal controls and processes to periodically inspect time-shifted tuning data for accuracy and to ensure complete processing. Part of this includes updating of known time-delays, channel maps, program lineups, equipment and household records, populations and geographic data and sample selection methods (where applicable). (Section 4.3.4)
- Inspection should include quality checking of guide data supplied to households with DVRs. (Section 4.3.4)
- Reporting attributes are similar to linear tuning, except that time-shifted tuning should be separated from "live" tuning as it may be assessed differently by the advertising community. The length of the time-shift (perhaps using various time intervals) should be indicated within the reported tuning records so that at minimum, the following time shifted intervals are distinguishable (Section 4.3.5):
  - Near live (within 15 minutes)
  - Within one hour of telecast
  - Same Day
  - Within 1, 2, 3 days reported separately
  - Within week
  - Total
- "Live" tuning data, not impacted by any trick-play or other time manipulation (other than standard insertion time delays), should be indicated. It is sufficient for "live" records to be indicated because they have no time-shifting or trick-play indicators present. (Section 4.3.5)

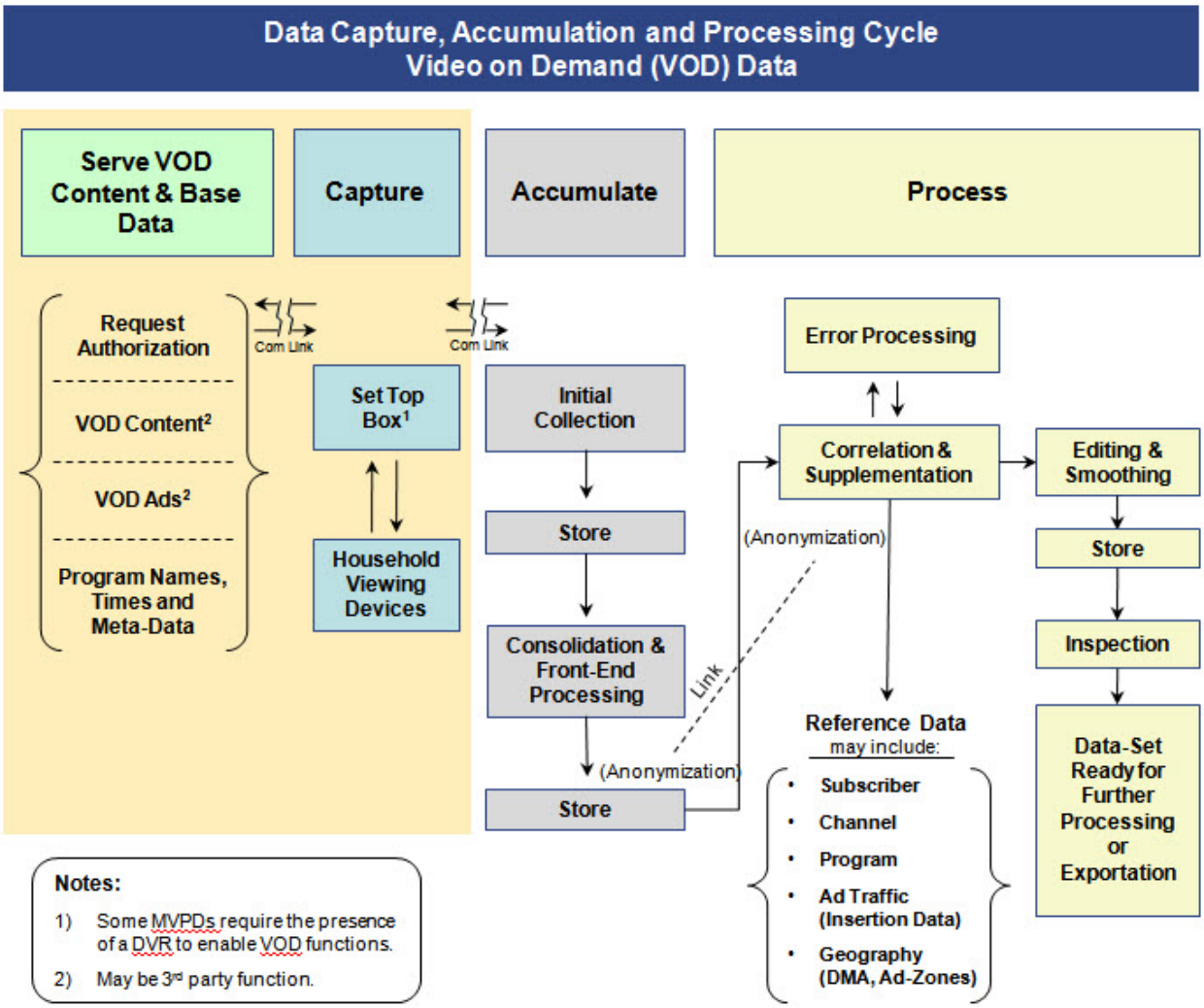
# 5 Guidelines for Video on Demand Data

This section of the Guidelines covers processing of Video on Demand Data. Information is presented to the extent it differs from the linear processing information and time-shifted information described previously – unless otherwise noted, the Guidelines for linear programming and time-shifted data apply to Digital Video on Demand data.

## 5.1 General Description of Video on Demand Data

The following chart illustrates the data-flow for the capture, accumulation and processing for Video on Demand Data, and associated activity.

### Figure 3. Video on Demand (For Illustrative Purposes as some processes may vary)




As illustrated in the figure above, the process of preparing Video on Demand Data for delivery is accomplished in several stages. The remainder of Section 5 addresses the distinct steps associated with each stage of the process.

Certain subscriber households with digital functions and often with DVRs (using the typical DVR back-channel communications) can access programming content on an as demanded basis – referred to as Video on Demand (VOD). This content may or may not include advertising, or if the demanded content has been previously aired in an ad-supported manner it may or may not include the original advertising messages. The nature of the advertising content associated with VOD transactions should be specified for users of the data by the MVPD.

