D3.2

Version	1.00
Date	2015/10/05
Author	IRT
Dissemination status	PU
Document reference	D3.2

D3.2 – Pilot-A Solution Integration and Trials

Grant Agreement nº:	621014
Project acronym:	HBB4ALL
Project title:	Hybrid Broadcast Broadband for All
Funding Scheme:	CIP-ICT-PSP
Annex I reference version:	2013/10/22
Project Duration:	2013/12/01 – 2016/11/01 (36 months)
Coordinator:	Universitat Autònoma de Barcelona (UAB)
Beneficiaries:	Rundfunk Berlin-Brandenburg (RBB) Institut fuer Rundfunktechnik GmbH (IRT) Rádio e Televisão de Portugal SA (RTP) Televisió de Catalunya SA (TVC) Schweizerische Teletext (TXT) Vsonix Gmbh (VSX) Fundación Centro de Tecnologías de Interacción Visual y Comunicaciones VICOMTECH (VIC) Screen Subtitling Systems Ltd (SCREEN) Holken Consultants & Partners (HC) People's Playground BV (PPG) Universidad Politécnica de Madrid (UPM)

This project is supported by funding from the 'Competitiveness and innovation framework programme 20017-2013' of the European Union.

Project no. 621014
HBB4ALL
Hybrid Broadcast Broadband for All

CIP- Pilot actions
Competitiveness and innovation framework programme 2007-2013

D3.2 – Pilot-A Solution Integration and Trials

Due date of deliverable: 2015/09/30
Actual submission date: 2015/10/05

Start date of project: 2013/12/01
Duration: 36 months

Lead beneficiary for this deliverable: Institut fuer Rundfunktechnik GmbH (IRT)

Revision 1.00

Project funded by the European Commission within the Seventh Framework Programme		
Dissemination Level		
PU	Public	**X**
PP	Restricted to other programme participants (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidential, only for members of the consortium (including the Commission Services)	

History Chart

Issue	Date	Changed page(s)	Cause of change	Implemented by
0.10	2015/05/15	-	Document structure	IRT
0.11	2015/05/26	-	Overall structure adapted to Pilot A	IRT
0.12	2015/05/28	-	Final draft structure	IRT
0.2	2015/07/31	all	Consolidated version with slightly changed structure. Moved technical details to annex.	ALL
0.3	2015/08/26	all	Updates by all partners. Reworking overview of technical details in annex	IRT
0.4	2015/09/03	all	Updates solving open issues in partners' contributions.	IRT, RTP, TVC, TXT, UAB, VIC
0.5	2015/09/29	all	Included feedback from internal review, Executive Summary	IRT, TVC, RBB
1.0	2015/10/05	all	Final version	IRT

Validation

No.	Action	Beneficiary	Date
1	Prepared	IRT	2015/05/15
2	Approved	UAB	2015/09/29
3	Released	UAB	2015/10/05

Table of Contents

Figures

Tables

Executive Summary

This document reports the progress of Pilot-A, part of the 3 year HBB4ALL project, funded under the Competitiveness and Innovation framework Programme 2007-2013 (CIP). The HBB4ALL project investigates accessibility in the hybrid broadcast-broadband TV (HbbTV) environment, where a prominent challenge is to add Access Services, in a cost-efficient manner, that is consistent with the services available on traditional broadcasts. In addition, there is an opportunity to offer HbbTV viewers the ability to customise the Access Services they are using to best meet their personal preferences or needs. Pilot-A investigates how broadcasters and service providers can provide subtitling services to support viewers who are hard-of-hearing or do not understand the language of the content. Pilot-A also investigates the automated generation of subtitles from program audio using automatic speech recognition and tests the feasibility of automatic real-time generation of multilingual subtitles from program audio (for the context of news programmes).

A principal objective of Pilot-A is to implement operational sub-pilots of HbbTV-based subtitling services that engage the complete chain of stakeholders. This will be done in three countries (Germany, Spain and Switzerland), while the Portuguese sub-pilot will implement customisable subtitles for multiple platforms. For the technical realisation of these sub-pilots, the partners have been developing and adapting the following tools, the so-called Service Components:

Authoring and Production components

- **Automatic Subtitling Component**, this generates real-time intra-lingual subtitles from English Breaking news content.
- **Machine Translation Component**, this automatically translates English subtitles into Spanish and generates EBU-TT-D subtitles.
- **Subtitling Format Conversion Framework**, this converts subtitles from legacy formats used in broadcast production into EBU-TT-D subtitles.
- **Subtitle Authoring Component**, this adds an EBU-TT-D export function as an option to an existing subtitle preparation tool.
- **Lightweight Subtitle Editor**, a low-cost native XML editor for EBU-TT-D subtitles that is customized for the re-purposing of subtitles for Internet distribution.
- **EBU-TT-D XML Schema**, this supports the validation of EBU-TT-D XML documents against the EBU-TT-D specification.

Processing / Middleware Components

- **Subtitle Contribution Component**, this receives EBU-TT-D documents over an IP connection and converts the EBU-TT data into a form that can be opaquely tunnelled within VBI in SDTV or VANC in HDTV.
- **Subtitle Distribution Component**, this supports the conversion of EBU-TT files to EBU-TT-D files.

- **Live Broadcast-Internet subtitle synchroniser**, this synchronizes subtitles to the A/V input of the DVB Transport Stream, from which audio was taken to generate the subtitles. It also generates MPEG-DASH assets for audio, video and subtitles.

- **IPTV subtitle publishing system**, this makes subtitles available in EBU-TT-D format for IP-based TV-systems.

- **HbbTV MPEG-DASH/ISOBMFF Segmenter and Packager**, an extension to the open source software MP4box which allows the generation of MPEG-DASH streams compliant with the MPEG-DASH profiles defined in HbbTV 1.5 and 2.0.

User Application and Rendering Components

- **Live Broadcast-Internet subtitle Viewer**, this un-packs the EBU-TT-D subtitles from MPEG-DASH segments and provides them as WebVTT subtitle cues to an HTML 5 browser for rendering.

- **Subtitle interpretation tool for browsers**, this extends the JW Player with the functionality to render EBU-TT-D subtitles.

- **HbbTV-based customisable subtitles for VoD playback**, this implements customisable subtitles for the VoD playback in the German HbbTV Mediathek.

- **HbbTV device and version detection**, this detects the specific device (e.g. manufacturer, type number) and the HbbTV version it implements, to ensure compatibility of the device with the application.

- **EBU-TT-D HbbTV personalised subtitle playback**, this allows subtitle customisation by the end user on his HbbTV device, for use in the Catalan sub-pilot.

- **vSubtitle**: Adaptive subtitle rendering, this is an add-on for vsonix' webcast vPlayer to handle synchronization, rendering and adaptation of subtitles and to allow users to adapt the subtitle display according to their user profile.

- **EBU-TT-D Rendering Tests**, these contain EBU-TT-D reference materials to ease a standard conformant implementation of EBU-TT-D.

- **EBU-TT-D Assertions**, these are assertions for EBU-TT-D implementation and adaptation in HbbTV 2.0, and can be used as a basis for a test suite to check conformance against the HbbTV 2.0 specification.

- **HbbTV 2.0 Demo App**, this is a prototype HbbTV 2.0 application, supporting live streaming on the Internet, using MPEG-DASH streaming and EBU-TT-D subtitles. This component has been specifically developed for presentations of showcases at IFA and IBC 2015 (by IRT and RBB, in close cooperation with industry partners, e.g. Samsung, who provided a prototype TV based on their current HbbTV development).

All components are overviewed in this document and technical implementation details are given in the Annex. Most of them are readily available for use in the respective sub-pilots. Updates and/or additional features and functionality are planned in some cases and will be added before the start of the sub-pilot.

For the technical realization of the sub-pilots, the partners have been integrating the above mentioned components in dedicated workflows for each sub-pilot. Functional and technical testing has been carried out. The HBB4ALL sub-pilots are scheduled to be carried out within the period of August 2015 – July 2016. The status of the planned sub-pilots in Pilot-A is as follows:

- **Customised HbbTV Subtitles for VoD in Germany** is planned to start in September 2015 as dedicated tests with a panel of users who will be requested to test the service for a period of approximately 12 weeks.

- **Customised HbbTV Subtitles for VoD in Spain**, for which a final technical release is planned in November 2015, followed by user tests in December 2015. The sub-pilot is scheduled to run from January to June 2016.

- **Customised HbbTV Subtitles for VoD in Switzerland**, in which HbbTV-based VoD subtitles have been realized in Italian and French language; for the German language this is planned in the first half of 2016.

- **Customised Subtitles for Wide focus multi-platform in Portugal** is planned to start in November 2015 after finalizing the integration work in October.

- **Automated HbbTV multilingual news broadcast subtitles in Spain**, for which user tests with specifically generated material have been finalised. The integration of the subtitles technologies within the HbbTV workflow is ongoing.

- **Automated HbbTV multilingual news broadcast subtitles in Switzerland**, which was planned at the time of writing the HBB4ALL proposal, will not be realized due to the lack of quality arising from the automatic translation.

