

# STANDARDS AND INFORMATION DOCUMENTS

AES69-2020



## **AES standard for file exchange - Spatial acoustic data file format**

Users of this standard are encouraged to determine if they are using the latest printing incorporating all current amendments and editorial corrections. Information on the latest status, edition, and printing of a standard can be found at:  
<http://www.aes.org/standards>

**AUDIO ENGINEERING SOCIETY, INC.**  
132 East 43rd St., Suite 405, New York, NY 10017, US.



The AES Standards Committee is the organization responsible for the standards program of the Audio Engineering Society. It publishes technical standards, information documents and technical reports. Working groups and task groups with a fully international membership are engaged in writing standards covering fields that include topics of specific relevance to professional audio. Membership of any AES standards working group is open to all individuals who are materially and directly affected by the documents that may be issued under the scope of that working group.

Complete information, including working group scopes and project status is available at <http://www.aes.org/standards>. Enquiries may be addressed to [standards@aes.org](mailto:standards@aes.org)

The AES Standards Committee is supported in part by those listed below who, as Standards Sustainers, make significant financial contribution to its operation.



Focusrite®



THE TELOS ALLIANCE\*



audio-technica



CLAIR



WEISS



This list is current as of  
2020/7/30

# **AES standard for file exchange – Spatial acoustic data file format**

Published by

**Audio Engineering Society, Inc.**

Copyright ©2020 by the Audio Engineering Society

## **Abstract**

Binaural listening is growing fast, because of growing sales in smartphones, tablets and other individual entertainment systems. The lack of a standard for the exchange of head-related transfer functions (HRTF) means each company keeps its binaural capture and rendering algorithms private. 3D audio is arising, and binaural listening could be the very first 3D audio vector with sufficient fidelity of HRTF.

The use of convolution-based reverberation processors in 3D virtual audio environments has grown with the increase in available computing power. Convolution-based reverberators guarantee an authentic and natural listening experience, but also depend on the acoustic quality of the applied spatial room impulse response (SRIR).

With a standardized file format for HRTF and SRIR data, each company can contribute its best algorithms, providing good personalized capture and/or rendering, allowing the consumer to choose a combination of technologies for the best quality of experience.

This document standardizes a file format to exchange space-related acoustic data, such as binaural listening parameters in the form of head related transfer functions. The format is scalable to match the available rendering process and is designed to include source materials from different databases.

An AES standard implies a consensus of those directly and materially affected by its scope and provisions and is intended as a guide to aid the manufacturer, the consumer, and the general public. The existence of an AES standard does not in any respect preclude anyone, whether or not he or she has approved the document, from manufacturing, marketing, purchasing, or using products, processes, or procedures not in agreement with the standard. Prior to approval, all parties were provided opportunities to comment or object to any provision. Attention is drawn to the possibility that some of the elements of this AES standard or information document may be the subject of patent rights. AES shall not be held responsible for identifying any or all such patents. Approval does not assume any liability to any patent owner, nor does it assume any obligation whatever to parties adopting the standards document. Recipients of this document are invited to submit, with their comments, notification of any relevant patent rights of which they are aware and to provide supporting documentation. This document is subject to periodic review and users are cautioned to obtain the latest edition.

---

**Audio Engineering Society Inc., 132 East 43rd St., Suite 405, New York, NY 10017, US.**

[www.aes.org/standards](http://www.aes.org/standards)    [standards@aes.org](mailto:standards@aes.org)