VOD transactions generally have associated indicators to identify the content and whether this VOD event was a fully complete program or a sub-section of that content such as a partial roll.

VOD content can also include movie-trailers or pre-views that are essentially unique smaller sections of content.

Similar to DVR based (time-shifted) transactions, the delivery method for the VOD programming data – IP, telephone or satellite delivered – can impact how quickly the requested program is available and the nature of the request for the event. In certain instances (slower non real-time connection methods) VOD transactions are available at predetermined intervals and channels within the programming environment of the household and can be accessed using the program guide and selection functionality.

Most VOD events can be paused, restarted, retained with the ability to view unviewed content later, in a manner similar to manipulation of time-shifted data. This essentially represents a form of "trick-play" related to the VOD data. These trick-play events have similar implications on VOD data as they did in the previous section of the Guideline on time-shifted data. VOD events can have stated time limits for the availability of content for tuning. If these limits are present, accumulation of data should reflect these time windows. Tracking the time tuned and extent of progress toward viewing VOD events is also important to the users of the data. MVPD replay features as described herein are more like time shifting of programming than video on demand, as they essentially represent a form of DVR trick play.

VOD data can either be purchased or free to the subscribing household. When purchasing content is involved in initiating the VOD, the appropriate authorization codes evidencing the purchase are required. Additionally, the nature of purchased content or other VOD transactions should be maintained on a secured and anonymous basis as well as age-appropriate checking functions.

VOD functionality can require the MVPD to have a management system at the head-end or other central location that receives content requests, processes requests and delivers content. These management systems can also be an integral part of the authorization process for content delivery and ultimately for the reconciliation of content delivered to internal records for billing and reporting purposes.

### 5.1.1 Impact of the Household Environment on VOD Viewership

VOD transactions and how these transactions are initiated and accessed depend largely on the type of connection with the household. Certain connection types, for example IP connections, can enable real-time requests and availability of VOD content. Others, for example telephone connections provide VOD content on a slower non-real-time basis to the household and may be available only on a predetermined basis through the program guide.