- **Customised Subtitles for Online learning (MOOC) in Germany** is planned to go online for PC and mobile platforms at the end of 2015 / beginning of 2016, after final user tests and refinements.

As part of Pilot-A complementary user experience testing of related aspects of subtitling has been performed. This testing has addressed different aspects of subtitle personalisation and perception, such as subtitle usability across devices and in different languages/alphabets. In addition the reception of subtitles displayed on a second screen device (e.g. a smartphone) has been tested followed by a determination of which factors may need to be adapted in subtitles for small screen devices. Also the timing of subtitle presentation on TV screens has been evaluated with viewers. The combination of this information and the reports resulted in guidelines and recommendations with respect to appropriate and desirable options that should be offered for subtitle customisation; the recommendations are part of this document. Subsequent further user testing, in parallel to the sub-pilots, is planned.

Specifically challenges with respect to the standard conformant usage of the novel EBU-TT-D subtitle format play an important role (relevant for all companies who plan to migrate to using EBU-TT-D). The project partners have participated significantly in standardisation bodies to ensure that the proposed subtitle documents and strategies used in the test services are compliant and reflected within new standards. Also, to ensure interoperable EBU-TT-D implementations, the HBB4ALL partners have cooperated very closely, e.g. by exchanging EBU-TT-D files for mutual testing and by using the EBU-TT-D XSD for verification of the XML documents. Similarly, regarding the verification of the correct rendering of EBU-TT-D subtitles, common reference material from the EBU-TT-D rendering test suite is being used.

In summary, the core objectives of the initial tasks of Pilot-A (tasks 3.1 and 3.2), have been largely completed and in some instances have exceeded the objectives described in the Description of Work (DoW). Progress is as anticipated. During the following, operational phase (task 3.3), the activities of Pilot-A will focus on evaluating the established services in the respective sub-pilots with dedicated user groups.

1. Introduction

The Hybrid Broadcast Broadband for All project (HBB4ALL) investigates accessibility services in the hybrid broadcast-broadband TV (HbbTV) environment. One of the most prominent challenges faced by broadcasters is the requirement to add Access Services, in a cost-efficient manner, also to audio-visual content delivered via Internet while remaining consistent with the Access Services available on traditional broadcasts and their respective workflows. A new, additional challenge is the desire to offer viewers the opportunity to customise the Access Services they are using to best meet their personal preferences or needs.

HBB4ALL will evaluate access services in four interlinked Pilots; Pilot-A: Multi-platform subtitle workflow chain; Pilot-B: Alternative audio production and distribution; Pilot-C: Automatic User Interface adaptation – accessible Smart TV applications; Pilot-D: Sign-language translation service.

The operational phase of the HBB4ALL project (Task X.3 – Operation Phase for all Pilots A to D, running from August 2015 – July 2016) will gather user feedback and assess the acceptance and quality of services in various delivery scenarios implemented using field user tests and also in complimentary qualitative lab tests. A number of different sub-pilots is scheduled to be carried out in the operational phase. In the specific case of Workpackage 3 / Pilot-A, novel subtitle services will be prototypically rolled out in the different HBB4ALL countries in order to tackle the above challenges and receive and evaluate user feedback for later improvement.

1.1. Purpose of the document

The HBB4ALL deliverable D3.1 – Pilot-A Progress Report (see [2]) provided an overview of the progress of Pilot-A during the first 11 months of the project timeline. The current document gives an update to D3.1, presenting the status of the sub-pilot preparation.

First, the technology and infrastructure, which has been implemented in support of the operational phase, is presented: chapter 2 provides an overview of the prototype service components and their state of completion; chapter 3 documents the integration of components and preparation of the test beds as end-to-end workflows for the sub-pilots. As Pilot-A is quite extensive, chapters 2 and 3 present outline summaries – additional details are presented in Annex 8.1 and 8.2 respectively.

Chapter 4 presents the current status of the envisioned sub-pilot preparation with respect to organisational issues, technical and functional setup, timing, anticipated users / user groups and the evaluation approach. Finally, chapter 5 documents preliminary user tests, which focused on improving the quality of subtitling services, specifically analysing different aspects related to subtitle personalization and perception by end users.

1.2. Acronyms and abbreviations

In this document, when necessary, identified partners within the project are referred to using the abbreviated names initially defined within the Consortium Agreement for HBB4ALL and reproduced on the cover sheet of this document. Abbreviations and acronyms are introduced in brackets in the text after the corresponding full text version. The introduced abbreviation or acronym will then be used in subsequent prose in the document.

1.3. Definitions and glossary

This document draws upon the terminology and definitions established by previous research projects (in particular ICT PSP/2007/1.)

- **Access Service** [UK] = **Accessibility service** [US] The provision of additional services or enhancements that improve the accessibility of TV services for viewers with disabilities or special needs.

- **Accessibility** The degree to which a product, device, service, or environment is available to as many people as possible. Accessibility can be viewed as the "ability to access" and possible benefit of some system or entity. Accessibility is often used to focus on persons with disabilities or special needs and their right of access to entities, often through use of Assistive technology or Access Services.

- **Audio-visual Content** All kinds of time-based content consisting of images and sounds.

- **Captioning** (North America). See *Subtitling*, *Intra-lingual*. A form of subtitles primarily intended as an access service for viewers with hearing impairments. Captions not only display words as the textual equivalent of spoken dialogue or narration, but they may include speaker identification, sound effects, and music description. Captioning aims to include as much of the original language as possible (i.e. a verbatim transcription). However, altering the original transcription may be necessary to provide time for the caption to be read and for it to be in synchronization with the audio.

- **Catch-up TV** A service that allows a viewer to see a TV program independent of when it was broadcast. This is usually a kind of on-demand service on the Internet, but may also be achieved via a Personal Video Recorder (PVR) on which the viewer has chosen to record the program, or through a push Video On Demand (VOD) subscription where the viewer receives the program via the Internet, his Set-top Box (STB) or his PVR.

- **Content Delivery Network** (CDN) is a large distributed system of servers deployed in multiple data centres across the Internet. These servers cache and store the content from Internet content providers that contract this service to enhance the availability and performance of the delivery of content to end-users, while reducing demand on the content provider's own servers.

- **Content Management System** (CMS) is a computer application that allows publishing, editing and modifying content, organizing, deleting as well as maintenance from a central interface. Content management systems typically provide procedures to manage workflows in a collaborative environment (see also MAM).

- **Control, Remote** (Remote Control) is also known as a remote, controller or sometimes channel changer. It is an electronic device used for the remote operation of a viewing device (*television set, set-top box or PVR*) often over very short distances within the home. The design of remote controls needs to consider their usability and accessibility. Blind and partially sighted persons and those with other disabilities often encounter difficulties with remote controls that render them inaccessible and that impairs their ability to switch on or configure access services.

- **DVB** Digital Video Broadcasting a set of technical guidelines, standards and specifications to benefit and advance digital media markets world-wide. It was originally European in origin but today is an alliance of 250-300 companies.

- **DVB subtitles** – bitmap or Teletext captions on digital television using DVB.

- **EBU** European Broadcasting Union.

- **EBU-TT** EBU Timed Text is a new Subtitling Format specification (EBU Tech 3350). The format is based on the W3C Timed Text Markup Language (TTML) specification and provides an easy-to-use method to interchange and archive subtitles in XML.

- **EBU-TT-D** EBU-TT part 'D' (EBU Tech 3380) is the format for the distribution of EBU-TT subtitles over IP.

- **HbbTV** Hybrid Broadcast Broadband TV is a major pan-European initiative, building on work in the Open IPTV Forum, to formulate standards aimed at harmonizing the broadcast and broadband delivery of entertainment to the end consumer through connected TVs and set-top boxes.

- **HTML** HyperText Markup Language, commonly referred to as HTML, is the standard markup language used to create web pages.Web browsers can read HTML files and render them into visible or audible web pages. HTML describes the structure of a website semantically along with cues for presentation, making it a markup language, rather than a programming language.

- **HTTP** the Hypertext Transfer Protocol is an application protocol for distributed, collaborative, hypermedia information systems. HTTP is the foundation of data communication for the World Wide Web.

- **Impairment, age-related** is a collection of sensory and cognitive impairments. In the general sense, it covers matters such as the deterioration of sight and hearing, memory impairment or memory loss. In the report, we look not only at persons who are elderly but also at the challenges facing children whose intellectual maturity has an impact on their ability to read subtitles. In principle, there can be other impairments that are related to stages in the person‘s life.

- **Impairment, hearing** is a generic term including both deaf and hard of hearing which refers to persons with any type or degree of hearing loss that causes difficulty working in a traditional way. It can affect the whole range or only part of the auditory spectrum. [For speech perception, the important region is between 250 and 4,000 Hz.] The term ‘deaf’ is used to describe people with such profound hearing loss that they cannot benefit from amplification, while the term ‘hard of hearing’ is used for those with mild to severe hearing loss but who can benefit from amplification.

- **Impairment, visual**. Visual impairment (*or vision impairment*) is vision loss (*of a person*) to such a degree as to qualify as an additional support need through a significant limitation of visual capability resulting from either disease, trauma, or congenital or degenerative conditions that cannot be corrected by conventional means, such as refractive correction, medication, or surgery. The loss may cover visual acuity, significant central or peripheral field defects or reduced contrast sensitivity.

