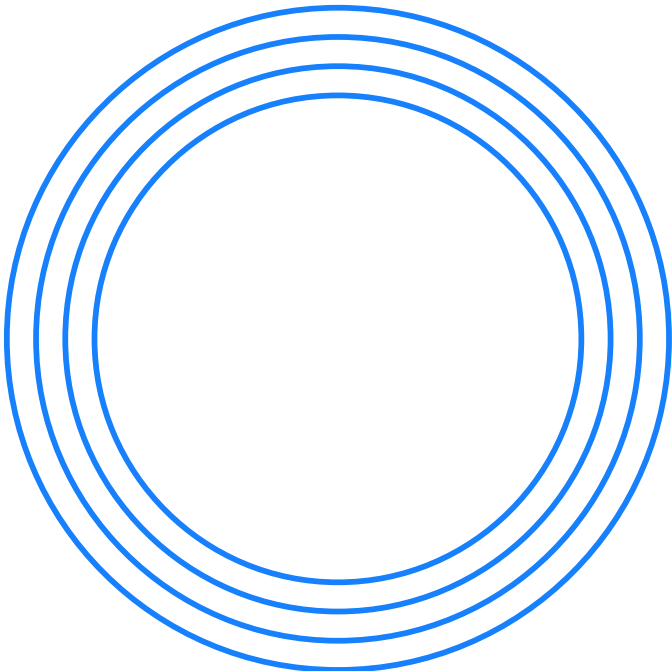
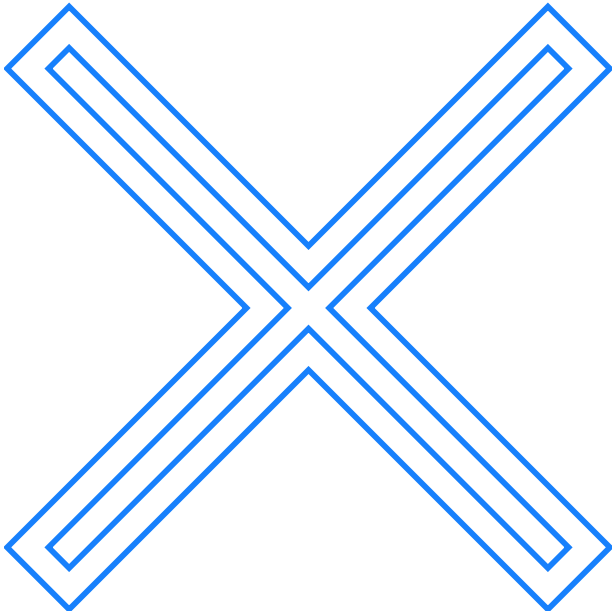
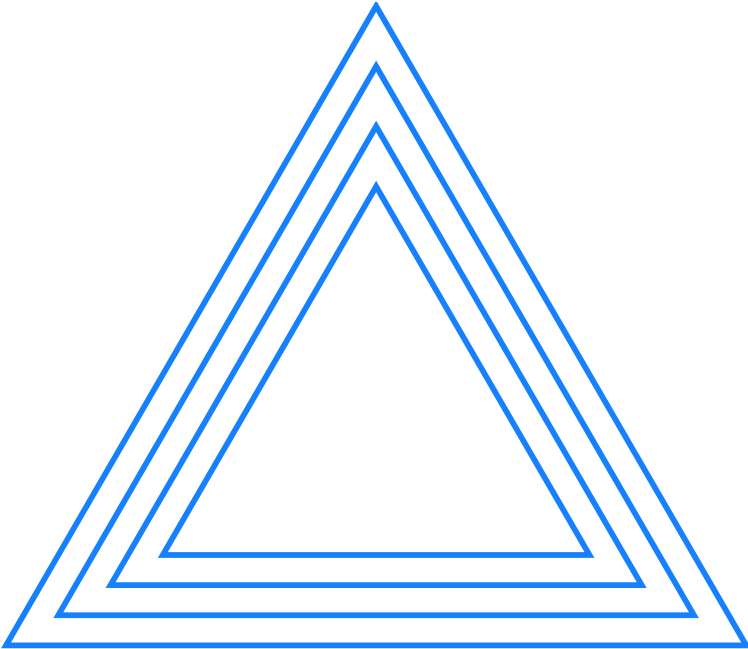
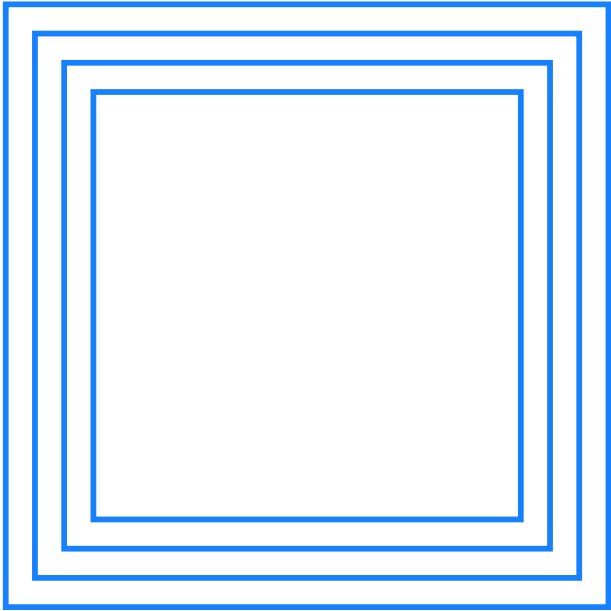




