

## 36 AKTUELL

40 Weshalb der Rundfunk einen offenen und  
herstellerunabhängigen Workflow für Next  
Generation Audio benötigt  
MICHAEL WEITNAUER



# WESHALB DER RUNDFUNK EINEN OFFENEN UND HERSTELLERUNABHÄNGIGEN WORKFLOW FÜR NEXT GENERATION AUDIO BENÖTIGT

MICHAEL WEITNAUER

Das Thema Next Generation Audio (NGA) nimmt im Rundfunk an Fahrt auf. Der Artikel erläutert die Grundlagen und greift Standardisierungsfragen hinsichtlich der Produktion und Verbreitung von NGA-Inhalten auf.

► Next Generation Audio (NGA) is gaining momentum in broadcasting. The article explains the basics and addresses standardisation issues regarding the production and distribution of NGA content.

## Einleitung

Medientechnologien werden kontinuierlich weiterentwickelt und sorgen so in regelmäßigen Abständen für Änderungen im Rundfunkumfeld. Leistungsfähigere Computer, effizientere Übertragungsmethoden und neue Forschungsergebnisse treiben die Evolution von und das Verlangen nach einem besseren medialen Erlebnis voran. Neben aktuellen Änderungen im Videobereich, wo Ultra High Definition (UHD), High Dynamic Range (HDR) und High Frame Rate (HFR) gerade auf dem Markt und im Rundfunk-Workflow Verbreitung finden, steht nun die Audioproduktion und -übertragung vor dem größten Evolutionsschritt seit Einführung des Stereotons. Sogenannte Next Generation Audio-(NGA-) Technologien und -Systeme sind dabei, sich auf dem Markt zu etablieren. Sie bringen großartige Verbesserungen für Zuschauer durch Personalisierungs- (wie etwa die Anpassung der Dialoglautstärke) und Interaktivitätsmöglichkeiten (siehe [1]) und vereinfachen die Produktion im komplexer werdenden Medioumfeld. Außerdem bieten sie elegante Lösungen für verschiedenste Probleme mit denen Medienorganisationen konfrontiert sind, wie beispielsweise die Sprachverständlichkeit.

Eine NGA-Produktion ist vielseitig anwend- und adaptierbar und kann von NGA-fähigen Endgeräten für jedes Ausgangsformat und für jede vorherrschende Umgebung passend prozessiert werden. Die Technologie dahinter wird *objektbasiertes* Audio genannt und stellt die Kerntechnologie hinter NGA dar. Objektbasiertes Audio setzt die Einführung von bestimmten Parametern und Metadaten (z.B. Position, Gain etc.), die dem Audiosignal angehängt werden, voraus. Dies befreit Anbieter von Medieninhalten davon, verschiedene Versionen bzw. Formate von einem Programm produzieren zu müssen. Das Grundprinzip der Audioproduktion jedoch bleibt sehr ähnlich zu den heutigen Verfahren, und ein Toningenieur kann die neue Technologie sehr leicht in seinen Arbeitsablauf übernehmen. Der zu erwartende Aufwand für eine NGA-Produktion ist vergleichbar mit einer traditionellen, kanalbasierten Produktion, jedoch mit dem großen Vorteil der skalierbaren Nutzung. In Anbetracht der stetig wachsenden Anzahl an relevanten Plattformen und Endgeräten, welche die Rundfunkanstalten bedienen müssen, besteht gerade bei der sogenannten objektbasierten Produktion eine erhebliches Einsparpotential.

Bild 1 vergleicht das Grundprinzip einer traditionellen, kanalbasierten Produktion mit dem objektbasierten Ansatz.

Ein hybrider Ansatz, bei dem kanal- und objektbasierte Mixe simultan verwendet werden, (siehe Bild 2) ist natürlich auch möglich und wird womöglich eines der meistgenutzten Szenarien werden.

NGA vereinfacht darüber hinaus ebenfalls die Verbreitung barrierefreier Dienste, wie spezielle Presets für Hörgeschädigte oder Audioinhalte mit individualisierbarer Länge, und kann auch dabei helfen, die Datenrate im Vergleich zu heute üblichen Codecs erheblich zu reduzieren. Kurz gesagt, der Ton in der Rundfunklandschaft ist dank NGA auf dem besten Weg, sich zu einem neuen, großartigen Erlebnis zu entwickeln, welches sowohl immersive als auch personalisierte Aspekte beinhaltet.