2020-12-06 printing

 **Documents**  
Document Sharing Hub

## Contents

|                                                                                          |           |
|------------------------------------------------------------------------------------------|-----------|
| <b>0 Introduction</b>                                                                    | <b>4</b>  |
| 0.1 General                                                                              | 4         |
| 0.2 Documentation conventions                                                            | 5         |
| <b>1 Scope</b>                                                                           | <b>5</b>  |
| <b>2 Normative references</b>                                                            | <b>5</b>  |
| <b>3 Terms, definitions, and abbreviations</b>                                           | <b>6</b>  |
| <b>4 General specifications</b>                                                          | <b>8</b>  |
| 4.1 Coordinate systems and spatial representations                                       | 8         |
| 4.2 Objects                                                                              | 10        |
| 4.3 Relation between the objects                                                         | 11        |
| 4.4 Numeric container                                                                    | 13        |
| 4.5 Dimensions                                                                           | 14        |
| 4.6 Data                                                                                 | 17        |
| 4.7 Metadata                                                                             | 18        |
| 4.8 Relation between files                                                               | 23        |
| <b>5 Conventions</b>                                                                     | <b>24</b> |
| 5.1 General                                                                              | 24        |
| 5.2 Convention-specific metadata (normative metadata)                                    | 24        |
| 5.3 Application-specific metadata (non-normative metadata)                               | 24        |
| <b>Annex A (informative) - NetCDF</b>                                                    | <b>25</b> |
| <b>Annex B (Informative) – Media, application programming interfaces, and networking</b> | <b>26</b> |
| B.1 Media type ‘SOFA’                                                                    | 26        |
| B.2 Application programming interface                                                    | 26        |
| B.3 Networking                                                                           | 26        |
| <b>Annex C (Normative) - Data types</b>                                                  | <b>27</b> |
| C.1 FIR                                                                                  | 27        |
| C.2 FIR-E                                                                                | 27        |
| C.3 TF                                                                                   | 27        |
| C.4 TF-E                                                                                 | 27        |
| C.5 SOS                                                                                  | 28        |
| <b>Annex D (Normative) – Convention sets</b>                                             | <b>30</b> |
| D.1 General                                                                              | 30        |
| D.2 GeneralFIR                                                                           | 30        |
| D.3 GeneralTF                                                                            | 30        |
| D.4 GeneralFIR-E                                                                         | 30        |
| D.5 GeneralTF-E                                                                          | 30        |
| D.6 SimpleFreeFieldHRIR                                                                  | 31        |
| D.7 SimpleFreeFieldHRTF                                                                  | 35        |
| D.8 SimpleFreeFieldHRSOS                                                                 | 35        |
| D.9 FreeFieldHRIR                                                                        | 35        |
| D.10 FreeFieldHRTF                                                                       | 37        |
| D.11 SimpleHeadphoneIR                                                                   | 37        |
| D.12 SingleRoomSRIR                                                                      | 40        |
| D.13 SingleRoomMIMOSRIR                                                                  | 43        |
| D.14 FreeFieldDirectivityTF                                                              | 45        |
| <b>Bibliography</b>                                                                      | <b>50</b> |

These forewords are not part of *AES69 – AES standard for file exchange – Spatial acoustic data file format*.

### Foreword

This standard builds upon an earlier project to define a spatially-oriented format for acoustics (SOFA). The SOFA project considered the requirements for a file format storing HRTF data and other spatial descriptions of acoustic systems. SOFA aimed at storing data representing HRTF data in a general way, capable of supporting any data measured with microphone arrays and loudspeaker arrays.

This project was developed as project AES-X212 by Task Group SC-02-08-E led by M. Parmentier. The principal authors were P. Majdak and M. Noisternig.

The following persons were involved in the development of this document in draft:

T. Ammermann, S. Bharitkar, F. Camerer, W. De Bruijn, T. Carpentier, M. Emerit, F. Fleischmann, K. Hamasaki, A. Harma, J-M. Jot, M. Kelly, R. Kessler, T. Knowles, P. Majdak, F. Melchior, M. Noisternig, B. Olson, G. Pallone, C. Par, M. Parmentier, A. Pereira, C. Pike, J. Plogsties, E. Ronciere, T. Sporer, B. van Daele, W. Woszczyk, H. Ziegelwanger

Mark Yonge

Chair, SC-02-08 Working Group on Audio-File Transfer and Exchange

### Foreword to second edition, 2020

This revision describes the AES69-2020, also known as SOFA 2.0. The revision includes:

- a new spatially continuous representation of emitters and receivers (by means of spherical harmonics, also known as Ambisonics)
- a clarification on the object “Data” consistently defining the dimension sizes of other variables
- information about ‘sofa’ as a registered media subtype for storage, transfer, and exchange
- new conventions describing directivity of musical instruments and loudspeakers (with flexibility not being covered by other AES standards)
- new conventions describing multiple-input and multiple-output measurements enabling complex interaction between sources and listeners (such as multiperspective representations)
- new conventions describing HRTFs represented in a spatially continuous way
- new general conventions in order to describe general measurements.
- updated description of room types.

In rare cases, this revision might break compatibility because it requires that the dimension sizes of the data object determine the dimension sizes in a file. In AES69-2015 (SOFA 1.0), this requirement was not described and the dimension sizes of other variables might have been used and enforced on the Data object.

This revision was developed by Task Group SC-02-08-E led by P. Majdak and M. Noisternig. The principal authors were P. Majdak, J. De Muynke, F. Zotter, and F. Brinkmann. Other contributors were J. Ahrens, S. Norcross, C. Pike, A. Farina, L. Neumann.

Piotr Majdak and Markus Noisternig

Chairs, SC-02-08, 2020-10-24

### Note on normative language

In AES standards documents, sentences containing the word “shall” are requirements for compliance with the document. Sentences containing the verb “should” are strong suggestions (recommendations). Sentences giving permission use the verb “may”. Sentences expressing a possibility use the verb “can”.

