

**Candidate Standard:
Revision B of ATSC Standard A/57, Content
Identification and Labeling for ATSC Transport**

Advanced Television Systems Committee

1750 K Street, N.W.

Suite 1200

Washington, D.C. 20006

www.atsc.org

The Advanced Television Systems Committee, Inc., is an international, non-profit organization developing voluntary standards for digital television. The ATSC member organizations represent the broadcast, broadcast equipment, motion picture, consumer electronics, computer, cable, satellite, and semiconductor industries.

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.

About the Candidate Standard

This specification is being put forth as a Candidate Standard by the TSG/S8 Specialist Group on Transport. ATSC members and non-members are encouraged to review and implement this specification and return comments to cs_amend_editor@atsc.org. ATSC Members can also send comments directly to the TSG/S8 Specialist Group. The ATSC believes this specification is stable. It is expected to progress to Proposed Standard within a period of time ending 31 December 2007.

Table of Contents

1. PURPOSE AND SCOPE	5
1.1 Foreword	5
1.2 Scope	5
2. REFERENCES	5
2.1 Normative References	5
3. DEFINITIONS	5
3.1 Compliance Notation	6
3.2 Treatment of Syntactic Elements	6
3.3 Terms	6
3.4 Acronyms	6
4. CONTENT IDENTIFIERS	6
4.1 General	6
4.2 ATSC Content Identifier	7
5. USE OF CONTENT LABELING DESCRIPTOR IN THE ATSC TRANSPORT	8
5.1 Constraints for ISAN	8
5.2 Constraints for the ATSC Content Identifier	8
6. EVENT LABELLING SEMANTICS	8
ANNEX A: ISAN GUIDE AND PRACTICE (INFORMATIVE)	9
1. INTRODUCTION	9
2. IDENTIFICATION OF CONTENT	9
3. REQUIREMENTS USED TO DESIGN THE ISAN STRUCTURE	11
4. ISAN ELEMENTS	11
4.1 Root Identifier	11
4.2 Episode Identifier	11
4.3 Version Identifier	11
5. ISAN IDENTIFIER SPACE	12
ANNEX B: OPERATIONAL ISSUES OF EVENT TRANSITIONS (INFORMATIVE)	13

Index of Tables and Figures

Table 4.1 ATSC Content Identifier Structure	7
Table A.1 ISAN Divisions by Number of Digits	12
Table A.2 ISAN Divisions by Count	12

Candidate Standard: Revision B of ATSC Standard A/57, Content Identification and Labeling for ATSC Transport

1. PURPOSE AND SCOPE

1.1 Foreword

The Advanced Television System Committee (ATSC) Technology and Standards Group (TSG) prepared this Standard. The document was approved by the members of TSG on 16 May 2007 as a Candidate Standard for a period of time ending 31 December 2007. Upon final approval by TSG and the ATSC membership, this Standard will supersede and replace A/57A.

1.2 Scope

This document defines mean to uniquely identify content (an audiovisual work). It specifies the use of an externally defined identifier, such as ISAN, and it defines an alternative house numbering system. This document also defines the semantics of the use of the MPEG-2 content labeling descriptor for labeling audiovisual content with these identifiers in the ATSC Transport.

2. REFERENCES

2.1 Normative References

The following documents contain provisions that through reference in this text constitute provisions of this Standard. At the time of publication, the editions listed were valid. All standards are subject to revision, and parties to agreements based on this Standard are encouraged to investigate the possibility of applying the most recent edition of the standards indicated below.

- [1] ISO: "ISO 15706-2:2006, Information and documentation — International Standard Audiovisual Number (ISAN)."
- [2] ISO: "ISO/IEC 13818-1:2000/Amd 1:2003, Carriage of metadata over ISO/IEC 13818-1 streams."
- [3] ISO: "ISO/IEC 13818-1:2000/Amd 4:2005, ISAN and V-ISAN use in the content labeling descriptor."
- [4] ATSC: "Program and System Information Protocol for Terrestrial Broadcast and Cable, with Amendment 1," Doc. A/65C, Advanced Television Systems Committee, Washington, D.C., 2 January 2006 (Amendment No. 1 dated 9 May 2006).
- [5] ATSC: "ATSC Digital Television Standard (A/53), Part 3 – Service Multiplex and Transport Subsystem Characteristics," Doc. A/53, Part 3:2007, Advanced Television Systems Committee, Washington, D.C., 3 January 2007.

