



**ATSC**

ADVANCED TELEVISION  
SYSTEMS COMMITTEE

## **Call for Proposals: ATSC 3.0 Audio System**

---

Doc. S34-112r4  
(Doc. S34-2-051r7)  
08 December 2014

**Advanced Television Systems Committee**  
1776 K Street, N.W.  
Washington, D.C. 20006  
202-872-9160

## Table of Contents

<b>1</b>	<b>INTRODUCTION .....</b>	<b>4</b>
1.1	Audio for ATSC 3.0 .....	4
<b>2</b>	<b>ATSC 3.0 SCOPE OF WORK .....</b>	<b>5</b>
2.1	Background.....	5
2.2	Layered Architecture .....	5
<b>3</b>	<b>GLOSSARY.....</b>	<b>6</b>
<b>4</b>	<b>AUDIO SYSTEM REQUIREMENTS.....</b>	<b>9</b>
<b>5</b>	<b>RESPONDING TO THIS CFP .....</b>	<b>12</b>
5.1	Overview of CfP Response Documentation .....	12
5.2	Overview of Audio System Testing, Evaluation and Demonstration.....	12
5.3	Stages of the ATSC 3.0 Audio System Selection Process.....	13
<b>6</b>	<b>AUDIO SYSTEM TESTING AND EVALUATION.....</b>	<b>13</b>
6.1	Proponent Registration .....	13
6.2	Precertification.....	14
6.2.1	List of Required Technical System Details.....	14
6.2.2	Precertification Listening Tests.....	16
6.3	Phase 1 – Formal End-to-end Codec Listening Tests and Feature Evaluations .....	17
6.3.1	Formal Listening Tests.....	17
6.3.1.1	Codec Performance Tests.....	18
6.3.1.2	Immersive Headphone Tests.....	19
6.3.2	Formal Feature Evaluations.....	19
6.3.2.1	Evaluation of Non-required Features.....	20
6.4	Proponent Demonstrations.....	20
6.5	Phase 2 – Decision on Audio System Selection .....	20
6.5.1	Decision-making Process .....	20
<b>7</b>	<b>SCHEDULE.....</b>	<b>21</b>
7.1	ATSC 3.0 Project Schedule .....	21
7.2	CfP Response Schedule .....	21
7.3	Audio Content Availability and Flow.....	21
7.4	Test and Evaluation Process Schedule .....	21
<b>8</b>	<b>FORM OF SUBMISSION .....</b>	<b>23</b>
8.1	Respondent Information Form.....	23
8.2	Contents of Proposal.....	23
8.3	Compliance Form .....	23
<b>9</b>	<b>CONSIDERATION PROCESS .....</b>	<b>23</b>
9.1	ATSC Due Process .....	23
<b>10</b>	<b>INTELLECTUAL PROPERTY .....</b>	<b>23</b>
10.1	ATSC Patent Policy .....	23
10.2	Copyright .....	23
10.3	Non-Disclosure .....	23
10.4	Information Sharing.....	24

**11 RESPONDENT RESOURCES ..... 24**

**12 SUBJECT TO CHANGE ..... 24**

**13 NO COMMITMENT..... 24**

**14 NO COMPENSATION ..... 24**

**15 SUBMISSION OF RESPONSES TO CFP..... 24**

**16 REFERENCES ..... 25**

**ANNEX A: RESPONDENT INFORMATION FORM ..... 26**

**ANNEX B: CFP COMPLIANCE FORM..... 27**

**ANNEX C: LOUDSPEAKER POSITIONS..... 28**

**ANNEX D: SUMMARY OF PHASE 1 LISTENING TESTS ..... 29**

**ANNEX E: PHASE 1 FEATURE EVALUATION CRITERIA ..... 31**

**ANNEX F: ATSC IMMERSIVE AUDIO LAYOUT DIAGRAM..... 32**

**ANNEX G: ENCODE/DECODE CONTENT CREATION FOR LISTENING TESTS..... 33**

**ANNEX H: CONTENT LICENSE..... 34**

## **Call for Proposals: ATSC 3.0 Audio System**

ATSC Technology Group 3 (TG3)

### **1 INTRODUCTION**

The Advanced Television Systems Committee (ATSC) is an international, non-profit organization developing voluntary standards for digital television systems. ATSC member organizations represent the broadcast, broadcast equipment, motion picture, consumer electronics, computer, cable, satellite and semiconductor industries. The ATSC DTV Standard (applicable documents include A/52 and A/53, available at <http://www.atsc.org/>) has been adopted for use in terrestrial broadcasting by the United States, Canada, Mexico, South Korea, Guatemala, Honduras and the Dominican Republic.

ATSC is in the process of developing a new terrestrial television broadcast standard, known as “ATSC 3.0,” with advanced performance and functionality made possible by new technologies and strategies. It may not be fully backward compatible to the current ATSC standard (retrospectively referred to as “ATSC 1.0”), including its interim extensions addressing connected TV (collectively referred to as “ATSC 2.0”). This next generation standard must provide improvements in performance, functionality and efficiency that are significant enough to warrant the challenges of a transition to a new system.

ATSC 3.0 should maximize the one-to-many (point-to-multipoint) attribute of broadcasting, which enables a highly efficient means for distribution of popular content to an unlimited number of receivers. ATSC 3.0 should provide robust mobile services to untethered devices that move, such as phones, tablets, laptops and personal televisions. Since these devices are likely to move across borders, it is highly desirable that the specification contains core technologies that will have broad international acceptance and enable global interoperability. ATSC should continue its efforts to facilitate cooperation among appropriate international organizations.

#### **1.1 Audio for ATSC 3.0**

The audio subsystem for ATSC 3.0 is expected to provide an enhanced feature set, improving upon the capabilities of the current ATSC audio system. In doing so, this new system will provide the listener with both a personalized and an immersive experience.

Personalization includes enhancement to the control of dialog, use of alternate audio tracks and mixing of assistive audio services, other-language dialog, special commentary, and music and effects. In addition, the system will support both the normalization of content loudness and contouring of dynamic range, based on the specific capabilities of a user’s fixed or mobile device and its unique sound environment.

Immersive audio functionality enables high spatial resolution in sound source localization in azimuth, elevation and distance, and provides an increased sense of sound envelopment. These features are supported over the listening area. Such a system might not directly represent loudspeaker feeds, but instead could represent the overall sound field.

ATSC 3.0 audio is expected to work with “home theater” audio-visual systems, with television sets (both with and without soundbars), and also with “personal” systems such as

tablets and handheld devices like smart phones (both with and without headphones). The level of immersive audio experience may vary depending on the platform in use.

The ATSC 3.0 audio system is expected to support television content including both video and audio, and also to support audio-only content.

The goal of this Call for Proposals (CfP) is to identify currently available audio technology that satisfies the audio requirements for the ATSC 3.0 Standard.

Systems proposed will be judged discretely and in their entirety, as comprehensive, end-to-end systems for emission of the ATSC signal. ATSC does not intend to develop the ATSC 3.0 audio system out of independent components from multiple sources. As such, this CfP solicits from proponents only complete audio solutions satisfying the system needs described herein.

## **2 ATSC 3.0 SCOPE OF WORK**

While this CfP is focused on ATSC 3.0 audio only, it may be helpful to proponents to understand the broader context of the organization's approach to the standard and the foundational work completed on the standard to date.

### **2.1 Background**

ATSC is in the process of developing a standard for a new method of delivery of real-time and non-real-time television content and data to fixed and mobile devices.

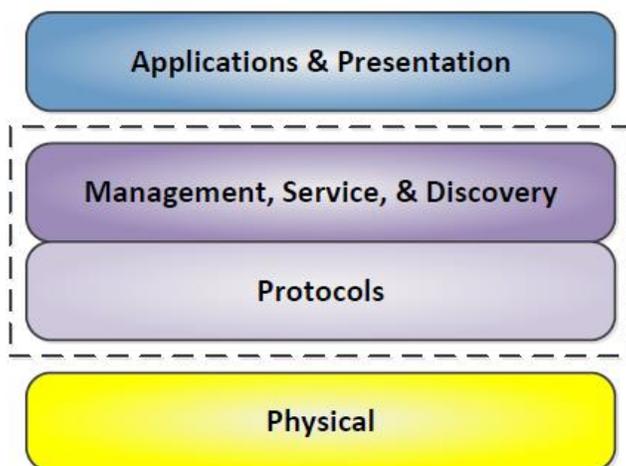
The project includes an assessment of technical requirements, research of possible solutions, and development of documentation to provide a complete specification for fixed and mobile services using new broadcast signals. Wherever practical, the standard shall utilize and reference existing standards that are found to be effective solutions to meet the requirements. Robustness of service for devices operating within the ATSC 3.0 service area should exceed that of current ATSC systems and that of cell phone and other devices enabling services similar to ATSC. Consideration will be given to technologies and proposals that enable a smooth transition from existing systems for both broadcasters and consumers. The initial scope of work for development of ATSC 3.0 is as follows:

“The ATSC 3.0 Technology Group (called TG3) will develop voluntary technical Standards and Recommended Practices for the next-generation digital terrestrial television broadcast system. ATSC 3.0 is likely to be incompatible with current broadcast systems and therefore must provide improvements in performance, functionality and efficiency significant enough to warrant implementation of a non-backwards-compatible system. Interoperability with production systems and non-broadcast distribution systems should be considered.”