Households access VOD content through functionality similar to that of a DVR so certain MVPDs, generally satellite providers, require the presence of a DVR to acquire VOD content. Others can provide VOD activity through standard digital set-top-box installations and remotes.

Acquiring VOD content may be subject to age-appropriate testing and authorization codes.

### 5.1.2 Other Digital Video Programming Delivery Streams

Certain VOD content can be acquired through alternative forms of digital video programming delivery systems, such as internet-based video delivery systems or internet-based DVR or Set-Top-Box controllers. In general, these systems require similar content identifiers for measurement and tracking methods to that of MVPD-based VOD. As these may be acquired through separate non-MVPD hardware and/or connections they may not be known to the MVPD, and therefore may not be measurable. Alternative delivery VOD programming may contain advertising messages, so the non-MVPD provider, if applicable, would be responsible for measurement of this content.

## 5.2 Supplying VOD Base Data and Video on Demand Accumulation

Prior to collecting and processing VOD data, the base data for consumer selection of the VOD content must be present and available to the consumer through the MVPD infrastructure in the household. This base data includes authorization functionality, VOD content, ad content (if applicable) and program and guide meta-data. This information is generally supplied through a back-channel to the set-top-box or DVR, depending on the MVPD and can be stored on these devices or real-time available through direct connections. Generally this information is searchable and selectable thorough a guide function supplied by the MVPD and selections of VOD content can be made directly from this guide using remote functions. Selection if VOD transactions then lead to initial processing stages of the content, such as authorization and staging of content and ads. Progress in tuning is also maintained as is the availability window for the content. Free VOD is processed similarly, however, authorization stages may or may not be executed.

VOD differs from Pay-per-view in that VOD content is delivered directly to that subscriber based on the transaction (essentially a unicast format), rather than Pay-per-view transactions generally being broadcast to the entire subscriber base with authorization allowing access by a particular subscriber.

VOD data is another source of tuning data along with linear and time-shifted DVR data. Most MVPDs will acquire and front-end process VOD data separately from other forms of data because of its unique attributes. In later stages of processing and reporting VOD data can be recombined with other data types by household, geographic region, programming source or other logical groupings. In order to process VOD transactions the household's set-top-box or DVR must be loaded with VOD "Client" software that enables the selection, processing and tracking of VOD content as described above. This Client software can be supplied by the MVPD or acquired from third-party suppliers.

Unique attributes of VOD data to be accumulated include:

- Authorization codes
- Program identifier
- Program provider
- Type (free versus Pay versus Subscription)

- Fee, If any
- Play window
- Time of request
- Time of tune
- Tuning progress (incomplete tunes, restart, etc.)
- Trick-play indicators

The MVPD should have adequate controls to ensure proper capture of complete VOD records and to ensure adequate identification of VOD content. VOD data should be de-identified from households or household members, although it may be combined in later stages with household and/or reference data for analysis purposes.

When data is included with the intent of measuring tuning audience, versus solely the transaction, accurate measurement of what portion of content that has been played is critical. Trick modes should be accurately captured and accounted for in calculating the portion played.

VOD data accumulation can be assessed and checked for completeness using the content management systems of the MVPD.

### 5.2.1 Initial Data Collection

MVPDs that process VOD data generally have two types of systems to enable these transactions -- Video/Ad servers and Back Office systems. Video/Ad servers provide the stream of content or ads to the household set-top-box or DVR. These can be owned by the MVPD or a third-party, similar to a third-party serving arrangement. Back-Office systems provide functions for business office functions (billing, etc.), meta-data and guide information, authorization, initial ingestion of content and ads by the MVPD from the content source, and purchase functions (credit card controls, etc.). These back-office systems are generally interfaced with the MVPD billing and subscriber systems or may simply be a subcomponent of these systems.