**iXML EXTENSION  
SPECIFICATION**  
**ASWG-G006**

v.1.1 11/2024



---

# Contents

- 3 — ABOUT ASWG
- 4 — INTRODUCTION
- 5 — <ASWG> VENDOR OBJECT
- 6 — METADATA FIELDS
- 11 — EXAMPLE

---

# ABOUT ASWG

---

The Sony PlayStation® Studios' Audio Standards Working Group (ASWG) consists of audio professionals representing Sony Interactive Entertainment (SIE) development studios from Japan, the United States of America and Europe.

Their role is to formulate and disseminate audio standards, recommendations and best practice documentation within SIE and to the wider development community working on SIE platforms.

The Group can be contacted at  
[pss.aswg@sony.com](mailto:pss.aswg@sony.com)

---

# INTRODUCTION

---

The ASWG iXML Extension is designed to provide developers of interactive audio content and audio researchers the ability to store production and research related metadata within the <BWFXML> chunk of a Broadcast Wave file, describing its contents and other related information.

The specification defines fields relating to metadata that can be used within interactive media development applications and workflows as well as machine learning and deep learning feature sets.

The extension contains fields covering sound effects, music, dialogue and audio-driven haptic content, as well as more general, project information.

This is an update to v1.0 first published in November 2021.



# VENDOR OBJECT

---

The <ASWG> vendor object is now included in the iXML Specification Revision v.3.01 dated October 2021. Details can be found at [ixml.info](http://ixml.info)

Anyone may include the <ASWG> object and any of the tags contained in this document, if they find them useful. Please do not include any tags not listed in this document within the <ASWG> object.

# METADATA FIELDS

The following metadata fields are used in this specification:

Metadata Field	Category	Description
contentType	General	Content type: sfx, music, dialogue, haptic, impulse, mixed
project	General	Project name asset was developed for
originator	General	Designer of the asset
originatorStudio	General	Group or studio of originator
notes	General	General information not covered in other fields
session	General	Application (Pro Tools/REAPER etc.) session name
state	General	File state; 'mastered', 'processed', 'raw', 'placeholder'
editor	General	Name of editor
mixer	General	Name of mix engineer
fxChainName	General	Name of any FX chain used on file. e.g. REAPER chain name
isGenerated	General	Content is AI generated, or contains elements/sections that are AI generated . true/false
masteringEngineer	General	Name of mastering engineer
originationDate	General	Date of original upload of asset. YYYY/MM/DD.
channelConfig	Format	Channel configuration of the file; mono, stereo, LCR, Quad, 5.0, 5.1, 7.0, 7.1, 5.1.2, 5.1.4, 7.1.2, 7.1.4, 12.2, ambisonic
ambisonicFormat	Format	Ambisonic format; #p, #h#p, #h#v. eg: 5p, 3h1v, 4h2p
ambisonicChnOrder	Format	Ambisonic channel order: fuma, acn
ambisonicNorm	Format	Ambisonic normalization: snd3, maxn, n3d
micType	Recording	Microphone(s) used. Where multiple mics used, prefix with channel number: 1-Neumann U87i, 2-AKG C414

Metadata Field	Category	Description
micConfig	Recording	Microphone configuration: Mono, AB, XY, ORTF, MS, other
micDistance	Recording	Microphone distance in meters OR head-mounted - 1, 2, 0.3, head, N/A
recordingLoc	Recording	Recording Location
isDesigned	Recording	Whether the sound is designed or a raw recording. 'true' if designed, 'false' if raw recording.
recEngineer	Recording	Name of recording engineer
recStudio	Recording	Name of recording studio
impulseLocation	Recording	Location of the recorded impulse
category	Sound Effects	UCS compliant SFX category name
subCategory	Sound Effects	UCS compliant SFX sub-category name
catId	Sound Effects	UCS compliant SFX category ID
userCategory	Sound Effects	UCS compliant user category
userData	Sound Effects	UCS compliant user data
vendorCategory	Sound Effects	UCS compliant vendor category
fxName	Sound Effects	UCS compliant FX name
library	Sound Effects	UCS compliant library
creatorId	Sound Effects	UCS compliant SFX creator/publisher
sourceId	Sound Effects	UCS compliant SFX Source ID
rmsPower	Audio Features	RMS power of file
loudness	Audio Features	Integrated loudness of file, measured with ITU-R BS1770-3 (or later) compliant metering, in LKFS/LUFS
loudnessRange	Audio Features	Loudness Range, in Loudness Unit (LU) measured with EBU 3342 compliant metering
maxPeak	Audio Features	Maximum sample value, measured in dBFS
specDensity	Audio Features	Spectral density of file - amount of power at each ERB frequency band*
zeroCrossRate	Audio Features	Zero Cross Rate. Average frequency of entire file