# AES standard for file exchange - Spatial acoustic data file format

## 0 Introduction

### 0.1 General

A spatial acoustic transfer function describes the spatial filtering of the incoming sound, due to the listener's anatomy for example. Previously, spatial-acoustic data have been stored in various incompatible formats, making an exchange of this data difficult. This document describes a format for storing spatial acoustic data with a focus on interchangeability and extensibility and provides a basis for a wider generalised interchange of space-related audio data.

In this document, the term 'transfer function' (TF) describes any filter, regardless the actual representation or interpretation. TF may, for example, refer to:

- The frequency domain representation of free-field head-related transfer functions (HRTFs);
- The time domain representation of free-field HRTFs, that is head-related impulse responses (HRIRs);
- The time domain representation of HRTFs measured in reverberant spaces, that is binaural room impulse responses (BRIRs);
- The Quadrature Mirror Filter (QMF) domain representation of free-field HRTFs, that is the set of QMF parameters; or, even more generally,
- The time domain representation of spatiotemporal room impulse responses, that is spatial room impulse responses (SRIRs).

If not otherwise stated, we do not distinguish between such terminologies and refer to all those filters as transfer functions.

Specifications given in 'conventions' (that is in-detail descriptions of the data exchange format for a given data set) shall use the correct terminology, for example using HRIR instead of HRTF when the data are represented in time domain.

The following requirements are identified:

- Description of a measurement setup with arbitrary geometry; that is, not limited to special cases like a regular grid, or a constant distance, or even spatial discrete points in space;
- Self-describing data with a consistent definition; that is, all the required information about the measurement setup must be provided as metadata in the file;
- Flexibility to describe data of multiple conditions (listeners, distances, and so on) in a single file;
- Partial file support and network support;
- Available as binary file with data compression for efficient storage and transfer;
- Predefined descriptions for the most common measurement setups, which are referred to as 'conventions'.

A TF measurement setup is described by various objects (4.2) and their relations (4.3); the information is stored in a numeric container (4.4) and structured by the measurement.

A measurement consists of data (4.6) describing a TF for a single condition, for example an HRIR, and is described by its corresponding dimensions (4.5) and metadata (4.7). All measurements are stored in a single data structure. Conventions for a consistent description of measurement setups are provided in Annex D of this document.

## 0.2 Documentation conventions

A **Courier** typeface is used in this document to identify computer listing examples to distinguish them from regular text.

## 1 Scope

This document standardizes a file format to exchange space-related acoustic data in various forms. The format is designed to be scalable to match the available rendering process. The format is designed to be sufficiently flexible to include source materials from different databases.

## 2 Normative references

The following referenced documents are indispensable for the application of this document. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

**ISO 8601:2004**, *Data elements and interchange formats – Information interchange – Representation of dates and times*, International Standards Organization, Geneva, Switzerland

**ISO 80000-2:2009**, *Quantities and units – Part 2: Mathematical signs and symbols to be used in the natural sciences and technology*, International Standards Organization, Geneva, Switzerland

**ISO 80000-5:2007**, *Quantities and units – Part 5: Thermodynamics*, International Standards Organization, Geneva, Switzerland

**ISO/IEC 10646:2012**, *Information technology – Universal Coded Character Set (UCS)*, International Standards Organization, Geneva, Switzerland

**ISO/PAS 17506:2012**, *Industrial automation systems and integration – COLLADA digital asset schema specification for 3D visualization of industrial data*. International Standards Organization, Geneva Switzerland.

**ITU-R TF.460**, *Standard-frequency and time-signal emissions*, International Telecommunications Union, Geneva, Switzerland.

**RFC 3339**, *Date and Time on the Internet: Timestamps*, Internet Engineering Task Force, [www.ietf.org](http://www.ietf.org).

**RFC 3986 (also IETF STD 66)**, *Uniform Resource Identifier (URI): Generic Syntax*, Internet Engineering Task Force, [www.ietf.org](http://www.ietf.org).

**netCDF-4** published and maintained by Unidata, Boulder, CO., US.

<http://www.unidata.ucar.edu/software/>

(See Annex A)

**sofa** media subtype as registered by the Internet Assigned Numbers Authority (IANA),

<https://www.iana.org/assignments/media-types/audio/sofa>

(See Annex B)

### 3 Terms, definitions, and abbreviations

For the purposes of this document, the following terms, definitions, and abbreviations apply.

#### 3.1

##### **Head-related transfer function**

##### **HRTF**

describes the spatial filtering of the incoming sound due to the listener's anatomy.

#### 3.2

##### **Head-related impulse response**

##### **HRIR**

describes the finite time-domain representation of an HRTF.