- **Internet Protocol Television, IPTV** is a system through which Internet television services are delivered using the architecture and networking methods of the Internet Protocol Suite over a packet-switched network infrastructure, e.g., the Internet.

- **Machine translation, MT** is an automatic translation of text from a source language to a target language by means of a computer programme.

- **Media Asset Management, MAM** systems are typically software systems to support management tasks and decisions surrounding the ingestion, annotation, cataloguing, storage, retrieval and distribution of audio, video and other media assets.

- **Metadata** is supplementary data about data, in this case information about television programs. E.g. This could be program listings or guides, or technical data delivered with the program to accomplish an access service.

- **Metric** is a criterion or measure of success in reaching a particular objective or goal.

- **Metric, Quality** is a measure of the perceived quality of a television picture or sound or associated service.

- **Multiplex or mux** is also called a virtual sub-channel in the United States and Canada, and Bouquet in France. It is a group of TV channels that are mixed together (*multiplexed*) for broadcast over a digital TV channel and separated out again (*de-multiplexed*) by the receiver.

- **Re-speaking** is a means to provide real-time intra-lingual subtitling for live events including television programs. It involves a captioner/subtitler re-speaking or dictating the captions, together with punctuation and other 'control words' that are transcribed using speech recognition software trained to the specific re-speaker's voice and that are automatically formatted for display.

- **Set-top box** is a device that enables an analogue television set to receive and decode digital television broadcasts.

- **Smartphone** is a mobile phone that offers more advanced computing ability and connectivity than a contemporary feature phone.

- **SMPTE** - The Society of Motion Picture and Television Engineers, SMPTE is a technical society for the motion imaging industry.

- **Spotting -** The offline determination of subtitle timing (i.e. when subtitles need to appear) using a proxy or copy of the associated video content.

- **Subtitling** is a generic term for the production of text as an alternative form of the audio content of Audio-visual Content. The term 'subtitling' is often interpreted as the process of converting the dialogue component of audio-visual content into text and displaying the text on the screen overlaid on the video image. [See also Captioning and Subtitling, Intra-lingual]. Translation subtitling, which involves a change in language between the spoken dialogue and the displayed text [See also Subtitling, Inter-lingual] aims to convey as much of the meaning of the original language as possible.

- **Subtitling, Closed** or User-selected Subtitles. The user has to select the service, c.f. Open Subtitling which is seen by all and cannot be turned on and off.

- **Subtitling, In-vision** is a synonym for Open Subtitling.

- **Subtitling, Inter-lingual** is also known as translation subtitling. The dialogue in the original audio content is translated into a different language in the text output.

- **Subtitling, Intra-lingual** see Captions. Also known as same-language subtitling, or subtitles for the deaf and hard-of-hearing (SDH).

- **Subtitling, Live.** Intra-lingual [most commonly] subtitles prepared at the moment of program broadcast or distribution. Usually live subtitles are created using stenography or re-speaking.

- **Subtitling, Open** Subtitling where the user does not have to do anything in order to see the subtitles, as they are an integral part of the picture.

- **Subtitling, Pre-prepared.** Subtitles prepared before the program is broadcast or distributed.

- **System Usability Scale** (SUS) provides a "quick and dirty", reliable tool for measuring usability. It consists of a 10 item questionnaire with five response options for respondents; from Strongly agree to Strongly disagree. Originally created by John Brooke in 1986, it allows you to evaluate a wide variety of products and services, including hardware, software, mobile devices, websites and applications[1].

- **Tablet or Tablet PC** is a device equipped with a touchscreen as the primary input device and designed for personal use.

- **Teletext** or broadcast Teletext is a television information retrieval service developed in the United Kingdom in the early 1970s. It offers a range of text-based information including closed subtitles and closed captioning. This service is typically available on page 888, but the actual page number depends on the broadcaster and country.

- **Transcription** is the representation of the dialogue and lyrics within the sound track of a TV program in written form. Written equivalents of sound effects (and song titles etc.) may also be included in a 'transcript' if the intention is to produce captions (intra-lingual subtitles).

- **Transcription, Verbatim** is a word-for-word representation of the dialogue and lyrics within the sound track of a TV program in written form.

- **Translation Subtitles** see Subtitling, Inter-lingual

- **TTML** Timed Text Markup Language is one of W3C's standards regulating timed text on the Internet. TTML is used in the television industry for the purpose of authoring, transcoding and exchanging timed text information and for delivering captions, subtitles, and other metadata for television material repurposed for the Web or, more generally, the Internet. There is partial and full

[1] http://www.usability.gov/how-to-and-tools/methods/system-usability-scale.html

support of TTML in components used by several Web browsers plugins, and in a number of caption authoring tools.

- **Vertical Blanking Interval**, (VBI) also known as the vertical interval or VBLANK, is the time difference between the last line of one frame or field of a raster display, and the beginning of the first line of the next frame. It is present in analogue television and can be used for data casting (*to carry digital data*), since nothing sent during the VBI is displayed on the screen; various test signals, time codes, closed captioning, teletext, CGMS-A copy-protection indicators, and other digital data can be sent during this time period.

- **World Wide Web Consortium, (W3C)** is an international standardisation organisation that develops Web standards to ensure the long-term growth of the Web.

- **XML** Extensible Markup Language is a markup language that defines a set of rules for encoding documents in a format which is both human-readable and machine-readable. It is defined by the W3C's XML 1.0 Specification and by several other related specifications, all of which are free open standards.

2. Service Components

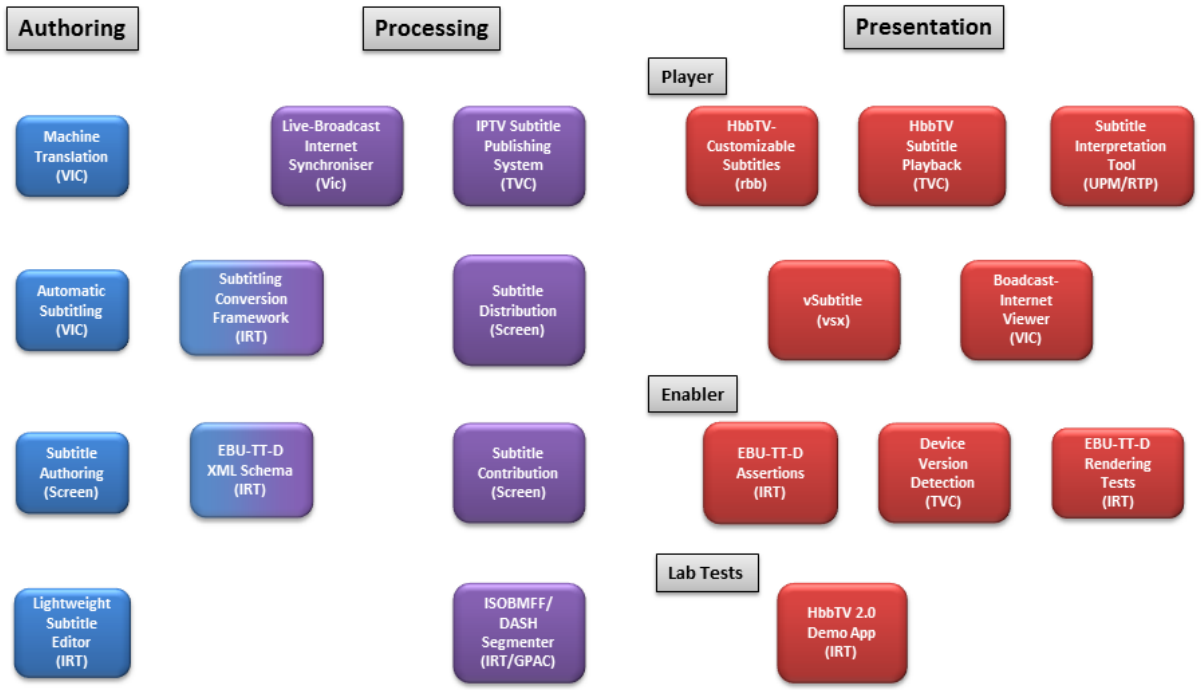

Figure 1. Overview of Pilot-A Service Components

Within Pilot-A a large number of service components are being developed and/or adapted for use in the respective sub-pilots. A graphical overview of the Service Components is shown in **Figure 1**. With respect to their role in a potential workflow, they are categorised as tools for Authoring / Production, Processing / Middleware and Presentation (where for the purpose of sub-categorisation some components are identified as part of the Player, some are an Enabler for the correct presentation of subtitles and one component is a lab test/showcase for HbbTV 2.0).

This chapter provides an overview of the prototype service components, their state of completion and the sub-pilots in which they are being used. Technical details can be found in Annex 8.1 (respective references are given).

2.1. Authoring and Production Components

2.1.1. Automatic Subtitling Component (VIC)

Summary

The automatic subtitling component generates intra-lingual subtitles in real-time from English ‘Breaking News’ content. The component is divided into five sub-modules: Audio Pre-Processing, Large Vocabulary Continuous Speech Recognition, Punctuation and Capitalization, English Sentences generation and EBU-TT-D Subtitle generation. The resulting generated English EBU-TT-D subtitles are integrated into the

HbbTV workflow and in addition provided to the *Machine Translation Component* for automated translation into Spanish.