Metadata Field	Category	Description
papr	Audio Features	Peak to Average Power Ratio
text	Dialogue	Transcript of the dialogue file
efforts	Dialogue	Whether the file contains efforts, dialogue or a mix of the two; 'true', 'false', 'mixed'
effortType	Dialogue	Effort type; e. g. 'strain', 'pain'
projection	Dialogue	Dialogue projection level. 1- whispered, 2- to self, 3- spoken, 4- projected, 5- shouted
language	Dialogue	Dialogue language - ISO639-1 Language Code
timingRestriction	Dialogue	Dialogue timing restriction; 'wild', 'time-sync', 'lip-sync', 'N/A'
characterName	Dialogue	Name of character
characterGender	Dialogue	Sex/gender of character
characterAge	Dialogue	Age of character
characterRole	Dialogue	Whether the character is a main (significant) character or a background character; 'significant', 'background'
actorName	Dialogue	Name of actor
actorGender	Dialogue	Sex/gender of actor; male, female, not specified
director	Dialogue	Name of dialogue director
direction	Dialogue	Director's notes, for context; explaining the scene and character motivation
fxUsed	Dialogue	General description of any effect present, e.g. radio, reverb
usageRights	Dialogue	Internal code denoting usage rights of the asset
isUnion	Dialogue	Whether the asset was recorded under a union contract; 'true', 'false'
accent	Dialogue	Regional accent of the spoken dialogue, if applicable
emotion	Dialogue	Emotional content present in the delivery of the dialogue
addresseeGender	Dialogue	Gender of addressee; male/female/malegroup/femalegroup/mixedgroup
isFormal	Dialogue	Either formal or informal, depending on the relationship between the speaker and the addressee. true/false.
devLanguage	Dialogue	Original language used by developer



Metadata Field	Category	Description
composer	Music	Name of composer
artist	Music	Artist name
songTitle	Music	Song title
genre	Music	Music genre
subGenre	Music	Music sub-genre
producer	Music	Producer name(s)
musicSup	Music	Music Supervisor name
instrument	Music	Instrument on track/stem
musicPublisher	Music	Name of publisher
rightsOwner	Music	Owner of the recorded work
isSource	Music	Whether the asset is as the composer delivered (source) or an edit of the source; 'true', 'false'
isLoop	Music	Whether the content is seamlessly loopable; 'true', 'false'
intensity	Music	Intensity of the music
isFinal	Music	Whether the cue is temp or final; 'true', 'false'
orderRef	Music	Internal order reference of cue, if applicable
isOst	Music	Whether the cue is part of the original soundtrack; 'true', 'false'
isCinematic	Music	Whether the cue is associated with a cinematic within a title; 'true', 'false'
isLicensed	Music	Whether the cue is licensed from or owned by a third party; 'true', 'false'
isDiegetic	Music	Whether the cue is used diegetically in the title; 'true', 'false'
musicVersion	Music	Internal version number
isrcId	Music	ISRC code
tempo	Music	Tempo in beats per minute (BPM). Multiple tempos may be included in this field, separated by commas

Metadata Field	Category	Description
timeSig	Music	Time signature of cue; 3:4, 4:4, 6:8. Multiple time signatures may be included in this field, separated by commas
inKey	Music	Which key the cue is in. Multiple keys may be included in this field, separated by commas
billingCode	Music	Internal project billing code

\* 'SpecDensity' is an array of 24 floats, separated by commas, corresponding to the average magnitude in each frequency band, uniformly spaced between 20-20kHz on an ERB scale. For a definition of ERB, see Moore, B. C. J., and Glasberg, B. R. (1983). "Suggested formulae for calculating auditory-filter bandwidths and excitation patterns," J. Acoust. Soc. Am. 74, 750-753.

---

## EXAMPLE

The following is an example of information that can be carried in the iXML <ASWG> field.

---

```
<?xml version="1.0" encoding="UTF-8"?>
<BWFXML>
  <IXML_VERSION>3.01</IXML_VERSION>
  <ASWG>
    <contentType>sfx</contentType>
    <project>My Game Project</project>
    <originator>A. Designer</originator>
    <originatorStudio>PlayStation Studios</originatorStudio>
    <session>C:\Program Files\REAPER\my_session.rpp</session>
    <channelConfig>ambisonic</channelConfig>
    <ambisonicFormat>3p</ambisonicFormat>
    <ambisonicChnOrder>acn</ambisonicChnOrder>
    <ambisonicNorm>sn3d</ambisonicNorm>
    <isDesigned>true</isDesigned>
    <loudness>-32.4</loudness>
    <loudnessRange>6.8</loudnessRange>
    <maxPeak>-8.4</maxPeak>
    <notes>3rd order ambisonic, Main Hall ambience.</notes>
  </ASWG>
</BWFXML>
```



---

## CONTACT

[pss.aswg@sony.com](mailto:pss.aswg@sony.com)