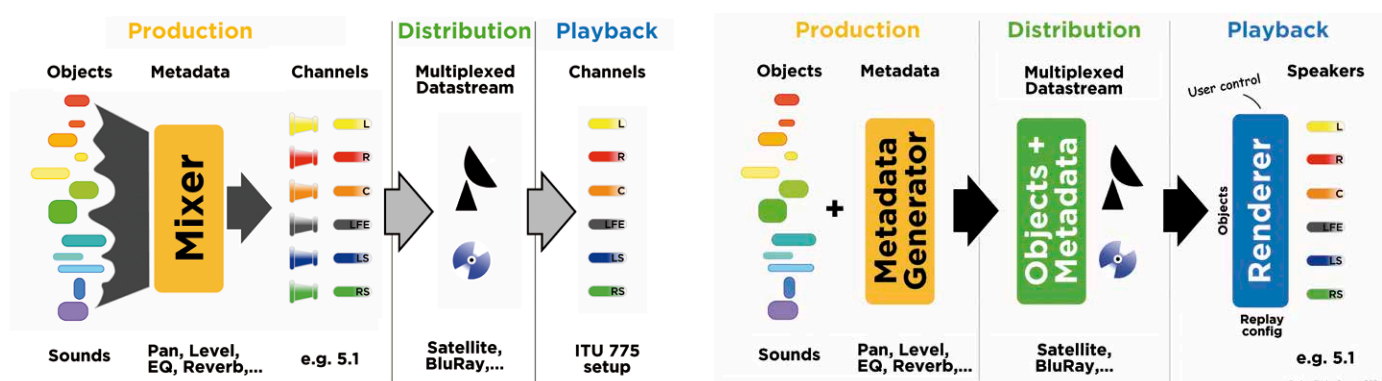


Bild 1. Grundprinzip der traditionellen kanalbasierten (links) und objektbasierten (rechts) Audioproduktion. Design: Florian Camerer, ORF

## Hintergrund

Die relevanten NGA-Systeme für die Übertragung zum Endkunden sind MPEG-H, Dolby AC-4 und DTS-UHD, die neben einem klassischen Codec noch weitere Komponenten für die Wiedergabe, wie einen Renderer, ein Metadatenmodell, Personalisierungs- und Interaktivitätsmöglichkeiten durch den Nutzer, Dynamic Range Control oder eine Lautheitsanpassung beinhalten. Jedes dieser Systeme nutzt ein eigenes Codec-spezifisches Metadatenformat und bietet Werkzeuge und Formate für die Produktion. Jedoch sind diese Formate untereinander nicht kompatibel. Die Integration von solchen Codec-spezifischen Formaten in aktuellen Produktionswerkzeugen hat deshalb bereits zu einem fragmentierten Markt geführt.

## NGA in DVB

Die aktuelle Spezifikation von DVB (Digital Video Broadcasting) ETSI TS 101 154 [2] für die Verwendung von Audio- und Video- (A/V) Codierverfahren beinhaltet unter anderem

- Eine Liste von A/V-Codern und Formaten, die in DVB unterstützt werden
- Ein DVB-Profil für jeden Codec
- Richtlinien für die Verwendung, das Encodieren, Decodieren und die Signalisierung

Die letzte Version der Spezifikation (Januar 2019) sieht für die Übertragung von NGA sowohl AC-4, MPEG-H Audio als auch DTS-UHD vor. Allerdings konnte im Standardisierungsprozess nicht sichergestellt werden, dass alle relevanten Komponenten der NGA-Systeme verpflichtend spezifiziert sind. Die Mindestanforderung an DVB-Endgeräte beinhaltet daher lediglich den Decoder für kanalbasierte Signale. Ein objektbasierter Renderer sowie die Möglichkeit zur Personalisierung der objektbasierten Inhalte ist nur als optional definiert. Daher muss bei den länderspezifischen Receiver-Spezifikationen unbedingt darauf geachtet werden, dass die relevanten Komponenten entsprechend verpflichtend spezifiziert werden. Andernfalls besteht die Gefahr, dass das enorme Potential von NGA den Zuschauern nicht zugänglich gemacht werden kann bzw. nur in den höherklassigen Modellen der Gerätehersteller enthalten sein wird.

