

# **ATSC Implementation Subcommittee Finding: Essential Information to be Carried in DTV Program Streams**

**Advanced Television Systems Committee**  
1750 K Street, N.W.  
Suite 1200  
Washington, D.C. 20006

The Advanced Television Systems Committee, Inc., (ATSC) is an international, non-profit membership organization developing voluntary standards for the entire spectrum of advanced television systems.

Specifically, ATSC is working to coordinate television standards among different communications media focusing on digital television, interactive systems, and broadband multimedia communications. ATSC is also developing digital television implementation strategies and presenting educational seminars on the ATSC standards.

ATSC was formed in 1982 by the member organizations of the Joint Committee on InterSociety Coordination (JCIC): the Electronic Industries Association (EIA), the Institute of Electrical and Electronic Engineers (IEEE), the National Association of Broadcasters (NAB), the National Cable Television Association (NCTA), and the Society of Motion Picture and Television Engineers (SMPTE). Currently, there are approximately 160 members representing the broadcast, broadcast equipment, motion picture, consumer electronics, computer, cable, satellite, and semiconductor industries.

ATSC Digital TV Standards include digital high definition television (HDTV), standard definition television (SDTV), data broadcasting, multichannel surround-sound audio, and satellite direct-to-home broadcasting.

## Table of Contents

<b>1. OVERVIEW .....</b>	<b>4</b>
<b>1.1 Discussion</b>	<b>4</b>
<b>2. ESSENTIAL INFORMATION GENERAL COMMENTS.....</b>	<b>6</b>
<b>3. PROGRAM RELATED COMMENTS.....</b>	<b>6</b>
<b>3.1 Content Advisory</b>	<b>6</b>
<b>3.2 Audience Measurement</b>	<b>6</b>
<b>3.3 Video Comments</b>	<b>6</b>
<b>3.4 Scanning Parameters</b>	<b>6</b>
<b>3.5 Aspect Ratio</b>	<b>6</b>
<b>3.6 Caption Distribution Packet</b>	<b>7</b>
<b>4. DISTRIBUTOR INFORMATION COMMENTS.....</b>	<b>7</b>
<b>4.1 Time of Day</b>	<b>7</b>
<b>4.2 PSIP Update</b>	<b>7</b>
<b>4.3 Private User Data</b>	<b>7</b>
<b>5. AUDIO COMMENTS.....</b>	<b>7</b>
<b>5.1 Audio Descriptor Packet</b>	<b>7</b>
5.1.1 “Strawman” Audio Descriptor Packet	8
<b>5.2 Audio/Video Synchronization</b>	<b>9</b>
<b>6. DATA BROADCAST COMMENTS.....</b>	<b>10</b>
<b>6.1 EIT/DET Data Service Descriptor</b>	<b>10</b>
<b>6.2 PMT/SLD information</b>	<b>10</b>
<b>6.3 Data Broadcast Rate</b>	<b>10</b>
<b>6.4 Data Service Start/Stop Time</b>	<b>10</b>
<b>6.5 DST 10</b>	<b>10</b>
<b>6.6 DST Repetition Rate</b>	<b>11</b>
<b>6.7 Data Essence (and Encapsulation Specific Metadata)</b>	<b>11</b>
<b>6.8 Binding between Metadata and Data Essence</b>	<b>11</b>
<b>6.9 Out of Band Delivery</b>	<b>11</b>

## Essential Information to be Carried in DTV Program Streams

### 1. OVERVIEW

There is a need for certain information related to digital television programs to be stored and transported within and between television facilities. The ATSC Implementation Subcommittee considers this information to be *essential* for the proper operation of both professional and consumer digital television equipment. Most of this associated information is metadata about the program stream elements, however, some of it is related data essence. Although the storage and/or transport of program streams may take many forms—i.e., videotape, data tape, servers, radio frequency transmissions or terrestrial circuits—there is a common need to preserve and transmit this information.

ATSC Implementation Subcommittee and its sub-groups S1, S2, S3, and DIWG have contributed to this document. The intent is to forward this document to SMPTE for the development of a Recommended Practice.

#### 1.1 Discussion

Digital television program elements may be video, audio, data or some combination. Much of the essential information described in this report is related specifically to one of these components of a complete program. For example, aspect ratio is related to the video. Some information is not specific to one component. For example, a content advisory relates to the entire program. In such cases, this information is associated with the video component since we are generally dealing with television programs.