3. DEFINITIONS

Terms, abbreviations, or units are defined in the cited normative references, or are described in Section 3.4 of this document.

3.1 Compliance Notation

As used in this document, “shall” denotes a mandatory provision of the Standard. “Should” denotes a provision that is recommended but not mandatory. “May” denotes a feature whose presence or absence does not preclude compliance, which may or may not be present at the option of the implementer.

3.2 Treatment of Syntactic Elements

This document contains symbolic references to syntactic elements used in transport coding subsystems. These references are typographically distinguished by the use of the underscore (e.g. TS_program_map_section) and may consist of character strings that are not English words (e.g. uimsbf).

3.3 Terms

ATSC Transport – The ATSC MPEG-2 Transport Stream as defined in ATSC A/53, Part 3 [5].

audiovisual work – A sequence of related images, with or without accompanying sound, which is intended to be made visible as a moving image through the use of devices, regardless of the medium of initial or subsequent fixation.

bslbf – Bit string, left bit first, where “left” is the order in which the bit strings are written in the Standard. Bits strings are written as strings of 1s and 0s within single quotation marks; e.g., ‘1000 1001’. Blanks within a bit string are for ease of reading and have no significance.

content – An audiovisual work.

forbidden – reserved and shall never be defined in the future.

reserved – a field or value that is not currently defined. If a field, shall be set to all ones. The values other than those indicated as reserved or all ones may be defined in the future.

uimsbf – Unsigned integer, most significant bit first.

work – A version of an audiovisual work.

3.4 Acronyms

DCC — Directed Channel Change

ISAN — International Standard Audiovisual Number

PMT — Program Map Table

TSID — Transport Stream Identification

UMID — Universal Material Identifier

4. CONTENT IDENTIFIERS

4.1 General

Content may be labeled with an International Standard Audiovisual Number (ISAN), an ATSC Content Identifier, or any other identifier. Constraints on use of these identifiers in the Transport Stream are in Section 5.

ISAN as used herein shall be as defined in ISO 15706-2 [1]. When an ISAN is assigned to a broadcast audiovisual work, it should be used as the content identifier. The ATSC Content Identifier alone should only be used when an ISAN is not available for the content. When both

exist, they both may be present. And, more than one ATSC Content Identifier may be present. See Annex A for background information on ISAN.

The ATSC Content Identifier is defined in Section 4.2.

4.2 ATSC Content Identifier

The ATSC Content Identifier is a structure that is composed of a TSID and a “house number” with a period of uniqueness. A “house number” is any number that the holder of the TSID wishes as constrained herein. Numbers are unique for each value of TSID. The syntax of the ATSC Content Identifier structure shall be as defined in Table 4.1.

Table 4.1 ATSC Content Identifier Structure

Syntax	No. of Bits	Format
ATSC_content_identifier() {		
TSID	16	uimsbf
reserved	2	bslbf
end_of_day	5	uimsbf
unique_for	9	uimsbf
content_id	var	uimsbf
}		

TSID – This 16 bit unsigned integer field shall contain a value of `transport_stream_id` per section 6.3.1 of A/65 [4]. Note: The assigning authority for these values for the United States is the FCC. Ranges for Mexico, Canada, and the United States have been established by formal agreement among these countries. Values in other regions are established by appropriate authorities.

end_of_day – This 5-bit unsigned integer shall be set to the hour of the day in UTC in which the broadcast day ends and the instant after which the `content_id` values may be re-used according to `unique_for`. The value of this field shall be in the range of 0-23. The values 24–31 are reserved. Note that the value of this field is expected to be static per broadcaster.