## NGA über HbbTV

Die aktuelle Version 2.0.2 der HbbTV-Spezifikation beinhaltet die Unterstützung von NGA-Decodern, welche derzeit in der DVB-Toolbox verfügbar sind. Dadurch können mit relativ geringem Aufwand NGA-Programme über IP-Verbreitung zum Endkunden übertragen werden, ohne die ganze DVB-Sendeinfrastruktur ändern zu müssen. Auch synchronisierte Szenarien, bei denen das ursprüngliche Audio aus dem DVB-Multiplex durch einen NGA-Stream aus dem Internet ersetzt und dann wieder synchron mit dem Broadcast-Video abgespielt wird, sind mit dieser Spezifikation umsetzbar.

Ein einheitliches User Interface zur Bedienung der NGA-Decoder ist momentan noch nicht spezifiziert. Dieses könnte über die HbbTV NGA API implementiert werden, um die Benutzerfreundlichkeit von Presets und weiteren Personalisierungsoptionen für die Zuschauer zu verbessern und zu vereinheitlichen.

Idealerweise kann dann beispielsweise auch über eine Taste auf der Fernbedienung ein Preset für bessere Sprachverständlichkeit ausgewählt werden. Weitere, automatisierte Lösungen sind ebenfalls denkbar, jedoch im Hoheitsbereich der Gerätehersteller.

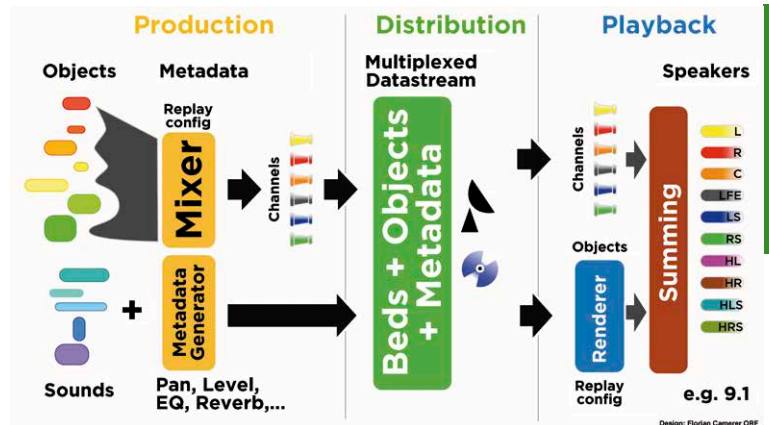


Bild 2. Grundprinzip eines hybriden NGA-Ansatzes. Design: Florian Camerer, ORF

## NGA in Endgeräten

Aktuell ist zu beobachten, dass verschiedene CE-zertifizierte Endgeräte wie TV-Geräte oder Set-Top-Boxen auf dem Markt sind, die NGA-fähig sind und mindestens einen NGA-Decoder implementiert haben. Für Mobilgeräte wie Smartphones wurden ebenfalls eine Reihe von Ankündigungen gemacht. Für keine Geräteart sind jedoch momentan belastbare Zahlen veröffentlicht. Es darf aber davon ausgegangen werden, dass die meisten aktuellen UHD-fähigen TV-Geräte bereits einen NGA-Decoder integriert haben.

## Zukunftsaussichten

Bei einer nüchternen Analyse eines zukünftigen NGA-Szenarios erscheint es sehr wahrscheinlich, dass eine Rundfunkanstalt mehr als ein NGA-System zur Verbreitung seiner Inhalte nutzen muss, da nicht alle NGA-Decoder auf allen relevanten Plattformen verfügbar sein werden. Weiterhin ist es sehr unwahrscheinlich, dass es einen Decoder geben wird, der auf allen Plattformen und Geräten verfügbar sein wird. Deshalb ist es umso wichtiger, in der Produktion unabhängig von herstellerspezifischen Formaten und Werkzeugen zu sein.

Abgesehen von dem oben beschriebenen Szenario birgt die Benutzung proprietärer Formate das Risiko einer wirtschaftlichen und technischen Abhängigkeit von einem bestimmten Hersteller. Weiterhin ist es ebenfalls nicht wünschenswert proprietäre Formate oder komprimierte Audiosignale in Archiven zu speichern, vor allem wenn man den Langzeit-Aspekt von Archivmaterial betrachtet. Des Weiteren muss sichergestellt sein, dass NGA-Programme zuverlässig zwischen verschiedenen Rundfunkanstalten und Produktionsfirmen sowie anderen Mitstreitern ausgetauscht werden können. Aus diesen und weiteren Gründen ist ein herstellerunabhängiges und offenes Produktionsformat für Next Generation Audio absolut essentiell.