### **2.2 Layered Architecture**

It has been envisioned that the ATSC 3.0 system will be designed with a “layered” architecture in order to leverage the many advantages of such a system, particularly pertaining to upgradability and extensibility. The specific layering architecture of the system is not predetermined but will be designed as the ATSC 3.0 standard is created. To set a proper course toward that end, however, and to efficiently assign the development work, a generalized layering model for ATSC 3.0 has been adopted, as shown in Figure 1. This model may be further subdivided and more tightly defined as work progresses, but it provides basic guidance for work

to begin on the system in discrete and clearly separable areas. Requirements are listed in this document with reference to the three generalized system layers shown here. Note that the middle two system layers are currently grouped into a single organizational layer, which is entitled the “Management and Protocols” (M&P) Layer. The ATSC 3.0 audio system is expected to reside in the upper layer (Applications & Presentation). Audio system signaling is expected to reside primarily in the middle layer (M&P).



**Figure 1: ATSC 3.0 Generalized Layer Architecture**

There are several specialist and ad-hoc groups within ATSC TG3 that are actively addressing the different layers of this architecture.

### 3 GLOSSARY

There are multiple potential receiver types and reception conditions envisioned for ATSC 3.0. To identify environments of receiving devices, four device types are defined to determine combined performance of a receiver and its antenna. Those are Fixed Device, Handheld Device, Vehicular Device and Portable Device. Further term definitions follow.

2.0	Nomenclature for stereo audio, with two audio channels (L, R), as found in legacy television audio systems.
5.1	Nomenclature for surround audio, with five full-range audio channels (L, C, R, LS, RS) and one low-frequency effects (LFE) channel, as found in the existing ATSC digital television audio system.
7.1+4	Nomenclature for a particular 11.1 loudspeaker arrangement suitable for immersive audio, consisting of 3 frontal loudspeakers (L, C, R) and 4 surround loudspeakers (left side [LS], left rear [LR], right side [RS], right rear [RR]) on the listener’s plane, and four speakers placed above the listener’s head height (arranged in LF, RF, LR and RR positions). (See Annex C: and Annex F:.) In ITU parlance [5], this arrangement is defined as 4+7+0, where the Upper layer is defined as 2/0/2, the Middle Layer as 3/2/2, and the Bottom layer as 0/0/0.1.

Broadcast	A distribution architecture (as defined in an IEEE/IETF sense: Broadcast/Multicast/Unicast – in other words, one-to-all, one-to-many, or one-to-one, as both Ethernet and IP architectures can be configured, for example). Broadcast architecture is supported in an over-the-air service (the only architecture supported there), and an online distribution service (wired or wireless).
Channel-based audio	A set of audio signals that is intended to be rendered directly on loudspeakers in a specific 2D or 3D physical arrangement (e.g., 22.2, 7.1+4, 5.1, stereo and the like).
Closed Captions	Text displayed on a television, video screen, or other visual display using words or symbols to provide additional or interpretive information, typically for the purpose of describing to the hearing impaired all significant audio content, including spoken dialogue and non-speech information such as the identity of speakers and, occasionally, their manner of speaking along with any significant music or sound effects. Closed Captions are not visible until activated by the viewer, usually via a remote control or menu option.
Content	One or several forms of Essence, each with its associated Metadata. Examples are Video Essence, Audio Essence, and Data Essence, plus the relevant Metadata. Thus, Content can include television programming, related or unrelated data, and software applications.
End-to-end Audio System	A complete audio system path including all components from the input of the broadcast encoder to the output of the consumer decoder.
Essence	Fundamental program material, including video, audio, graphics, data and the like, that, together with Metadata, constitute content. Unlike Metadata, Essence has inherent stand-alone value. Essence often is described in terms of a specific type of program material, e.g., video essence, audio essence, data essence, and the like.
Fixed Device	A stationary receiving device with a separate high-mounted (10 m AGL) antenna.
Handheld Device	A small form factor receiving device suitable for carrying in hand, purse or pocket. The antenna is built-in, either internal or deployable. Normal operation is either at pedestrian speeds walking or at vehicular speeds in a moving vehicle.
Higher-Order Ambisonics (HOA)	A technique in which each produced signal channel is part of an overall description of the entire sound scene, independent of the number and locations of actually available loudspeakers.
Immersive Audio	An audio system that enables high spatial resolution in sound source localization in azimuth, elevation and distance, and provides an increased sense of sound envelopment. These features are supported over the listening area. Such a system might not directly represent loudspeaker feeds but instead could represent the overall sound field.

LFE	Low-frequency effects channel. A limited frequency response channel that carries only low frequency (e.g., 100 Hz and below) audio. In advanced systems where more than one such channel is included, the nomenclature LFE1, LFE2, etc., is used.
BRIR	Binaural Room Impulse Response. A binaural room impulse response is measured by using an artificial head inside a listening room with microphones in both ear canals. A BRIR characterizes the linear transfer function from a loudspeaker to the left and right ear microphone in form of an impulse response. The impulse response should be long enough to capture the complete decay of the room reverberation tail. The BRIR database contains 30 (loudspeakers) x 2 (ears) impulse responses (see Annex C:). The BRIRs will be used to generate the reference for headphone tests via convolution with the original Channel-based items and reference-rendered Object- and HOA-based items.
LR-BRIR	Low-Resolution Binaural Room Impulse Response. A binaural room impulse response is measured by using an artificial head inside a listening room with microphones in both ear canals. A BRIR characterizes the linear transfer function from a loudspeaker to the left and right ear microphone in form of an impulse response. The impulse response should be long enough to capture the complete decay of the room reverberation tail. "Low-resolution" refers to the fact that only 30 loudspeakers at the position listed in Appendix 3 are measured, as opposed to a much finer sampling grid. So the LR-BRIR database contains 30 (loudspeakers) x 2 (ears) impulse responses. The LR-BRIRs will be used to generate the reference for headphone tests via convolution with the original Channel-based items and reference-rendered Object- and HOA-based items.
Metadata	Data describing other data. Unlike Essence, Metadata has no inherent stand-alone value.
MUSHRA	MUlti-Stimulus test with Hidden Reference and Anchor, a subjective audio quality test specified in ITU BS.1534 [1].
Non-real Time (NRT)	Generally, content that is delivered in advance of its use and stored. May refer to content that is delivered faster than real-time, such that buffering is required in the receiving device.
Object-based audio	An audio signal with parametric metadata that in combination represent an audio source that is intended to be rendered at a designated spatial position, independent of the number and location of actually available loudspeakers. Audio objects may also be used for optional or adjustable audio elements for purposes such as dialog enhancement, alternate language, or other personalizable aspects. The object metadata may also control other parameters of the audio signal, such as volume, adjustment constraints and equalization.

Portable Device	A receiving device that uses a built-in or set-top antenna, transportable to different locations but stationary during use.
Real Time (RT)	Content that is consumed concurrently with its reception.
Rendering	In the context of audio, the realization of aural content for acoustical presentation. In the context of this document, it often specifically refers to a method of aural presentation that differs from the native format of audio content, but attempts to retain the essential characteristics of the source, e.g., immersive content with a native format of 22.2 channels “rendered” for presentation in a 5.1 channel format, or Object-based audio “rendered” for presentation in a 7.1+4 channel format
Scalable Audio	An audio system that uses hierarchical techniques to: (1) Provide increasing quality of service with improving reception conditions, and/or (2) Provide different levels of service quality as required by different device types or presentation environments.
Service	A collection of media components presented to the user in aggregate; components can be of multiple media types; a Service can be either continuous or intermittent; a temporally continuous Service can consist of a sequence of events.
Sweet spot	The listening position at the coordinate origin of the loudspeaker setup (see Annex F:).

#### 4 AUDIO SYSTEM REQUIREMENTS

This section presents the ATSC 3.0 System Requirements pertaining to audio. Ultimately each of these Requirements will need to be satisfied by the selected ATSC 3.0 audio system. They are excerpted from the complete set of ATSC 3.0 System Requirements [3].

Per ATSC convention, one of the following terms is used within each requirement below:

- **shall** – This word indicates specific provisions that are to be followed strictly (no deviation is permitted).
- **shall not** – This phrase indicates specific provisions that are absolutely prohibited.
- **should** – This word indicates that a certain course of action is preferred but not necessarily required.
- **should not** – This phrase means that a certain possibility or course of action is undesirable but not prohibited.

The Requirements are grouped here based on general functionality categories, which are listed in the rightmost column in the table below. For further clarification or context, refer to [3].