Initial flow of VOD content comes from a content provider and sometimes reaches a MVPD through a content distribution company that "pitches" the content to various MVPDs. VOD content is received by the MVPD, generally at a central location, and is distributed through various means to decentralized locations, such as head-ends. In the case of some satellite or telco MVPDs, this can be sent directly to the households.

When VOD content is received and "ingested" by the MVPD, and therefore ready for consumption by subscribers, it needs to be added to various menu and guide systems to enable selection. During this entire flow of capturing and staging VOD content (and ads) for the consumer, there should be adequate internal controls to ensure received content is not lost and is in fact made available to the consumer with appropriate selection functions.

VOD records can be collected at varying intervals depending on the back-channel communication method associated with the subscribing household, similar to time-shifting data. These varying intervals require processes and controls to ensure alignment of data (for multi-part tunes to content).

VOD data can be transactional (initiation, event based, close, etc.) or second-by-second. Additionally they can be free or pay transactions.

The mere request and/or purchase of VOD content may not be sufficient for users of this data, since they may be interested in determining when content was associated with actual household tuning (presumably therefore viewed by someone in the household). Purchase of VOD content does not ensure it was tuned, just as recording of time-shifted data on a DVR does not necessarily mean it is eventually tuned.

The MVPD should track VOD transactions in a manner that facilitates payment for the use of content, however for viewing data purposes whether the content was tuned is important regardless of how the MVPD is compensated for a VOD transaction. Linkage of tuning to associate request transactions is critical and should be a key component of collecting VOD data.

Storing VOD transactions should follow similar principals of de-identification as previous sensitive digital tuning data and should be retained for a sufficient period to facilitate sales, advertising and activity measurement and error follow-up.

The purchase of VOD content generally allows the consumer to use full trick-play functionality during the viewing of the content, however, in certain ad-supported situations trick-play functions can be disabled by the MVPD systems so the tuning of the ad cannot be altered. If the VOD content and the advertising are bound together by the MVPD into a single streamed unit, this restriction of trick-play is generally not possible; limitations of trick-play are generally applicable in situations where content and ads are served as separate units by the MVPD or a third-party.

When VOD content and ads are served separately as separate streams, the MVPD or a third party will generally build a play-list up front when the VOD transaction is initiated by the consumer so all applicable streams are identified and organized into the overall VOD session.

Control of time alignment among the various hardware and software components of the VOD stream is a critical aspect of the MVPD's role in processing this data. NTP services are generally necessary to keep content servers, ad servers, headend equipment, etc., aligned.

Counting on the "client" side, when assurance is present that content and ads actually aired, rather than just upon request or loading is desirable for accurate measurement.

Significantly pausing a VOD session may cause the MVPD to dismantle the play lists or otherwise stop the VOD session for internal purposes. These facilities will need to be reconstructed upon restart of the VOD session, as long as the availability window requirements are met.

Logging of VOD sessions may be accomplished in local time, however, eventually this information should be converted or available in a universal equalized time, such as a UTC or GMT to facilitate aggregation across geographies.

### 5.2.2 Initial Data Store

Data should be stored by the MVPD after collection. This may include accumulating data over varying time periods depending upon collection method from each household. Appropriate internal controls should exist to ensure alignment of resulting data sets accurately by date and time.

Strong consideration should be given to de-identifying sensitive data early in the data-capture and storage process. Specific VOD transaction and tuning data is generally considered sensitive, so therefore specific household identifying information should be linked with this data only on an anonymous basis.

Because tuning of VOD data can occur over periods of time of varying lengths, data storage techniques should provide for this potential time lag between VOD transaction and tuning.

### 5.2.3 Consolidation & Front End Processing

Front end processing can include aligning data received to ensure that data is captured completely for each VOD transaction, even considering that certain events may be delayed. Adequate controls should exist to ensure that VOD transactions are recognized and recorded only for households with the associated VOD capability. Generally, VOD transactions are measured on a census basis; however, should sampling be employed, the same rigors surrounding sampling that are noted in the time shifted section of the guidelines should be adhered to.