Some of this essential information is essence, not metadata. For example, Time of Day is not metadata about the program or its elements, however, it is critical to the operation of an affiliate station when broadcasting a network feed.

And, some of this essential information is a hybrid. For example, the Caption Distribution Packet (CDP) contains the captioning essence as well as captioning metadata. In this case, the CDP is considered information related to the video program stream. It is a widely accepted practice to associate captioning with the digital video stream. (Perhaps this is a holdover from the use of line 21 captioning in analog television.) Indeed, SMPTE has standardized a transport for closed captioning in digital video VANC space (see SMPTE 334M).

It should be noted that the original goal of this ad hoc study was to determine the essential metadata that a network or syndicator is required to convey to its affiliates for digital television. It was understood that the most likely method to transport this essential information would be in the VANC space of SMPTE 259M or SMPTE 292M serial digital video. It is assumed that the essential information listed in Table 1.1 will be carried in this manner.

**Table 1.1** Essential Information to be Carried in DTV Streams

Description	Data Payload	Delivery	Distribution	Emission	Comments
<b>Program Related:</b>					
Content Advisory	2 bytes/frame	O	X	X	See note 1.
V-ISAN	12 bytes/frame	X	X	X	For audience measurement.
AMOL II	12 bytes/frame	O	X	X	
<b>Video Related:</b>					
Scanning parameters, aspect ratio & "A" frame indication	24 bytes/frame	X	X	X	See note 2.
Caption distribution packet	Variable (max 161 bytes/frame)	O	X	X	See note 1. Also see EIA-708B.
<b>Audio Related:</b>					
Audio descriptor packet	Variable (max 20 bytes/frame)	O	X	X	Estimate, awaiting Dolby proposal.
A/V synchronization	3 bytes/frame	X	X	n/a	
<b>Data Broadcast Related:</b> (See notes 3 and 4)					
EIT/DET	Variable (likely to be small)	X	X	X	Only for standalone data service.
EIT/DET data service descriptor (profile and level)	5 bytes	X	X	X	DS bound to TV program. Additional info needed over EIT for TV.
PMT/SLD information	N*5 Bytes (N>=2)	X	X	X	N = number of data service program elements.
Data broadcast bit rate	4 bytes	X	X		1 to 1 map to bit rate.
Data service start / stop time	8 bytes	X	X		Schedules broadcast time relative to TV program.
Data service table (DST)	Variable (10-4K bytes)	X	X	X	
DST repetition rate	Small (integer)	X	X		
Binding between data, meta-data & TV program	12 bytes	X	X		Assumes V-ISAN.
<b>Distributor Information:</b>					
Time of day	16 bytes/frame	n/a	X	n/a	
PSIP EIT updates	2 bytes/frame	n/a	X	n/a	
Private user data	36 bytes/frame	n/a	X	n/a	
n/a = not applicable; O = optional; X = required if applicable					
Notes:					
1 This information (or portions thereof) is intended for emission to consumers as a component of PSIP.					
2 This information ( or portions thereof) is intended for emission to consumers in the MPEG-2 transport stream.					
3 It is assumed that data essence travels with this metadata in VANC or "in-band." See Data Broadcast comments for the "out of band" case.					
4 It is assumed that the data broadcast is constant bit rate (CBR).					
5 Total payload is typically less than 5K bytes/frame however is dependent upon number of program elements.					

## 2. ESSENTIAL INFORMATION GENERAL COMMENTS

The Delivery, Distribution, and Emission columns in Table 1.1 show where metadata should be carried through the process of television program creation and transmission. This implicitly indicates where, in the process, metadata items would ideally be created or transformed: during content creation (before delivery), at the network or syndicator (for distribution), and at stations (before emission).

Considerable discussion has transpired as how to code and transport this essential information within the television plant. Both SMPTE 334M Vertical Ancillary Data Mapping for Bit-Serial Interface and SMPTE 336M KLV coding have been suggested. There are pros and cons for each, and the IS was unable to reach consensus on a recommending a preference.

## 3. PROGRAM RELATED COMMENTS

### 3.1 Content Advisory

EIA-744 “Transport of Content Advisory Information Using Extended Data Service (XDS)” and EIA-766 “US Region Rating Table (RRT) and Content Advisory Descriptor for Transport of Content Advisory Information Using ATSC A/65 Program and Information Protocol (PSIP)” specify 54 possible combinations of content advisory information (8 MPAA and 45 TV Parental Guidelines).