<i>Req. #</i>	<i>Requirement Description</i>	<i>Audio Category</i>
65	The system shall support Immersive Audio with enhanced performance when compared with existing 5.1 or 7.1 channel-based systems.	Immersive & legacy support requirements

<b>Req. #</b>	<b>Requirement Description</b>	<b>Audio Category</b>
66	The system shall support delivery of Content from mono, stereo, 5.1 channel and 7.1 channel audio sources, as well as sources supporting Immersive Audio.	Immersive & legacy support requirements
70	The system shall enable Immersive Audio on a wide range of loudspeaker configurations.	Next-gen system flexibility & renderer requirements
71	The system shall enable Immersive Audio on loudspeaker configurations with suboptimum loudspeaker locations.	
72	The system shall enable Immersive Audio for headphones	
73	The system shall enable audio reproduction on loudspeaker configurations not designed for Immersive Audio such as 7.1 channel, 5.1 channel, two channel and single channel loudspeaker configurations.	
74	The system should enable user control of certain aspects of the sound scene that is rendered from the encoded representation, (e.g., relative level of dialog, music, effects, or other elements important to the user).	Personalization & interactive control
75	The system shall enable user-selectable alternative audio tracks to be delivered via Terrestrial Broadcast or via Broadband and in Real Time or Non-real Time. Such audio tracks could be intended to replace the primary audio track or be mixed with the primary audio track and shall be delivered for synchronous presentation with the corresponding video content.	
76	The system shall enable mixing in the receiver of alternative audio tracks (e.g., assistive audio services, other language dialog, special commentary, music and effects) with the main audio track or with each other, with relative levels and position in the sound field and with receiver adjustments suitable to the user.	

<b>Req. #</b>	<b>Requirement Description</b>	<b>Audio Category</b>
87	The system shall enable broadcasters to provide users with the option of varying the loudness of a TV Program's dialog relative to other elements of the audio mix, for purposes of increased intelligibility.	Personalization & interactive control (in combination with Next-gen loudness management & DRC)
80	The system shall support information and functionality to normalize and control the loudness of reproduced audio Content.	Next-gen loudness management & DRC
81	The system shall enable adapting the loudness and dynamic range of audio Content as appropriate for the receiving device and environment of the Content presentation.	
77	The system shall have sufficiently low latency to support live broadcasts (e.g., live news), taking account of the need for audio and video to remain synchronized.	Broadcast system & infrastructure requirements
78	The system shall enable compliance with the FCC rules relating to the Twenty-First Century Communications and Video Accessibility Act.	
79	The system shall enable Immersive Audio in combination with any ATSC 3.0 video format.	
82	The system shall support Closed Captions and video description.	
142	The system shall allow the deployment of new or extended Services, Content types, protocols, and data structures, including signaling and announcement, such that devices not built to support them continue to function normally, e.g. with no adverse effects.	
147	The system shall provide the capability to maintain audio/video synchronization within +/- 100 microseconds when measured from the input to the ATSC 3.0 emission encoding subsystem to the output of the ATSC 3.0 decoding subsystem.	
67	System shall support state-of-the-art audio codecs.	

<i>Req. #</i>	<i>Requirement Description</i>	<i>Audio Category</i>
68	The system shall support highly efficient audio compression for Immersive Audio. When delivering mono, stereo, 5.1 channel or 7.1 channel sources, the system efficiency shall be comparable to, or better than, the most efficient current systems used for delivery of these formats to the consumer, when evaluated at equivalent quality levels.	Compression requirements: Relative performance: efficiency/quality
69	The system shall enable audio reproduction for mono, 2, 5.1, and 7.1 channel and Immersive audio sources, of a quality equal to or better than that specified in Recommendation ITU-R BS.1548-4, 2013 (User requirements for audio coding systems for digital broadcasting), Appendix2, Section 2.1.1.1, as determined via the methodology described in Recommendation ITU-R BS.1116-1 extended to cover the 7.1 channel and Immersive Audio cases.	

## 5 RESPONDING TO THIS CFP

Proponents will be required to submit a documented response to the CFP as described below, and must be prepared to submit their systems to subjective testing and other evaluation by ATSC. The process will be multi-staged, following the schedule detailed in Section 7.

### 5.1 Overview of CFP Response Documentation

An initial response to this CFP (the *Registration Document*) shall be required by the Registration deadline in Table 3 below (Section 7.4). Required contents of the Registration Document are described in Section 6.1 and details on the form of submission are presented in Section 9.

For registered proposals that are accepted by ATSC (see Section 6.1), a detailed response to this CFP (the *Full Proposal*) shall be required by the Precertification deadline in Table 3 (Section 7.4). Required contents of the Full Proposal are described in Section 6.2, and include the required technical system details (Section 6.2.1) and the test report from the precertification listening tests (Section 6.2.2).

### 5.2 Overview of Audio System Testing, Evaluation and Demonstration

Proponents shall be prepared to submit their systems for testing and evaluation by ATSC. This process will include both mandatory and optional elements, as detailed in Section 6.

The process will include a request for further system documentation (the Full Proposal, see Table 1 in Section 6.2.1) as well as subjective listening tests and evaluation of feature-support elements, in a series of hierarchical steps as detailed in Sections 6.2.2 and 6.3. A detailed set of evaluation criteria for system features that will be applied to proposed systems follows in Section 6.3.2.

A final, optional element will allow proponents to demonstrate the features and operational elements of their systems to ATSC members in a free-form manner, as described in Section 6.4.

### 5.3 Stages of the ATSC 3.0 Audio System Selection Process

There will be five steps to the ATSC Audio System selection process, as follows:

1. Proponent Registration
2. Precertification
3. Formal Testing (“Phase 1”)
  - a. Subjective Listening
  - b. Feature Evaluation
    - i. Required Features
    - ii. Non-required Features (optional)
4. System Demonstrations (optional)
5. System Selection (“Phase 2”)

ATSC approval will be required after steps 1 and 2 above for a proponent to proceed to the next steps. Each of the above steps is detailed in Section 6.

## 6 AUDIO SYSTEM TESTING AND EVALUATION

Proposed audio systems shall be evaluated in their entirety, both individually and comparatively, according to the process described below.

### 6.1 Proponent Registration

All proponents are required to submit their intention to participate (Registration Document) by the Registration Date specified in Section 7. The following information must be included with this submission:

- A completed Respondent Information Form (see Annex A:).
- Description of proponent’s company and technology being proposed to ATSC).
- An overview describing how the technology meets the System Requirements for ATSC 3.0 Audio (enumerated in Section 4).
- An overview of any features that *exceed* the capabilities enabled by the System Requirements for ATSC 3.0 Audio (enumerated in Section 4).
- A statement that proponent is financially capable to undertake the process of presenting the audio system described herein.
- A statement of proponent’s access to and description of appropriate test and evaluation facilities compliant with ITU BS.1116 [2], including access to qualified test subjects (see Sections 6.2.2 and 6.3.1).
- Proponent’s agreement to abide by all relevant ATSC rules and willingness to comply with all items listed in Annex B:.
- Executed copies of all licenses required to access audio test content materials (see Annex H:).

A full review of the form of submission for registration is presented in Section 8.

ATSC will evaluate each proponent's declaration of intent to participate, and judge its ability to plausibly fulfill the requirements of the process. ATSC will advise proponents of its decision by the Acceptance date in Section 7 below.

Accepted proponents may proceed immediately to the Precertification phase of system evaluation. Any content licensing issues encountered by proponents will be handled on a case-by-case basis.

## 6.2 Precertification

Precertification is a process of evaluation of baseline system performance and system features submitted by each accepted respondent to this Call for Proposals. Respondents are required to submit a combination of materials describing their proposals in significant detail (the Full Proposal), along with the results of self-administered subjective assessments of their proposed systems. Proponents will have the opportunity to present their proposals to ATSC, to explain, clarify, and answer questions about their systems. Presentations to the Audio Experts will occur during the period between the Precertification deadline and the Certification date (see Section 7.4). The submissions will be evaluated to determine which proposals will be certified to continue to the stage of formal evaluation and independent subjective testing.

The submissions required of all respondents will be evaluated by a group of audio experts within ATSC<sup>1</sup> (hereafter, the *Audio Experts*) in an open process following ATSC rules and procedures. The results of the self-administered subjective assessments will be required to satisfy the quality requirements (see Section 4, Requirement 69) to qualify for acceptance and further evaluation of the associated coding systems. (See section 6.2.2 below for information on the subjective assessment methods to be used, the results to be submitted, and the threshold of acceptable performance.)

After the presentations, the Audio Experts will evaluate the plausibility of the proposed systems achieving the quality level reported by the test results at the coded bit-rate claimed for the system.

Section 6.2.1 below lists the information required to be included in descriptions of proposed systems. The required information will be used by the Audio Experts in arriving at their determination of plausibility of the submitted subjective assessment results being achieved by the systems proposed, and the system's suitability for use in ATSC 3.0. System descriptions and other data that are found to plausibly support the results reported will result in the associated systems being certified. The information submitted also will be used throughout the subsequent evaluation process for systems that continue beyond Precertification, to understand and evaluate the operation of the related proposals and any of their characteristics that may become significant during evaluation.

### 6.2.1 List of Required Technical System Details

The proponent's documentation shall include system diagrams in block diagram or single-line form that illustrate projections of the complete end-to-end audio system paths, from content origin to the consumer experience, for multiple production/presentation formats such as stereo, 5.1 and immersive audio. These diagrams shall include audio paths, data paths and tables that

---

<sup>1</sup> ATSC TG3/S34-2 Specialist Group on ATSC 3.0 Audio

document the control of features available on the system in a given mode, and for reproduction on all consumer platforms and devices required to be supported by ATSC 3.0.