This component is used and combined with other components in the Automated HbbTV multilingual news broadcast subtitles sub-pilot (see Section 3.6). A detailed description of the Automatic Subtitling Component can be found in Annex 8.1.1.

Status of completion

A. Development of all the modules

The development of the following modules has been finished and the modules have been integrated into the Automatic Subtitling system:

- Audio Pre-Processing
- Punctuation and Capitalization
- English Sentences
- EBU-TT-D Subtitles Generation

The following module will be further improved:

- Large Vocabulary Continuous Speech Recognition: An initial version of the recognition engine was integrated in the system. Currently, both Acoustic and Language models of the engine are being improved through new paradigms and techniques in order to achieve better recognition results. The improvement of the recognition engine includes advances in the techniques used for modelling (new Deep Neural Networks) and the inclusion of more training corpus adapted to the specific domain. The integration of new models will be done iteratively once the error rates have been improved. We expect to have the final models by March 2016.

B. Integration of the system in the HbbTV workflow and with the Machine Translation Component.

The Automatic Subtitling Component has been successfully integrated into the HbbTV workflow and the communication with Machine Translation Component was developed and resolved.

2.1.2. Machine Translation Component (VIC)

Summary

The Machine Translation Component is implemented as part of a Service Oriented Architecture (SOA). It receives as input the resulting English subtitles from the *Automatic Subtitling Component*, translates these into Spanish, generates EBU-TT-D subtitles and makes the subtitles available to the HbbTV workflow via an HTTP upload.

This component is used and combined with other components in the Automated HbbTV multilingual news broadcast subtitles sub-pilot (see Section 3.6). A detailed description of the Machine Translation Component can be found in Annex 8.1.2.

Status of completion

The component is fully functional and further work would only involve either model refinement for better translation quality, or fixing remaining processing issues that may surface during further functional testing.

2.1.3. Subtitling Format Conversion Framework (IRT)

Summary

The Subtitle Conversion Framework (SCF) uses open standard technologies to convert subtitles from legacy formats used in broadcast production into EBU-TT-D subtitles. To encourage a wide adoption, this framework has been published as Open Source (used in pilots Customised HbbTV Subtitles for VOD portal - Germany (see section 3.2), Customised Subtitles for Wide focus multi-platform – Portugal (see section 3.5) , and in pilot Customised HbbTV Subtitles for VOD portal - Spain (see section 3.3)). A detailed description of the SCF can be found in Annex 8.1.12.

Status of completion

Although the SCF was released in alpha status, it is already being used in production. The stable release is planned for spring 2016.

2.1.4. Subtitle Authoring Component (SCREEN)

Summary

Screen has developed an EBU-TT –D export function as an option in a subtitle preparation tool. This export functionality is currently available within a standard commercial product from Screen (WINCAPS QU4NTUM). This preparation tool is in use within the Testing Environment for End-To-End Subtitle Workflows described in section 3.1.

Screen has also developed a prototype renderer that decodes EBU-TT-D documents to investigate the feasibility of decoding EBU-TT-D documents within a pure browser environment (HTML 5 and JavaScript). It is intended that this prototype will be extended further to provide a preview capability for EBU-TT-D files and that it may be used within an authoring environment to give visual feedback as to the anticipated presentation of generated EBU-TT-D files on viewers devices. A more detailed description of the EBU-TT-D prototype renderer can be found in Annex 8.1.17.

Status of completion

The EBU-TT–D export function is currently available within a standard commercial product as a configurable option and is in use by two customers. The subtitle preparation tool is available to partners within the HbbTV project for use in project related activities.

The EBU-TT-D prototype renderer has successfully demonstrated the feasibility of JavaScript based decoding and rendering of simple EBU-TT-D documents within a pure browser environment. The prototype demonstrates strategies for converting EBU-TT_D syntax into HTML5 and CSS style data that can be rendered by a standard browser. JavaScript functions have been implemented to support line wrapping and text positioning. Further work is required to process complex EBU-TT-D documents, breaking a multiple subtitle file into Intermediate Synchronic Documents (each representing a single subtitle) and into overlaying

these documents onto a corresponding video playback at the appropriate presentation times. It is anticipated that this extended prototype will be made available to HBB4ALL partners within the project for use in project related activities by the end of 2015 following these further development efforts and improvements to the user interface.

2.1.5. Lightweight Subtitle Editor(IRT)

Summary

The Lightweight Subtitle Editor is a low-cost XML editor for EBU-TT-D subtitles that is customized for the re-purposing of subtitles for Internet distribution. It is built on the XML editor oXygen. It is planned to have this authoring tool used and tested by multiple HBB4ALL partners. A detailed description of the Lightweight Subtitle Editor can be found in Annex 8.1.13.

Status of completion

The technology to implement the editor became available after the project had already started. A first prototype is expected by the end of 2015 and further improvements should be available in spring 2016.

2.1.6. EBU-TT-D XML Schema (IRT)

Summary

The EBU-TT-D XML Schema provides the possibility to validate EBU-TT-D XML documents against the specification EBU-TT-D (EBU Tech 3380 [6]). The work was contributed to the EBU and will be further maintained by IRT as part of the EBU standardization work. The schema may be used by all HBB4ALL partners who are using EBU-TT-D subtitles in their respective tool and/or sub-pilot (currently already in use by TVC and VIC). A detailed description of the EBU-TT-D XML Schema can be found in Annex 8.1.14.

Status of completion

The XML Schema is currently published as alpha release and is planned to be finalized in 2016.

2.2. Processing and Middleware Components

2.2.1. Subtitle Contribution Component (SCREEN)

Summary

Screen is investigating the transport of EBU-TT data in the VBI / VANC spaces of an SDI video signal. A data insertion component for use within a commercially available Screen subtitle transmission system is under development that receives EBU-TT-D documents over an IP connection and converts the EBU-TT data into a form that can be opaquely tunnelled within VBI in SDTV or VANC in HDTV.

Screen is developing a second subtitle data to EBU-TT conversion component for use within a commercially available Screen subtitle transmission system (Polistream Black) for the real-time conversion of (live) subtitle content in proprietary formats into EBU-TT-D format for onward delivery to HbbTV distribution systems. This component will receive subtitle information (in a Screen internal format) over an IP connection and convert the subtitle data into an EBU-TT document that can be forwarded again over IP.

Additional Screen components already exist within the Screen subtitle transmission system that can accept other forms of subtitle data and convert the subtitles to the Screen internal format, thus the conversion of a wide range of other formats of subtitle data is possible by using these two components in sequence.

For clarification, this component differs from the Subtitle Conversion Framework (SCF) developed by IRT within this project. The SCF component is designed for the off-line conversion of subtitle files (each containing all the subtitles for complete programs). The Subtitle Contribution Component from SCREEN converts individual subtitles in real time as they are processed and transmitted by the subtitle transmission system. This allows any subtitles handled by the transmission system to be converted and made available to any connected HbbTV service (including live subtitles supplied by steno or re-speaking systems). A more detailed description of the Subtitle Contribution Components can be found in Annex 8.1.18.

Screen is also researching the real time delivery of subtitle synchronization information and subtitle content to MPEG2-TS multiplexers from a Screen subtitle transmission system, with a specific focus on the delivery of EBU-TT documents to MPEG-DASH systems.

Status of completion

The EBU-TT data insertion component that inserts data into the VBI / VANC spaces of an SDI video signal is in early prototype form awaiting completion of the live Subtitle Contribution Component for testing. A set of standard components for data extraction of the EBU-TT data from the VANC is already available within the Screen subtitle transmission system (these do not need to be EBU-TT specific).

The subtitle data to EBU-TT conversion component is under development and the first prototype is anticipated by early Q4 2015. This conversion component will be capable of supporting both live subtitling trials (where subtitles are generated using the audio of a live broadcast) and for trials involving pre-prepared subtitle files in any file format supported by a Screen subtitle transmission system (where the prepared subtitles are sequentially fed through the conversion component by timed release).

2.2.2. Subtitle Distribution Component (SCREEN)

Summary

Screen is developing functionality within a Screen data conversion tool that will support the conversion of EBU-TT files to EBU-TT-D files. The design for this conversion capability is split into two components, one component that can read EBU-TT format data and a second component that can write EBU-TT-D data. As this functionality has been split into two parts, and these parts are implemented within a commercially available Screen data conversion tool, it is possible to convert any supported format of subtitle file into an EBU-TT format file and vice versa.

These components will be used to test the conversion pathways between legacy subtitle formats and EBU-TT format files, with particular focus on the mapping of subtitle text, timing, style and position between formats and any losses in conveyed information that result from a mismatch in the support for various subtitling features in legacy and EBU-TT formats. A more detailed description of the Subtitle Distribution Component can be found in Annex 8.1.19.

Status of completion

The conversion component for writing EBU-TT-D data is complete and is being used by a broadcaster (not a project partner). A number of minor issues have been identified and a program of further development is underway. The component for reading EBU-TT data is in development and is anticipated for prototype release by end of Q3 2015.