The MVPD should have automated alarms to notify proper operations personnel for significant collection and processing failures. Failures of DVR equipment or other set-top-box equipment can generally indicate that VOD transactions will also fail.

### 5.2.4 Storing Consolidated Data

Storing consolidated VOD data is similar in process to other types of digital tuning data, but there can be extra sensitivity around making sure data is de-identified. Stored data generally includes any trick-play records associated with that transaction.

## 5.3 Video on Demand Data Processing

Processing of Video on Demand data is similar to that of time-shifted data because of the complexities surrounding linking all appropriate events together with the VOD event, accounting for delays and trick-play activity. Also, authorization for the VOD transaction itself drives additional processing. It is common that VOD transactions are processed in different job streams from linear tuning, although similar to time-shifting, these transactions can be recombined with linear data for reporting purposes by household.

VOD processing should account for the following major functions and/or attributes:

- Initiating the VOD transaction
- Authorization and payment (payment not always applicable)

- Identification of VOD content and content type (program, sub-set of program, preview, trailer, etc.)
- Playback of VOD content
- Identification of trick-play events
- Consumer alerts or action initiators
- PIP Activity and PIP Switches
- HDMI on/off
- Obstruction of screen

In addition to VOD events, the following data should also be used in processing:

- Guide data for VOD

### 5.3.1 Data Correlation and Supplementation

Similar to linear tuning, VOD tuning records need to be correlated and supplemented to include household, subscriber and other reference data. Generally this reference data is appended to tuning data records after collection, front-end processing and initial storage.

### 5.3.2 Data Editing and Smoothing

Editing VOD records generally surround accumulating transactions from the same VOD events. MVPDs should have adequate controls to ensure that transactions are only accumulated for the same VOD event within a household. For analyses that look across households, similar controls should apply to ensure the same VOD event is the basis for accumulation.

### 5.3.3 Storing Processed Video on Demand Data

Same as linear and time-shifted data.

### 5.3.4 Time-Shifted Data Inspection

Same as linear and time-shifted data.

### 5.3.5 Final Video on Demand Data Processing and Reporting

Same as linear and time-shifted data.

## 5.4 Guidelines Requirements for Video On Demand Data Set Processing (Additive to Requirements in Section 3.5 and Section 4.4)

The following is a summary of the requirements for Video On-Demand Data Sets that should be considered in conjunction with the prior requirements (i.e. Linear and Time Shifted) outlined in this document. These requirements are contained within the narrative portion of Section 5 of the guidelines and are summarized below for ease of reference,

### 5.4.1 Accumulation

- In order to process VOD transactions the household's set-top-box or DVR must be loaded with VOD "Client" software that enables the selection, processing and tracking of VOD content as described above. This Client software can be supplied by the MVPD or acquired from third-party suppliers. (Section 5.2)
- Unique attributes of VOD data to be accumulated include (Section 5.2):
  - Authorization codes
  - Program identifier
  - Program provider
  - Type (free versus Pay versus Subscription)
  - Fee, If any
  - Play window
  - Time of request
  - Time of tune
  - Tuning progress (incomplete tunes, restart, etc.)
  - Trick-play indicators
- The MVPD should have adequate controls to ensure proper capture of complete VOD records and to ensure adequate identification of VOD content. VOD data should be de-identified from households or household members, although it may be combined in later stages with household and/or reference data for analysis purposes. (Section 5.2)
- When data is included with the intent of measuring tuning audience, versus solely the transaction, accurate measurement of what portion of content that has been played is critical. Trick modes should be accurately captured and accounted for in calculating the portion played. (Section 5.2)
- VOD data accumulation can be assessed and checked for completeness using the content management systems of the MVPD. (Section 5.2)
- The MVPD should track VOD transactions in a manner that facilitates payment for the use of content, however for viewing data purposes whether the content was tuned is important regardless of how the MVPD is compensated for a VOD transaction. Linkage of tuning to associate request transactions is critical and should be a key component of collecting VOD data. (Section 5.2.1)
- Storing VOD transactions should follow similar principals of de-identification as previous sensitive digital tuning data and should be retained for a sufficient period to facilitate sales, advertising and activity measurement and error follow-up. (Section 5.2.1)
- Control of time alignment among the various hardware and software components of the VOD stream is a critical aspect of the MVPD's role in processing this data. NTP services are generally necessary to keep content servers, ad servers, headend equipment, etc., aligned. (Section 5.2.1)