#### 3.3

##### **Spatial room impulse response**

##### **SRIR**

describes the finite time-domain representation of spatial room impulse responses. Sometimes referred to synonymously as directional room impulse responses (DRIR).

#### 3.4

##### **Transfer function**

##### **TF**

#### 3.5

##### **Impulse response**

##### **IR**

#### 3.6

##### **Time-domain finite IR**

##### **FIR**

#### 3.7

##### **Time-domain infinite IR**

##### **IIR**

#### 3.8

##### **Time of arrival**

##### **TOA**

describes the time required by the sound from the source to reach the listener; broadband delay.

#### 3.9

##### **Spatially oriented format for acoustics**

##### **SOFA**

#### 3.10

##### **Second-Order Sections**

##### **SOS**

describes high-order IIR filters implemented as series of second-order sections, i.e. serially-cascaded biquadratic (second-order) filters.

**3.11****Digital asset exchange file****DAE**

describes an open standard interchange file format for interactive 3-D applications according to ISO/PAS 17506:2012 that is usually identified with a .dae filename extension.

**3.12****Application programming interface****API**

specifies the interface of different software components.

**3.13****Uniform resource identifiers****URI**

defines a string of characters used to identify a name of a web resource or local file. RFC 3986 defines the generic syntax to be used in all URI schemes. URI's are further clarified in "URIs, URLs, and URNs: Clarifications and Recommendations 1.0" published by the World Wide Web Consortium (W3C).

**3.14****Identifier****ID**

is a lexical token (that is a string of one or more characters) that uniquely identifies an entity. A typical example is a Subject ID replacing the subject's name by a unique identifier to assure privacy in publicly available databases.

**3.15****Coordinated Universal Time****UTC**

defines the primary time standard by which the world regulates clocks and time. The current version of UTC is defined by International Telecommunications Union Recommendation (ITU-R TF.460), "Standard-frequency and time-signal emissions". RFC 3339 defines the date and time format for use in Internet protocols that is a profile of the ISO 8601 standard for representation of dates and times.

**3.16****Conventions**

define recommendations on the naming of AES69 attributes, variables, and dimensions for discipline-specific data structures, that is particular measurement setups. In other standards a set of 'conventions' is often referred to as a 'profile'.

**3.17****Universal Character Set Transformation Format – 8 bit****UTF-8**

defines the encoding that can represent every character in the Unicode Character Set (UCS). The standard ISO/IEC 10646:2012 specifies the UCS, which is applicable to the representation, transmission, interchange, processing, storage, input and presentation of the written form of the languages of the world as well as additional symbols.

**3.18****Internet Assigned Numbers Authority****IANA**

is a standards organization that oversees global internet-protocol address allocation, autonomous system number allocation, root zone management in the domain name system, media types, and other internet-related symbols and numbers, see <https://www.iana.org/>.

## 4 General specifications

### 4.1 Coordinate systems and spatial representations

#### 4.1.1 General

Objects and data described in this document are represented in the Cartesian and spherical coordinate systems depicted in figure 1. Spatially discrete representation directly describes points in space by using one of these coordinate systems. Spatially continuous representation uses the spatial frequency domain represented by a finite series of harmonics converted from one of these coordinate systems. The client application should convert the geometry information to whatever fits best to its internal data representation format.

#### 4.1.2 Cartesian

Cartesian coordinates shall comprise orthogonal linear values,  $(x, y, z)$  as defined in ISO 80000-2:2009.

All units shall be given in metres.

#### 4.1.3 Spherical

Spherical coordinates shall comprise angular values for azimuth and elevation, and a linear radius,  $(\phi, \theta, r)$  corresponding to:

$$\begin{aligned}x &= r \cdot \cos\left(\frac{\pi}{180}\theta\right) \cdot \cos\left(\frac{\pi}{180}\phi\right) \\y &= r \cdot \cos\left(\frac{\pi}{180}\theta\right) \cdot \sin\left(\frac{\pi}{180}\phi\right) \\z &= r \cdot \sin\left(\frac{\pi}{180}\theta\right)\end{aligned}$$

where the azimuth angle  $\phi \in [0, 360)$  is the horizontal angle in degrees measured counter-clockwise from the positive x-axis, the elevation angle  $\theta \in [-90, 90]$  is the vertical angle in degrees measured from the x-y plane, and  $r \geq 0$  is the distance in metres from the origin. Elevation angles are positive when the angle is measured above the x-y plane (that is in the upper half-space).