unique_for – This 9-bit unsigned integer shall be set to the number of days, rounded up, measured relative to the hour indicated by `end_of_day`, during which the `content_id` value is not reassigned to different content. The value shall be in the range 1 to 511. The value zero shall be forbidden. The value 511 shall have the special meaning of “indefinitely”. Note that the value of this field is expected to be essentially static per broadcaster, only changing when the method of house numbering is changed. Note also that decoders can treat stored `content_id` values as unique until the `unique_for` fields expire, which can be implemented by decrementing all stored `unique_for` fields by one every day at the `end_of_day` until they reach zero.

content_id – This variable length field shall be set to the value of the identifier according to the house number system or systems for the value of TSID. Each such value shall not be assigned to different content within the period of uniqueness set by the values in the `end_of_day` and `unique_for` fields. The identifier may be any combination of human readable and/or binary values and need not exactly match the form of a house number, not to exceed 242 bytes¹.

¹ This a very large number space, but this limit value was set because it is the number of bytes that can be placed in this field when the `ATSC_content_identifier()` is used in the `MPEG-2_content_labeling_descriptor()` per Section 5.

5. USE OF CONTENT LABELING DESCRIPTOR IN THE ATSC TRANSPORT

When content is labeled in the ATSC Transport, it shall be done using the ISO content_labeling_descriptor() defined in ISO/IEC 13818-1, Amendment #1 [2].

Note: ISO explicitly permits multiple copies of this descriptor to be in the same descriptor loop.

If more than one house number applies, or both an ISAN and a house number apply, or more than one house number applies, then multiple instances of the content_labeling_descriptor() may be used.

5.1 Constraints for ISAN

When using ISAN as defined in [1], the content labeling descriptor defined in [2] shall be encoded according to [3] and further constrained as follows:

- metadata_application_format shall be set to 0x0011
- content_reference_id_record_flag shall be set to '1'
- content_time_base_indicator shall be set to '0000'
- content_reference_id_record_length shall be set to 8
- content_reference_id_byte shall be set to the binary encoding of the ISAN

5.2 Constraints for the ATSC Content Identifier

When using the ATSC Content Identifier defined here, the content labeling descriptor of [2] shall be encoded and constrained as follows:

- metadata_application_format shall be set to 0xFFFF
- metadata_application_format_identifier shall be set to 0x4741 3934 ("GA94")
- content_reference_id_record_flag shall be set to '1'
- content_time_base_indicator shall be set to '0000'
- content_reference_id_record_length shall be set to the length of the instance of the ATSC_content_identifier() structure in the content_reference_id_byte field.
- content_reference_id_byte shall be set to the ATSC_content_identifier() structure as defined in Section 4.

6. EVENT LABELLING SEMANTICS

Content identifiers are metadata for a PSIP event. The use of the ISO Content Labeling Descriptor in the ATSC Transport is optional. When present, the descriptor shall be in the EIT event descriptor loop. Additionally, the descriptor may be present in the PMT in the descriptor loop immediately following the program_info_length field in the TS_program_map_section() generally following the guidelines of A/53, Part 3 [5].

If the content labeling descriptor is present for an event at any time, it shall be present in the event descriptor loop of each EIT-0 broadcast starting no later than one second after the beginning of the transmission of the essence of the event, and it shall remain present in the event descriptor loop of each EIT-0 broadcast throughout the broadcast of the essence.

See Annex B for a discussion of operational issues on event transitions.

Annex A: ISAN Guide and Practice (Informative)

1. INTRODUCTION

The television and film industry has many places where it is critically important that a particular movie, series episode, commercial, or other material (and their many versions) be succinctly and uniquely identified. These places include just about all aspects of production, post-production, network and broadcast operations:

Advertising sales	Production
Scheduling	Content Storage Management
Program guide management	Broadcasting
Audience measurement	Quality control
Asset management	Automation

Different program versions may be created by combining different data with the same audio and video. Each such version must be properly identified.

Different aspects of operations and different types of content each have their own identification schemes. Converting between these schemes is tedious, error-prone, and wasteful. An ISAN scheme meets this need and enables the integration of operations.