2.2.3. Live Broadcast-Internet Subtitle Synchroniser (VIC)

Summary

This server-based component synchronises the subtitles created through the *Automatic Subtitling Component* and the *Machine Translation Component* to the A/V input of the DVB Transport Stream, which was used as audio input from the *Automatic Subtitling Component*. The generated output includes Audio, Video and Subtitle Segments according to the MPEG-DASH specifications for "adaptable" broadband distribution.

This component is used and combined with other components in the Automated HbbTV multilingual news broadcast subtitles sub-pilot (see section 3.6). A detailed description of the Live Broadcast-Internet synchroniser can be found in Annex 8.1.3.

Status of completion

The Live Broadcast-Internet synchroniser component is fully functional and its implementation has been finalised. The integration in the whole pipeline system creating a live streaming service from a DVB signal including automatically generated subtitles provided by the Automatic Subtitling and the Machine Translation components has been successful.

2.2.4. IPTV Subtitle Publishing System (TVC)

Summary

This component makes subtitles available in EBU-TT-D format for all IP-based systems. It is an engine that fulfils the specific requirements regarding publishing for each particular platform, maximizing the functional reusability.

This component is closely related with the Catalan sub-pilot “Customised HbbTV Subtitles for VoD portal” (see section 3.3). Additional details of the IPTV subtitle publishing system can be found in Annex 8.1.7.

Status of completion

A first version of this component has been developed. With it, TVC can access the XML files containing the generated EBU-TT-D subtitles and their characteristics and descriptions, such as timestamps for each subtitle. As the development of the customisation component is still ongoing, the subtitles published currently are with colors and sizes pre-defined in the XML file, without personalisation available. It is anticipated that the option of personalised subtitle customisation will be added into the app by November 2015.

2.2.5. HbbTV MPEG-DASH/ISOBMFF Segmenter and Packager (IRT)

Summary

This component is an extension to the open source software MP4box, which was developed by GPAC under subcontract to IRT. As a result, MP4box now directly supports generation of MPEG-DASH streams compliant with the MPEG-DASH profiles defined in HbbTV 1.5 and 2.0.

This component may be used for any sub-pilot in which HbbTV devices are used; specifically it has been used for the realization of the HbbTV 2.0 Demo App (see Annex 8.1.16). Additional details of the HbbTV MPEG-DASH/ISOBMFF Segmenter and Packager can be found in Annex 8.1.11.

Status of completion

The component HbbTV MPEG-DASH/ISOBMFF Segmenter and Packager has been completed.

2.3. Presentation Components (User Application and Rendering)

2.3.1. Live Broadcast-Internet Subtitle Viewer (VIC)

Summary

The Live Broadcast-Internet Subtitle Viewer receives the MPEG-DASH segments created from the *Live Broadcast Subtitle Synchronizer*, un-packs the EBU-TT-D subtitles from the ISOBMFF packaging format and provides them as WebVTT subtitle cues to the HTML 5 browser for rendering.

This component is used and combined with other components in the Automated HbbTV multilingual news broadcast subtitles sub-pilot (see section 3.6). Additional details of the Live Broadcast-Internet Viewer can be found in Annex 8.1.4.

Status of completion

The adoption/parsing of EBU-TT-D formatted subtitles encapsulated in ISOBMFF fragments is finished and has been successfully integrated in the whole system.

2.3.2. Subtitle Interpretation Tool for Browsers (UPM)

Summary

The Subtitle Interpretation Tool for Browsers builds upon the platform independent plugin JW Player. It enriches this very widespread Video player with the functionality to render EBU-TT-D subtitles.

This services component is focused on the Portuguese sub-pilot, entitled "Customised subtitles for wide focus multi-platform" as described in sections 3.5, 4.4 and 8.2.3 of this deliverable. A detailed description of the Subtitle Interpretation Tool for Browsers can be found in Annex 8.1.5.

Status of completion

The start of a new development to integrate the subtitling tool as a plugin to the JW Player has caused a delay from the initial planned date for completion. Currently, the tool is able to recognize and interpret the EBU-TT-D format. Further work is planned to integrate the JavaScript code in JW Player as a plugin, avoiding other subtitling mechanism supported by JW Player, which are not based on EBU-TT-D and which do not support styles. The next step will be the integration of the plugin in RTPplay (September/October 2015), the JW Player deployed by RTP.

2.3.3. HbbTV-Based Customisable Subtitles for VoD Playback (RBB)

Summary

These components implement customisable subtitles for the VoD playback in the German HbbTV Mediathek (catch-up TV) service. The components are divided into a server-based EBU-TT-D subtitles parser module, the enhancement of the existing Mediathek video player to support customisable subtitles and the adaptation of the HbbTV Mediathek graphical user interface (GUI).

The service component HbbTV-Based Customisable Subtitle for VoD Playback is used in the German sub-pilot (section 3.2), in combination with the *Subtitling Format Conversion Framework* described in section 2.1.3. Additional details of this component can be found in Annex 8.1.6.

Status of completion

The service component HbbTV-Based Customisable Subtitle for VoD Playback was rolled out for the German sub-pilot on 14 April 2015. Since then it has been openly available in the RBB Mediathek application on RBB TV and in the ARD Mediathek application since 12 May 2015, see section 3.2. In addition to allowing the selecting and de-selecting of subtitles, it has also integrated the functionalities of allowing a viewer to adapt the font size and to adapt the position of subtitles for display on the screen. Subtitles are rendered with an underlying 67% opaque black box. The user adaptation of subtitle background (outline, box etc.) will be integrated at a later stage. Although the service has been available to users of the aforementioned application it has not been communicated or promoted, to allow testing. In September 2015 RBB will start to publically promote the service.

2.3.4. HbbTV Device and Version Detection (TVC)

Summary

This JavaScript based component detects the specific device (e.g. manufacturer, type number) and the HbbTV version it implements.

This component will be widely used in the Catalan sub-pilot (see section 3.3). As this sub-pilot will be performed integrating this HbbTV subtitle service in TVCs VoD service, the possible test users are all the ones in Catalonia that have a suitable HbbTV device. In principle we won't take much advantage of this component, as from the results obtained in the tests performed in our labs all the TVs can correctly support the subtitling features used. At some point, however, when there are more HbbTV 2.0 devices in the market, the version detection could be useful for detecting the devices with that higher version of HbbTV implemented, and would allow the users to decide if they want to use our subtitle personalization service

with our player or if they prefer to let the HbbTV 2.0 device player to take care of the subtitling. Additional details of the component HbbTV device and version detection can be found in Annex 8.1.8.

Status of completion

A JavaScript component has been developed which takes the user-agent of an HbbTV client device and from that detects the device and the HbbTV version implemented on it. This component ensures compatibility of television with the application.

2.3.5. EBU-TT-D HbbTV Personalised Subtitle Playback (TVC)

Summary

This component allows subtitle customisation by the end user on his HbbTV device. The application "3 a la carta" is HbbTV 1.0 compliant, using the television's native HbbTV player for video and this JavaScript component for subtitle playback. The subtitles are provided and rendered in the EBU-TT-D format. By default, the subtitles initially presented to the viewers are the non-personalized "basic" subtitles. As no MPEG-DASH transport is used, the EBU-TT-D subtitle file is transported independently ('out-of-band') of the MPEG-4 A/V file. This independent transport of the subtitles implies some synchronization problems related to the display of the subtitles with the video. This issue is solved using JavaScript time-based synching playback.

This component is the primary key of the Catalan sub-pilot "Customised HbbTV Subtitles for VOD portal" (see section 3.3 and 8.2.2), and will be widely used during its lifetime. Additional information of the component EBU-TT HbbTV Personalised Subtitle Playback can be found in Annex 8.1.9.

Status of completion

The initial phase of the development had as its objective the display of baseline non-customizable subtitles, while a following phase featured storing custom subtitle settings for each user. TVC has successfully completed the first phase, accomplishing the correct synchronization and a presentation of the basic subtitles for VoD content, without any customisation.

TVC is still progressing the use of JavaScript to develop a customisation component that will allow end-user personalization of the subtitles and the recovery of the end-user's default settings using cookies. The customisation component itself, the primary objective of this system, is currently in development and planned to be available by November 2015.

2.3.6. vSubtitle: Adaptive Subtitle Rendering (VSX)

Summary

The vSubtitle component was developed within HBB4ALL as a standard add-on for the Vsonix vPlayer, a webcast player that was designed for the playback of video lectures e.g. for the realization of MOOCs or to be used in corporate learning platforms. The component handles synchronization, rendering and adaptation of subtitles and allows users to adapt the subtitle display according to their user profile. The component will be used in the pilot "Customised Subtitles for Online learning (MOOC)" (see sections 3.8 and 8.2.5). A detailed description of vSubtitle can be found in Annex 8.1.10.

Status of completion

A first version of the vSubtitle component has been implemented and integrated with the UI adaptation service.