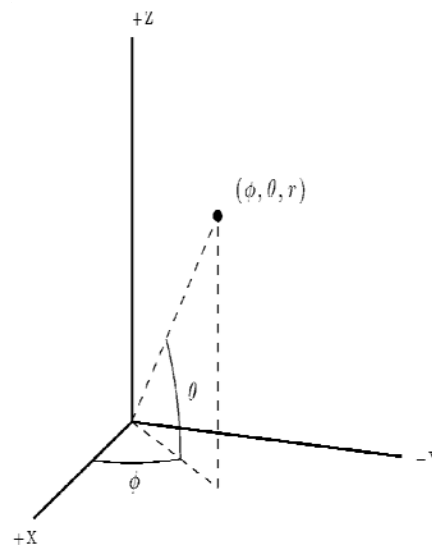


Figure 1: Cartesian and spherical coordinate systems.

NOTE The applied spherical coordinate system does not correspond to ISO 80000-2:2009, which defines the spherical coordinate system with the 'zenith' angle  $\delta$  measured from the positive z-axis. The zenith is the complement of elevation, that is  $\theta = 90^\circ - \delta$ . AES69 implements the elevation  $\theta$  that is typically used in virtual audio applications and perceptual studies (and not the zenith  $\delta$ ).

The units shall be given in  $(\phi, \theta, r) = (\text{degree}, \text{degree}, \text{metre})$ .

#### 4.1.4 Spherical Harmonics

Spatially continuous data in spherical coordinate system shall be represented as a series of spherical-harmonic functions coefficients  $S_{l,m}(\phi, \theta)$  corresponding to:

$$S_{l,m}(\phi, \theta) = \begin{cases} \sqrt{2} N_{l,|m|} P_{l,|m|}(\sin \theta) \cos(|m|\phi), & m > 0 \\ N_{l,0} P_{l,0}(\sin \theta), & m = 0 \\ \sqrt{2} N_{l,|m|} P_{l,|m|}(\sin \theta) \sin(|m|\phi), & m < 0 \end{cases} \quad (1)$$

with the order  $l \in \{0, 1, \dots, L\}$ , the degree  $m \in \{-l, \dots, 0, \dots, l\}$ , the normalization factor

$$N_{l,m} = \sqrt{\frac{(2l+1)}{4\pi} \frac{(l-m)!}{(l+m)!}}, \quad (2)$$

and the Legendre associated polynomials  $P_{l,m}(x)$  that can be defined via

$$P_{l,m}(x) = \frac{1}{2^l l!} (1-x^2)^{m/2} \frac{d^{l+m}}{dx^{l+m}} (x^2-1)^l. \quad (3)$$

Here,  $L$  defines the maximal order of the series and thus the spatial accuracy of the representation.

NOTE This representation is based on Blenko et al. (1997) and is similar to Ambisonics widely used in the audio applications, e.g., Zotter and Frank (2012).

NOTE The terms 'degree' and 'order' correspond to the common usage in the field of audio engineering. Their definitions may vary in other fields, especially, in quantum physics, the meaning of degree and order is the other way round.

NOTE The normalization from Eq. 2 is often called 'orthonormalized', e.g., Wiczorek and Meschede (2018), or 'N3D' in the field of audio engineering. Other normalization methods for the calculation of spherical harmonics used in other fields are not considered in AES69.

The units apply to the radius only and shall be given in metre.

Spatially continuous data available up to  $L$  and stored in a linear arrangement corresponding to a given combination of  $l$  and  $m$  shall be arranged such that the  $i$ -th data element corresponds to:

$$i = l(l+1) + m \quad (4)$$

with  $l \in \{0, 1, \dots, L\}$  and  $m \in \{-l, \dots, 0, \dots, l\}$ .

NOTE This ordering is the same as the Ambisonics Channel Number (ACN).

Figure 2 shows the visual representations of the first few real-valued spherical-harmonic functions according to Eq. 1 and depicts their index in a linear order according to Eq. 5.