2. IDENTIFICATION OF CONTENT

The preferred identification for the unique and international identification of audiovisual works is the ISAN, which is defined in ISO 15706-2 [1]. For the purposes of ISAN, an audiovisual work is: “*a sequence of related images, with or without accompanying sound, which is intended to be made visible as a moving image through the use of devices, regardless of the medium of initial or subsequent fixation*”. The term “audiovisual work” is used interchangeably with the term “work” and “content”. Additional quotes from ISO 15706-2 [1] are in italics as follows:

Some examples of audiovisual works to which an ISAN may be issued are:

- *motion pictures (e.g., a feature film);*
- *short films;*
- *trailers (i.e., previews);*
- *productions for television or other means of delivery, including individual episodes of serial audiovisual works such as television series;*
- *industrial, educational and training films;*
- *commercials;*
- *broadcasts and recordings of live events such as sports events and newscasts;*
- *composite and multimedia works if they contain a significant audiovisual component, including non –linear (e.g. interactive) audiovisual components.*

Some examples of works to which an ISAN is not issued are:

- *any non-audiovisual element of an audiovisual work (e.g. the soundtrack, the screenplay, or a single image from an audiovisual work);*
- *sound recordings;*
- *still photographs, slide sets and similar static images;*
- *multimedia works that do not contain a significant audiovisual component.*

Versions of an audiovisual work may be identified as such by using a supplementary identifier or code in conjunction with the ISAN for the audiovisual work. Some examples of versions and other kinds of changes for which [a] new ISAN is not assigned are:

- *different language versions whether subtitled or dubbed;*
- *a change in the rights or ownership of an audiovisual work;*
- *different formats or resolution (e.g. analogue or digital; wide screen or “pan & scan”);*
- *editing for the purposes of television broadcast;*
- *changes in the physical carrier(s) for an audiovisual work.*

An ISAN consists of three parts: a root segment, an episode, and a version, consisting of 6, 2, and 4 bytes respectively for a total of 12 bytes, or 96 bits. Whenever a ISAN is represented in human readable form, it is displayed as sequence of hexadecimal digits (0–9, A–F) and alphanumeric (for the check digits) as follows:

ISAN RRRR-RRRR-RRRR-EEEE-X-VVVV-VVVV-Y

Where R represents the root segment, E represents the episode and V represents the version. X and Y are alphanumeric check digits computed over parts of the number as described more fully in [1].

Because the ISAN identifier is a globally unique identifier, it may be used as an index for linking elements within the ATSC Transport that refer to a work. For example, any additional transport information, such as more advanced program guide data, would use the ISAN as the index between the EIT-k event and that additional information. The same would be true for external guide data available over an alternate channel. Thus, while an ISAN identifier is optional in general, it may be used for any such linkage.

Any instance of a work identified with an ISAN identifier may be announced by including the ISAN identifier as a descriptor in an event. This includes the current event and future events. Only one ISAN identifier is permitted per event. Since each event may comprise more than one work, the meaning of the ISAN identifier is changed when sent in an ATSC Transport. In this case, all material in the event is identified with the single ISAN identifier chosen for that event. For example, the event containing a television series episode with commercials could have the unique ISAN identifier for the particular root segment, episode, and version, but the episode may be accompanied by different commercials when broadcast at different times or by different broadcasters. So, the receiver may not use this identifier to uniquely identify a particular event, but may use it to find all such events that are announced with the content identified by the ISAN. Because the identifier is unique to a work, it may be used as an index for linking other metadata to events, including private as well as other ATSC metadata. Simply labeling the additional metadata with the same ISAN identifier does this. When the receiver has a return channel, this might also include linkage to external metadata such as external program guide information or the data base associated with a particular TSID.

Since the ISAN identifier is parseable into Root Program, Episode, and Version fields, a receiver can use it to unambiguously identify:

- All episodes and versions of a root segment
- All versions of root segment R, episode E
- A particular root segment R, episode E, and version V

3. REQUIREMENTS USED TO DESIGN THE ISAN STRUCTURE

The core function is to provide a unique identifier that for a single selected piece of completed content. For example, each combination of episode and each version of that episode of a specific program would have an associated ISAN identifier. The ISAN identifier is intended to be applied after post-production, which differentiates it from a Universal Material Identifier (UMID)—a numbering system developed specifically for production and post-production use.