2.3.7. EBU-TT-D Rendering Tests (IRT)

Summary

Reference material is being provided by IRT for the use of HBB4ALL partners and manufacturers to facilitate a standard conformant implementation of EBU-TT-D. All partners may use the material to verify their EBU-TT-D implementations (RBB and UPM already have the material in use). A detailed description of the EBU-TT-D Rendering Tests can be found in Annex 8.1.15.

Status of completion

A first set of test material was provided to partners and manufacturers. It will be completed by 2016.

2.3.8. EBU-TT-D Assertions (IRT)

Summary

HbbTV 2.0 assertions build the basis for a test suite to check conformance against the HbbTV 2.0 specification. IRT has provided assertions for EBU-TT-D implementation and adaptation in HbbTV 2.0. These assertions define the expected rendering given EBU-TT-D features and provide the necessary references to the specification. Also, the assertions have been used by IRT in combination with the EBU-TT-D Rendering Tests (see Annex 8.1.15) and for the standard conformant development of the HbbTV 2.0 Demo App (see Annex 8.1.16).

Status of completion

The work on the HBBTV assertions has been completed.

2.3.9. HbbTV 2.0 Demo App (IRT)

Summary

This component implements an HbbTV 2.0 application, supporting live streaming on the Internet, using MPEG-DASH streaming and EBU-TT-D subtitles. This component will specifically be used for the presentation of showcases at IFA and IBC 2015 (by IRT and RBB, in close cooperation with industry partners). Additional details of the HbbTV 2.0 Demo App can be found in Annex 8.1.16).

Status of completion

At the time of writing, the component HbbTV 2.0 Demo App is being completed.

3. Sub-Pilot integration

This chapter provides an outline of the sub-pilots end-to-end workflows planned for Pilot-A, with respect to the following parts in the chain: Authoring and Production, Contribution and Publishing and User Application and Rendering. Detailed descriptions, including pictures of the workflow parts as they are planned for use in the sub-pilots may be found in Annex 8.2 (respective references are given).

3.1. Testing Environment for End-To-End Subtitle Workflows (IRT)

The components and services developed in the HBB4ALL project often require subtitles in the EBU-TT-D format, whereas in traditional TV production environments, subtitles are provided in EBU STL (pre-prepared), Nufor (live) or proprietary formats (for both pre-prepared and live). Thus it is important to understand and test how HBB4ALL components can be integrated into an existing system, while still maintaining all existing services like Teletext and DVB subtitles. For example, integration can be realized by applying subtitle format conversion at different points in the system so that new services can access all information that is provided by EBU-TT while existing services can still be fed with their compatible formats.

This testing environment focuses on the aspect of integrating new subtitle workflows into existing production environments. It comprises a complete end-to-end subtitle scenario distributing Teletext and DVB subtitles within a valid MPEG-Transport Stream. The setup reproduces a typical broadcaster's production environment as might exist within many German public broadcasters.

This environment forms the basis for testing the integration of new distribution channels as well as other HBB4ALL components that potentially can be integrated into a running broadcasters system. One candidate for test implementations is a migration to EBU-TT as the master and editing format for subtitles. Another candidate is the MPEG-DASH distribution as provided by the *Live Broadcast-Internet subtitle synchroniser*.

Authoring and Production

Authoring will be done using the Screen subtitle preparation software (WINCAPS QU4NTUM). For pre-prepared content exported STL files will additionally be transformed into EBU-TT using the *Subtitle Format Conversion Framework*.

All existing subtitle services (live and pre-prepared) are then managed by the Screen subtitle transmission system (Polistream Black). That includes cueing pre-prepared subtitles, embedding subtitles into HD-SDI and the rendering of DVB subtitles.

Contribution and Publishing

For the traditional broadcast signal, a modulated DVB-T output will be generated carrying an MPEG Transport Stream including Teletext and DVB subtitles. Additionally, web encoders will create a stream to test the playout for new subtitle services (used for the development of the *HbbTV 2.0 Demo App*, see section 8.1.16). For on demand HbbTV services the Conversion Tool developed in HBB4ALL will also be used to create EBU-TT subtitle files.

Reception and User Application

For reception of the broadcast TV production a standard Set Top Box is used. Additional to these well-established user devices, the test environment is used for the development of the *HbbTV 2.0 Demo App* (see section 8.1.16).

A detailed description of the Testing Environment can be found in Annex 8.2.6.

3.2. Customised HbbTV Subtitles for VOD portal - Germany (RBB)

Authoring and Production

At RBB the subtitle editor works with current subtitle preparation software. To produce the EBU-TT-D file the subtitler has to follow a twostep approach:

1) First the file produced for broadcast is exported in EBU STL format.

2) Second the exported EBU-STL file is converted into an EBU-TT-D format file (through the use of the *Subtitling Conversion Framework*).

The created file will be used for the HbbTV VoD application and for VoD application portals for PC and mobile devices alike.

The additional time for the subtitler to create a file in the correct format is minimal (possibly 1-2 minutes per file).

Contribution and Publishing

The subtitler uploads the created EBU-TT-D file to the existing Content Management System (CMS) used by RBB. This upload feature and an automated linking between video asset and the subtitle file were implemented. This integration work was very specific to the existing infrastructure at RBB and the other public regional broadcasters RBB cooperates with through the ARD.

User application and Rendering

The customized subtitle feature was developed as an extension of the HbbTV client application for the RBB VoD portal. For the integration the HbbTV client has to request the relevant subtitle information. Although the sever uses EBU-TT-D as the file format for the subtitles, the exchanged information between the web server and the HbbTV application is not an EBU-TT-D XML file but instead the subtitle information is extracted from the XML file and sent as a JSON data structure.

A detailed description of this sub-pilot's integration can be found in Annex 8.2.1.

3.3. Customised HbbTV Subtitles for VOD portal - Spain (TVC)

Authoring & Production

For the production of broadcast subtitles TVC uses both manual and automated strategies. For both of these ‘creation modes’ a server based solution handles the transformation into the different subtitle formats

required. The created format from this solution depends on the distribution channel. To create EBU-TT-D files an additional feature had to be implemented for this server and the production is now fully automated. Through this integration a subtitler does not need any additional time to support the provision of subtitle files for the HbbTV VoD application.

Contribution and Publishing

TVC integrated the publishing of the subtitle information in several different existing management systems (Product Management, Media Asset Management and Programming System). Furthermore the communication to the server based solution that creates the subtitle file was implemented. Through an intelligent automation mechanism the need for manual interaction was minimized.

User Application and Rendering

TVC took a general approach to align the server API's that deliver the subtitle information to VoD clients on different platforms. The HbbTV VoD application is seen as one of them. A JavaScript plugin was implemented for the VoD application to render the EBU-TT-D subtitles. It accepts the EBU-TT-D XML file and translates the relevant information into HTML before rendering it. As EBU-TT-D (and the TTML language that EBU-TT-D is based on) has a very similar syntax to HTML the transformation is very straightforward.

A detailed description of this sub-pilot's integration can be found in Annex 8.2.2.

3.4. Customised HbbTV Subtitles for VOD portal - Switzerland (TXT)

Authoring and Production

No specialised authoring or production tool is being used for the generation of subtitles for the HbbTV case. The existing broadcast subtitles are being used as they are.

Contribution and Publishing

At broadcast the subtitles are scanned in a fully automated process, without human intervention. For use in the HbbTV application, the subtitle files are converted from TTML into JSON which can be processed easier for HbbTV purposes.

User Application and Rendering

The HbbTV application was extended with a subtitling feature which was developed for the player component. With this extension the player allows the display or hiding of the subtitles in video content from the SRG SSR[2] VoD portal on HbbTV client devices.

[2] The Swiss Broadcasting Corporation, owner of SWISS TXT

3.5. Customised Subtitles for Wide focus multi-platform – Portugal (RTP/UPM)

Authoring and Production

RTP has fully automated the production for the EBU-TT-D subtitles by using the *Subtitling Conversion Framework (SCF)*. The subtitler uploads a file he created for broadcast production to a database. This is the common workflow for broadcast subtitle production. The subtitles are then automatically converted through a scheduled task.

Contribution and Publishing

RTP implemented an automated mechanism that schedules tasks for conversion, contribution and publishing of subtitle files. Through this mechanism no additional manual time resource is needed for the addition of the EBU-TT-D subtitle files.

User Application and Rendering

RTP decided to extend the player that it currently uses for VoD content for PC and mobile. UPM is implementing the subtitle plugin for that player. The player used is the JW player and one of the most widespread players on the web. The plan is the integration work can be adopted not only by RTP but also by other companies.

A detailed description of this sub-pilot's integration can be found in Annex 8.2.3.

3.6. Automated HbbTV multilingual news broadcast subtitles- Spain (VIC/UAB)

Authoring and Production

The *Machine Translation Component*, *Automatic Subtitling component* and the *Live Broadcast-Internet subtitle synchroniser component* take a real-time audio/video broadcast stream as input and create a stream of subtitles in the EBU-TT-D format. The complete creation process is fully automated and no manual intervention is needed.

Contribution and Publishing

The *Machine Translation Component* and the *Automatic Subtitling Component* are contributing the created subtitle stream to the *Live Broadcast-Internet subtitle synchroniser*. This component synchronises the subtitle stream with the A/V Content and creates an MPEG-DASH Stream containing Audio, Video and Subtitles.