System descriptions shall describe the audio buffer model<sup>2</sup> included in the design and used for testing. Proponents shall report the effect on bit rate at the required audio performance level of required system features and on non-required features. Proponents are required to make note of any mutually exclusive features.

Table 1 below lists the information that must be provided by all proponents as part of their submissions in the precertification process.

ITEM	DETAILS
Block Diagram(s) showing all significant processes in coding & decoding functions	From inputs to outputs of emission encoder
	From inputs to outputs of reference consumer decoder
Narrative description of all processes included in functional block diagram(s)	
Narrative description of overall operation of coding and decoding functions	Include details on codec latency performance
Technical specifications of proposed system	Input formats (channel count, bit depth, sample rates) <ul style="list-style-type: none"> <li>• Channel-based</li> <li>• Object-based</li> <li>• Soundfield-based</li> <li>• Other</li> </ul>
	Data stream <ul style="list-style-type: none"> <li>• Possible bit-rate ranges</li> <li>• Combined bit rates of audio data and metadata at typical operating points and associated audio performance</li> </ul>
	Output formats (channel count, bit depth, sample rates) <ul style="list-style-type: none"> <li>• Loudspeaker feeds</li> <li>• Headphone feeds</li> <li>• Rendering to smaller numbers of channels than coded</li> </ul>
Organization of coded data for streaming (including broadcast and broadband) transmission	Description of method for initiating decoding upon acquiring the stream*
Organization of coded data for file storage	Description of method to initiate decoding at a random point within the file*

<sup>2</sup> Characterize the methods used to enable fixed or variable encode and decode rates, and avoid buffer over- or under-flows.

ITEM	DETAILS
Synchronization	Description of method for establishing and maintaining synchronization of audio with other content components (e.g., video), delivered via broadcast and/or broadband paths*
Narrative description of methods and mechanisms to meet each ATSC 3.0 audio requirement	Include audio system latency characteristics and A/V synchronization performance*
Parameter ranges and defaults	Describe the ranges and default settings of all user-controllable features of the system (by both broadcasters and consumers)

\* It is acknowledged that these items are dependent on other non-audio elements of ATSC 3.0 and relevant decisions yet to be made. Proponents should respond with their view of an optimal context for these items.

**Table 1. System Information to be provided by proponents in Precertification phase.**

### 6.2.2 Precertification Listening Tests

All proponents will conduct self-administered listening tests of their respective systems per ITU-R BS.1116 [2], with a minimum of 20 validated listeners each, using the precertification test content materials provided by ATSC (listed in a separate document), and under the schedule detailed in Table 3.

Mandatory elements include content at 2.0 (to be evaluated on loudspeakers<sup>3</sup>), 5.1 and 7.1+4 channels (see Annex C: and Annex F:). Tests for HOA and 22.2 channels are optional. For each output configuration test set the proponent will report to ATSC the bit rate at which its system provides mean values consistently higher than 4 on the BS.1116 impairment 5-grade scale at the reference listening position, as cited in Section 2.1.1.1 of Annex 2 of ITU-R BS.1548 [4].

Mandatory and optional tests will be evaluated separately. If a proponent is judged to produce an acceptable bit rate for all mandatory tests, it will be certified to proceed to formal testing for Phase 1 mandatory tests. If such a system also participates in any optional tests, and is also judged to produce an acceptable bit rate in these tests, it will also be approved to proceed to formal testing for the corresponding Phase 1 optional tests.

Preparation of all test content shall follow the requirements detailed in Annex G:. These requirements ensure that systems utilized for testing are representative of embodiments capable of being implemented in day-to-day ATSC 3.0 applications. The coded bit rate will be measured as the total number of bits used to represent a content item divided by the content's duration, expressed as bits/sec.

Proponents shall also perform BS.1116 tests on a set of 5.1-channel content using the AC-3 codec [7] at 384 kbps, and report the resulting scores with their proposal submission. ATSC will supply all certified proponents with the source reference *and* the decoded AC-3 content elements to be used in these tests. Results of these tests will be used to set a baseline quality level for the current ATSC audio system, for assessment of proposed codecs' efficiency gain, per

<sup>3</sup> ITU-R BS.1116 permits stereo (2.0) content to be tested either on speakers or on headphones; for ATSC 3.0 Audio testing of 2.0 content, loudspeakers will be used.

Requirement 68. They will also be used to validate consistency among precertification listening rooms, test execution and test subject populations.

All tests shall be performed using software that conforms to the BS.1116 testing methodology, such as Audio Research Labs *Subjective Training and Evaluation Program* (STEP) software, or equivalent testing software. Proponents shall provide a report on their testing, including all information as specified in ITU-R BS.1116 [2] and adhering to the statistical analysis detailed in a format to be specified by a separate document that will be distributed subsequently by ATSC to registered proponents. As a part of the submitted test report, proponents also shall upload the coded and decoded audio files to the ATSC ftp site, as well as deliver a reference decoder (with appropriate licensing to ATSC) on a hardware platform of the proponent's choice, in conformance with the schedule shown in Section 7.4. The reference decoder must enable ATSC Audio Test Coordinators, without proponent assistance, to decode proponent-supplied coded audio files and store the decoded audio files for reference and comparison with submitted decoded audio files.

ATSC will validate results provided by all proponents, and advise them of their certification status according to the schedule in Section 7.2.

### 6.3 Phase 1 – Formal End-to-end Codec Listening Tests and Feature Evaluations

All Certified proponents may proceed to Phase 1 testing, which will include comparative subjective listening tests and evaluation of system functionalities. Regarding feature evaluations, each proponent's system will be examined independently, at a time and place of mutual agreement between the proponent and ATSC, but remaining within the bounds of the schedule in Section 7.

Each proponent participating in Phase 1 testing shall be expected to provide a facility suitable for hosting of formal listening tests (as described in Section 6.3.1 below), which shall be listed in the proponent's Registration Document. The facility must support testing of both mandatory and optional elements shown, regardless of whether the proponent itself is participating in any optional tests.

As an initial step in the process, proponents will be required to encode audio test material provided by ATSC (via an ftp site) through their proposed system (following the requirements of Annex G:), creating coded and decoded files following a format to be supplied to Certified proponents by ATSC. Proponents shall upload the coded and decoded results to the ATSC ftp site, in conformance with the schedule shown in Section 7.4. The reference decoder submitted during the Precertification phase (see Section 6.2.2) must enable ATSC Audio Test Coordinators, without proponent assistance, to decode proponent-supplied coded results, and to store the decoded streams for reference and comparison with submitted decoded results.<sup>4</sup>

#### 6.3.1 Formal Listening Tests

Each certified proponent shall host comparative listening tests of all certified proponents' systems per ITU-R BS.1534 [1] in a facility compliant with such a testing regime described in the proponent's Registration Document, using the Phase 1 test content materials provided by ATSC (to be listed in a subsequently distributed document to certified proponents), and under

---

<sup>4</sup> Exceptions to this (e.g., if the proponent wishes to provide a new reference decoder for bug fixes) will be handled on a case-by-case basis.

the schedule detailed in Section 7.4. If the proponent wishes to use a listening room different from that used for its precertification testing, the proponent shall re-run the BS.1116-2 tests of 5.1-channel AC-3 content noted in Section 6.2.2 in the new facility, and include these results in its Phase 1 test report.

To ensure an adequate total number of listeners' results, once the number of proponent testing sites has been identified, ATSC will determine the number of validated listeners that each site will be required to provide, and will inform the proponents of this number. In any case, the number of listeners shall not be less than 10 per site.

All test content should be prepared per the requirements in Annex G:. All source audio files will be verified as identical across all testing sites prior to each set of listening tests using MD5 hashes. See Annex D: for test audio content flow plan.

Tests shall be administered under the guidance of ATSC's Audio Test Coordinators, and be open to observation by other ATSC members. All test subjects will be selected per ITU BS.1534 [1], and provided with uniform instructions to be subsequently provided by ATSC to all Certified proponents.

All tests shall be performed using software that conforms to the ITU BS.1534 [1] testing methodology, such as Audio Research Labs *Subjective Training and Evaluation Program* (STEP) software, or equivalent testing software, and test results shall be provided to all ATSC S34-2 members and other Certified proponents, in a format to be specified in a separate document that will be distributed subsequently by ATSC to Certified proponents. Results of the listening tests will be reviewed by the Audio Test Coordinators, and presented to ATSC TG3/S34-2 as detailed in Section 6.5, and according to the schedule in Section 7.2.

Two sets of listening tests shall be performed:

- Codec performance tests (see Section 6.3.1.1)
- Immersive headphone tests (see Section 6.3.1.2)

These tests are described in the following sections, and summarized in Annex D:.

#### 6.3.1.1 Codec Performance Tests

These tests are intended to evaluate the quality vs. bit-rate performance of all systems comparatively over a range of operating points appropriate for commercial applications. All certified proponents' codecs shall be tested using the mandatory Phase 1 content elements. These content elements include channel-based material in 2.0 (evaluated on loudspeakers), 5.1 and 7.1+4 channel formats, which will be tested in their respective native speaker arrangements (i.e., without up/downmixing or rendering of any content). Speaker position details are presented in Annex C:, and 7.1+4 speaker layout is presented in Annex F:.