## Offene Standards für Codec-unabhängige Arbeitsabläufe

Verschiedene NGA-relevante Standards sind über die letzten Jahre im Rahmen der EBU (European Broadcasting Union), ITU-R (International Telecommunication Union, Radiocommunication Sector) und SMPTE (Society of Motion Pictures and Television Engineers) mit der Unterstützung des IRT entstanden.

## Audio Definition Model

Der grundlegende Standard für einen offenen Workflow ist das sogenannte Audio Definition Model (ADM), welches bereits jetzt bzw. in naher Zukunft von allen kommerziellen

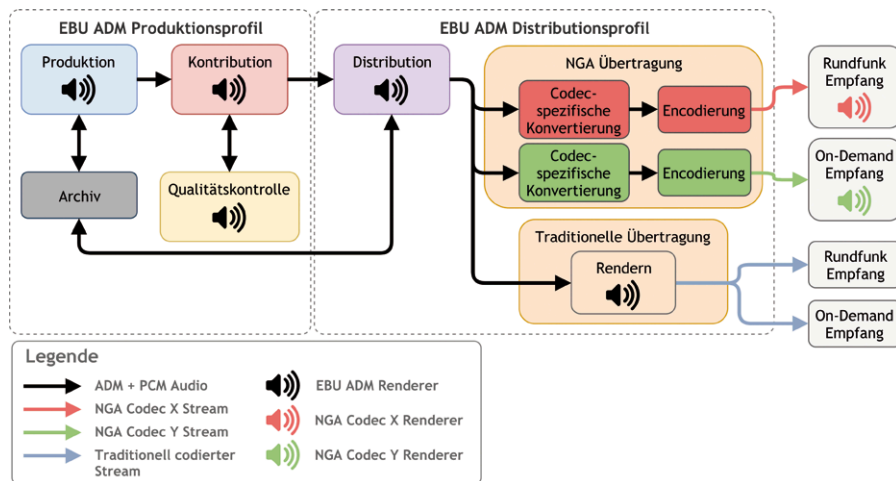


Bild 3. Vereinfachte Darstellung eines idealen NGA-Arbeitsablaufs auf Grundlage von offenen und unabhängigen Standards und Formaten in der Produktion  
Quelle: Michael Weitnauer, IRT

NGA-Systemen als Schnittstelle unterstützt werden soll. Das ADM-Format wurde als ITU-R Recommendation BS.2076 [3] standardisiert und ist ein offenes XML-Metadatenmodell für alle Next Generation Audio-Technologien wie kanalbasierter, szenenbasierter und objektbasierter Ton. Darüber hinaus können binaurale Signale abgebildet und Matrix-Instruktionen für einen Down-/Upmix von kanalbasiertem Audio beschrieben werden. Das ADM beinhaltet einen Satz an definierten Metadaten und Parametern mit dem Audiosignale für NGA vollumfänglich beschrieben werden können. Als offenes und unabhängiges Format ist es daher auch geeignet, die Interoperabilität zwischen verschiedenen herstellerspezifischen NGA-Formaten sicherzustellen.

Das ADM ist bewusst sehr flexibel gestaltet und versucht alle möglichen Anwendungen zu unterstützen. Dementsprechend ist es vergleichsweise komplex. Um bei dieser Komplexität die Kontrolle zu behalten und die Interoperabilität zwischen verschiedenen Implementierungen sicherzustellen, ist es daher sinnvoll, eine entsprechende Untermenge, die für eine bestimmte Anwendung sinnvoll sind, in Form von Profilen zu spezifizieren. Kürzlich hat die EBU ein ADM-Profil für die Produktion veröffentlicht (EBU Tech 3392) [6]. Weitere Profile, z.B. für die Distribution von NGA-Inhalten, sind in Planung.

#### Dateiformate für ADM

Die ADM-Daten können innerhalb eines Datenblocks des BWF (EBU Tech 3285) [4] und dessen Nachfolger BW64 (ITU-R BS.2088) [5] gespeichert werden, welche speziell für Inhalte ohne Video geeignet sind. Für den Austausch von NGA-Inhalten zusammen mit Video sind die Formate IMF und MXF besser geeignet. Momentan gibt es jedoch noch keinen standardisierten Weg, ADM in MXF oder IMF zu speichern. Ein Vorschlag wie dies umgesetzt werden kann, soll jedoch bereits 2019 von der EBU veröffentlicht werden.