An ISAN identifier may be used for a variety of purposes and transported in a number of ways. It can serve as the key to index a variety of databases that reference the content that it identifies. Further examples are within program guide information to distinguish integrated programs and within non-public identification information to distinguish program continuity elements or interstitials.

An ISAN identifier can be associated with any complete program, commercial, promotional announcement, or other finished content.

Preferably, an ISAN identifier is assigned shortly before creation of the content, but after it is certain that the creation will take place. This timing allows all related systems and parties to prepare for the distribution of the new content. However, an ISAN identifier may also be assigned during or after creation.

4. ISAN ELEMENTS

The three elements described below are collectively referred to as an “ISAN identifier,” which together identify a unique piece of content. Each element is a binary number. When presented in human readable form, they are presented in a compact, consistent, and easily readable manner.

To the extent possible, the methods provided for identification of episodes and versions do not waste the number space.

The first two elements also comprise the ISAN identifier, which identifies the work as a whole and does not vary with the version. As such, it is not appropriate for broadcast identification purposes, as it is always the case that a particular version is broadcast. The ISAN is defined in ISO 15076-2 [1].

4.1 Root Identifier

A binary number assigned to a piece of content. Closely related content is distinguished by version and episode. For example, an episodic program has a common root identifier for all episodes and versions.

The Root Identifier is assigned by the registration system after a system user provides the mandatory descriptive data.

4.2 Episode Identifier

A binary number assigned by the registration system after the content owner provides the registration system with adequate episode descriptive information. All episodes of a series should have the same Root Identifier.

Programs that do not have episodes (e.g., movies), use an episode identifier of zero.

4.3 Version Identifier

A binary number assigned by the registration system after the content owner provides adequate version descriptive information. All versions that are editorially related to the original content have the same Root Identifier and if episodic have the same episode identifier. All versions are

uniquely distinguished by their having different version numbers. Assignment of different version identifiers may be based on language, aspect ratio, subtitling, or other similar characteristics.

Each editorially-significant variation in the content receives a different version identifier.

5. ISAN IDENTIFIER SPACE

The basic ISAN identifier can be represented in 16 hexadecimal digits. The divisions are presented in Tables A.1 and A.2. Note that the term “digits” means *hexadecimal* digits.

Table A.1 ISAN Divisions by Number of Digits

ISAN		
ISAN		
Digits of Root	Digits of Episode	Digits of Version
12	4	8

Table A.2 ISAN Divisions by Count

ISAN		
ISAN		
Number of Root Identifiers	Number of Episode Identifiers	Number of Versions
281,474,976,710,660	65,536	4,294,967,296

Annex B: Operational Issues of Event Transitions (Informative)

An EIT instance with the `content_labeling_descriptor` is referred to as a labeled event below. The timing requirements placed on the EIT-0 for the current event bound the period of time that an event could be mislabeled to a well-known period. This “uncertainty period” is set to 1 second, and may involve any of the following event and channel transitions:

- Switching from one labeled event to a subsequent labeled event on the same Virtual Channel
- Switching from a labeled event to a non-labeled event on the same Virtual Channel
- Switching from a non-labeled event to a labeled event on the same Virtual Channel
- Switching from a labeled event on one virtual channel to a non-labeled event on a different Virtual Channel
- Switching from a non-labeled event on one virtual channel to a labeled event on a different Virtual Channel
- Switching from a labeled event on one virtual channel to a labeled event on a different Virtual Channel

In all cases, the receiver should assume that the last identifier received has a one second uncertainty. Note that the switching scenarios above are independent of how the switch occurs—whether by the viewer, the broadcaster (via a real time MPEG transport switcher, for example), or by Directed Channel Change (DCC). See ATSC A/65 [4]. In the event that the receiver switches signal sources or has reason to know that such a switch occurred, it may elect to consider the new source to be non-ISAN-labeled for the duration of the uncertainty period. In addition, the timing of the uncertainty periods starts over with every such switch or knowledge of a switch.