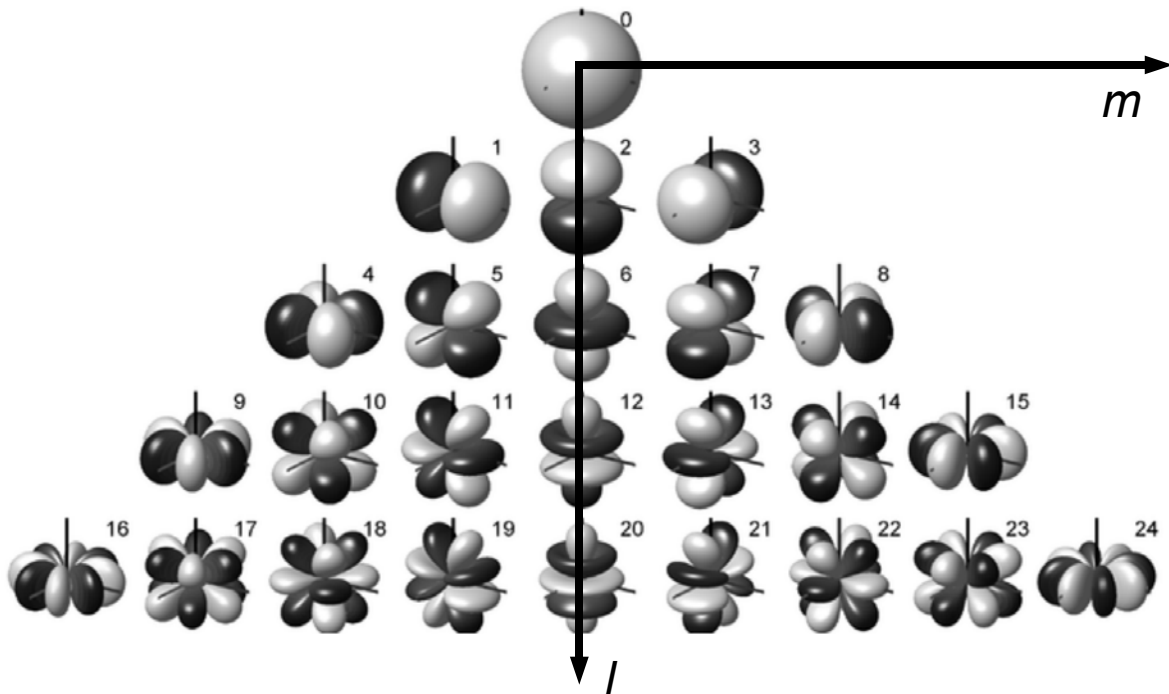


**Figure 2: Visual representation of the first few real-valued spherical harmonics and their index  $i$  when stored in a linear vector as shown in Eq. 4. The vertical axis shows representations for the order  $l$  up to the order of  $L=4$ . The horizontal axis shows representations for the degree  $m$  increasing to the right.**

## 4.2 Objects

### 4.2.1 General

The contribution of a particular object is described by the metadata.

### 4.2.2 Receiver

A receiver object shall represent any acoustic sensor. The number of receivers shall be a positive integer number of unlimited size.

### 4.2.3 Listener

A listener object shall represent a logical unit of all receivers (or their spatially continuous representation), incorporating all the acoustic sensors used for a measurement setup. An AES69 file shall not contain more than one listener object.

**EXAMPLE** A listener object can represent a human head equipped with miniature in-ear microphones, a dummy-head or any microphone-array structure such as a sphere or a frame.

**NOTE** Combining all receivers in a single logical listener object is important. In measurements the orientation and/or position of the listener usually vary without substantial changes in the geometric relations of receivers. For example, in HRTF measurements done for multiple positions in a room, the position of the head varies and the relation between the head and the microphones does not change.

#### 4.2.4 Emitter

An emitter object shall represent any acoustic source. The number of emitters shall be a positive integer number of unlimited size.

#### 4.2.5 Source

The source object shall represent a logical unit of all emitters (or their spatially continuous representation), incorporating all the acoustic sources used for a measurement setup. An AES69 file shall not contain more than one source object.

EXAMPLE A source object can represent a single-driver loudspeaker (with one single emitter), a multi-driver loudspeaker (with the particular drivers as emitters), a loudspeaker array (with the particular loudspeakers as emitters), or – for music recordings – a choir (with the particular singers as emitters).

NOTE Combining all emitters in a single logical source object is important. In measurements the orientation and/or position of the source usually vary without substantial changes in the geometric relations of emitters. For example, in room acoustic measurements done for multiple positions in a room, the position of the (multichannel) loudspeaker varies and the relation between the different single loudspeaker drivers does not change.

#### 4.2.6 Room

The room object shall represent the volume enclosing the measurement setup. A free-field measurement setup shall be considered as a special case of the room object. An AES69 file shall not contain more than 1 room object.

A verbal room description may be added for measurements performed in reverberant spaces. A room object may also be described as a simple shoebox model, or with a link to a digital asset exchange (DAE) file for a more complex description.

Any additional descriptive files shall be linked using URIs. The URI syntax shall correspond to RFC 3986. In the case of multiple entries, each URI shall be separated by the end-of-line token (hexadecimal **0A<sub>16</sub>**).

The room model for each measurement setup shall be described in the corresponding convention (5).

### 4.3 Relation between the objects

#### 4.3.1 Global coordinate system

The global coordinate system shall be Cartesian as defined in 4.1.2 or spherical as defined in 4.1.3 and shall define the position of the source and listener objects within the room. Source and listener objects shall share the same global coordinate system. The origin of the global coordinate system may be chosen arbitrarily.