- Counting on the "client" side, when assurance is present that content and ads actually aired, rather than just upon request or loading is desirable for accurate measurement. (Section 5.2.1)
- Logging of VOD sessions may be accomplished in local time, however, eventually this information should be converted or available in a universal equalized time, such as a UTC or GMT to facilitate aggregation across geographies. (Section 5.2.1)
- Appropriate internal controls should exist to ensure alignment of resulting data sets accurately by date and time. (Section 5.2.2)
- Strong consideration should be given to de-identifying sensitive data early in the data-capture and storage process. Specific VOD transaction and tuning data is generally considered sensitive, so therefore specific household identifying information should be linked with this data only on an anonymous basis. (Section 5.2.2)
- Because tuning of VOD data can occur over periods of time of varying lengths, data storage techniques should provide for this potential time lag between VOD transaction and tuning. (Section 5.2.2)
- Generally, VOD transactions are measured on a census basis; however, should sampling be employed, the same rigors surrounding sampling that are noted in the time shifted section of the guidelines should be adhered to. (Section 5.2.3)
- The MVPD should have automated alarms to notify proper operations personnel for significant collection and processing failures. Failures of DVR equipment or other set-top-box equipment can generally indicate that VOD transactions will also fail. (Section 5.2.3)

### 5.4.2 Processing

- VOD processing should account for the following major functions and/or attributes (Section 5.3):
  - Initiating the VOD transaction
  - Authorization and payment (payment not always applicable)
  - Identification of VOD content and content type (program, sub-set of program, preview, trailer, etc.)
  - Playback of VOD content
  - Identification of trick-play events
  - Consumer alerts or action initiators
  - PIP Activity and PIP Switches
  - HDMI on/off
  - Obstruction of screen
- Guide data for VOD should also be used in processing (Section 5.3)
- MVPDs should have adequate controls to ensure that transactions are only accumulated for the same VOD event within a household. For analyses that look

across households, similar controls should apply to ensure the same VOD event is the basis for accumulation. (Section 5.3.2)

# 6 Privacy Considerations

The following privacy requirements are applicable to MVPDs and other organizations processing multi-channel digital video data who comply with these guidelines.

Protecting the privacy of subscribers is a critical objective of MVPDs and their data accumulation and processing vendor-agents (agents). In this respect, MVPDs and their agents comply with legal requirements to protect subscriber privacy. Each MVPD (and agents) maintains a Privacy Policy, and good faith efforts are made to engineer privacy considerations into systems as they are developed. MVPD (and agents) employees who engage with data accumulation and processing systems are informed of the requirements of the Privacy Policy. Since privacy laws are very complex with evolving regulations, and may vary by jurisdiction, each MVPD (and agents) administers a process to monitor and learn about new developments. New privacy-related legal requirements are analyzed and reflected into systems and communicated to applicable employees in a timely manner, as they become required by law.

A key aspect of privacy measures taken by MVPDs (and agents) is physical and logical security controls over collection, accumulation, de-identification and processing of digital set-top-box data. Each MVPD (and agents) has internal controls to provide reasonable assurance that access to digital set-top-box data and subscriber records is appropriately restricted to internal authorized users and programs and authorized agents in accordance with the Privacy Policy. Access granted to outside individuals and organizations is restricted based on the requirements of the Privacy Policy. Individual personally identifiable information of subscribers (and their household members), inclusive of tuning activity linked with subscriber identity, is not provided to entities outside of the MVPD (and agents).