All mandatory content elements shall be encoded by each proponent at three different bit rates ("Low," "Medium" and "High"). If a proponent is certified for any optional content formats, these also shall be tested at three bit rates, also without any pre-rendering or downmixing of content (except for rendering to loudspeaker channels as required for HOA playback; see footnote 4). Optional content shall be tested with 22.2 speakers. The bit rates to be

used for testing for the various content formats are shown in Table 2 below.<sup>5</sup> A tolerance of  $\pm 2\%$  around these values shall be permissible.

Whereas mandatory formats shall be tested in discrete channel mode, the 22.2-channel optional format may be encoded either as discrete channels or as channels-plus-objects.

CONTENT FORMAT	LOW	MEDIUM	HIGH
Stereo	32	64	96
5.1	80	144	208
7.1+4	144	256	384
22.2 <sup>6</sup>	288	512	768
HOA <sup>7</sup> +LFE [ch=(N+1) <sup>2</sup> ]	10 kbps/ch +4 kbps/LFE	20 kbps/ch +8 kbps/LFE	30 kbps/ch +16 kbps/LFE

**Table 2. Bit rates for Phase 1 codec testing (in kbps),  $\pm 2\%$ .**

### 6.3.1.2 Immersive Headphone Tests

This test evaluates the certified proposed systems' ability to meet Requirement 72 by evaluating the impact of coding artifacts on immersive headphone rendering. Immersive listening on headphones will be tested using selected 7.1+4 content encoded at 256 kbps with each system's immersive virtualization employed during decoding.<sup>8</sup> The reference element for these tests will be a binauralized representation of the immersive original content convolved with the binaural room impulse response (BRIR) of a selected BS.1116-compliant room, which is not a room used by any proponent.<sup>9</sup>

Those proponents that were certified for any optional formats (22.2-channel and/or HOA) shall also be tested using selected Phase 1 optional elements, at a bit rate of 512 kbps for 22.2-channel content, and  $20(N+1)^2 + 8$  kbps for HOA+LFE content, where  $N$  is the HOA Order. In either case, selected content shall be rendered for immersive headphone listening by each proponent's renderer.

A summary of all subjective listening tests (both mandatory and optional) to be performed is contained in Annex D:.

### 6.3.2 Formal Feature Evaluations

Certified proponents shall present their systems to ATSC's Audio Test Coordinators and other interested ATSC members for evaluation of the required features listed in Annex E:.

<sup>5</sup> The coded bit rate for an item is the total number of bits used to represent a content item divided by the content's duration, expressed as bits/sec.

<sup>6</sup> Optional format; tests of such content are only applied to those systems that are certified for this mode.

<sup>7</sup> Optional format; tests of such content are only applied to those systems that are certified for this mode. HOA content is characterized by a certain order,  $N$ . The number of corresponding "HOA-channels" is given by  $(N+1)^2$ .

<sup>8</sup> Proponents are free to use the reference BRIR for their rendering, but shall inform ATSC if they do so.

<sup>9</sup> ATSC will select the room used to acquire the BRIR model, and provide the BRIR and the convolved reference content to all proponents as part of the Original Phase 1 content.

These feature evaluations will take place concurrently during the time window indicated in Table 3 and collocated<sup>10</sup> with the Phase 1 listening tests described in Section 6.3.1 above.

The Audio Test Coordinators will judge whether each of the performance criteria for required elements are met on a binary (“pass/fail”) basis, and report the results to ATSC TG3/S34-2.

#### 6.3.2.1 Evaluation of Non-required Features

During this feature evaluation process, proponents also may present to the Audio Test Coordinators and other ATSC observers any features of their systems that are not required by this Call for Proposals. Of particular interest are suggestions for scalable audio processes. The test coordinators will present their findings from such presentations to S34-2 in a brief narrative form.

### 6.4 Proponent Demonstrations

Certified proponents will be encouraged to present demonstrations of their systems to all ATSC members. The demonstration opportunity will be a single event for all proponents interested in participating, at a place and time to be announced by ATSC at least 60 days in advance. This demonstration will be wholly separate from the testing and evaluation described in Section 6.3 above. Details and associated costs of the demonstration event will be contained in a separate document issued by ATSC.

### 6.5 Phase 2 – Decision on Audio System Selection

ATSC’s Audio Test Coordinators will present TG3/S34-2 with results of Precertification testing (for all certified proponents), Phase 1 Formal Listening Tests, and Phase 1 Formal Feature Evaluations (including pass/fail results on required features, narrative descriptions of end-to-end system operation and narrative descriptions of features exceeding requirements).

#### 6.5.1 Decision-making Process

The ATSC TG3/S34-2 Ad Hoc Group will base its recommendation for standardization on multiple factors. All data collected in the proposal, testing and evaluation process along with unique, individual needs of the group’s members are critical components in this decision making process.

Based on the foregoing procedures, and per ATSC process, the S34-2 group will work to create a consensus recommendation for standardization. This consensus recommendation will be reviewed in the parent group TG3/S34 (Specialist Group on Applications and Presentation), and if agreed therein, the recommendation will be presented to ATSC TG3, Technology Group on ATSC 3.0 for standardization. However, the ATSC reserves the right to combine various technologies into a final audio system, which will then be documented as an ATSC Standard.

Since ATSC 3.0 devices are likely to move across borders, it is highly desirable that the specification contains core technologies which will have broad international acceptance and enable global interoperability.

---

<sup>10</sup> In the same general location; a BS.1116-compliant listening room is not required.

## 7 SCHEDULE

### 7.1 ATSC 3.0 Project Schedule

The overall schedule for ATSC 3.0 development currently calls for completion of a Candidate Standard by mid-2015. To meet this target, the decision on an audio system should be completed by mid-2015 or sooner. The schedule for responding to this CfP and subsequent steps in the selection process are based on meeting these overall targets, and are detailed below.<sup>11</sup>

### 7.2 CfP Response Schedule

Initial responses to this CfP are due 30 calendar days after the issuance of this CfP—the “Registration Date” referenced in Section 6.1 above. (See also Sections 8 and 15 for additional details.) Notice of acceptance of CfP responses will be given by ATSC to respondents no later than 35 days after issuance of this CfP. Accepted respondents can proceed to the Precertification phase immediately upon acceptance and execution of all required content licenses, as described in Section 6.1 above.

### 7.3 Audio Content Availability and Flow

Audio content for Precertification will be made available by ATSC to accepted proponents immediately upon their acceptance. Accepted proponents will be provided with ftp credentials at that time, which will allow them to access the content for downloading and testing.

Additional Phase 1 audio test content will be made available by ATSC to certified proponents via ftp according to the schedule in Section 7.4. Certified proponents will be required to return Phase 1 test content via ftp to ATSC appropriately coded (see Annex D:) and formatted by the deadline noted in Section 7.4.

See Annex D: for a diagram detailing the flow of audio test content for Phase 1 testing.

### 7.4 Test and Evaluation Process Schedule

Precertification results are due to ATSC by 90 days after the release of this CfP. Presentation of proposals will occur within two weeks of this deadline. Notice of certification will be issued to proponents by ATSC no later than 110 days from the issuance of this CfP.

Proponents will be required to encode audio test content files according to the parameters described in Section 6.3.1 (and summarized in Annex D:), following a format provided by ATSC, and return them to ATSC via ftp within two weeks. Proponents failing to meet this deadline will be eliminated from consideration.

Coded test content for Phase 1 formal listening tests will be made available to Certified proponents 145 days after release of this CfP. Proponents can then proceed to Phase 1 testing (per Section 6.3). Phase 1 listening tests (per Section 6.3.1) and presentation of features for evaluation, both required and non-required (per Section 6.3.2), must be completed no later than 200 days after the release of this CfP.

---

<sup>11</sup> The text throughout Section 7 lists approximate elapsed time windows. True deadlines showing actual calendar dates (avoiding major holidays and weekends) are listed in Table 3.

Optional system demonstrations, as described in Section 6.4 above, may take place at any time within 170 to 200 days of the release of this CfP. The exact date and location of this event will be announced by ATSC no later than 110 days from the issuance of this CfP.

The Audio Test Coordinators will deliver all results of Precertification and Phase 1 testing and feature evaluation to ATSC S34-2 as soon as possible after the completion of Phase 1, but no later than 220 days from the release of this CfP.

A summary of this schedule including exact calendar dates is shown in Table 3.