NOTE In the case of 'free field' conditions (see room types defined in 4.7.7), the listener should be positioned at the origin of the coordinate system. The default orientation of the listener should be in the direction of the positive x-axis (that is in the horizontal plane facing the frontal direction). Assuming the emitter in the centre of the source, the source position then directly reflects the relation in between the source and the listener objects (that is the direction of arrival of the sound wave). See the convention **SimpleFreeFieldHRIR** (Annex D.6) for an example of this measurement condition.

#### 4.3.2 Local coordinate systems

The source and listener objects located within the global coordinate system shall define the source-related and listener-related local coordinate systems, respectively.

NOTE The system of axes of the local and global coordinate systems do not have to be parallel. They may be arbitrarily chosen.

The local coordinate systems shall define the positions of emitters within the source object, and the positions of receivers within the listener object, respectively. All emitters shall share the same source-related local coordinate system; all receivers shall share the same listener-related local coordinate system.

The relation of global and local coordinate systems is shown in figure 3.

Spatial information about all receivers shall be stored either spatially discrete or spatially continuous. The spatially discrete representation describes the position of each receiver within the local coordinate system, which in that case shall be Cartesian as defined in 4.1.2 or spherical as defined in 4.1.3. The spatially continuous representation describes the spatial information about the receivers as a finite series of harmonic functions as defined in 4.1.4.

Spatial information about all emitters shall be stored either spatially discrete or spatially continuous. The spatially discrete representation describes the position of each emitters within the local coordinate system which in that case shall be Cartesian as defined in 4.1.2 or spherical as defined in 4.1.3. The spatially continuous representation describes the spatial information about the emitters as a finite series of harmonic functions as defined in 4.1.4.

For spatially discrete emitters and receivers, the directivity pattern may be considered. Where relevant, the directivity pattern of an emitter shall be defined in an emitter-related coordinate system; where relevant, the directivity pattern of a receiver shall be defined in a receiver-related coordinate system. The emitter objects and receiver objects located in the source-related and listener-related coordinate systems shall define the emitter-related and receiver-related local coordinate systems, respectively. The emitter-related and receiver-related local coordinate systems shall be of the same type than the source-related and listener-related local coordinate systems, respectively.

### 4.3.3 Object orientation

The orientation of the objects source, listener, emitters, and receivers may be considered. Where relevant, the orientation of an object shall be defined by the rotation of the respective local coordinate systems using global coordinates.

Two orthogonal vectors, the 'view' and 'up' vectors, shall define the orientation of a local coordinate system. The 'view' vector shall define the direction of the positive x-axis of the respective local coordinate system, that is the unit vector  $u_x = (1, 0, 0)$  in Cartesian coordinates (4.1.2) or the unit vector  $u_x = (0^\circ, 0^\circ, 1)$  in spherical coordinates (4.1.3). The 'up' vector shall define the direction of the positive z-axis of the respective local coordinate system, that is the unit vector  $u_z = (0, 0, 1)$  in Cartesian coordinates or the unit vector  $u_z = (0^\circ, 90^\circ, 1)$  in spherical coordinates. The direction of the 'view' and 'up' vectors shall be defined in the parent coordinate system of the respective local coordinate system, that is the global coordinate system for source and listener, and the source-related and listener-related local coordinate systems for emitters and receivers, respectively.

If not otherwise defined, the orientation of the respective local coordinate system shall be that of the corresponding global coordinate system (4.3.1).

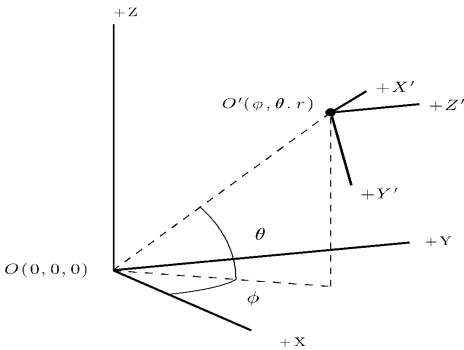


Figure 3: Global and local coordinate systems.

NOTE 1 In spherical coordinates, the 'view' vector describes the azimuth and elevation angles of the source/listener orientation and the 'up' vector describes the roll.

NOTE 2 The origin of the local coordinate system is at position  $(\phi, \theta, r)$  relative to the global coordinate system, which in this example is in spherical coordinates. The orientation of the local coordinate system is given by the view and up vectors, which define the direction of the positive X' axis and Z' axis, respectively.

4.4 Numeric container

The information shall be stored in a single file by serializing the data into a binary stream. The binary data format shall be as defined by netCDF-4.

The file names shall have the extension '.sofa' (see Annex B.1).