User Application and Rendering

VIC integrated the rendering component in the Open Source JavaScript based DASH IF player[3]. This integration work allows the player to now receive an MPEG-DASH stream and supports client-based rendering and synchronisation of EBU-TT-D subtitles from that MPEG-DASH Stream.

A detailed description of this sub-pilot's integration can be found in Annex 8.2.4.

[3] http://dashif.org/software/

3.7. Automated HbbTV multilingual news broadcast subtitles - Switzerland (TXT)

At the time of writing the HBB4ALL proposal, Swiss TXT had planned to implement an automated generation of HbbTV multilingual new broadcast subtitles. However, due to a lack of quality it was decided not to implement an automated translation. Therefore, this feature will not be realised for a sub-pilot by Swiss TXT.

3.8. Customised Subtitles for Online learning (MOOC) - Germany (VSX)

Authoring and Production

VSX uses the open source tool 'Subtitle Edit' for the manual creation of their subtitle files in the WebVTT format. As the sub-pilot targets content that is only distributed over broadband no conversion from previous created broadcast subtitles is necessary.

Contribution and Publishing

VSX integrated the subtitle production with their internal CMS and CDN architecture.

Reception & User Application

VSX integrated the rendering component of their player application with the AccessGuide application. Through this integration work it is now possible for users to store their preferred style settings for subtitles for future viewing sessions.

A detailed description of this sub-pilot's integration can be found in Annex 8.2.5.

4. Service pilots

The HBB4ALL sub-pilots are scheduled to be carried out within the period of August 2015 – July 2016. This section presents the status of each sub-pilot preparation with respect to organisational preparation, technical and functional setup, timing, envisioned users / user groups and the evaluation approach.

4.1. Customised HbbTV Subtitles for VOD portal - Germany

4.1.1. Status of work

The complete subtitle production workflow is ready and set-up as described in section 3.2, and currently processes approximately 5 video items per week. Subtitles are converted into EBU-TT, published and delivered to the end user terminals for a synchronised display.

At the time of writing, two out of three desired subtitle adaptation aspects have been implemented: font size and text positioning. The sub-pilot will run with these features. The remaining feature for implementation is the adaptable subtitle background, i.e. the way a subtitle text silhouettes against the underlying TV picture. RBB plans to add an outline option to the currently implemented box option, as well as an option to scale-down the TV-picture with the display of the subtitles pushed below that picture, so that no occlusion of the video can occur.

4.1.2. Timeframe

In September 2015 (calendar week 39) RBB plans to start its dedicated tests with a panel of users. The users will be requested to test the service for a period of approximately 12 weeks. After the end of the sub-pilot, and following the evaluation, RBB plans to add the remaining features and GUI adaptations according to the findings of the sub-pilot.

4.1.3. Test users and evaluation approach

The German sub-pilot can potentially target ~6 million households with HbbTV-enabled devices in Berlin-Brandenburg area and beyond in Germany for a quantitative analysis. For a qualitative validation, RBB will again work closely with organisations for the deaf and hard-of-hearing in Berlin and Brandenburg. 30 users from the target group will be chosen for an evaluation of the sub-pilot. It is required that all the 30 test users must have an HbbTV-enabled device ready for the sub-pilot phase. RBB aims at organising a closing workshop at the end of the operational phase with selected test users.

To monitor user experience, the HBB4ALL consortium has agreed upon the System Usability Scale (SUS) as a standardized instrument to measure usability as a core part of user experience. In the RBB context, it will be used to assess the overall system usability of the subtitle applications. The survey is limited to a maximum of 3 rounds in order to assess both novices to intermediate users of the application, while avoiding the risk of repetition when people have to fill the same survey too often.

The field pilots shall also be used to gain insights into possible system improvements and new features. In order to gather related data, we will follow an approach combining a visual logging technique with a follow-up co-creation workshop.

Therefore RBB has prepared a two-fold qualitative evaluation approach, utilising an adapted SUS questionnaire complemented by a visual logging strand.

a) The SUS questionnaire will be provided to the test users at the beginning and end of the test phase. It comprises of ten standard assertions to be rated, targeting the overall usability of the system.

b) During the pilot run time, the participants will get a form sheet with key screenshots of the application, where they can mark and comment any issues encountered during the pilot period. As a large part of the target group is communicating by fax, these forms can be sent as support requests during the pilot or compiled at the end of the pilot period. Based on these visual issue reports the RBB team will conduct a final co-creation workshop where they will prepare new solution prototypes and paper prototype building blocks, addressing the prevailing issues from the pilot. During the workshop the team will work with the participants in small groups to ideate, visualize and prioritize possible solutions.

Optionally, a mid-term SUS survey may be enhanced with a standardized survey covering user experience aspects beyond usability, such as attractiveness and hedonic quality.

As an additional measure, the development of the SUS is triangulated with actual usage numbers acquired through quantitative data.

As well as qualitative tests, RBB will gather and analyse quantitative data. Of general interest will be how many times a video was played with subtitles switched on, compared to without subtitles. To gather this data RBB will use the standard system within the ARD for quantitative tracking and analysis of online and HbbTV services.

In addition to the end user tests RBB will conduct a series of interviews with the editors and managers involved in the subtitle production and provisioning process to gather feedback on aspects of the workflow and costs involved.

4.1.4. Challenges, problems and potential solutions

As the service is already up and running no new technical problems are anticipated. Any foreseen problems are more of an organizational nature. RBB aims to recruit a demographically representative sample of testers. However as the testers need to have an HbbTV device this many prove difficult. If this is the case RBB will look into the option of providing testers with equipment and adjusting the sample group demographics.

A further potential risk is that the test users do not understand the application or the methodology. For both these issues RBB is being advised by local user associations thus minimizing this risk.

4.2. Customised HbbTV Subtitles for VOD portal - Spain

4.2.1. Status of work

TVC has working developments of the different components needed to perform this sub-pilot. The *IPTV Subtitle Publishing System* and *HbbTV Device and Version Detection Components* have a finalised version which can execute the desired functions, and which are still to be integrated into our workflow. For the *EBU-TT-D HbbTV Personalized Subtitle Playback System*, as explained before, various phases of development were defined. The initial phase aiming at showing baseline non-customizable subtitles for VOD contents has

been accomplished. The next phase of development is about reading and applying different options present in the EBU-TT-D file, e.g. display regions and colours, and the final phase has the objective of including user customization options for the subtitles. These two phases are just being started and have to be further developed.

TVC has generated EBU-TT-D subtitles for some VoD content. These subtitles have been successfully tested in TVCs labs. TVC has analysed the playback of EBU-TT-D format subtitles in the TVC HbbTV video player. TVC is also working with an in-app event marking concept (using Adobe Omniture – now part of Adobe Marketing Cloud) to be used as a quantitative validation method.

4.2.2. Timeframe

During **June 2015 - TVC** have fulfilled the first phase, which just offer subtitles in white over a transparent background, at the bottom of the image.

Anticipating continuous integration activity, the present roadmap plans to strengthen previous releases and fix possible issues, if any, from production deployment feedback. Finally, TVC aims to meet the following targets on the dates listed:

- **July 2015** for phase 2. To launch subtitles parsing and rendering display regions and colour as specified in the EBU-TT-D file.
- **November 2015** for phase 3. Release user customisable subtitles upgrade.
- **December 2015** User test with UAB completed
- **January to June 2016** Carrying out of sub-pilot.

4.2.3. Test users and evaluation approach

For the customised subtitle service pilot it is planned that the possible users were any user in Catalonia, who has an HbbTV device which can run the desired functionalities of subtitle personalization of TVCs client application in a correct way. For this service pilot, the evaluation will be done measuring and analysing the audience activity on that HbbTV application (using Adobe Omniture).

Among these test users, only the ones who actually benefit from the subtitle features would be the potential final users of the personalised subtitles service, the deaf or hard-of-hearing group. Additionally, tests with participants from this specific user group will be performed under lab conditions by UAB, involving 30 Spanish users. These tests will be evaluated by end-user questionnaires and interviews.

4.2.4. Challenges, problems and potential solutions

There are several challenges involved in the performance of the customized HbbTV subtitles for the VoD portal pilot. These can be caused by both the subtitles format required and personalisation difficulties.

The EBU-TT-D subtitle format is not supported by versions 1.1 and 1.5 of the HbbTV specification. The last HbbTV 2.0 standard has native support for subtitling services and recommends the usage of EBU-TT-D. This affects the vast majority of devices on the market, which run mostly HbbTV 1.1/1.5. For that reason, to accomplish the display of EBU-TT-D subtitles on HbbTV devices currently available in the market, it was

necessary to develop a playback system based on an independent player, not the HbbTV device's 'built in' one.

The TVC website has a license to use a commercial player, JW Player v.6, which supports the rendering of closed captions or subtitles in the video display. Initially, on previous versions this player supported TTML, but regrettably it is not EBU-TT-D compatible, even though EBU-TT-D is a TTML subset. Also, the roadmap for JW Player v.6 focusses on WebVTT (a plain text format that's part of the HTML5 standard), and SRT (a plain text format that's popular amongst file sharers). So, it does not seem that this solution can be exploited on the corporate website as planned.