### 3.2 Audience Measurement

Versioned International Standard Audiovisual Number (V-ISAN) replaces the Universal Program Identifier (UPID) as a unique program identifier. See ATSC document IS/214. Nielsen Media Research has indicated that it will make use of the V-ISAN for audience measurement in the future.

The current industry standard for ratings measurement is Nielsen Media Research’s AMOL II (automated measurement of line-ups). All major broadcast and cable networks in the U.S. support the use of AMOL.

### 3.3 Video Comments

In order to optimize receiver presentation, it is essential that aspect ratio, scanning parameters, and any film cadence information be captured during content creation and carried throughout the production and distribution processes. The payload shown in Table 1.1 assumes Video Index (SMPTE RP-186) as the model.

### 3.4 Scanning Parameters

Scanning parameters include frame rate, film cadence, speed-ups and slow-downs, interlace or progressive, active horizontal pixel count, active vertical line count, horizontal offset from SAV, and vertical offset from EAV. Colorimetry information, specifically Y equations, should be indicated.

### 3.5 Aspect Ratio

Common aspect ratios used in motion picture and television production include 4:3, 16:9, 14:9, 1.85:1, and 2.39:1. Aspect ratio metadata should include any pan and scan information necessary to adapt one aspect ratio to another, whether for processing or display. Video Index (SMPTE RP-

186) and Active Format Description (ETR-154) have been proposed as mechanisms to achieve this. For SMPTE 259M facilities, several manufacturers have implemented Video Index.

The ATSC IS/S2 has recommended that the Active Format Descriptor be adopted by the ATSC (T3/S6) and included in future standards. This transport level work is underway.

### 3.6 Caption Distribution Packet

Closed Captioning should be carried in accordance with EIA-708B utilizing the DTV Closed Caption Distribution Packet. The necessary payload can be determined from EIA-708B Section 11.1.5, which specifies the interface data rates as follows:

“The data rate required to convey a CDP stream is dependent on the frame rate of the stream and on the amount of service information included within each packet. The worst-case data rate would occur at a 60 Hz frame rate, with 16 services all full described within each CDP packet. In this case, the CDP packet size would be 161 Bytes...”

## 4. DISTRIBUTOR INFORMATION COMMENTS

### 4.1 Time of Day

Time of day time code (SMPTE 12M) and user bits are necessary as a reference to network time for an affiliate.

### 4.2 PSIP Update

IS/S2 held several discussions about how a network or syndicator could inform an affiliate of “last minute” program changes. This is particularly common during live events, such as Sports programming, when regular network programming may be “joined-in-progress” or “slid” if a game runs long. For syndication, a last minute episode change can occur. This could result in considerable data being sent to the affiliate.

It has been proposed that this information could take the form of a flag and an address, or perhaps a URL. The affiliate’s management system and/or PSIP generator would know that new or updated information was available and where it could be found. It was felt that each network would utilize a unique system, however, a common interface should be developed among PSIP generator manufactures.

### 4.3 Private User Data

Examples of use for private user data are logo triggers (ABC, CBS), automation segment identification (CBS), and up-converter control (FOX).

## 5. AUDIO COMMENTS

### 5.1 Audio Descriptor Packet

Dolby AC-3 encoders require professional metadata for proper operation. Although initially many encoders will be left in a preset configuration, Dolby strongly recommends that metadata from the program producer be utilized to ensure optimum reproduction at a consumer receiver.

It is expected that network distribution of multi-channel digital sound will utilize a compressed system such as Dolby E or Leitch Diamond Audio/APTX. These systems provide the capability for professional metadata to be embedded within the compressed audio stream. In this case, separate audio metadata is not required.

If “discrete” audio is distributed to the affiliate, a provision to distribute audio metadata is desirable. Craig Todd, of Dolby, has suggested that an Audio Descriptor Packet, similar to the Closed Caption Distribution Packet, be created for this purpose.

#### 5.1.1 “Strawman” Audio Descriptor Packet

ATSC IS/ S1 & S2 initially suggested that a minimum audio metadata set should consist of:

- Number of channels transmitted
- Channel configuration; i.e., L, R, C, LFE, and so on.
- Dynamic range parameter (dynrng) for AC-3 encoders
- Dialog normalization parameter (dialnorm) for AC-3 encoders

At the fall 2000 SMPTE Conference, Jeffrey Riedmiller of Dolby Laboratories and Tom du Breuil of Motorola Broadband presented a paper entitled “Audio in Cable, Broadcast, and Satellite Distribution—Issues and Solutions for the Digital Transition via Audio Metadata.” It contained a discussion of the “key” metadata that was desirable for Dolby Digital surround transmission to a cable head-end IRD. This “key” metadata is listed below:

- acmod
- bsmode
- cmixlev
- compr
- dialnorm
- dsurmod
- dynrng
- lfeon
- surmixlev

In general, this compares favorably with the IS/S1 and S2 suggested list.

After a review of ATSC Standard A/52, it is clear that the Riedmiller and du Breuil suggested metadata is roughly the AC-3 bit stream information (see A/52, 5.3.2) with the addition of dynrng. Using this as a guide, a “strawman” audio descriptor packet is suggested below. Since dual mono transmissions (acmod = 0) are a reasonable possibility, second channel parameters are included.

**Table 5.1** “Strawman” Audio Descriptor Packet

Metadata Item	A52/AC-3 Syntax	Data Payload	Notes
Type of service	bsmod	3 bits	bit stream mode
Channel coding	acmod	3 bits	audio coding mode
Center mix level	cmixlev	2 bits	
Surround mix level	surmixlev	2 bits	
Dolby surround mode	dsurmod	2 bits	
Low frequency effects channel status	lfeon	1 bit	
Dialogue normalization level	dailnorm	5 bits	
Compression gain word exists	compre	1 bit	
Compression gain word	compr	8 bits	
Language code exists	langcode	1 bit	
Language code	langcod	8 bits	
Dialogue normalization level, chan 2	dailnorm2	5 bits	
Compression gain word exists, chan 2	compr2e	1 bit	
Compression gain word, chan 2	compr2	8 bits	
Language code exists, chan 2	langcod2e	1 bit	
Language code, chan 2	langcod2	8 bits	
Dynamic range gain word exists	dynrange	1 bit	
Dynamic range gain word	dynrng	8 bits	
Dynamic range gain exists	dynrng2e	1 bit	
Dynamic range gain word	dynrng2	8 bits	
Time code first half exists	timcod1e	1 bit	
Time code first half	timecod1	14 bit	
Time code second half exists	timecod2e	1 bits	
Time code second half	timecod2	14 bits	
Total payload		107 bits	
Note: The data payload could be reduced slightly by assuming langcod always defaults to English and is not sent. Lancod2 remains variable. Since dual mono is mutually exclusive with surround, lfeon, cmixlev, surmixlev, and dsurmod are not sent with xxx2 elements. Therefore, the data payload can be a low as 91 bits.			

## 5.2 Audio/Video Synchronization

Some ability to track the relative latency of audio/video was considered essential by IS/S1. This is critical to maintain correct “lip sync” during downstream processing or playback. IS/S1 was considering various proposals to handle this including a combination of time stamps and some sort of audio envelope watermark. With a time stamp system, it will be important to determine exactly how audio and video programs are linked. There may be problems with voice-overs used in news and sports broadcasts which may be completely unrelated (in terms of A/V dync) to the video program and other audio associated with the program.

Their other approach, of using an audio envelope system in the form of a watermark, may be more useful, since it does not require linkage of the various audio and video sources in a

complete program. Of course in this case, the “synchronizing information” is imprinted in the video causing metadata to be superfluous.

## 6. DATA BROADCAST COMMENTS

General notes with regard to data broadcast comments include the following:

- It is assumed that the essential information travels with data essence in VANC or “in-band.”
- Metadata items are not necessarily needed on a per/frame basis. Use within a television plant will determine the necessary rate.
- Editing operations must be aware of data-related metadata presence. If not, metadata may need to appear repetitively in every frame or a metadata bridge must be provided around the edit station.
- This also assumes a continuous playout; i.e., no stops for commercials.

Table 1.1 represents the minimum set of required metadata, not necessarily the most desirable.

### 6.1 EIT/DET Data Service Descriptor

The Data Service Descriptor is used in either the EIT (Event Information Table) or DET (Data Event Table) to signal the maximum transmission bandwidth (profile) and associated buffer requirements (level) for the total data service. This signaling allows the receiver to prepare for dealing with the data service.