The focus of this guideline is the collection, accumulation, de-identification and processing of multi-channel digital video data in stages prior to production of audience estimates, however, we recognize that further uses of this data (in later stages) may include analyses such as audience levels and consumer behavior. The MVPD (and agents) should have adequate privacy controls at the processing stages that are within the scope of this guideline, and the MVPD (and agents) should strive to ensure that these functions are maintained with similar rigor in later stages beyond this guideline.

## 6.1 Consumer Notice and Interactions

Each MVPD (and agents) considers the legal requirements of consumer notice as they evolve; most regulations in this area are only now being discussed. Transparency and consumer disclosures related to notice and interactions therefore is critical in this early stage of evaluation. Each MVPD maintains, within its privacy policy, a description of any active consumer notice requirements or protections employed, which includes opt-out functionality where utilized.

### 6.2 Privacy Policy Requirements

The following is a summary of the Privacy requirements that are contained within the narrative portion of Section 6 of the guidelines and are summarized below for ease of reference,

- Each MVPD (and agents) maintains a Privacy Policy, and good faith efforts are made to engineer privacy considerations into systems as they are developed. (Section 6)
- MVPD (and agents) employees who engage with data accumulation and processing systems are informed of the requirements of the Privacy Policy. (Section 6)
- Each MVPD (and agents) administers a process to monitor and learn about new developments in privacy laws which are complex with evolving regulations and may vary by jurisdiction. (Section 6)
- New privacy-related legal requirements are analyzed and reflected into systems and communicated to applicable employees in a timely manner, as they become required by law. (Section 6)
- Each MVPD (and agents) has internal controls to provide reasonable assurance that access to digital set-top-box data and subscriber records is appropriately restricted to internal authorized users and programs and authorized agents in accordance with the Privacy Policy. Access granted to outside individuals and organizations is restricted based on the requirements of the Privacy Policy. (Section 6)
- Individual personally identifiable information of subscribers (and their household members), inclusive of tuning activity linked with subscriber identity, is not provided to entities outside of the MVPD (and agents). (Section 6)
- Each MVPD maintains, within its privacy policy, a description of any active consumer notice requirements or protections employed, which includes opt-out functionality where utilized. (Section 6.1)

### 6.3 Privacy Policy Links and Contacts

The Privacy Policy of the MVPD (and agents) should be available for public review -- which includes subscribers, regulatory and legal authorities and users of multi-channel digital video data.

## 7 Auditing/Validation Guidance

Third-party auditing and validation is encouraged for data capture, accumulation and processing applications where multi-channel digital video tuning data is anticipated to be used by organizations external to the owner MVPD. Auditing is recommended to include both the accumulated data itself (i.e., actual physical collection testing) as well as the internal controls/process used to capture, accumulate and process MVPD data. Auditing should include agents used by the MVPD if that MVPD seeks external validation.

Although auditing reports can be issued as infrequently as once per year, some form of interim testing should be performed for more than one period during the year to assure internal

controls and compliance with these guidelines is maintained. Audit reports should clearly state the periods covered by the underlying audit testing and the period covered by the resulting compliance certification.

## 8 Conclusions and Contacts

This document represents the combined effort of the MRC, NCC Media, participant MVPDs and many other organizations to bring consistency and increased understanding to the methods for capture, accumulation and processing of multi-channel digital video tuning data, and associated transactions. We encourage adoption of these guidelines by all MVPDs that intend to externally report this data and we encourage the use of these guidelines by all users of this data to facilitate common understanding.

For further information or questions, please contact the following individuals:

MRC:

Mr. Anthony Torrieri

420 Lexington Avenue, Suite 343

New York, NY 10170

atorrieri@mediaratingcouncil.org

212-972-0300

NCC Media:

Mr. Nick Garramone

405 Lexington Avenue, 6th Floor

New York, NY 10174

nick.garramone@nccmedia.com

212-299-0756