

Barrierefreiheit in TV & Streaming

Oktober 2026

Deutsche TV-Plattform e.V. (DTVP)

Bockenheimer Landstraße 31
60325 Frankfurt/Main, Germany
mail@tv-plattform.de
T +49 69 870 03 29-10

Geschäftsführung:
Nicole Ludwig
Amtsgericht Frankfurt
VR-Nr. VR9797

Vorsitzende:
Nicole Agudo Berbel
Stellvertretende Vorsitzende:
Sascha Molina · Dr. Dietmar Schill

Inhaltsverzeichnis

1	Einleitung	4
2	Ebenen der Barrierefreiheit.....	4
3	Komponenten der Barrierefreiheit	6
3.1	Audiodeskription	6
3.1.1	Beschreibung.....	6
3.1.2	Technische Umsetzung.....	6
	Implementierung über Audiospur/Audioversion	6
	Implementierung der Audiodeskription über eine getrennte, zusätzliche App	6
	Implementierung über eigene VoD-Inhalte	7
3.2	Klare Sprache	7
3.2.1	Beschreibung.....	7
3.2.2	Technische Umsetzung.....	7
	Implementierung über die Audiospur	7
	Implementierung über Next Generation Audio (NGA).....	8
3.3	Untertitel.....	9
3.3.1	Beschreibung.....	9
	Ausprägungen.....	9
	Zielgruppen	10
	Qualitätsmerkmale	10
3.3.2	Technische Umsetzung.....	11
3.3.2.1	Verbreitungsformen im linearen Broadcast	11
3.3.2.2	Verbreitungsformen im IP-Streaming	12
3.3.2.3	Zulieferung und Unterstützung TV Endgeräte	13
3.3.2.4	Zulieferung und Unterstützung bei IP-basierten Streaming-Plattformen	13
3.4	Deutsche Gebärdensprache (DGS)	15
3.4.1	Beschreibung.....	15
	Zielgruppen	15
3.4.2	Technische Umsetzung.....	15
	Verbreitungsformen.....	15
	Zulieferung und Unterstützung Geräte.....	16

	Zulieferung und Unterstützung Plattformen.....	16
4	Auffindbarkeit und Signalisierung	17
5	Benutzeroberflächen.....	21
6	Nutzervoreinstellungen.....	22
7	HbbTV	23
8	Bibliographie.....	24
9	Impressum.....	26
	Herausgeber:.....	26
	Redaktion:.....	26
	Über die Deutsche TV-Plattform e.V. (DTVP):	26
	Haftungsausschluss:	26

1 Einleitung

Die Deutsche TV-Plattform (DTVP) hat sich die Aufgabe gestellt, den Stand der Vorbereitung des Marktes auf das Inkrafttreten der neuen Anforderungen an die Barrierefreiheit von Medienangeboten durch das Barrierefreiheitsstärkungsgesetz (BFSG) und dem Medienstaatsvertrag (MStV) zusammenzutragen. In der DTVP sind sowohl Inhalteanbieter und Medienplattformen (als Zugangsdienst) als auch Hersteller von Endgeräten aktiv. Dies ermöglicht, die jeweiligen Barrierefreiheitskomponenten und auch deren Zusammenspiel zu diskutieren. Der Bericht ist entlang der abgrenzbaren Komponenten und Instrumente gegliedert. Das Dokument ist das Ergebnis der Taskforce (TF) Barrierefreiheit der DTVP.

Der Markt hat sich inzwischen gut auf die neuen Vorgaben eingestellt. Jeder Anbieter für sich hat seine Bausteine überprüft, wo nötig im Sinne der Barrierefreiheit ergänzt und auch am Zusammenspiel der Beteiligten mitgearbeitet.

Im November 2025 wurden die Ergebnisse zur Barrierefreiheit bei TV und VoD (Video on Demand) –Angeboten einigen Vertretern und Vertreterinnen von Betroffenenverbänden vorgestellt. Ziel war es, Anmerkungen und Vorschläge für weitere Produktentwicklungen zu sammeln. Vertreten waren der Deutsche Blinden- und Sehbehindertenverband, der Deutsche Schwerhörigenbund, der Sozialverband VdK sowie das Büro des Beauftragten der Bundesregierung und der Bundesfachstelle Barrierefreiheit.

2 Ebenen der Barrierefreiheit

Die folgende Grafik zeigt schematisch das Zusammenwirken verschiedener Ebenen, die für die Funktionen der Barrierefreiheit relevant sind. Unterschieden werden dabei:

- die GUI-Ebene (grafische Benutzeroberfläche), über die die Nutzerinteraktionen abgebildet werden,
- die funktionale Ebene, die die Aufbereitung von Metadaten übernimmt und Funktionen für die Nutzung bereitstellt, sowie
- geräteseitig implementierte Einstellungsoptionen, die dem Nutzer über eine grafische Benutzeroberfläche des Geräts eine Einstellung von Präferenzen zur Barrierefreiheit erlauben.

Dabei wird unterschieden nach Ebenen, die durch den Gerätehersteller realisiert werden und solchen, die Teil von Apps Dritter sind, die auf dem Gerät laufen.

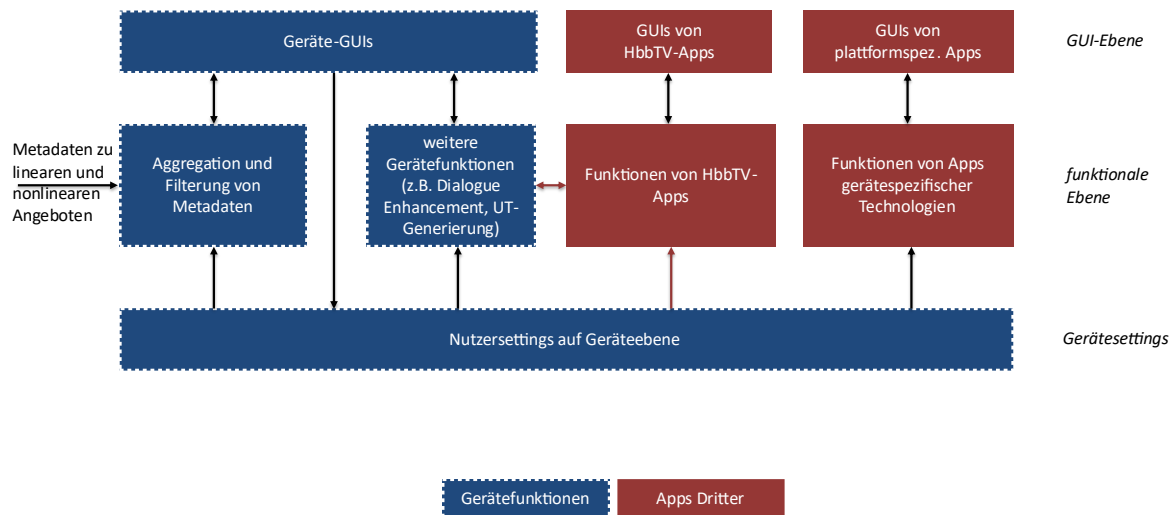


Abbildung 1: Ebenen der Barrierefreiheit auf einem TV-Gerät: GUI-Ebene, funktionale Ebene und Geräteeinstellungen, unterteilt in Gerätefunktionen und Apps Dritter

Auf dieser Grundlage können die Bereiche Auffindbarkeit und Signalisierung, Benutzeroberflächen, Nutzervoreinstellungen und HbbTV voneinander unterschieden werden:

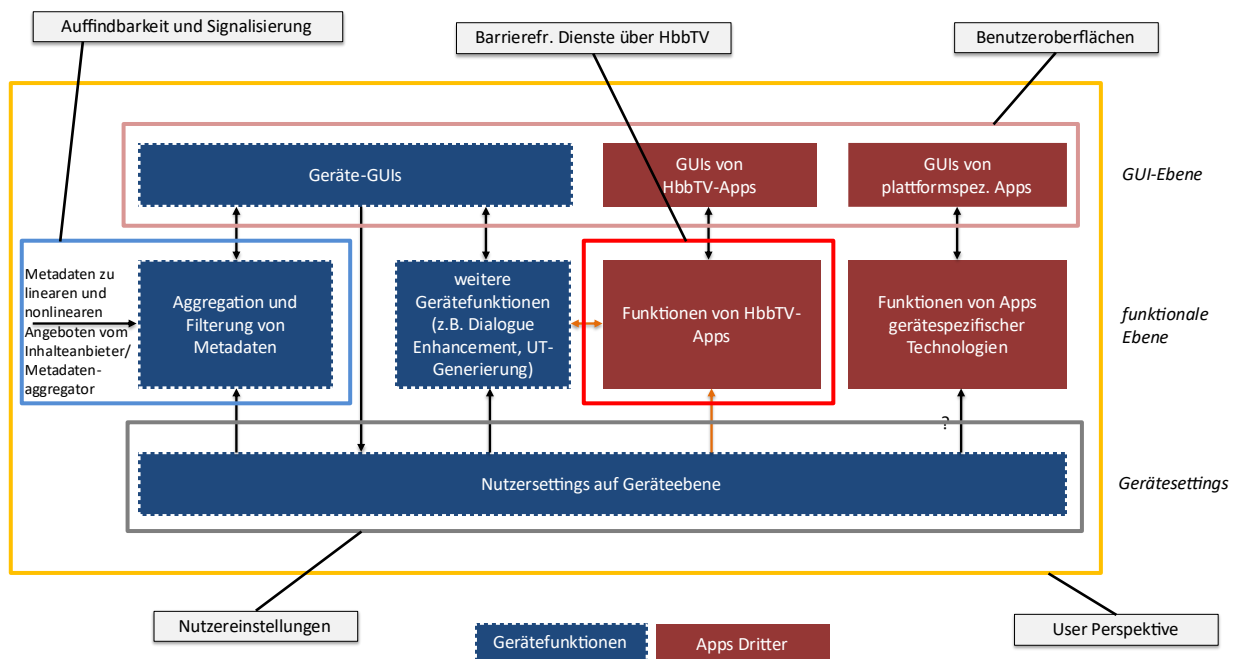


Abbildung 2: Zuordnung der Berichtskapitel zum Ebenen-Modell

3 Komponenten der Barrierefreiheit

3.1 Audiodeskription

3.1.1 Beschreibung

- Audiodeskription beinhaltet eine verbale Bildbeschreibung für Szenen eines Inhalts und dient der Teilhabe von blinden und sehbehinderten Menschen, um einem visuellen Inhalt folgen zu können.
- Inhalteanbieter liefern (derzeit) diese Bildbeschreibung (bei VoD als alternative Version des Inhalts mit inkludierter Audiodeskription), Medienplattformen liefern diese alternative Version oder lediglich den Zusatzinhalt synchron zum Bild zu den Haushalten, Endgeräte bieten die Aktivierung dieses Dienstes als Option in der Benutzeroberfläche dar.

3.1.2 Technische Umsetzung

Implementierung über Audiospur/Audioversion

- Heute stellen Inhalteanbieter wie Sender oder VoD-Anbieter in der Regel eine (zusätzliche) Audiospur bereit, die vom Kunden am Endgerät ausgewählt werden kann.
- Die Inhalte der Audiodeskription müssen zusätzlich durch eine Redaktion erstellt und eingesprochen werden.
- Die Audiospur wird für die Verbreitung i.d.R. als Teil des Signalstroms bereitgestellt, könnte aber auch aus einer getrennten Quelle herangezogen werden (s. Verbreitung über Apps).

Derzeit wird in einem Projekt "IP-Audiosync" eine zweite Implementierung Audiodeskription durch die ARD über einen separaten Online Audiostream pilotiert:

- Das Projekt ermöglicht die Bereitstellung weiterer Tonkanäle zur Unterstützung der Barrierefreiheit und insbesondere auch deren individuelle Nutzung über Mobilgeräte zur Stärkung der Inklusion.
- Die Wiedergabe der zusätzlichen Tonversionen kann dabei sowohl über das TV-Gerät als auch – zur individuellen Nutzung – über ein Mobilgerät erfolgen.
- Eine Smartphone-App synchronisiert den Online-Audiostream mit dem linearen Broadcast Signal.
- Ziel ist es, diese Technik betriebstauglich auszubauen, so dass die Tonkanäle auch bei Live-Sendungen ohne großes Delay zur Verfügung stehen.

Implementierung der Audiodeskription über eine getrennte, zusätzliche App

- Beispiele sind RTL über die separate „Greta“-App (die eine vorhergehende Registrierung erfordert) oder ProSieben/ Sat.1 mit der Audio2-App.
- Die App wird auf Smartphone oder Tablet installiert.

- Ein Beitrag muss vorab heruntergeladen werden. Nach „Starten“ synchronisiert die App mit dem Originalton der laufenden Sendung, so dass der Nutzer synchron zum Bild die Audiodeskription nutzen kann.
- ProSieben/ Sat.1: Audiodeskription läuft mit leichter Verzögerung zum linearen TV.

Implementierung über eigene VoD-Inhalte

- Audiodeskription und Ton des Bildes sind hier bereits zusammen gemischt.
- Bei ProSieben/ Sat.1 kann Audiodeskription z. B. so über einen eigenen Kanal (Channel) bei Joyn aufgerufen werden.

3.2 Klare Sprache

3.2.1 Beschreibung

- Klare Sprache beschreibt ein Signal, in dem der Hintergrund (Musik, Atmo, Geräusche, ...) abgesenkt und die Dialoge evtl. angehoben werden, um es den Nutzerinnen und Nutzern zu ermöglichen, den Dialogen mit weniger Höranstrengung folgen zu können.
- Insgesamt wird hiermit die Sprachverständlichkeit für Menschen mit Hörbeeinträchtigungen verbessert. Sender liefern (derzeit) eine zusätzliche Audiospur, Medienplattformen senden diesen Zusatzinhalt synchron zum Bild weiter zu den Haushalten, Endgeräte bieten diesen in einem Menü zur Sprachauswahl an.

3.2.2 Technische Umsetzung

Implementierung über die Audiospur

- „Klare Sprache“ wird z.B. bei ARD und ARTE für lineares Fernsehen dauerhaft automatisiert erzeugt. Dabei kommt eine KI-basierte Technologie zum Einsatz, die z.B. im Sendezentrum in den SDI-Signalweg eingefügt wird.
- Auch bei On-Demand-Inhalten kann eine alternative Audiofassung im Vorfeld vom Inhaltenanbieter automatisch mit KI-basierter Technologie erstellt und von Nutzerinnen und Nutzern als akustisch optimierte „Klare Sprache“-Version ausgewählt werden. Innerhalb der ARD erfolgt dies einheitlich. Für andere Sender / Inhalte wird dies unterschiedlich gehandhabt.
- Alle Verfahren beruhen auf einer Trennung der Komponenten Dialog und Hintergrund und anschließendem Zusammensetzen in einer neuen Mischung, um den Dialog mit weniger Anstrengung verstehen zu können. Dialog und Hintergrund können dabei jeweils angehoben oder abgesenkt werden. Das Mischungsverhältnis ist in der aktuellen Implementierung bei der ARD fest vorgegeben und kann nicht von den Nutzern und Nutzerinnen verändert werden.

- Heute stellen Sender, Streaming-Plattformen oder vereinzelt auch VoD-Anbieter „Klare Sprache“ als eine weitere, zusätzliche Audiospur zur Verfügung, die von Nutzerinnen und Nutzern am Endgerät ausgewählt werden kann.
- Die Signalisierung der Klaren Sprache erfolgt im DVB Transport Stream (TS) in der Program Map Table (PMT) und in der Event Information Table (EIT). Die ARD setzt zur Übertragung der „Klaren Sprache“ auf eine zusätzliche Tonspur und kennzeichnet diese in der PMT mit dem Language Descriptor „QKS“; ZDF setzt zur Übertragung der „Klaren Sprache“ die mit „MUL“ gekennzeichnete Tonspur ein. Eine Umstellung auf „QKS“ ist geplant. In der EIT erhält der Component Descriptor den Eintrag „Klare Sprache“.
- Ggf. wären hierfür Standards (bzw. deren einheitliche Anwendung) sinnvoll, z.B. als Teil von DVB; im DVB-I-Standard [1] wurde bereits eine entsprechende Metadatenstruktur definiert.
- Die Audiospur könnte auch aus einer getrennten Quelle herangezogen werden, die im Endgerät mit dem Videosignal synchronisiert wird (siehe auch [IP-Audiosync bei Audiodeskription](#)).

Endgeräteseitig gibt es alternativ verschiedenste Technologien, die darauf abzielen, die Sprachverständlichkeit in Echtzeit zu verbessern, etwa in TVs, Kopfhörern, oder speziellen Geräten, die in den Signalweg integriert werden. Die genutzten Technologien reichen über eine einfache Entzerrung des Audiosignals bis zu KI-basierter Dialogseparation.

Implementierung über Next Generation Audio (NGA)

Zukünftige, auf NGA basierende Implementierungen ermöglichen zusätzliche Features für barrierefreie Angebote:

- NGA ist ein Sammelbegriff für objektbasierte Audiocodecs.
- NGA-Codecs bieten u.a. erweiterte personalisierte Funktionen für ein „Klare Sprache“-Angebot.

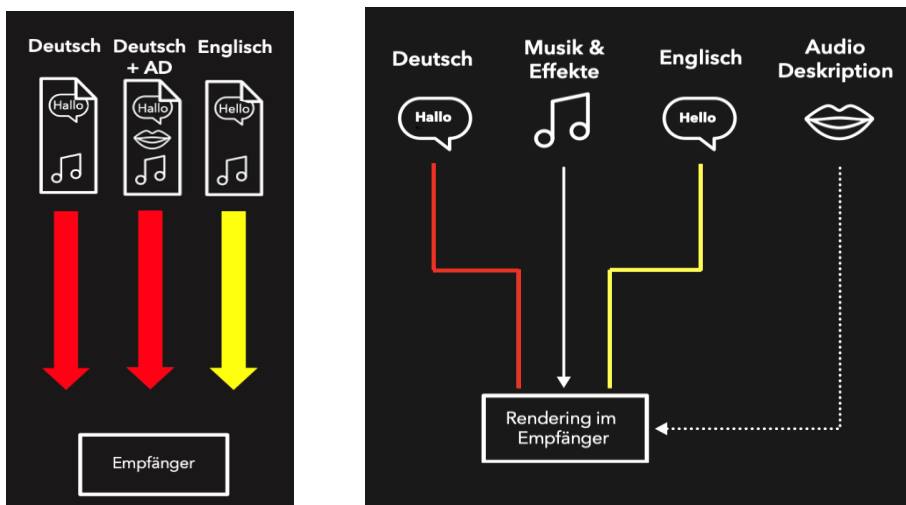


Abbildung 3: Audiodeskription heute (fertig gemischte Tonfassungen) und mit Next Generation Audio (Komponenten werden im Empfänger zusammengesetzt)

Die Anforderung an die Sprachverständlichkeit durch die Zielgruppen ist individuell sehr unterschiedlich. Eine individuelle Anpassung der Hintergrundabsenkung abhängig von Wiedergabesituation und dem eigenen Hörvermögen ist daher wünschenswert. Allerdings ist eine exakte Trennung des Dialogs vom Hintergrund nicht einfach. Eine größere Absenkung verursacht daher hörbare Artefakte, z.B:

- „Jaulen“, „Zischen“,
- verzerrt klingende Sprache,
- Veränderung des Klangs,
- Musik im Hintergrund ist hinsichtlich der Trennung besonders kritisch.

3.3 Untertitel

3.3.1 Beschreibung

Ausprägungen

Untertitel werden auf vielfältige Weise genutzt und ermöglichen den Nutzerinnen und Nutzern das Mitverfolgen von gesprochenem Wort und anderen akustischen Geschehnissen in Textform. Entsprechend unterschiedlicher Anforderungen haben sich folgende Untertitelformen ausgeprägt:

- Komplett-Untertitel für Menschen ohne oder mit vermindertem Hörvermögen MIT Geräuschbeschreibungen,
- Komplett-Untertitel für Menschen mit vermindertem Hörvermögen, vermindertem sprachlichem Verständnis oder situativem Audioproblem (bspw. U-Bahn) OHNE Geräuschbeschreibungen,
- Teil-Untertitel, die nur fremdsprachige oder schwer verständliche Passagen abbilden,
- Übersetzende Untertitel, die komplett das gesprochene Wort in eine andere Sprache oder die gleiche Sprache übersetzen (ohne Geräuschbeschreibungen).

Auch in technischer Hinsicht gibt es hilfreiche Einordnungen. Bei der Produktion kann unterschieden werden zwischen:

- Vorproduzierten Untertiteln,
- Live-Untertiteln¹.

Derzeit werden Untertitel i.d.R. redaktionell erstellt. Eine (teil-)automatisierte Erstellung (insb. mit KI-Tools) wird von den Inhaltenanbietern derzeit getestet und inselweise eingeführt. Bislang

¹ Bei der Produktion von Live-Untertiteln ist zu beachten, dass es zu einem signifikanten Delay von mehreren Sekunden kommt, da das Live-Signal verzögert transkribiert wird – wenn das Gesamtsignal nicht um dieses Delay verzögert wird. Das Delay bei Nachrichten kann durch die Einbindung von Redaktionssystemen oder Teleprompter-Daten oftmals reduziert werden. Auch Künstliche Intelligenz Technologien können hier längerfristig für eine Verbesserung sorgen.

wird die redaktionelle Verantwortlichkeit in Deutschland für diese, automatisiert erstellten Inhalte teilweise problematisch gesehen. Eine automatisierte Erstellung würde es grundsätzlich Inhalteanbietern, Medienplattformen oder Endgeräteherstellern ermöglichen, Untertitel für solche Inhalte anzubieten, für die sonst keine Untertitel vorhanden sind. Eine durch den Einsatz von KI-Tools erzeugte Kürzung der wortwörtlichen Transkription muss gekennzeichnet werden, da dies keine reine Abbildung von Sprache nach Text, sondern KI-generierten Inhalt darstellt.

Für die Verbreitung kann folgende Unterscheidung getroffen werden:

- fest im Bild eingebrannte Untertitel (sog. Open Captions) und
- optional zuschaltbare Untertitel (sog. Closed Captions).

Zielgruppen

Untertitel werden v.a. für folgende Zielgruppe aufgrund der Anforderung der Barrierefreiheit erstellt:

- Menschen mit Hörbeeinträchtigung oder Hörverlust.

Sie werden aber u.a. auch von folgenden Zielgruppen genutzt:

- Menschen, die aufgrund von Fremdsprachigkeit Unterstützung beim Verstehen von Inhalten wünschen,
- Menschen, die situativ auf Ton verzichten möchten.

Qualitätsmerkmale

Es können unterschiedliche Qualitätskriterien auf Untertitel angewendet werden²:

- Untertitel sollen das gesprochene Wort und inhaltlich relevante Geräusche und Musiken so wiedergeben, dass auch ohne Audio-Nutzung der Inhalt der Sendung verständlich ist.
- Um die durchschnittliche Lesegeschwindigkeit einzuhalten, können Kürzungen des Textes im Untertitel nötig sein. Dabei sollte der Inhalt so gut wie möglich beibehalten werden. Wird diese Kürzung automatisiert vorgenommen, ist der Hinweis auf die KI-Generierung notwendig.
- Die Position der Untertitel soll so gestaltet werden, dass vorhandene Bildinformationen nicht überdeckt werden.
- Für eine Zuordnung der Sprechenden können Farben oder andere Merkmale – wie Sonderzeichen oder Namen in Klammern vorangestellt – eingesetzt werden.
- Eine Synchronität von Untertitel und Bildinhalt sollte stets gewährleistet sein.

² Qualitätskriterien von ARD, ORF, SRF und ZDF finden sich z. B. unter <https://www.daserste.de/specials/service/untertitel-standards100.html>

3.3.2 Technische Umsetzung

3.3.2.1 Verbreitungsformen im linearen Broadcast

Die Verbreitung von Untertiteln im Fernsehen erfolgt über verschiedene technische Standards, wobei die gängigsten Methoden DVB-Bitmap, Teletext und zunehmend HbbTV sind. Diese unterschiedlichen Formate bieten jeweils spezifische Vor- und Nachteile, insbesondere in Bezug auf die verwendeten Verbreitungswege.

DVB-Bitmap-Untertitel

DVB-Bitmap-Untertitel [2] zeichnen sich durch ihre hohe Qualität aus, da Schriftart und Layout vom Sender festgelegt werden. Die Untertitel können – abhängig vom Gerät – über die Fernbedienung aktiviert werden.

Teletext-Untertitel

Teletext-Untertitel [3][4] befinden sich auf einer Seite des Teletext-Angebots. Die Untertitel können ebenfalls über die Fernbedienung aktiviert werden. Im Vergleich zu DVB-Bitmap ist die typographische Darstellung jedoch qualitativ limitiert. Durch die eingeschränkt verfügbaren Zeichensätze eignen sich Teletext-Untertitel nicht für alle Sprachen. Aktuell nutzen etwa 10 Millionen Menschen regelmäßig Teletext. Allerdings nehmen die Angebote der Sender in diesem Bereich kontinuierlich ab, weshalb die Zukunft der Teletext-Untertitel ungewiss ist.

HbbTV-Untertitel

Hybrid Broadcast Broadband (HbbTV) [5] bietet eine hohe Qualität und ermöglicht es den Nutzerinnen und Nutzern, das Erscheinungsbild der Untertitel teilweise selbst zu konfigurieren. Die Umsetzung der HbbTV-Untertitel im linearen Broadcast/Livestream erfolgt allerdings meist noch nicht nach einem festgelegten Standard, sondern applikationsbasiert. Der Sender muss dafür eine eigene HbbTV Applikation implementieren. Je nach Implementierung durch die Sender (z.B. ARD) können diese auch ohne Internetverbindung genutzt werden. Allerdings ist HbbTV nicht auf allen Endgeräten und IPTV-Plattformen verfügbar. Die Bedienung erfolgt derzeit ausschließlich über die HbbTV-Menüstruktur, was die Benutzerfreundlichkeit erheblich einschränkt.

DVB TTML-Untertitel

DVB TTML-Untertitel [6] erlauben auch im linearen Broadcast die Übertragung von textbasierten Untertiteln im XML-Format. Es können alle Sprachen und üblichen Darstellungsformen eingesetzt werden. Wie bei DVB-Bitmap-Untertiteln können Schriftart und Layout vom Sender festgelegt werden, so dass eine hohe Darstellungsqualität erzielt werden kann.

3.3.2.2 Verbreitungsformen im IP-Streaming

Im IP-Streaming werden vor allem HLS mit WebVTT und MPEG DASH mit einem Profil von TTML (z.B. EBU-TT-D) verwendet. WebVTT und TTML sind zwei Verbreitungsformate für Untertitel, die vom Standard her alle Darstellungsanforderungen erfüllen:

Web Video Text Tracks (WebVTT) [7] ist ein vom World Wide Web Consortium (W3C) entwickeltes Format für die Darstellung von Untertiteln und anderen textbasierten Daten (z.B. Beschriftungen). Es wurde in Anlehnung an den HTML5-Standard entworfen, um Untertitel für Web-basierte Plattformen und Streaming-Dienste effizient zu unterstützen. Die Formatierung unterstützt grundlegende Textstile sowie die Positionierung auf dem Bildschirm. Die Zeitcodierung erlaubt eine präzise Synchronisation mit der Video- oder Audiospur. Es ist einfach lesbar und editierbar, da es auf reiner Textsyntax basiert. WebVTT findet u.a. bei dem Segment-basierten Streaming-Protokoll HTTP Live Streaming (HLS) [8] Anwendung.

Timed Text Markup Language (TTML) [9] ist ebenfalls ein Standard der W3C (World Wide Web Consortium) für die Erstellung, Darstellung und Verteilung von zeitgesteuertem Text wie Untertiteln, Transkripten und Metadaten in audiovisuellen Medien. TTML wird oft in professionellen Anwendungen verwendet, wo präzise Steuerung und erweiterte Formatierungsfunktionen erforderlich sind. TTML ist eine XML-basierte Sprache, die unter starker Mitwirkung von Seiten der Inhaltenanbieter entwickelt wurde, um Text mit präzisen Zeitinformationen zu versehen und für verschiedene Plattformen nutzbar zu machen. TTML kommt insbesondere bei Streaming-Diensten eine hohe Bedeutung und Verbreitung zu. Wie WebVTT enthält TTML eine Millisekunden genaue Zeitcodierung und gestattet eine äußerst flexible Formatierung von Textstilen sowie eine genaue Positionierung. Die European Broadcast Union (EBU) hat mit EBU-TT [10] und EBU-TT-D [11] Profile von TTML definiert. EBU-TT dient u.a. dem Austausch und der Langzeitarchivierung von Untertiteln. EBU-TT-D wurde für die Online-Verbreitung definiert (siehe unten). Im Streaming findet TTML vor allem bei dem Segment-basierten Streaming-Protokoll MPEG Dynamic Adaptive Streaming over HTTP (DASH) [12][13] Verwendung. Durch unterstützende Standards von MPEG ist TTML auch auf Präzision, Flexibilität und Skalierbarkeit im MPEG DASH-Streaming optimiert.

Es werden sowohl bei WebVTT als auch bei TTML verschiedene Untermengen mit eingeschränktem Funktionsumfang eingesetzt. In Deutschland verbreitete Untermengen von TTML sind insbesondere EBU-TT-D und EBU-TT-D-Basic-DE [14].

EBU-TT-D ist ein Standard der European Broadcasting Union für die Darstellung und Verteilung von Untertiteln in digitalen Medien, insbesondere für IPTV-Plattformen wie VoD-Streaming-Dienste. Der Standard basiert auf XML und ist eine spezialisierte Version von TTML, die von der W3C entwickelt wurde. Zugunsten der Kompatibilität mit modernen Streaming-Technologien wie DASH, handelt es sich um eine "lightweight"-Variante von TTML, die speziell für Distribution und einfache Integration entwickelt wurde. EBU-TT-D wird von TV-Geräten ab der Version HbbTV 2.x unterstützt.

EBU-TT-D-Basic-DE [14] ist eine sehr reduzierte Form des EBU-TT-D-Standards, die vom Institut für Rundfunktechnik (IRT) in Zusammenarbeit mit den Sendern der ARD für die ARD-Mediathek definiert wurde. Ziel war es bei der Definition 2012, Online-Untertitel möglichst einfach, schnell und praxisorientiert umzusetzen. EBU-TT-D-Basic-DE ist als ein Mindeststandard für die Bereitstellung von Untertiteln in Deutschland zu sehen, der die wesentlichen Merkmale wie Schriftgröße, Schrift-Positionierung, Zeilenanzahl, Farben und Kontraste sowie präzise Synchronisation umfasst.

3.3.2.3 Zulieferung und Unterstützung TV Endgeräte

Linearer Broadcast

DVB-Bitmap-Untertitel und Teletext-Untertitel werden auf allen Geräten, die in den letzten Jahren in den Markt gekommen sind, unterstützt. Da die Standards seit langem etabliert sind, ist auch die Umsetzung sehr zuverlässig.

DVB TTML wird bereits von einigen TV-Endgeräten unterstützt, in Deutschland aber bisher wenig bis gar nicht eingesetzt.

Eine Unterstützung von HbbTV ist auf fast allen handelsüblichen Geräten vorhanden. Die Verantwortung für die Unterstützung von HbbTV-Untertiteln liegt aber durch die Implementierung einer eigenen Applikation auf Seiten der Sender.

IP-Streaming

Ein Großteil der Untertitel wird von Videoplayern wiedergegeben, die von Inhaltenanbietern und Plattformbetreibern implementiert werden. Dabei kommen in der Regel zwei Hauptformate zum Einsatz: WebVTT und TTML. Beide Formate haben eine weite Verbreitung bei Streaming-Plattformen.

Auf TV-Empfangsgeräten, die HbbTV 2.0 oder höher implementieren, gibt es sowohl bei der Video on Demand Verbreitung als auch für Live Streams die Möglichkeit, die „native“ Unterstützung von EBU-TT-D-Untertiteln zu nutzen. Allerdings ist hier die Interoperabilität auf Endgeräten noch nicht so hoch wie bei Teletext-Untertiteln und DVB-Bitmap-Untertiteln. An einer Verbesserung wird gearbeitet.

3.3.2.4 Zulieferung und Unterstützung bei IP-basierten Streaming-Plattformen

Bei der Ausstrahlung müssen die Untertitel auf sämtlichen Plattformen dargestellt werden. Die Synchronität der Untertitel variiert je nach Verbreitungsart:

- DVB-Bitmap: Muss für jeden Verbreitungsweg separat synchronisiert werden, was zusätzlichen Aufwand bedeutet:

- Im DVB-Transportstrom werden Video, Audio und Untertitel vom Sender synchron bereitgestellt. Die Medienplattformen müssen jedoch bei einem erneuten Encoding/Transcoding die Untertitel neu synchronisieren.
- Teletext: Dieser wird bereits mit Audio und Video verkoppelt übertragen, was zu einer geringen Asynchronität führt. Die Sender müssen die Untertitel daher lediglich durchleiten, ähnlich wie die Audiospuren.

Auf den Medienplattformen werden in der Regel nicht die im Broadcast genutzten Technologien Teletext und DVB-Bitmap für die Verbreitung genutzt. Bei der Umsetzung in ein anderes Format erzeugen DVB-Bitmap-Untertitel einen erhöhten Aufwand, da die Untertitel mittels Bilderkennungsverfahren (OCR – Optical Character Recognition) erst ausgelesen werden müssen. Heute im Einsatz befindliche Packager-Lösungen gestatten die synchrone Transformation der DVB-Bitmap-Untertitel in textbasierte Formate wie WebVTT und TTML bzw. die Einbettung der Untertitel in Form von Grafiken in den Text-Stream (wird jedoch nicht von HLS unterstützt).

Auch die synchrone Transformation von Teletext-Untertiteln in die textbasierten Formate WebVTT und TTML ist mit heutigen Packagern möglich. Dies ist aufgrund der geringeren Komplexität (Rechenaufwand) gegenüber der Verwendung von Bilderkennungsverfahren klar zu bevorzugen. Untertitel im Teletextformat können zuverlässig dekodiert werden.

Grundsätzlich ist die Bereitstellung von Untertiteln in einem zukunftssicheren Anlieferungsformat entscheidend, um den Anforderungen moderner Medieninfrastrukturen gerecht zu werden. Diese Infrastrukturen entwickeln sich stetig weiter, insbesondere in Hinblick auf Interoperabilität, Barrierefreiheit, Mehrsprachigkeit und einer plattformübergreifenden Nutzung. Untertitel sollten auch bei technologischem Fortschritt kompatibel bleiben und damit langfristig nutzbar sein. Als mögliche zukunftssichere Untertitelformate bieten sich an:

- EBU-TT-D oder EBU-TT-D-Basic-DE,
- TTML-Profil, die speziell für die Anforderungen moderner Medieninfrastrukturen entwickelt wurden.

Fazit

Die Wahl des Untertitelverfahrens hängt von verschiedenen Faktoren ab, darunter die Zielgruppe, die technischen Möglichkeiten der Verbreitungswege und die gewünschten Qualitätsmerkmale. DVB-Bitmap eignet sich für Verbreitungswege über DVB. Teletext-Untertitel hingegen erweisen sich besonders nützlich als Quelle für Transkodierungen auf IP-basierten Medienplattformen.

Gleichwohl könnten bei Verwendung eines modernen Untertitel-Standards wie den TTML-Derivaten Inhalte auf neue Plattformen migriert werden, ohne dass eine vollständige Neukonvertierung erforderlich wird. Ein einheitliches Format reduziert den Aufwand für unterschiedliche Workflows und Konvertierungen.

3.4 Deutsche Gebärdensprache (DGS)

3.4.1 Beschreibung

Die Deutsche Gebärdensprache (DGS) wird seit 2002 als eigenständige Sprache anerkannt. Für gehörlose Menschen, die die DGS zur Kommunikation nutzen, ist sie als Muttersprache zu betrachten mit eigener Grammatik und regionalen Dialekten. Die deutsche Schriftsprache wird wie eine Fremdsprache zusätzlich erlernt. Das Lesen von Texten ist für Gehörlose oft schwierig und eine Untertitelung deshalb für viele Gehörlose nicht ausreichend. DGS ist insbesondere für gehörlose Kinder, die noch nicht lesen können, ein wichtiges Angebot. Die Verfügbarkeit von Medieninhalten mit DGS ist somit ein wichtiger Baustein für die gesellschaftliche Integration und sozialer Teilhabe.

Zielgruppen

- Gehörlose oder im Gehör stark eingeschränkte Erwachsene,
- Gehörlose oder im Gehör stark eingeschränkte Kinder, die noch keine Untertitel lesen können,
- Menschen, die DGS lernen,
- Umfeld von Menschen mit Einschränkungen.

3.4.2 Technische Umsetzung

Die Übersetzung der gesprochenen Sprache in DGS erfolgt bisher durch Dolmetscher und Dolmetscherinnen, sowohl für Live- als auch für vorproduzierte Sendungen. Die Verfügbarkeit von Gebärdensprachdolmetschern ist begrenzt. Um barrierefreie Angebote mit DGS auszuweiten, werden aktuell KI-basierte Techniken entwickelt, die ausgehend von der Spracherkennung die Erstellung von gebärdenden Avataren ermöglichen. Die Herausforderungen sind komplex: Gebärdensprache hat eine eigenständige Struktur und Grammatik und ist ein Zusammenspiel von Bewegungen mit den Händen, Mimik, Kopf- und Körperhaltung sowie Mundbewegungen. Die bisherigen Entwicklungen und die Qualität der Avatare sind bisher für eine vollständig automatisierte Generierung von DGS noch nicht ausreichend. Das Angebot von Inhalten mit DGS ist deshalb begrenzt. Insbesondere bei einer Ausweitung des Angebots gilt es, die Erfordernisse und qualitativen Ansprüche der Zielgruppe zu berücksichtigen.

Verbreitungsformen

Gebärdensprache wird nicht wie bei Untertiteln über ein vereinbartes technisches Verfahren als Zusatzsignal zum ursprünglichen AV-Signal verbreitet. In der Regel wird vom Inhaltenanbieter ein zusätzliches AV-Signal mit einem ins Bild eingefügten DGS-Dolmetscher produziert und als eigenständiges Streamingsignal zur Verfügung gestellt. Man spricht hier auch vom „Einbrennen“ des Gebärdensprachdolmetschers. Eine Sendung / ein Inhalt mit Gebärdensprache wird so in zwei Versionen angeboten: einmal mit und einmal ohne

Gebärdensprache. Die Version mit Gebärdensprache kann auf einem separaten Broadcast-Kanal ausgestrahlt werden (derzeit die Ausnahme) oder (über HbbTV) als IP-Stream vom Nutzer dynamisch hinzugeschaltet oder in einigen VoD-Angeboten gezielt ausgewählt werden.

Die Variante, dass ein Gebärdensprachdolmetscher fest ins Originalbild eingefügt wird und dann für alle, auch die nicht-Gehörlosen, sichtbar wird, ist die Ausnahme. Es macht jedoch alle Nutzer darauf aufmerksam, dass es auch Medienangebote mit DGS gibt.

Die Möglichkeit, einen freigestellten DGS-Dolmetscher oder einen entsprechenden Avatar separat zu übertragen und bei Bedarf auf Nutzerseite zusätzlich ins Videosignal einzufügen, wird erprobt, bisher aber nur außerhalb Deutschlands betrieblich genutzt.

Die Nutzung eines Medieninhalts mit oder ohne DGS muss aus Nutzersicht komfortabel und einfach möglich sein. Dies gilt sowohl dafür, die Verfügbarkeit eines DGS-Angebotes zu signalisieren, als auch dafür, ein entsprechendes Angebot auswählen und nutzen zu können. Die HbbTV-Technologie bietet eine Möglichkeit, eine entsprechende Nutzerschnittstelle zu realisieren.

Über HbbTV können Broadcast-Ausstrahlungen ohne Gebärdensprache mit einem IP-Stream mit Gebärdensprache verknüpft werden. Über die Einblendung einer Farbtaste ins lineare Broadcastsignal wird ein DGS-Angebot signalisiert, durch Drücken der gleichen Farbtaste auf der Fernbedienung kann das entsprechende Streamingsignal mit der DGS komfortabel gestartet werden.

Ähnliches gilt auch für VoD-Angebote in der HbbTV-ARD-Mediathek. Neben einer Übersicht, welche Angebote insgesamt mit DGS verfügbar sind, bietet HbbTV bei Vorhandensein von DGS ein entsprechendes Auswahlmenü in den Spracheinstellungen des VoD Inhalts an.

Sollen DGS-Inhalte ohne Nutzung von HbbTV angeboten werden, können diese entweder unter entsprechenden Menüpunkten zusammengefasst, über Auswahl-Links beim Inhalt als Alternative angeboten oder über eine Voreinstellung des Nutzers zur Regel gemacht werden (soweit der Inhalt auch mit DGS verfügbar ist).

Zulieferung und Unterstützung Geräte

Inhalte mit eingebranntem Gebärdensprachdolmetscher stellen keine besonderen technischen Anforderungen an die Darstellung und werden daher von allen Geräten unterstützt. Dabei ist es unerheblich, ob diese Inhalte über das Broadcast Signal, über Broadband oder über Anwendungen wiedergegeben werden.

Zulieferung und Unterstützung Plattformen

Eine Zulieferung von Gebärdensprachdiensten, die direkt im Bild eingebrannt sind, können von den Medienplattformen direkt an den User durchgereicht werden und stellen auch hier keine technische Herausforderung dar. Anders ist dies bei Gebärdensprachdiensten über HbbTV, da auf Seiten der Medienplattformen die HbbTV Technologie weder bei der Entgegennahme der

Signale noch bei der Weiterverbreitung unterstützt wird. Hier wird noch nach Lösungen gesucht.

4 Auffindbarkeit und Signalisierung

Signalisierung und Auffindbarkeit werden hier als zwei miteinander korrespondierende Faktoren betrachtet

Die Auffindbarkeit reflektiert die Frage der Anzeige von Informationen über das Vorhandensein von Barrierefreiheitskomponenten. Hier geht es darum, wie Nutzerinnen und Nutzer audiovisuelle Inhalte mit Barrierefreiheitskomponenten finden können.

Bei dem Aspekt der Signalisierung geht es um die Übermittlung von Informationen über das Vorhandensein von Barrierefreiheitskomponenten (z.B. Untertitel, Audiodeskription und Gebärdensprache) in einem bestimmten audiovisuellen Medieninhalt (im Rundfunk sowie in linearen und nicht-linearen Mediendiensten wie Mediatheken und Apps) vom Inhaltenanbieter an die Medienplattform oder das Endgerät.

Im Kapitel Auffindbarkeit und Signalisierung wird somit der Frage nachgegangen, ob Informationen über das Vorhandensein von Barrierefreiheitskomponenten so signalisiert werden, dass sie von den empfangenden Endgeräten, Plattformen etc. ausgelesen und angezeigt werden können und in der Konsequenz für die Nutzerinnen und Nutzer auffindbar sind.

Die Auffindbarkeit von Barrierefreiheitskomponenten kann vorbereitet werden, indem mittels Metadaten zu den einzelnen Medieninhalten für jede einzelne Sendung Informationen übergeben werden, welche Komponenten zur Barrierefreiheit (Untertitel, Audiodeskription, Gebärdensprache, Klare Sprache, Leichte Sprache) verfügbar sind.

Die Auffindbarkeit kann auf dieser Basis über folgende Funktionen unterstützt werden:

- In der Programmvorschau (EPG-Electronic Program Guide) und im laufenden Programm oder in einem VoD-Katalog werden diese Funktionen für jede einzelne Sendung angezeigt,
- In der Programmvorschau oder im VoD-Katalog können Sendungen, die einzelne Funktionen zur Barrierefreiheit anbieten, gefiltert werden; diese Filterung kann/soll durch die Voreinstellung von Nutzerpräferenzen unterstützt werden,
- Während des Konsums einzelner Programme / Inhalte können die dazu verfügbaren Funktionen zur Barrierefreiheit:
 - per default entsprechend der auf dem Gerät eingestellten Nutzerpräferenzen aktiviert werden³,
 - leicht in einem Info-GUI des Geräts angezeigt und aktiviert/deaktiviert werden.

³ Die Berücksichtigung von Nutzervoreinstellungen ist auch Gegenstand des folgenden Kapitels Nutzervoreinstellungen.

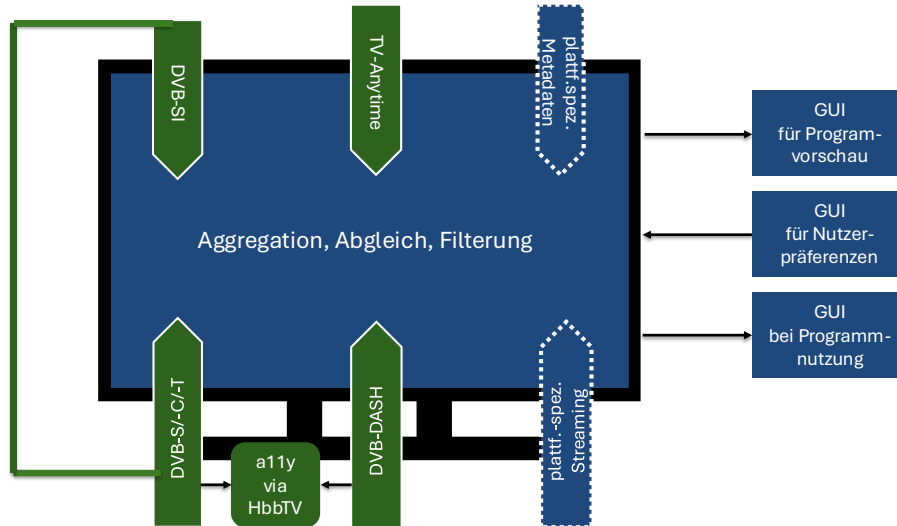
Aus diesen Zielen leiten sich für die Signalisierung nachfolgende Anforderungen ab (dies berücksichtigt die Voraussetzung, dass eine Funktion zur Barrierefreiheit nur verfügbar ist, wenn diese tatsächlich auch angezeigt werden kann, in den angelieferten Streams tatsächlich unterstützt wird und das Endgerät über die nötigen technischen Funktionen verfügt):

- Die zugelieferten EPG-Metadaten für Vorschau und laufende Sendung enthalten alle für die Auffindbarkeit nötigen Informationen,
- Die zugelieferten AV-Metadaten zur technischen Konfiguration der Medienströme enthalten alle für die Aktivierung während der laufenden Nutzung von TV-Programmen nötigen Informationen,
- Die zugelieferten AV-Metadaten enthalten Informationen, welche Funktionen in den zugehörigen AV-Streams grundsätzlich und mit welchen technischen Parametern verfügbar sind, um diese zum Beispiel in einem EPG darstellen zu können.

Die Zulieferung dieser beiden Kategorien von Daten ist in der nachfolgenden Grafik dargestellt. Für den konventionellen Broadcast sind die EPG-Metadaten in dem gleichen Datenstrom enthalten wie auch die AV-Daten und -Metadaten. HbbTV-Applikationen, die ebenfalls Funktionen zur Barrierefreiheit beinhalten können, werden sowohl beim Broadcast wie auch im Over-the-top Segment (wie auch von DVB-I abgedeckt) über die AV-Metadaten verlinkt und gestartet.

Falls Metadaten plattformspezifisch bereitgestellt werden, liegt die Verantwortung, alle Funktionen zur Barrierefreiheit in diesen Daten korrekt und vollständig abzubilden beim Inhaltenanbieter, deren Zugänglichmachung für den Nutzer beim Plattformanbieter.

Zulieferung von EPG-Metadaten für Vorschau und laufende Sendung



Zulieferung der TV-Streams mit AV-Metadaten zur technischen Konfiguration

Abbildung 4: Zulieferung von EPG-Metadaten und TV-Streams mit AV-Metadaten an das Endgerät

Für diese Signalisierungsanforderungen stehen konkret drei Standards zur Verfügung, die grundsätzlich alle benötigten Aspekte abdecken:

- hinsichtlich der EPG-Metadaten für Vorschau und laufende Sendung:
 - für reine Broadcast-Nutzung: EIT in DVB-Service Information (DVB-SI),
 - für reine IP-Streaming-Nutzung: TV-Anytime [15],
 - für hybride Nutzung: DVB-SI soweit verfügbar, aber bevorzugt TV-Anytime.
- und im Kontext der AV-Metadaten für Medienstreams:
 - Broadcast: PMT und EIT p/f in DVB,
 - IP-Streaming: Service List in DVB-I, mpd in DVB-DASH, HLS Master Playlist.

Eine Analyse dieser drei Standards ergibt, dass hinsichtlich der benötigten Features nur die beiden marktwert aktuell am meisten genutzten Features Untertitel und Audiodeskription umgesetzt sind. Einfache/Leichte Sprache⁴ wird von Broadcastern bereits auch für AV-Angebote genutzt, ist aber noch in keinem der Standards zur Signalisierung verfügbar.

In der folgenden Tabelle steht "x" für nicht im Standard umgesetzt und "-" dafür, dass keine Signalisierungsoption im Standard vorhanden ist.

Funktion	TV-Anytime	DVB-SI	DVB-DASH/HLS
Untertitel	x	x	x
Audiodeskription	x	x	x
Klare Sprache	x	[nur für NGA]	x
Gebärdensprache	x	-	x
Einfache/Leichte Sprache	-	-	-

Zusätzlich zu der statischen Betrachtung der verfügbaren Metadaten ist für lineares TV der Aspekt der Triggerung des Sendungswechsels zu beachten, damit das Gerät zuverlässig zu jeder Sendung die korrekten Metadaten anzeigen kann. Während hier in DVB-SI mit der EIT present/following (p/f) für den Broadcast ein bewährter Push-Mechanismus zur Verfügung steht, ist im OTT-Segment kein gleichwertiger Push vorgesehen. Allerdings machen sich viele Plattformbetreiber die EIT p/f Daten zunutze, um in ihren EPGs und Videoplayern den Sendungsstatus berücksichtigen und anzeigen zu können. Dies ist insbesondere bei Sendezeitverschiebungen relevant.

⁴ Angebote in Einfacher oder Leichter Sprache sollen u.a. Menschen mit Lernschwierigkeiten das Verstehen von Texten erleichtern. In Deutschland leben mehr als sieben Millionen funktionale Analphabeten, die Texte lesen können, aber die Zusammenhänge eines komplexen Textes nicht begreifen können. Um diesen Menschen Zugang zu Informationen zu bieten, gibt es die Einfache und die Leichte Sprache. Die Angebote werden auch von Sprachanfängern, Menschen mit anderen Lernschwierigkeiten, mit Demenzerkrankungen oder Hirnverletzungen genutzt. Einfache Sprache ist eine sprachlich vereinfachte Version von Standardsprache oder Fachsprache und der Sprachstil ist betont einfacher, verständlicher und klarer. Leichte Sprache geht noch weiter. Sie folgt einem festen Regelwerk, das neben Sprachregeln auch Rechtschreibregeln sowie Empfehlungen zu Typografie und Mediengebrauch umfasst und dadurch das leichteste Verständnis von Texten ermöglicht.

5 Benutzeroberflächen

Bei der Analyse der barrierefreien Gestaltung von Benutzeroberflächen wurden Benutzeroberflächen von Endgeräten sowie von App-Anwendungen über HbbTV und von herstellerspezifischen Apps in den Portalen in die Betrachtung miteinbezogen, sofern sie Zugang zu audiovisuellen Mediendiensten ermöglichen.

Nach dem hier zugrunde gelegten Verständnis umfasst eine Benutzeroberfläche alle Komponenten – sei es Software oder Hardware –, die den Nutzerinnen und Nutzern Informationen bereitstellen und ihnen Steuerelemente anbieten (vgl. Norm ISO 9241-110 [16]).

In erster Linie umfassen die Ausgabemöglichkeiten visuelle Elemente wie Texte, Schaltflächen, Symbole und Menüs, über die Informationen dargestellt und Aktionen ermöglicht werden. Zusätzlich können akustische Ausgaben über Lautsprecher erfolgen, etwa zur akustischen Rückmeldung oder zur sprachlichen Nutzerführung.

Für die Eingabe stehen den Nutzerinnen und Nutzern unterschiedliche Möglichkeiten zur Verfügung. Klassische Methoden sind etwa Fernbedienungen oder Tasten am Endgerät, über die Nutzerinnen und Nutzer mit dem System interagieren können. Darüber hinaus bieten Benutzeroberflächen auch sprachgesteuerte Nutzerführung an, bei der Steuerungsbefehle durch gesprochene Sprache gegeben werden. Eine weitere Form der Interaktion ist die Berührungs- oder Gestensteuerung, bei der das System natürliche menschliche Bewegungen, Berührungen oder Gesten erkennt und in entsprechende Aktionen umsetzt.

Zielvorstellung: Die Benutzeroberflächen, über die Zugang zu audiovisuellen Mediendiensten geschaffen werden kann, sind weitgehend von den Regelungen des European Accessibility Act (EAA) [17], des Medienstaatsvertrags (MStV) [18] und des Barrierefreiheitsstärkungsgesetzes (BFSG) [19][20] erfasst.

Das BFSG legt verbindliche Anforderungen an die Gestaltung von Produkten und Dienstleistungen fest. Unternehmen müssen ihre Produkte und Dienstleistungen so entwickeln und herstellen, dass sie den Anforderungen an die Barrierefreiheit entsprechen – das heißt: Sie müssen auch von Menschen mit Behinderungen ohne besondere Erschwernis und grundsätzlich ohne fremde Hilfe genutzt werden können.

Konkret bedeutet das:

- Wahrnehmbarkeit: Inhalte müssen so aufbereitet sein, dass sie von allen Nutzerinnen und Nutzern erfasst werden können, beispielsweise durch Textalternativen für Bilder oder ausreichende Farbkontraste sowie anpassbare Schriftgrößen.
- Bedienbarkeit: Die Größe von Schaltflächen muss eine angemessene Bedienbarkeit berücksichtigen.
- Verständlichkeit: Inhalte und Navigation sollen logisch aufgebaut und leicht verständlich sein.

- Robustheit: Die Technik muss mit verschiedenen Endgeräten und Hilfsmitteln (wie z.B. Screenreadern) zuverlässig funktionieren.

Die Mitglieder der DTVP arbeiten daran, die Barrierefreiheit ihrer Benutzeroberflächen zu erhöhen. Im Einzelnen unterscheiden sich die Methoden und Darstellungen, mit denen die Anforderungen an die Barrierefreiheit auf der Benutzeroberfläche umgesetzt werden. Barrierefreiheit ist beispielsweise hinsichtlich der Einstellungsmöglichkeiten von Barrierefreiheitsanforderungen auch ein Abgrenzungsmerkmal gegenüber Wettbewerbern. Die Möglichkeit, Benutzeroberflächen-spezifische Lösungen zuzulassen, schafft Raum für innovative Lösungen der Benutzerführung für mehr Barrierefreiheit. Dadurch werden Unternehmen motiviert, auch über das Maß des rechtlich Geforderten barrierefreie Lösungen zu entwickeln, um sich von ihren Wettbewerbern abzuheben. Es werden mithin Anreize zur Entwicklung von „Leuchtturmprojekten“ gesetzt.

6 Nutzervoreinstellungen

Sowohl bezüglich der Barrierefreiheitskomponenten des Medieninhalts (Untertitel, ...) wie auch bezüglich der Einstellungen der Benutzeroberfläche (Schriftgröße, Screenreader, Sprachsteuerung,...) wäre es aus Endkundensicht naheliegend, dass eine gewählte Komponente oder Funktion auch ausgewählt bleibt, wenn er/sie den Sender oder vom integrierten Receiver in eine App wechselt. Dies setzt eine **VOR**-Einstellung der einzelnen Barrierefreiheitskomponenten oder Benutzerschnittstellen voraus. Innerhalb der Benutzeroberfläche eines Endgerätes oder einer auf dem Endgerät installierten App sind solche Nutzer**VOR**einstellungen in der Regel möglich.

Für ein abgestimmtes Zusammenspiel zwischen verschiedenen HbbTV-Apps auf einem Endgerät wurde inzwischen eine Lösung entwickelt: Zum Accessibility Framework mit HbbTV 2.0.4 gibt es einen Proof of Concept zwischen der ARD und einigen Geräteherstellern für eine Schnittstelle zwischen Geräte-Einstellungen und ARD-Einstellungen. Erste Tests auf den Testgeräten waren erfolgreich.

Für ein Zusammenspiel zwischen beliebigen Anbietern wäre ein einheitliches Set an Metadaten hilfreich, mit denen Funktionen und Komponenten der Barrierefreiheit abgebildet werden. Da es potenziell eine unbegrenzte Zahl möglicher Einstellungen gibt, wäre es zunächst sinnvoll, sich am Markt über Kernfunktionen zur Barrierefreiheit abzustimmen, für die einen Austausch von Voreinstellungen angestrebt werden sollte. Ansatzpunkte hierfür zumindest bezüglich der Barrierefreiheitskomponenten der Inhalteanbieter liefern die Vorschläge des ANSI / CTA Accessibility Preferences Data Model [21] aus den USA.

Dieses Modell bildet europäische Anforderungen allerdings nur teilweise ab. Innerhalb der Task Force wurde als ein erstes Beispiel für mögliche Mindestanforderungen für einige Dienste genannt:

- Audio Description (Yes/No),

- Sign Language (Yes/No),
- Subtitles:
 - Subtitles for Deaf and Hard of Hearing (Yes/No),
 - Subtitles for translation (Yes/No),
 - Subtitle Language:
 - First preferred Subtitle Language (Text Field),
 - Second preferred Subtitle Language (Text Field).

Auch das für HbbTV standardisierte Datenmodell zum Austausch der Einstellungen zwischen Gerät und Applikation kann als Blaupause auch für andere Gerätetypen dienen.

7 HbbTV

HbbTV bietet ein universelles technisches Framework für Applikationen auf SmartTVs. Dieses ist bei ETSI standardisiert und heute herstellerübergreifend auf etwa 16 Mio. aktiven Geräten im deutschen Markt vertreten. Auf Big Screens ist dies der technische Standard⁵ für SmartTV-Apps mit der größten Reichweite.

HbbTV ist speziell entwickelt zum Einsatz im Kontext linearer Programme und wird auch ein Hauptbestandteil vom neuen Standard DVB-I sein, dessen Markteinführung in Deutschland aktuell vorbereitet wird. HbbTV kann sowohl mit Programmen über DVB als auch mit IP-Livestreams verbunden werden. Neben den „Red Button“-Apps zum Start aus den linearen Programmen heraus können HbbTV-Apps auch von Geräteportalen oder von einem Mobilgerät aus gestartet werden.

Durch seine flexible Einsatzmöglichkeit erlaubt HbbTV auf vielfältige Weise die Unterstützung von Funktionen zur Barrierefreiheit:

- Zugang zu AV-Streams mit Barrierefreiheits-Inhalten im Kontext Linear und Mediathek
- Individualisierungsoptionen (z. B. Untertitel, GUIs)
- Abgleich mit Geräteeinstellungen und -funktionen

HbbTV wird zur Unterstützung der Barrierefreiheit im deutschen Markt bereits seit Jahren genutzt. Aktuelle im Betrieb befindliche Beispiele dafür sind:

- Individuell personalisierbare Darstellung (Größe, Position) von Untertiteln
- Anpassbares GUI für den Teletext
- Geräteübergreifend einheitliche Aktivierung der Audiospur "Klare Sprache" im Broadcast
- Aktivierung von Videos mit Gebärdensprachdolmetscher
- Filterung auf Videoinhalte mit Untertiteln/Audiodeskription/DGS in der Mediathek (EPG und OnDemand)

⁵ Überblick über die verfügbaren HbbTV-Standards: <https://www.hbbtv.org/resource-library/>

Die EBU hat 2021 den Technical Report TR065 (Guidelines for delivering Accessibility Services using HbbTV) [22] veröffentlicht, indem weitere Informationen und Beispiele zu finden sind.

8 Bibliographie

- [1] ETSI TS 103 770: Digital Video Broadcasting (DVB); Service Discovery and Programme Metadata for DVB-I
- [2] ETSI EN 300 743: Digital Video Broadcasting (DVB); Subtitling systems
- [3] Recommendation ITU-R BT.653: "Teletext systems".
- [4] ETSI EN 300 706: "Enhanced Teletext specification".
- [5] ETSI TS 102 796: Hybrid Broadcast Broadband TV
- [6] ETSI EN 303 560: Digital Video Broadcasting (DVB); TTML subtitling systems
- [7] W3C Candidate Recommendation WebVTT: The Web Video Text Tracks Format
- [8] IETF RFC 8216: HTTP Live Streaming
- [9] W3C Recommendation (TTML): Timed Text Markup Language
- [10] EBU Tech 3350: EBU-TT, Part 1 Subtitling Format Definition
- [11] EBU Tech 3380: EBU-TT-D Subtitling Distribution Format
- [12] ETSI TS 103 285: Digital Video Broadcasting (DVB); MPEG-DASH Profile for Transport of ISO BMFF Based DVB Services over IP Based Networks.
- [13] ISO/IEC 23009-1:2022 Information technology – Dynamic adaptive streaming over HTTP (DASH) – Part 1: Media presentation description and segment formats
- [14] XML-Format für die Distribution von Untertiteln in den ARD Mediatheken (EBU-TT-D-Basic-DE)
- [15] ETSI TS 102 822-3-1: Digital Video Broadcasting (DVB); Broadcast and On-line Services: Search, select, and rightful use of content ("TV-Anytime"); Part 3: Metadata; Sub-part 1: Phase 1 – Metadata schemas
- [16] ISO 9241-110: Ergonomics of human-system interaction — Part 110: Interaction principles
- [17] European Accessibility Act (EAA): RICHTLINIE (EU) 2019/882 DES EUROPÄISCHEN PARLAMENTS UND DES RATES vom 17. April 2019 über die Barrierefreiheitsanforderungen für Produkte und Dienstleistungen
- [18] Medienstaatsvertrag (MStV)
- [19] Barrierefreiheitsstärkungsgesetz (BFSG): Gesetz zur Umsetzung der Richtlinie (EU) 2019/882 des Europäischen Parlaments und des Rates über die

Barrierefreiheitsanforderungen für Produkte und Dienstleistungen
(Barrierefreiheitsstärkungsgesetz – BFSG)

[20] BFGSGV: Verordnung über die Barrierefreiheitsanforderungen für Produkte und Dienstleistungen nach dem Barrierefreiheitsstärkungsgesetz (Verordnung zum Barrierefreiheitsstärkungsgesetz – BFGSGV)

[21] ANSI/CTA-2115: Accessibility Preferences Data Model

[22] EBU TR O65 Guidelines for delivering Accessibility Services using HbbTV

9 Impressum

Herausgeber:

Deutsche TV-Plattform e.V. (DTVP)

Bockenheimer Landstrasse 31

D-60325 Frankfurt/Main.

www.tv-plattform.de

E-Mail: mail@tv-plattform.de

Telefon: +49 69 870 03 29-10

Amtsgericht Frankfurt/M.

Vereinsregister Nr. 9797

Geschäftsführung: Nicole Ludwig

Vorsitzende: Nicole Agudo Berbel

Stellvertretende Vorsitzende: Sascha Molina, Dr. Dietmar Schill

Redaktion:

Task Force Barrierefreiheit der Arbeitsgruppe Medien over IP (AG MoIP) in der DTVP

Leiter der AG MoIP: Frank Heineberg, RTL Deutschland

Leitende der Task Force Barrierefreiheit: Katrin Heyeckhaus, ZVEI e. V., Patrick Krisam, Telekom, Andreas Tai, ARD

Über die Deutsche TV-Plattform e.V. (DTVP):

Von Sendern über Infrastrukturbetreiber, von Geräteherstellern bis zu Forschung und Bundesbehörden: Die Deutsche TV-Plattform (DTVP) bringt seit 1990 die Medienbranche zusammen – für offene Standards, gemeinsame Projekte und die Zukunft von TV & Streaming in der DACH-Region. Die über 50 Mitglieder der DTVP repräsentieren das gesamte Ökosystem des digitalen Fernsehens und haben nur ein Ziel: durch die Einführung digitaler Technologien auf Grundlage offener Standards ein perfektes Medien-Erlebnis zu bieten – für alle Inhalte, auf allen Endgeräten, zu jeder Zeit, an jedem Ort.

Haftungsausschluss:

Die Informationen in dieser Handreichung wurden gründlich recherchiert und nach bestem Wissen zusammengestellt, basierend auf dem unvoreingenommenen Ansatz der AG MoIP in der DTVP/ Task Force Barrierefreiheit. Alle hierin enthaltenen Informationen spiegeln den Status quo zum Redaktionsschluss am 31. März 2026 wider. Die Mitglieder der Arbeitsgruppe und der DTVP können jedoch ihre Qualität und/oder die Aktualität, Korrektheit und Vollständigkeit der zusammengestellten Informationen nicht garantieren. Daher kann die DTVP als Herausgeberin dieser Broschüre keine Haftung für wesentliche oder immaterielle Verluste oder Schäden übernehmen, die durch die Nutzung dieser Publikation oder deren Inhalte oder infolge ungenauer oder unvollständiger Informationen entstehen.