### 6.2 PMT/SLD information

In order to construct a valid and useful PMT (Program Map Table) and SLD (Service Location Descriptor) for the data service, the PSIP generator/emission mux needs both the `stream_type` and `association_tag` values for each Program Element making up the data service. The `stream_type` is used to identify the type of stream for the Program Element (allowing the appropriate strategy for decoding of the Program Element by the receiver). The `association_tag` value is used to uniquely bind a TAP in the DST to the Program Element carrying the data being pointed to by the TAP (survives PID remapping by remultiplexors). The `association_tag` must be unique within the PMT/SLD (meaning within the MPEG-2 Program and the Virtual Channel).

### 6.3 Data Broadcast Rate

The total broadcast rate for the data service needs to be known by the emission station provisioning/management system, in order to reserve room in the output multiplex for the data service (assuming all business and contractual arrangements have been made). The profile in the Data Service Descriptor provides only an upper bound to the Data Broadcast Rate. Note: Opportunistic data broadcast may require a different type of characteristics.

### 6.4 Data Service Start/Stop Time

The provisioning/control system needs to know when the metadata related to data broadcast is valid.

### 6.5 DST

The Data Service Table provides information to allow receivers to associate applications with references to data consumed with them. The DST is the main mechanism to bind together all of the components making up a data service. Essential fields within the DST are as follows:

- Protocol Encapsulation: This field identifies the A/90 encapsulation protocol, allowing the receiver to set up the strategy to extract the payload and other information from the data element.
- AppId (and related fields): identifies the application that consumes the data.
- TapId: The Tap is the construct that carries the association tag, which connects the description in the DST to the actual Program Element carrying the data. The TapId identifies the particular structure in the DST.
- Association Tag: As above.
- Selector: The selector field is used to provide higher-level protocol information to identify the specific resource identified by the Tap.

## 6.6 DST Repetition Rate

If the DST is not pre-encapsulated with the Data Essence, then the entity that emits the DST must be informed of the desired repetition rate. A/90 requires that the DST must be broadcast at least once during a data service. In practice, the repetition rate for the DST represents a compromise between bandwidth consumption and acquisition latency. The choice of the DST repetition rate is likely to be done by the data service provider (or author).

## 6.7 Data Essence (and Encapsulation Specific Metadata)

The Data Essence represents the content for the Data Service. The form of encapsulation (and the necessary meta-data) will depend upon the characteristics of the Data Service and the desires of the author/provider. For the purposes of creating the minimum set of metadata for data broadcast, the assumption has been made that the Data Essence has been pre-encapsulated and carries any necessary metadata.

## 6.8 Binding between Metadata and Data Essence

A/V content, Data Essence, and metadata may travel through the distribution system intimately bound or may take separate paths. Some mechanism to reconnect the metadata, A/V content, and Data Essence is needed, either implicit (for example, the VANC mechanism for attaching data to A/V) or some form of explicit connection. For this exercise, it is assumed that the V-ISAN (formerly UPID) is used as the linkage.

## 6.9 Out of Band Delivery

The assumption made in Table 1.1 is that the essential information (metadata) travels with data essence in VANC. There is a possibility that the amount of available space in VANC could be inadequate. In this case, the data essence can be delivered “out of band” (not in VANC) as an independent MPEG-2 transport stream, pre-encapsulated/pre-packetized with the Data Service Table (DST). The metadata necessary to allow the data broadcast is delivered, along with the audio and video, in VANC.

Table 6.1 represents the modification to the Data Broadcast section of Table 1.1 for this situation. The amount of metadata is not large, probably 40–50 bytes, depending upon the number of program elements comprising the data service and the length of the title string for the EIT/DET entry.

**Table 5.1** Essential Information in “Out of Band” Data Broadcasting Case

Description	Data Payload	Delivery	Distribution	Emission	Note
Information necessary to create the EIT/DET entry for this data event (title string)	Variable (likely to be small)	X	X	X	Only present if standalone data service.
EIT/DET data service descriptor (profile and level)	2 bytes	X	X	X	DS bound to TV program. Additional info needed over EIT for TV.
PMT/SLD information	N*5 bytes (N>=2)	X	X	X	N = number of data service program elements.
Data broadcast bit rate	4 bytes	X	X		1 to 1 map to bit rate.
Data service start/stop time	8 bytes	X	X		Used to schedule broadcast time for data service (relative to TV program).
Binding between data, metadata and TV program	12 bytes	X	X		Assume V-ISAN for all binding.
Notes:					
1 It is assumed that the data broadcast is constant bit rate (CBR).					
2 Metadata items are not necessarily needed on a per/frame basis. Use within a television plant may determine the necessary rate.					
3 This represents the minimum set of required metadata, not necessarily the most desirable set.					