#### Serial ADM

Während das Dateiformat BW64 kombiniert mit dem ADM, die Möglichkeit zum file-basierten Austausch von objekt-, kanal- und szenenbasierten Audioinhalten bietet, sind BW64 und ADM nicht ohne weiteres für Streaming-Anwendungen, wie beispielsweise Live-Produktionen geeignet. Damit das ADM auch für Streaming-Anwendungen genutzt werden kann, wurde kürzlich eine serielle Form des ADM in der ITU-R veröffentlicht [7]. Bei Serial ADM werden die Metadaten in Segmente eingeteilt und es wird dadurch ermöglicht, jeweils immer nur die Metadaten zu übertragen, die aktuell zur Wiedergabe notwendig sind. Damit S-ADM mit relevan-

ten Protokollen wie etwa AES3, MADI, SDI oder SMPTE 2110 transportieren werden kann, laufen bereits weitere Standardisierungsbemühungen in der SMPTE.

#### EBU ADM Renderer (EAR)

Mit ADM kann man zwar Audiosignale vollumfänglich beschreiben, es liefert jedoch keine Anweisungen, wie die Parameter zu interpretieren sind. Dies übernimmt der sogenannte Renderer, welcher ein essentieller Baustein für die NGA-Produktion ist und die ADM-Metadaten und Audiosignale für die gewünschte Lautsprecheranordnung bzw. Kopfhörerprozessiert. Solch ein Renderer wird in jedem Schritt des Produktions-Workflows benötigt, in dem Ton für das Editieren, Monitoring oder zur Qualitätskontrolle abgespielt werden soll. Der Höreindruck des Toningenieurs sollte dabei selbstverständlich möglichst identisch mit der Zuhörerfahrung sein. Außerdem müssen alle Tools, die entlang des Workflows eingesetzt werden, dasselbe Hörergebnis liefern. Dies mag trivial und offensichtlich erscheinen, war jedoch bis vor kurzem auf Grund des fragmentierten Marktes nicht gesichert.

Erst durch die gemeinschaftliche Entwicklung eines offenen Standards für einen Produktionsrenderer durch das IRT zusammen mit der BBC, France TV und dem Forschungs- und Entwicklungsinstitut bcom im Rahmen einer EBU-Arbeitsgruppe wurde dies möglich. Der „EBU ADM Renderer“ (EAR) ist nicht nur in EBU Tech 3388 spezifiziert [8], sondern wurde auch als open-source Referenzimplementierung veröffentlicht [9] und ist im vergangenen Jahr mit dem EBU Technology & Innovation Award ausgezeichnet worden. Der EAR ist eine vollständige Interpretation des ADM-/S-ADM-Formates und ist grundsätzlich kompatibel mit den Renderern von allen NGA-Decodern in den handelsüblichen Endgeräten. Die EBU und die Autoren des EAR arbeiten mit Herstellern wie Dolby, Fraunhofer IIS und DTS zusammen, um den Renderer als internationalen Standard bei der ITU-R zu veröffentlichen.



**MICHAEL WEITNAUER**

ist F&E-Ingenieur und Projektleiter für den Bereich „Next Generation Audio“ am Institut für Rundfunktechnik.

www.irt.de

## Idealer Workflow

Aus der Sicht einer Rundfunkanstalt basiert ein idealer Workflow auf offenen Formaten und Standards, unabhängig von Übertragungscodecs, unter Verwendung von unkomprimierten Audiosignalen während der ganzen Produktionskette. Dies erlaubt die Wahl zwischen verschiedenen Herstellern für etwa Digital Audio Workstations (DAW), Plugins oder Mischkonsolen und verhindert eine Abhängigkeit von proprietären Technologien. Darüber hinaus wird so die bestmögliche Qualität der Audiosignale sichergestellt. Eine vereinfachte Lösung für diese wesentlichen Anforderungen könnte wie folgt gestaltet werden und ist in Bild 3 dargestellt.

Das ADM-Format bzw. das serielle ADM (S-ADM) für einen Live-Workflow wird in allen Produktionsschritten bis hin zur Encodierung für die Ausstrahlung verwendet. Erst dort wird das ADM in ein Encoder-spezifisches ADM-Profil konvertiert und dann in einem komprimierten NGA-Stream übertragen. Alle Produktionstools, die vor der Encodierung benutzt werden, sollten den EBU ADM Renderer integriert haben, um ein zuverlässiges und Codec-unabhängiges Monitoring des ADM-Formats gewährleisten zu können. Die Nutzung der EBU ADM-Profile für die Produktion und Distribution stellt sicher, dass alle Werkzeuge untereinander kompatibel sind und dass eine Rundfunkanstalt die Freiheit hat, sich für ein passendes Produkt zu entscheiden.