TEST ELEMENT	RESULT/DELIVERABLE	CfP	DEADLINES <sup>12</sup>
Registration	CfP response from prospective proponent	6.1	<b>12-Jan-2015</b>
Acceptance	Notice to proponent from ATSC	6.1	<b>16-Jan-2015<sup>13</sup></b>
	Precertification content made available		
Precertification	Proponent self-test report due	6.2	<b>09-Mar-2015</b>
	Detailed system descriptions due		
System Presentations	Proponent F2F presentations to ATSC	6.2	<b>11-Mar-2015<sup>14</sup></b>
Certification	Notice to proponent of ATSC acceptance	6.2	<b>20-Mar-2015</b>
Phase 1 Original Content Available	Original test content to proponents	6.3	<b>23-Mar-2015</b>
Demo Date/Location Set	Proponents advised of Demo Event details	6.4	<b>23-Mar-2015</b>
Coded Audio Uploaded	Proponents deliver coded content to ATSC	6.3	<b>10-Apr-2015</b>
Phase 1 Test Content Available	All Phase 1 test material to proponents	6.3	<b>01-May-2015</b>
Phase 1 Tests	Listening test scores (per BS.1534)	6.3	<b>26-Jun-2015</b>
	Required features (Pass/Fail)		
	Non-required features (Narrative)		
Demo Event	At Proponent's option	6.4	<b>13/14-Jul-2015<sup>15</sup></b>
Phase 2	Delivery of test results to S34-2	6.5	<b>15-Jul-2015</b>
Recommendation	Audio System decision delivered to S34	6.5	<b>14-Aug-2015</b>

**Table 3: Summary of ATSC 3.0 Audio CfP Process Schedule**

<sup>12</sup> Deadlines are 11:59 p.m. U.S. Eastern Time on the dates shown below.

<sup>13</sup> Proponents may be notified earlier than this date, depending on the submission date of their Registration. ATSC will strive to notify proponents of the Acceptance decision as soon as possible after their Registration. Proponent's access to audio content for Precertification testing will be provided immediately upon Acceptance.

<sup>14</sup> Tentative date of S34-2 meeting within ATSC F2F meeting set (09-12 March 2015), San Jose, CA.

<sup>15</sup> Tentative date(s) for Demo Event, coinciding with scheduled ATSC F2F meeting set, Washington, DC.

## 8 FORM OF SUBMISSION

As a guideline, the ATSC recommends that responses to this CfP be of an adequate length to ensure that key information is conveyed to the readers in a concise fashion. In order to expedite the review process, the most important information is to be contained in the body of the document with all supplemental information contained in an annex of secondary information that is referenced in the body. Content included that does not specifically pertain to this CfP might not be considered by the review group.

Transmittal information is provided in Section 15.

A proposal of acceptable form responding to this CFP must include the following:

### 8.1 Respondent Information Form

Responses to this CFP must include a completed Respondent Information Form, provided in Annex A:.

### 8.2 Contents of Proposal

Responses to the CfP shall fulfill all items described in Section 6.1.

### 8.3 Compliance Form

The Compliance Form for this CfP is provided in Annex B:.. A completed copy of this form, and all respondent statements that it specifies, must be included as part of a respondent's submission.

## 9 CONSIDERATION PROCESS

### 9.1 ATSC Due Process

Determination of whether a proposed methodology is incorporated into an ATSC Standard or other technical document shall be made in accordance with the due process of ATSC as described in the [ATSC Bylaws](#) and [ATSC Procedures for Technology Group and Specialist Group Operation](#). Respondents to this CFP must state that they have reviewed and agree to abide by these ATSC rules.

## 10 INTELLECTUAL PROPERTY

All respondents to this CFP must follow the guidelines detailed in the following sections.

### 10.1 ATSC Patent Policy

Respondents to this CFP must state that they will comply with the [ATSC Patent Policy](#).

### 10.2 Copyright

Respondents to this CFP must provide ATSC with the right to publish, copy, and distribute their proposed specifications as required by section 15 of the [Operational Procedures for Technology Groups and Subcommittees \(B/3\)](#).

### 10.3 Non-Disclosure

Consideration of proposals will take place in ATSC technical meetings, which are open to individuals with a direct and material interest in the work. Therefore, ATSC cannot enter into

non-disclosure agreements. Respondents must be willing to provide ATSC with enough technical detail to enable the development of standards without a non-disclosure agreement.

#### 10.4 Information Sharing

ATSC reserves the right to share responses to this CFP with other organizations supporting the development of next generation DTV standards.

### 11 RESPONDENT RESOURCES

Respondents must provide a statement that they have the financial ability and resources to participate in the ATSC evaluation process and, if selected, to fully develop their proposal into a working audio solution.

The end result of the work of TG3 will be to produce an ATSC Standard. Accomplishing this goal will require testing—both laboratory and field tests are planned. This testing process may involve certain costs to respondents that—at the date of issue of this CFP—could not be estimated.

### 12 SUBJECT TO CHANGE

ATSC reserves the right to modify or withdraw this CFP without notice.

### 13 NO COMMITMENT

ATSC reserves the right not to revise existing standards or to develop new standards based upon this CFP.

### 14 NO COMPENSATION

ATSC is a voluntary standards organization. ATSC will not provide compensation for responses to this CFP that result in specifications embodied in our Standards.

### 15 SUBMISSION OF RESPONSES TO CFP

All submissions should be made in both an electronic and printed form. Send an electronic version (in Adobe Acrobat format) to:

Mark Richer, President, ATSC: [mricher@atsc.org](mailto:mricher@atsc.org)

Jerry Whitaker, Vice President, Standards Development, ATSC: [jwhitaker@atsc.org](mailto:jwhitaker@atsc.org)

In addition, send two printed copies of each submission to:

Mark Richer  
President  
Advanced Television Systems Committee  
1776 K Street NW  
Washington, D.C. 20006  
+1 202 459 6690 (voice)

Questions relating to this CfP or the work of ATSC TG3 should be directed to Mr. Whitaker or Mr. Richer.

## 16 REFERENCES

- [1] ITU-R Recommendation BS.1534-2, “Method for the Subjective Assessment of Intermediate Sound Quality (MUSHRA)”, International Telecommunications Union, Geneva, Switzerland, 2014.
- [2] ITU-R Recommendation BS.1116-2, “Methods for the Subjective Assessment of Small Impairments in Audio Systems Including Multichannel Sound Systems,” International Telecommunications Union, Geneva, Switzerland, 2014.
- [3] ATSC TG3-S31-087r10, “ATSC 3.0 System Requirements,” Advanced Television Systems Committee, Washington, DC, 2014.
- [4] ITU-R Recommendation BS.1548-4, “User requirements for audio coding systems for digital broadcasting”, International Telecommunications Union, Geneva, Switzerland, 2013.
- [5] ITU-R Recommendation BS.2015-0, “Advanced Sound System for Programme Production,” International Telecommunications Union, Geneva, Switzerland, 2014.
- [6] JTC1/SC29/WG11/N13411, “Call for Proposals for 3D Audio,” ISO/IEC, Geneva, Switzerland, 2013.
- [7] ATSC A/52:2012, “Digital Audio Compression (AC-3) (E-AC-3) Standard,” Advanced Television Systems Committee, Washington, DC, 2012.

**ANNEX A: RESPONDENT INFORMATION FORM**

<b>Respondent Name:</b>	
<b>Primary contact name:</b>	
Address	
Mail stop or internal designation	
City, State (or Province)	
Postal Code and Country	
e-mail address	
Voice phone number	
Fax number	
<b>Secondary contact name:</b>	
Address	
Mail stop or internal designation	
City, State (or Province)	
Postal Code and Country	
e-mail address	
Voice phone number	
Fax number	

**ANNEX B: CFP COMPLIANCE FORM**

<b>Respondent Name:</b>		
<b>Required Item</b>	<b>CFP Section</b>	<b>Response</b>
Respondent agrees to support ATSC in its efforts to create and evaluate complete systems up to and including hardware implementation.	1	<input type="checkbox"/> Yes <input type="checkbox"/> No
Respondent Information Form Submitted	Annex A:	<input type="checkbox"/> Yes <input type="checkbox"/> No
System Description	5.1	<input type="checkbox"/> Yes <input type="checkbox"/> No
Submission of statement regarding Bylaws and Procedures Review and agreement	9.1	<input type="checkbox"/> Yes <input type="checkbox"/> No
Submission of statement indicating intent to comply with the ATSC Patent Policy	10.1	<input type="checkbox"/> Yes <input type="checkbox"/> No
Submission of statement indicating intent to comply with the ATSC Copyright and Reference Policy	10.2	<input type="checkbox"/> Yes <input type="checkbox"/> No
Submission of statement Regarding Respondent Resources	11	<input type="checkbox"/> Yes <input type="checkbox"/> No

**ANNEX C: LOUDSPEAKER POSITIONS**

No.	Per MPEG N13411 / ITU-R BS.2051					Layout			
	Position Label	Az. °	Az. Tol.	El. °	El. Tol.	2.0	5.1	7.1+4	22.2
1	M+000	0	±2	0	±2		C	C	X
2	M+030	30	±2	0	±2	L	LF	LF	X
3	M-030	-30	±2	0	±2	R	RF	RF	X
4	M+060	60	±2	0	±2				X
5	M-060	-60	±2	0	±2				X
6	M+090	90	±5	0	±2			LS	X
7	M-090	-90	±5	0	±2			RS	X
8	M+110	110	±5	0	±2		LSur		
9	M-110	-110	±5	0	±2		RSur		
10	M+135	135	±5	0	±2			LR	X
11	M-135	-135	±5	0	±2			RR	X
12	M+180	180	±5	0	±2				X
13	U+000	0	±2	35	±10				X
14	U+045	45	±5	35	±10			ULF	X
15	U-045	-45	±5	35	±10			URF	X
16	U+030	30	±5	35	±10				
17	U-030	-30	±5	35	±10				
18	U+090	90	±5	35	±10				X
19	U-090	-90	±5	35	±10				X
20	U+110	110	±5	35	±10				
21	U-110	-110	±5	35	±10				
22	U+135	135	±5	35	±10			ULR	X
23	U-135	-135	±5	35	±10			URR	X
24	U+180	180	±5	35	±10				X
25	T+000	0	±2	90	±10				X
26	L+000	0	±2	-15	+5, -25				X
27	L+045	45	±5	-15	+5, -25				X
28	L-045	-45	±5	-15	+5, -25				X
29	LFE1	45	±15	-15	±15		LFE	LFE	X
30	LFE2	-45	±15	-15	±15				X