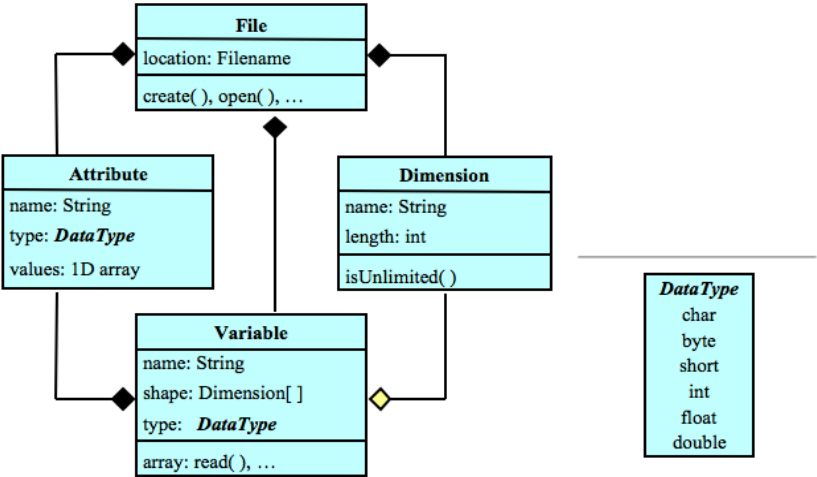


Figure 4: Classic netCDF data model.

The data format should use the classic netCDF-4 data model (figure 4) or the enhanced netCDF-4 data model (figure 5). Conventions shall define whether the classic or enhanced netCDF-4 data model is used..

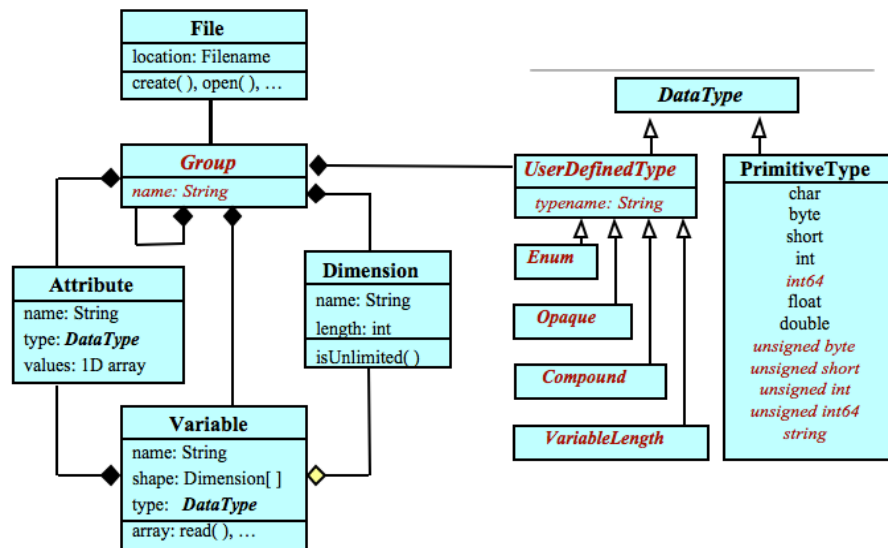


Figure 5: Enhanced netCDF-4 data model.

NOTE netCDF-4 is based on the classic netCDF data model shown in figure 4. Since the enhanced data model is more complex and not well spread in various computer systems yet, AES69 mostly uses the classic data model and parts from the enhanced model. This way allows a simple data representation but still full flexibility in the future. A deeper knowledge of netCDF format details is not required to read or write netCDF datasets.

According to the netCDF terminology, information shall be stored in **variables** and **attributes** using pre-defined **dimensions** (4.5).

A **variable** shall represent an array of values of the same type. A variable shall be either an array of the netCDF type 'double' (referred to as 'double' throughout this document) or an array of the netCDF type 'character' (referred to as 'string' throughout this document). A variable can be a scalar value (that is a zero-dimensional array), a single-dimensional array, or a multi-dimensional array. Variables shall use one of the dimensions defined in 4.5. A variable may have associated attributes.

All **attributes** shall be netCDF attributes of the netCDF type 'string' (referred to as 'attribute' throughout this document).

Arrays of attributes (that is string arrays) shall not be used. Where relevant, an array of attributes may be defined as a variable of type 'string' and shall be given the same name as the attribute.

Recommendations on the naming of attributes, variables, and dimensions within a netCDF file are called **conventions** (5). This standard defines dimensions and general rules for creating conventions. Sets of conventions related to particular measurement setups are described in Annex D.

#### 4.5 Dimensions

The dimensions of a variable shall be defined by the corresponding conventions. A variable shall use the predefined dimensions provided in table 1.

NOTE 1 Throughout this document, the matrix sizes are denoted by  $[A_1, A_2, \dots, A_k]$  where  $A_k$  represents the size of the k-th dimension of a K-dimensional matrix.