Dieser vereinfachte und ideale Arbeitsablauf setzt voraus, dass (serielles) ADM und PCM ebenfalls in der Kontribution von NGA-Signalen benutzt wird. Eine Konvertierung von proprietären oder herstellerspezifischen Kontributionsformaten wie Dolby ED2 oder MPEG-H Contribution Mode ist jedoch möglich und sollte von den Rundfunkanbietern auch angewendet werden, um eine Kompatibilität mit der ADM-basierten Infrastruktur zu gewährleisten.

Das mittelfristige Ziel sollte sein, ADM und S-ADM zusammen mit dem EAR durch die ganze Audio Kette hinweg zu benutzen: von der Aufnahme zur Produktion, Archiv, Distribution und Reproduktion mit Audio Tools (z. B. DAWs, Konsolen, Plugins, etc.). Für eine transparente und kompatible Umsetzung sollten ebenfalls standardisierte ADM Profile und offene Encoder-spezifische Schnittstellen genutzt werden.

Ein Workflow ähnlich dem skizzierten idealen Workflow wurde im ORPHEUS-Projekt [10] erfolgreich verwirklicht und im EBU Technical Report TR 042 veröffentlicht [11].

## Zusammenfassung

Ein offener und Hersteller-unabhängiger Workflow für Next Generation Audio basierend auf internationalen Standards ist möglich und in Anbetracht der wahrscheinlichen Verfügbarkeit von Decodern auf relevanten Plattformen absolut erstrebenswert, wenn nicht sogar notwendig. Auf Basis von ADM, S-ADM und dem EBU ADM Renderer (bzw. dem ITU Renderer) lässt sich ein solches Szenario realisieren was Rundfunkanstalten sowie Produktionsfirmen zu Gute kommt und den horizontalen Markt fördert. Die Benutzung der genannten Formate und eines standardisierten Produktionsrenderers ist eine essentielle Grundlage für einen erfolgreichen Umstieg in der Produktion bzw. Distribution von NGA-Programmen und stellt ebenfalls sicher, dass eine Rundfunkanstalt zukunftssicher produzieren kann. ◀

### Referenzen

- [1] Shotton, M., F. Melchior, M. Evans, et al. "Object-based Broadcasting – Curation, Responsiveness and User Experience". IBC. 2014. Amsterdam, The Netherland
- [2] DVB Spezifikation ETSI 101154 v2.5.1 (2019-01): [https://www.etsi.org/deliver/etsi\\_ts/101100\\_101199/101154/02.05.01\\_60/ts\\_101154v020501p.pdf](https://www.etsi.org/deliver/etsi_ts/101100_101199/101154/02.05.01_60/ts_101154v020501p.pdf)
- [3] ITU-R Recommendation BS.2076 (Audio Definition Model): <https://www.itu.int/rec/R-REC-BS.2076/en>
- [4] EBU Tech 3285 (Broadcast Wave Format): <https://tech.ebu.ch/publications/tech3285>
- [5] ITU-R Recommendation BS.2088 (BW64): <https://www.itu.int/rec/R-REC-BS.2088/en>
- [6] EBU Tech 3392 (EBU ADM Production Profile): <https://tech.ebu.ch/docs/tech/tech3392.pdf>
- [7] ITU-R Recommendation BS.2125 (Serial ADM): <https://www.itu.int/rec/R-REC-BS.2125/en>
- [8] EBU Tech 3388 (EBU ADM Renderer): <https://tech.ebu.ch/publications/tech3388>
- [9] EBU ADM Renderer Referenzimplementierung: [https://github.com/ebu/ebu\\_adm\\_renderer](https://github.com/ebu/ebu_adm_renderer)
- [10] The ORPHEUS Project: <https://orpheus-audio.eu/>
- [11] EBU Tech Report 042: <https://tech.ebu.ch/publications/tr042>
- [12] EBU Tech Report 045: <https://tech.ebu.ch/publications/tr045>
- [13] EBU Recommendation 151: <https://tech.ebu.ch/publications/r151>