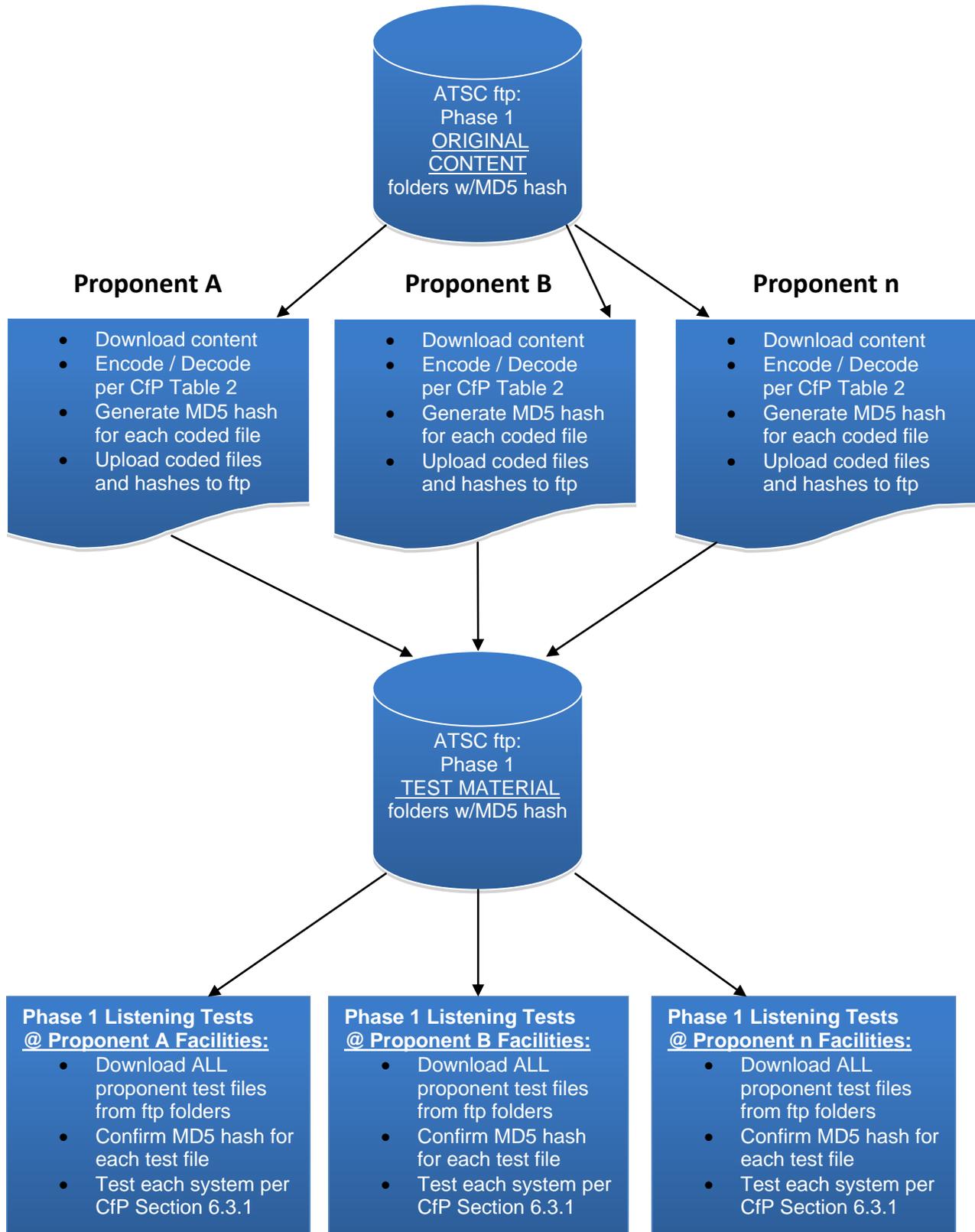
The table above shows speaker positions to be used for all listening tests, adapted from ITU [5] and ISO/IEC [6]. Negative azimuth values indicate clockwise rotation from vertical center axis (i.e., to the right side of soundstage); negative elevation values indicate downward rotation from horizontal center axis. (See also Annex F: for 7.1+4 layout diagram.)

**ANNEX D: SUMMARY OF PHASE 1 LISTENING TESTS**

The following table summarizes the subjective listening tests to be conducted by all proponents. Details are presented in Section 6.3.1.

TEST	MODE	MANDATORY /OPTIONAL	CODED BIT RATES	REFERENCE	COMMENT
<b>Codec Performance</b>	Stereo	Mandatory	32, 64, 96 kbps	Original Stereo	Channels only
	5.1	Mandatory	80, 144, 208 kbps	Original 5.1	Channels only
	7.1+4	Mandatory	144, 256, 384 kbps	Original 7.1+4	Channels only
	22.2	Optional	288, 512, 768 kbps	Original 22.2	C only or C+O
	HOA+LFE [ch=(N+1) <sup>2</sup> ]	Optional	10, 20, 30 kbps/ch, +4, 8, 16 kbps/LFE	Original HOA rendered to 22.2	All HOA Orders @ 22.2 playback
<b>Immersive Headphone</b>	7.1+4	Mandatory	256 kbps	Orig. content convolved through BRIR of BS.1116- compliant room	BRIR provided to proponents with source content (N=HOA Order)
	22.2	Optional	512 kbps		
	HOA	Optional	20(N+1) <sup>2</sup> + 8 kbps		

The diagram below shows the method of content flow that will be used for proponents to obtain encoded content for their respective listening tests, and to encode and deliver the audio test content material.

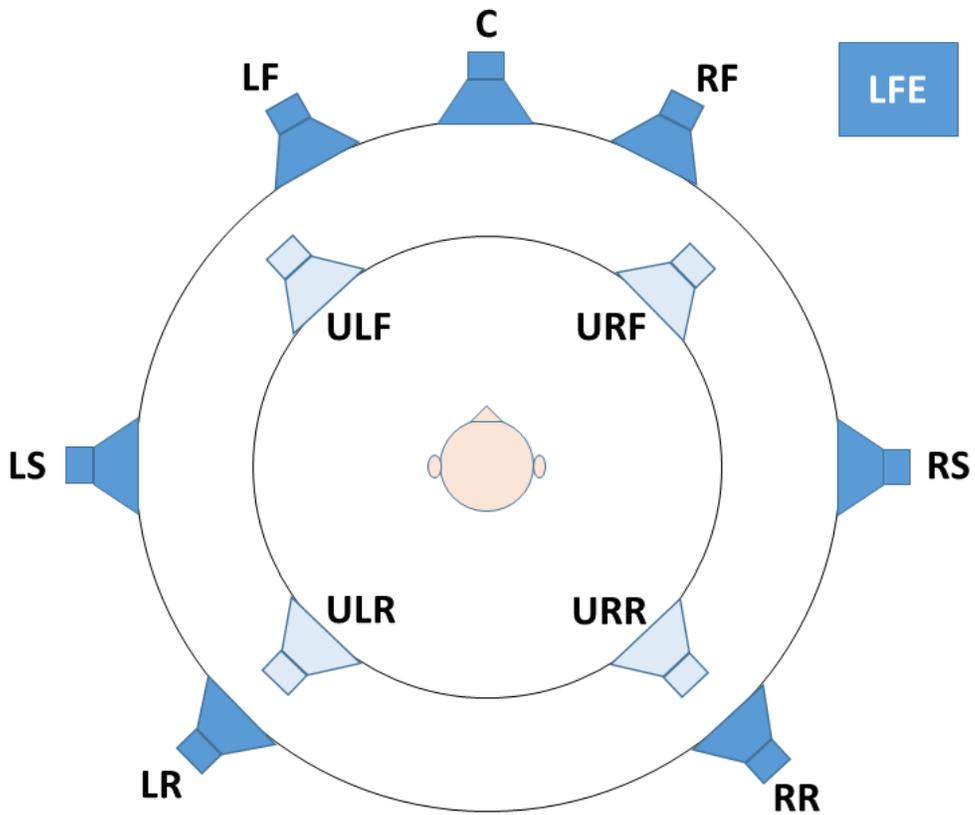


**ANNEX E: PHASE 1 FEATURE EVALUATION CRITERIA**

Each certified proponent will be asked to demonstrate its system's performance on the following required capabilities:

1. Demonstrate the system's ability to present audio objects in a channel-based environment. This shall include presenting multiple objects simultaneously, turning objects on and off, adjusting their volume, placement within the sound stage, and any other adjustable parameters available (e.g., dialog enhancement). The presentation of objects should be demonstrated with multiple channel-based environments (e.g., 2.0, 5.1 and 7.1+4 channels), and the required bit rate should be identified and described.
2. If proponent was certified for HOA, demonstrate the system's ability to present audio objects in an HOA-based environment. This shall include presenting multiple objects simultaneously, turning objects on and off, and adjusting their volume, placement within the sound stage, and any other adjustable parameters available (e.g., dialog enhancement). The presentation of objects should be demonstrated at one or more of the HOA bit rates in Table 2 (Section 6.3.1.1).
3. Demonstrate the system's ability to normalize and control output loudness.
4. Demonstrate the system's ability to control dynamic range, and adapt dynamic range to different receiver types.
5. Present examples of broadcaster control of features 1 through 4 above.
6. Demonstrate the system's ability to present immersive audio when the listener is not in a "sweet-spot" position.
7. Demonstrate the system's ability to synchronize and combine extra sound elements delivered via broadband with main soundtrack delivered via broadcast (e.g., alternate language dialog delivered via broadband replacing main dialog in broadcast soundtrack. Proponents are free to use a transport of their choice.
8. In case any system feature results in the use of different encoding parameters (e.g., audio block size, latency, etc.), demonstrate the effect of the different operating modes on audio system performance.
9. Demonstrate the system's ability to acceptably downmix coded 5.1-channel content to stereo.
10. Demonstrate the system's ability to acceptably downmix coded 7.1+4-channel content to 5.1-channel audio.
11. *If certified for optional formats only:* Demonstrate the system's ability to acceptably downmix coded HOA and/or 22.2-channel content to 7.1+4-channel audio.
12. Demonstrate the system's ability to present immersive content on a variety of speaker configurations, including improperly placed speakers.

**ANNEX F: ATSC IMMERSIVE AUDIO LAYOUT DIAGRAM**



**Figure 2: Plan view of ATSC Immersive Audio (7.1+4) speaker layout.**

## ANNEX G: ENCODE/DECODE CONTENT CREATION FOR LISTENING TESTS

This annex defines requirements and outlines the steps for the preparation of all test material for listening tests included in this Call for Proposals. This includes tests performed by the proponent as part of the Precertification phase. These requirements ensure that systems utilized for testing are representative of embodiments capable of being implemented in day-to-day ATSC 3.0 applications.

All tests shall be performed using single-pass encoding in a constant or average bit rate mode at the system's nominal latency. The coded bit rate will be measured as the total number of bits used to represent a content item divided by the content's duration, expressed as bits/sec.

The following outlines the steps needed to provide input content for encoding and decoding by the proponents and subsequent preparation of listening test material from the original and decoded content. The aim is to have each set of items to be used for testing concatenated and encoded/decoded as a single file (i.e., one concatenated file of all audio content items in a given channel format [e.g., a 2.0 file, a 5.1 file, etc.]).

This ensures that any effect of initial bit-buffer fullness is minimized and that no item-by-item tuning takes place. The decoded output of each file is subsequently split into the original sequence of separate items under test for use in the listening tests.

The following steps shall be performed:

1. Audio Test Coordinators concatenate each set of test items into a single file (one file per channel format, e.g., 2.0, 5.1, etc.).
2. Proponents encode and decode the files, and provide to ATSC the decoded output, the encoded bitstream, and a reference decoder on a hardware platform of their choice.
3. Test Coordinators verify the provided output waveform by decoding the supplied encoded bitstreams with the supplied reference decoder, comparing the result to the supplied output waveform.
4. Proponents split the provided output waveforms into the individual test items using a split-up script provided by the Test Coordinators.
5. Test coordinators set up listening test scripts<sup>16</sup> and provide the package of test items and listening test scripts to the listening test labs.

Note: In the case of Precertification testing, step 3 is performed after submission of the precertification results, and steps 4 and 5 are performed by proponents as part of their internal testing.

---

<sup>16</sup> The test script should be an \*.asi script, such as that used by Audio Research Labs *Subjective Training and Evaluation Program* (STEP) software.

## ANNEX H: CONTENT LICENSE

The following is a sample license for obtaining release of audio material from content owners to proponents for use in ATSC 3.0 Audio System testing.

### FOOTAGE LICENSE PROPONENT COMPANY AGREEMENT

This License Agreement (the "Agreement") is entered into and is effective as of September 12, 2014, between \_\_\_\_\_ ("Content Provider"), and \_\_\_\_\_ ("Proponent Company") located at \_\_\_\_\_, regarding the use of certain Footage (as defined below).

WHEREAS, Content Provider owns or has the right to exploit digital entertainment audio content, and

WHEREAS, ATSC is developing a new standard currently referred to as ATSC 3.0 (the "Standard"), and

WHEREAS, Proponent Company wishes to use Content Provider's audio content for certain non-public testing purposes to assist ATSC in its development of the audio system to be selected for the Standard (the "Testing");

NOW, THEREFORE, the following is agreed to and understood between the parties hereto:

**1. The Footage** Content Provider shall employ commercially reasonable efforts to provide audio or audio-visual materials chosen by Content Provider in its sole discretion, which may include digital files of motion pictures or television programs listed in **Exhibit A** attached hereto (the "Footage"), and encoded as set forth in **Exhibit A**. Delivery of the Footage to Proponent Company shall be via secure electronic delivery or as otherwise agreed to by the parties.

**2. Ownership** As between the parties to this Agreement, Proponent Company acknowledges that Content Provider is the sole owner of the Footage, including all copyright and other exclusive rights in and to the Footage and any derivative works of the Footage that may be created in relation to the Testing. Proponent Company shall make no claim as to any ownership interest in the Footage, such derivative works, or any portion thereof.

**3. Rights Granted** Content Provider grants to Proponent Company a non-exclusive license to use the Footage solely (i) for the purpose of the Testing solely at its facility located at \_\_\_\_\_ and according to the rules set forth by the Advanced Television Systems Committee, Inc. ("ATSC") through its appointed test coordinators (ii) in connection with reporting the results of the Testing (including through technical demonstrations) in non-public ATSC meetings where attendance is restricted to ATSC members, other ATSC approved participants and Proponent Companies ("Reporting"). Unless otherwise specified in writing between the parties, the License shall expire 210 days from delivery.

**4. Limitations of Use** The Footage is being provided to Proponent Company solely for the purpose of the Testing and Reporting, and any exhibition, broadcast, distribution, public performance or display or other exploitation of the Footage is strictly prohibited. Access by Proponent Company representatives to the Footage shall be strictly limited to Proponent Company employees, ATSC staff and ATSC's appointed test coordinators whose access is necessary to conduct the Testing.

**5. License Fees** The Footage is being provided to Proponent Company without payment of any monetary license fee. In consideration of Content Provider providing the Footage to Proponent Company, Proponent Company agrees to share with Content Provider all results obtained from the testing and analysis of the Footage by Proponent Company.

**6. Return of Material** Upon expiration of the License, all Footage and any copies of it must be either (1) returned to Content Provider at Proponent Company's expense, or (2) destroyed, with a written certification to such effect being delivered to Content Provider. Under no circumstances may Proponent Company retain any Footage, or conduct further testing using the Footage, after the expiration of the License.

**7. Warranty/Representation/Indemnity** Content Provider makes no warranty or representation, express or implied, with respect to the Footage. Proponent Company agrees to indemnify and hold Content Provider (and its parent and subsidiary companies, partners, participants and assigns, and their respective officers, directors, agents and employees) harmless from any actions, claims, liabilities, damages or costs (including reasonable attorney fees) of any kind or nature whatsoever which may arise out of (a) Proponent Company's use of the Footage; (b) the exhibition or distribution of any of the Footage to any third party (other than as expressly authorized and set forth herein); (c) Proponent Company's breach of this Agreement; or (d) Proponent Company's failure to have all necessary rights to any Proponent Company-owned or licensed equipment or technology used to exercise its rights and perform its obligations hereunder. This indemnity shall survive the termination of this Agreement.

**8. Miscellaneous** The parties to this Agreement are independent contractors with respect to each other, and nothing in this Agreement shall create any association, partnership, joint venture or agency relationship between the parties. This Agreement shall be construed in accordance with the laws of \_\_\_\_\_. The Parties agree to submit to the jurisdiction of the United States courts located in \_\_\_\_\_.

**An executed copy of this agreement must be sent by the Proponent Company to ATSC as follows:**

Mark Richer  
President  
Advanced Television Systems Committee, Inc.  
1776 K Street NW  
Washington, DC 2006  
[mricher@atsc.org](mailto:mricher@atsc.org)

**ACCEPTED AND AGREED:**

**Content Provider**

By: \_\_\_\_\_  
Name:  
Title:  
Date:

**Proponent Company**

By: \_\_\_\_\_  
Name:  
Title:  
Date:

**EXHIBIT A**

Title/Event:  
Date:  
Elements:  
Format:  
Duration:

Title/Event:  
Date:  
Elements:  
Format:  
Duration:

Title/Event:  
Date:  
Elements:  
Format:  
Duration:

Title/Event:  
Date:  
Elements:  
Format:  
Duration: