ATVOD Working Group on Access Services (WGAS)

Report prepared by Nick Tanton

Introduction

The perceived ease with which viewers can now get access services with linear TV in the UK [1] makes it hard for users/customers to understand why on-demand services generally lag behind in accessibility.

The end-to-end chain (content acquisition, versioning, scheduling, delivery and presentation) for on-demand generally differs substantially from that of linear TV. Some processes are (or could be) the same or similar whilst others are not; some also may apply to one on-demand service provider but not to another. In particular the species (and sometimes generational variants) of end-user equipment are diverse and some may not have been designed a priori with accessibility in mind.

For on-demand content there are no insuperable barriers to providing subtitles for Hard-of-Hearing people or Audio Description for Blind and Partially Sighted users. However there are various practical technical and commercial challenges which a regulator or service provider may need to recognise and to address if these access services are to be made available in a realisable and usable form to the end-user. For any particular service provider, the mix and impact of these challenges depends on the delivery platform(s) on which they have a presence, their role within a particular delivery chain and their business model.

The content of this report represents a distillation of understanding from within the ATVOD Working Group on Access Services (WGAS) and input from individual companies who have participated in those discussions. It sets out to identify the technical challenges (some of which also impact commercial sensitivities … workflow etc.), to signpost both what has been and is already being done to overcome them and to identify what further steps might be taken by ATVOD and by its service providers.

It focusses on subtitles (STs) for the Hard-of-Hearing (HoH) with pre-recorded content and on Audio Description (AD) for Blind and Partially Sighted people (B&PS). Throughout, the abbreviation "STs" should be taken to mean subtitles for the Hard-of-Hearing in the language of the main audio track (i.e. in English, Welsh, Scots or Irish Gaelic) - we are not here considering translation subtitles.

The report addresses the provision of access services associated with assets which have been prepared for linear platforms but not, as yet, those specifically prepared for on-line delivery.

---

[1]   after much concerted work by broadcasters, industry, the regulator and lobby groups

The processes involved in providing subtitles & AD

The ecology of on-demand services is potentially more complex and diverse than that of linear television. The end-to-end chain (from original content provider to end-user) however involves broadly similar overall processes which are also applicable to the provision of access services.

These can be categorised as :
- content acquisition,
- authoring/editing,
- versioning,
- scheduling and playout &
- delivery and presentation.

For some on-demand (OD) service-providers (SPs) all of these processes have a direct impact on their business, for others perhaps only a subset. An organisation which has a significant presence both in linear and in on-demand such as one of the UK Public Service Broadcasters or Sky, may be able to take strategic control of most of the end-to-end chain. Other content providers, content aggregators and OD SPs work under different commercial and technical constraints (commercial scale, leverage in acquisition and in delivery negotiations, being only part of the service chain etc.). It's thus unwise to make overly-simplistic direct comparisons between the accessibility of services (and their associated economics) on BBC iPlayer or 4OD and those of all OD SPs.

Content acquisition

OD content is typically derived from material developed directly for the UK or imported as film or video usually from the US. Recorded content originated for the UK linear TV market and available on catch-up services is likely to have some form of subtitle data or at the very least may offer access to a script from which subtitles may be authored. It may also have an AD track. Other UK content may or may not have pre-existing subtitle data and is unlikely to have an AD track.

For imported content if there is any such data available it needs to be explicitly and pro-actively procured with the Audio Visual (AV) content. Furthermore, that data, where available, is likely to be in a format (and of a style) that requires further processing - see below. The imminent implementation of FCC [2] mandates for newly-made and archive AV material in the US will make it much easier in the near future to acquire content with subtitle data (even, it is to be hoped, at no extra cost) but is unlikely to obviate some form of format conversion and post-processing.   It is suggested that SPs always endeavour to procure US content with STs and AD when available and that the UK administration actively work in Europe towards mandating

---

[2]   US Federal Communications Commission

that AS components for AV content be free (or at the very least low cost) at the point of exchange and that those components be catalogued. Equally the UK administration should take steps to ensure that business-unit boundaries within the procurement chain present no barriers to access service provision.

Here we need to note that for historical reasons there are a number of formats for subtitle data and that these tend to be region-specific (eg. US vs Europe) and to have been based less on content exchange and more on maintaining a simple relationship with legacy delivery standards for linear TV. Thus for example the EBU [3] subtitle "STL" file format (which dates from 1991) is directly related to and constrained by World System Teletext. This multiplicity of formats has been a barrier to implementing access services on imported content.   However, recent and ongoing development in W3C [4] and the EBU relating to the XML [5]-based "Timed Text" successfully decouples the data structure from delivery method and seeks to make content exchange and any subsequent processing (editing, versioning and archiving for example) considerably more straightforward.

If content is procured without STs or AD in any form, they must be prepared subsequently from scratch in what is a labour-intensive albeit well-understood process.   Some metadata associated with the AV content (finished script, shot-timings etc.) may be available from the content provider (if negotiated for) and can go some way to reducing the effort and cost of authoring STs or AD.

Authoring/editing/versioning/archiving

In the event that content has been procured with STs or AD there may be a need to convert the data from one format to another. Generally this is not an insuperable or necessarily expensive problem – input from Screen Subtitling Systems (see Annex 1) and reports from IBC2013 indicate that suitable tools are (or can be made to be) available to do this with only modest manual intervention (see below). There may of course be consequent changes required to associated workflows. A wider knowledge of such tools would be helpful.

The input content may be "hard-parted" (e.g. partitioned episodically for advert breaks etc.) and video editing will be required to reconstruct a continuous narrative – the subtitle data will then need to be conformed to the new edit using conventional subtitle authoring tools.

Whether for linear or for on-demand services, different versions of AV content are typically prepared for different outlets and purposes. OD service providers will

---

[3] European Broadcasting Union

[4] World Wide Web Consortium – specifically the Timed Text Working Group W3C TTWG

[5] the e**X**tensible **M**arkup **L**anguage implementation of Timed Text (**TT**) is TTML

typically make "long-form" programmes, "short-form" programmes and "clips". This effectively involves an edit for each version and so the access service components (STs and AD) will also need to be edited (or re-authored) to match the revised AV content. A particular challenge comes with managing fast turn-round content (e.g. from a contemporary event) especially when the original was captured (and subtitled) live [6]. Managing these different versions is also a challenge especially to ensure that the correct ST or AD data is associated with each version.

The editorial/stylistic conventions to which UK viewers are accustomed for STs and for AD in linear TV have been derived with considerable input from 3rd sector bodies representing HoH and B&PS people over a number of years. These are reflected in OFCOM & ATVOD guidance [7] [8].

However whenever STs are imported from other access services cultures the input subtitle data may be subject to different technical constraints (eg. monochrome "closed captions"), another frame-rate, a different stylistic approach (eg. not highlighting relevant sounds "telephone rings"), positioning (including avoiding occlusion of a speaker's mouth), line length, spelling, slang/adult language and reading speed.

To fully respect the resulting current user expectations in the context of all OD services will require considerable intervention whenever STs and/or AD are imported from other access services cultures. It may be appropriate to apply best endeavours using what has been procured and to move to full "compliance" with these guidelines over a timescale to be defined.

Editing a pre-existing AD track isn't feasible if that track is already a mix of narrative effects and description (a remix will be required). Simply editing a separate description track may be impractical [9] unless entire description passages either side of the new edit points are omitted if necessary (which may have a significant deleterious impact on the user experience for B&PS people).

Importing AD may require managing a different frame rate, a different pace of narrative, vocabulary (elevator/lift, sidewalk/pavement …) or the tone of delivery, voice or accent. In many cases re-voicing the AD may be more time-efficient and cost-effective.

If content has not been procured with STs or AD then it becomes necessary to author them as one would with newly commissioned content. As noted above any

---

[6] It seems reasonable to suggest that, in the short term, clips shorter than say 30s and fast turn-round short-form content be excused from carrying STs and that priority be accorded to long- and then to short-form content.

[7] http://stakeholders.ofcom.org.uk/binaries/broadcast/guidance/guidelines.pdf

[8] http://www.atvod.co.uk/uploads/files/Access_Services_best_practice_guidelines_FINAL_120912.pdf

[9] as it would involve re-voicing part of a description in the same or a "matching" voice

procured associated metadata may help in this process but will add additional complication with time-codes if the content was originated at a different frame-rate.

If the SP expects to deliver the resultant content to a number of platforms or to archive the content for re-use, it makes considerable sense to "author once/deliver many" – authoring in a future-proof easily exchangeable format (such as EBU-TT) which can support the diverse requirements of each platform and from which the appropriate physical signal can be derived for that or any other platform. There's nothing particularly new in this – the BBC has been transcoding (automatically and on-the-fly) subtitle data and AD in different signal forms for DTT and for DSAT since the start of digital television in the UK in 1998 [10].

Scheduling and playout

A necessary part of the user experience for HoH and B&PS people is their knowing whether to expect access components with the requested content or a variant of it. Early experience with DTV demonstrated the importance of accurate (and timely) metadata particularly when not all content has STs and/or AD. For this user community consistency of availability of access services during a series or for a programme strand is also important – to have one episode of a series unsubtitled or undescribed is almost as bad as having no STs or AD on the whole series. All user terminals should (and some may) have mechanisms for using this metadata to inform the user. The exact data format for delivering it (and whether it's used) will depend on the platform but it should be accurate.

Maintaining separate instances of complete content with and without ASs has a significant adverse impact on work-flow, head-end architecture, storage capacity, bandwidth and content management. Whilst for in-vision signed versions of a "programme" this is currently inevitable, for STs and AD the provision of a single stream containing "closed" access service components to be selected by the user and presented (or not) by the user equipment (as per linear TV in the UK) is much simpler, cost-effective and reliable for the VOD operator and more user-friendly. The alternative is to deliver the STs or AD via a side-channel which presents an additional set of challenges and complications (management, navigation, validation, entitlement, synchronisation etc.) and adds risk for successful delivery and user-satisfaction.

Delivery and Presentation

A particular challenge for the OD SP comes in the variety and diversity of end-user equipment to which on-demand AV content is delivered.

---

[10] The BBC has also adopted EBU-TT as an authoring, exchange and archiving format for linear and OD delivery.

To cover these bases, content-assets including their accessibility components will typically have to be coded and delivered in a number of forms and according to a variety of standards. This further reinforces the need for an "author once, deliver many" approach to subtitles and AD. As with conversion of various input formats to a "universal" standard for editing, archiving and exchange such as EBU-TT, conversion from a TTML form to a particular delivery format (eg. webVTT) is relatively straightforward [11] [12] [13] (see Annex 1).

Such user equipment may be standardised according to a particular platform operator's prescription or might involve the use of a browser whose functionality has been implemented either without fully understanding the user and service requirements of STs or AD or by compromising those requirements for other commercial or market considerations [14]. In practice the SP may (or may not) have control of the perceived result and the user experience will probably differ between different platforms or different instances of end-user equipment.

In addition the manner in which the user interacts with such services (eg. selects STs, whether a selection is maintained between sessions etc.) often differs in some degree between platforms or user equipments. A lack of consistency in the user experience (and the increased importance for HoH and B&PS viewers of these ASs) has an impact on the perceived quality of service and is typically blamed on the SP with consequent management and cost overheads. This effect will be amplified by the heightened expectations that HoH and B&PS viewers now have from experiencing mainstream UK linear channels. Here, when access services are more widely available on UK OD services, ATVOD could usefully play a part in educating users to understand that there will inevitably be a diversity of approach between different delivery systems.

Standardisation

As noted above, considerable progress has been and continues to be made on the standardisation of subtitle data by W3C and the EBU both for business-to-business and business-to-consumer. A wide set of totally different subtitle formats can now be replaced by a managed set of profiles of TTML – same basis, different constraints and extensions. There is considerable cross-fertilisation between W3C-TTWG and

---

[11] Several browser manufacturers are initially supporting WebVTT ahead of TTML – W3C TTWG's new charter includes in the group's deliverables a recommendation for a mapping document for WebVTT. See http://www.w3.org/TR/ttml1/ and also http://www.w3.org/2014/03/timed-text-charter.html.

[12] TTML is believed to have a US legal mandate, so increased support is to be expected thanks to the 21st Century Act.

[13] BBC has been publishing TTML subtitles on iPlayer for some years.

[14] In some instances there are more obvious technical constraints – when subtitling AV on a very small screen (eg. on a mobile phone) and where subtitles may have to compete with other visual information.

the EBU, with EBU-TT developments expected to be incorporated in the next revision of TTML and the use of EBU expertise for defining a streaming TTML [15] [16].

In January 2014, the EBU published a draft specification for the carriage of EBU-TT-D in ISO Base Media File Format (BMFF, ISO/IEC 14496) as Tech 3381. This defines how EBU-TT-D subtitles can be encapsulated in ISO BMFF and thus facilitate the provision of EBU-TT-D subtitles using MPEG DASH [17].   ISO MPEG has now published a stable version of ISO/IEC 14496-30, and the EBU is currently in the process of publishing a final version of Tech 3381.

EBU-TT-D was developed partly in response to a request from the HbbTV association to provide a suitable subtitle format for its HbbTV 2 connected television standard. HbbTV 2 also depends on DVB's DASH profile. Work is ongoing within HbbTV and DVB to specify how EBU-TT-D will be adopted within these standards. A group of German broadcasters has already adopted a narrow "ARD" subset of EBU-TT for distribution [18].

The UK's Digital Production Partnership (DPP) is in the process of finalising as an open document a variant of EBU-TT Part 1 to be the delivery format for prepared subtitles, captured live subtitles and audio description scripts. No specific timeline for adoption of this format will be specified but it is expected that UK broadcasters will require delivery from access services providers in the DPP format in supply contracts as and when they are procured or refreshed. The DPP format adds additional metadata (currently present in EBU-TT) to support broadcast workflows. EBU is considering these additions and may adopt a subset of them in the upcoming revision to EBU-TT Part 1 [19].

The EBU have a series of open EBU-TT documents that are published or planned.

EBU Tech 3350   EBU-TT Part 1: Subtitling format definition
(Introduction to EBU-TT and normative definition of the XML based format)
https://tech.ebu.ch/docs/tech/tech3350.pdf

---

[15] https://tech.ebu.ch/groups/pdfxp

[16] Note that AD scripts and relevant metadata could in principle also be conveyed using EBU-TT if suitable extensions were defined.

[17] MPEG Dynamic Adaptive Streaming over HTTP (ISO/IEC 23009-1), an adaptive bitrate system for use both for live streaming and for on-demand content.

[18] http://www.irt.de/webarchiv/showdoc.php?z=NjM0MCMxMDA2MDE4I3BkZg==

[19] http://dpp-assets.s3.amazonaws.com/wp-content/uploads/2014/03/Subtitle_Ex_Format_Final.pdf

EBU Tech 3360	Mapping EBU-TT (Tech 3264) to EBU-TT Subtitle Files *(draft for comments)*
	(Part 2: How EBU-TT provides backwards compatibility with EBU STL)
	https://tech.ebu.ch/docs/tech/tech3360.pdf

EBU Tech 3380	EBU-TT-D Subtitling Distribution Format *(business-to-consumer)*
	https://tech.ebu.ch/docs/tech/tech3380.pdf

EBU Tech 3381	EBU-TT Carriage of EBU-TT-D in ISO/BMFF *(draft for comments)*
	https://tech.ebu.ch/docs/tech/tech3381.pdf

In preparation	EBU-TT Part 3: Live Subtitling - Authoring and Contribution
	(How to use EBU-TT for the production and contribution of live subtitles)
	EBU-TT Part 4: Intentional Subtitling (for annotation)
	(How EBU-TT can be used in future scenarios for 'authoring of intent')
	EBU-TT Part 5: User Guide (How to use EBU-TT)

Conclusions and recommendations

- There are no fundamental *technical* barriers to providing subtitles and where appropriate audio description on on-demand content. For service/content providers, most other barriers are related to workflow issues which can be significantly complicated by the variety of delivery standards (even within company groups).

- Providers of regulated on-demand services should be strongly encouraged to provide subtitles and audio description with their content. This might be phased over time, guided by commercial and technical considerations (eg. with long-form content taking priority to be followed eventually by short-form content).

- On-demand service providers should always endeavour to procure content with subtitles and audio description or generate their own (subject to commercial practicalities).

- ATVOD should encourage content owners and distributors to make subtitles and audio description available with all content, in the recommended format and free (or at a nominal charge).

- ATVOD service providers should be encouraged to adopt EBU-TT at all relevant parts of the delivery chain as soon as is practicable. This should simplify their work-flows.

- ATVOD service providers should be encouraged to provide audio description using the simplest practicable method for that platform [20].

- ATVOD and service providers should prioritise the provision of subtitles and audio description to future platforms and to existing platforms that can already deliver them. Note that the reliable provision of subtitles on pre-recorded material is probably more pressing than for live or near-live content.

- ATVOD and ATVOD service providers should work with others in Europe towards mandating that pre-existing access service components for acquired AV content be free at the point of exchange and that those components are catalogued.

- All new on-demand services/platforms should design-in provision of subtitles and audio description in a form that respects the user requirements.

---

[20] There is no standard content exchange format for AD and ATVOD service providers could usefully encourage and contribute to suitable standardisation.

- It is always appropriate to apply best endeavours using what has been procured to offer access services even if only in a sub-optimum style on the understanding that a move to fully-matching established guidelines will follow over a timescale to be defined.

- ATVOD should take steps to strongly encourage platform operators, equipment manufacturers and browser writers to support the extraction and presentation of in-stream access service components and to discourage any strategies which would complicate or otherwise risk access service provision.

- With the various stakeholder organisations ATVOD could help educate users to understand that there will inevitably be some diversity of approach between different delivery systems and some systems might better serve a particular set of users.

- The ATVOD WGAS should be maintained not least as an active forum amongst ATVOD service providers for the sharing of experience and current practice in implementing subtitles and audio description. This should include sharing understanding about EBU-TT and providing input to other bodies (eg. UK DPP, UK DTG etc.) to speed the development of representative test scenarios and content files of EBU-TT data designed to validate service provider implementations.

List of contributors

Action on Hearing Loss, BBC, BBC Worldwide, BT, Channel 4, Channel 5, Community Channel, DeLuxe, Digital TV Group (DTG), Discovery, Fox, IMS, ITV, Ofcom, Red Bee Media, RNIB, Screen Library, Screen Subtitling Systems, Sense, Sky, STV, TalkTalk, Technicolor, Turner, UTV, Viacom, Virgin Media

Nigel Megitt (BBC) as chair of EBU XMLSubs group – aka EBU-TT)

Nick Tanton (consultant) as chairman of ATVOD WGAS

Annex 1

Contribution on subtitle format conversion by Screen Subtitling Systems [21]

Web delivered (OTT) subtitles

A wide range of input subtitle files can be used to create web based subtitle outputs. Alternatively where media has already been subtitled in any delivery format the subtitles can be extracted and re-purposed for web delivery in a fully automatic process.

Suitable input formats:

- All file based subtitle file formats including but not limited to:
  PAC, RAC, CHK, STL, OVR, SIM, ZHP, CAP, 890, EBU-TT, SMPTE-TT
  and multiple text and XML based file format

- Teletext subtitles
  As used on old analogue UK services and also present on standard definition (SD) tape and some video file formats or in some high-definition (HD) video formats.

- DVB bitmap to EN 300 743
  As used on all UK Freeview and Freesat services and in some Transport Stream (TS) file formats.

- Closed Captions to EIA 608B and EIA-708
  US SD and HD formats used in all North American broadcasts and present in a range of formats in tape and file based media.

---

[21] www.screensystems.tv

Subtitle Function Chart

These charts and notes show all the input and output options for both file based and real-time conversions.

Look down the left hand side of the table to select the source of subtitle data. This could be from pre-prepared files or existing signals on a tape, video server or received from a satellite. Then select the output type required on the top line and the intersecting cell will show if the conversion is possible.

Output Styles / Input Formats	Subtitles for VOD on computers, tablets & mobiles (9)	Subtitles for live streamed video on computers, tablets & mobiles (10)
Subtitle Files (1)	Implemented & available	At prototype stage
Live Subtitles (2)	N/A	At prototype stage
Teletext Subtitles (3)	Implemented & available	At prototype stage
Imitext or SCTE-27 bit map data (4)	Implemented & available	At prototype stage
Teletext Drive Data (5)	*Possible in principle*	*Possible in principle*
DVB Bitmap Subtitles for SD and HD (6)	Implemented & available	At prototype stage
EIA- 608 Closed Captions (7)	Implemented & available	At prototype stage
EIA- 708 Closed Captions (8)	Implemented & available	At prototype stage

Notes to Table

Inputs

(1)   Pre-prepared Subtitle files

Physical Interface: PC files

Formats: PAC, RAC, CHK, STL, OVR, SIM, ZHP, CAP, 890, EBU-TT, SMPTE-TT and multiple text and XML based file formats

(2)   **Subtitles created live at transmission time for live events, news and sport**

Live subtitling styles include:

     Live to air typing using standard or stenographic keyboards
     Voice recognition systems with or without a verification stage
     Pre-prepared subtitle files cued to air manually.

Physical Interface: Ethernet or Serial

Formats: Polistream, Nufor and XML based formats

(3)   **Teletext inputs**

Physical Interface: Analogue, SDI, ASI or DVB over IP

Formats: PAL or DVB-Teletext or any video file format containing VBI data in picture or VBI track, SD or HD

(4)   **Imitext is a Screen and Scientific Atlanta bit map Subtitle format**

SCTE-27 is a derivative defined by Cable-Labs

Physical Interface: ASI or DVB over IP

Formats: Transport Streams

(5)   **From remote Polistream System only.** Provides remote control of Polistream components and some Legacy Screen products via the broadcast signal in a private format so that Teletext subtitle coding constraints are avoided.

Physical Interface: Analogue or SDI, ASI or DVB over IP

Formats: PAL or NTSC

(6)   **DVB Bit Map subtitles used as an input**

Physical Interface: ASI or DVB over IP

Formats: Transport Streams

(7)   **Line 21 Closed Captions to EIA 608B**

Physical Interface: Analogue or SDI or file based formats with Closed Caption data as VBI or embedded data

Formats: PAL, NTSC or MPEG-2 video

(8)   **HD Closed Captions to EIA 708**

Physical Interface: HD-SDI or multiple file formats supporting VANC data tracks in MPEG-2 files Transport Streams over ASI or IP

Formats: HD

Outputs

(9) **Subtitles and captions for Web delivered (OTT) as VOD**

Interface: Via web server

Format: PNG + XML files for bitmap subtitles or SMIL, SAMI, XML Timed Text or Web VTT
formats for code based subtitles

(10) **Subtitles and captions for Web delivered (OTT) as live or streamed services**

Interface: Via web server
Format: PNG + XML files for bitmap subtitles or SMIL, SAMI or XML Timed Text
formats for code based subtitles