Also, the HbbTV 2.0 subtitling services limitations are currently unknown, as there is no end-to-end implementation of EBU-TT-D based subtitling on a hybrid HbbTV distribution system. Thus, the subtitle producers cannot know what constraints they may need to take into account to make the generated subtitles suitable for HbbTV distribution. There is also a lack of knowledge on the customisation options that will be available on that HbbTV 2.0 subtitling services. This gives more relevance to the development of TVC's own system to perform the subtitle personalisation, covering the different customisation options that we think it may require.

The possible adaptation options of the subtitle content to meet the requirements of each user that needs this service are a challenge too, and may imply significant changes on the subtitle production process.

4.3. Customised HbbTV Subtitles for VOD portal - Switzerland

4.3.1. Status of work

TXT had already implemented the subtitles in the VoD services of the Italian language HbbTV service offering in Switzerland (RSI+) by December 2014. Meanwhile TXT has also implemented the subtitles in the French language for the HbbTV VoD service (RTS+), beginning in May 2015.

4.3.2. Timeframe

TXT plans to realise subtitles for the German language HbbTV VoD service in Switzerland (SRF+) by the first half of 2016.

4.3.3. Test users and evaluation approach

No user survey is planned for this sub-pilot.

4.3.4. Challenges, problems and potential solutions

HbbTV 1.1 does not natively support the rendering of any subtitle format. Therefore, TXT has created an API to convert the TTAF[4] subtitle files into a JSON representation. To verify the structure and content, the data will be validated by the *EBU-TT-D XML Schema* during the conversion process.

[4] Timed Text Authoring Format

4.4. Customised Subtitles for Wide focus multi-platform – Portugal

4.4.1. Status of work

The technical component is being developed at the time of writing this deliverable. A first functional version has been validated with RTP content and with content provided by other HBB4ALL partners, including EBU-TT-D subtitles. UPM is currently working on the interpretation of some subtitle characteristics (e.g., alignment). The start of the operational phase will require the integration of the component in the player used by RTP in its website (RTP Play). The service pilot will be available for any user using this player via RTP's website[5]. Moreover, the service pilot will be validated by Portuguese user group associations contacted by RTP.

4.4.2. Timeframe

This is the currently considered timeframe for this sub-pilot:

- September 2015: the required component (plugin for JW Player) is complete and ready to use, including customisation features (UPM)
- September 2015 / October 2015: the component / plugin is integrated in RTP Play, the player deployed by RTP in its website (based on JW Player) (RTP)
- End of October 2015/November 2015: the component is integrated and the operational phase starts.

4.4.3. Test users and evaluation approach

4.4.3.1. Lab Tests

By the end of September 2015, after the integration of the plug-in in RTP Play, RTP will contact the Portuguese users association FPAS (Federação Portuguesa de Associações de Surdos [Portuguese Deaf and Hard-of-hearing Association]) to set up a meeting to define a plan to test the new JWPlayer plug-in developed by UPM.

The main goal is to have 10 (maximum 15) potential users testing our sub-pilot in RTP facilities, with the support of access services experts from UAB.

Evaluation will be focused on technical items and interactivity with different devices, and will consider the personalization of the layout of subtitles (black box, bigger characters vs. "standard" layout, font size, etc.).

4.4.3.2. Large scale tests

The Portuguese sub-pilot will be evaluated by the users online. For this the System Usability Scale (SUS) will be used as adapted by UAB for use within Pilot-B, without any major changes, because the user will be

[5] http://www.rtp.pt

manipulating and interacting with their own device during the test. The SUS questionnaire will need to be translated into Portuguese.

The questionnaire link will be available on RTP's website and open to all viewers that use the RTP Play service.

4.4.4. Challenges, problems and potential solutions

During the test period a currently unsupported feature was detected when using the Subtitle Conversion Framework (SCF), specifically with conversion of subtitles for a program that is published in several parts (as several videos). When an EBU STL subtitle file was converted for the second part of a published program, we needed to synchronize the first subtitle in the subtitle file that is for the second part of the program with the beginning of the second part of the program. We did this by applying a negative offset to all the subtitles in the EBU STL file with a duration equal to the length of the first part of the program. In this way the subtitles of the first part assume negative timecodes (and these subtitles should be ignored) and the second part will start with positive timecodes as the video. We have verified that the Subtitle Conversion Framework (SCF) does not convert negative timecodes.

This situation will not allow the subtitling of the second parts of a program when they are published in the web.

Also, a problem was found related to the evaluation method. Having the questionnaire link available on RTP's website, and thus being open to all people that use the RTP Play service – including people without hearing problems and the deaf and hard-of-hearing – may cause the results to deviate from the expected.

One possible solution is to present this idea to the deaf and hard-of-hearing association and ask them if they are available to link the questionnaire on their own website. In this case, the results may be more precise, by narrowing down the participation of users to just deaf community (potential users of deaf association website).

4.5. Automated HbbTV multilingual news broadcast subtitles- Spain

At the time of writing this deliverable, the user tests for the Automated HbbTV multilingual news broadcast subtitles service has been concluded. The service has been designed to integrate two different threads:

- On the one hand, the service benefits from the technologies and mechanisms to generate subtitles automatically in different languages using speech recognition techniques and machine translation technologies.

- On the other hand, the service integrates the previous automatically generated subtitles in a real-time HbbTV workflow. It creates an MPEG-DASH stream with a broadcasted TV channel and feeds it with automatically generated subtitles to be consumed by a DASH client.

The user tests designed at the beginning of the project aim to provide a feedback about the quality of the automatically generated subtitles, testing only the first thread described above. Even though the technologies to create multilingual subtitles from speech are still being continuously improved, they are mature enough to perform the user tests and measure their enhancement to the end user.

This service has been working in parallel on three main activities during the execution of the project, until the time of writing this deliverable:

a) The improvement of the generation of multilingual subtitles automatically from speech.

b) Performing user tests with specifically generated material. For this purpose, videos with automatically generated subtitles in different languages have been created.

c) The integration of the subtitle technologies within the HbbTV workflow.

To provide a feedback about the quality of the automatically generated subtitles activity b) has already finished (UAB), while activity c) has continued the integration of the tested technologies in the HbbTV workflow (by VIC). The results obtained with the user tests are described in section 5.2.

This service was defined at the beginning of the project as a lab-test service to be deployed in VIC's Digital TV laboratory and consequently it is not going to be used for the sub-pilots described in the document.

4.6. Automated HbbTV multilingual news broadcast subtitles- Switzerland

As already mentioned in section 3.7, this sub-pilot will not be implemented by Swiss TXT.

4.7. Customised Subtitles for Online learning (MOOC) – Germany

4.7.1. Status of work

As described in sections 2.3.6 and 3.8, the different components of the Massive Open Online Course (MOOC) service integrating the vPlayer, the vManager CMS as well as *vSubtitle Component* for adaptive subtitle rendering are implemented and are being tested at the time of writing for the preparation of the sub-pilot. The sub-pilot will be a MOOC on media accessibility containing a number of video lectures on topics related to Pilot A, including subtitle authoring, subtitle adaptation as well as audio descriptions. The content is actually prepared by UAB in collaboration with Vsonix and an external partner and will be available at the end of August 2015. The complete service will be hosted via the Vsonix CDN to be integrated on the HBB4ALL website. The different components of the MOOC, as well as the AccessGUIDE application, have been integrated to be used as a complete prototype to be tested in the final lab test for the sub-pilot. This will take place early September 2015.

As a cross-pilot activity, the MOOC service will integrate components that have been integrated and trialled in WP3 (Pilot A) for sub-title rendering as well as components that have been realized and trialled in WP5 (Pilot C) for the UI adaptation part of the application.

In October 2014 UAB and Vsonix planned and realised user tests for the AccessGuide software, which is used in the MOOC application to determine the user preferences and needs for the subtitle adaptation as well as for the UI adaptation part. The software was tested with several participants from the main target group. This includes deaf and hard of hearing people as well as people with visual disabilities with ages ranging from 16-66. We used the thinking loud methodology, where the test participants were individually interviewed and thereafter were presented the updated wireframes and a simulation of the user interaction flow. The participant was free to comment and review the current version of the application, followed by an open discussion with all the participants about the problematic parts of the application.

The results indicated that the AccessGuide software could be improved by increasing the contrast of UI elements and by simplifying the instructions. The results also indicated that an additional button for readjusting the settings could be very helpful. Most results confirmed the conceptual approach and suggested modifications, which were part of the concept, but couldn't be realized in this version of the prototype. Overall the conceptual direction was approved by the participants.

4.7.2. Timeframe

The MOOC application integrating the vSubtitle component as well as the AccessGUIDE application for subtitle personalisation has been integrated in a first complete version that was ready at the beginning of July 2015. This version is being tested at the time of writing and refined to be prepared for the final user tests that will take place in September 2015.

Based on the test results, further refinements will be made before the MOOC service is operational, including the video course on media accessibility, which will go online for PC and mobile platforms at the end of 2015 / beginning of 2016. UAB has planned to finalize the MOOC content by the beginning of September 2015.

4.7.3. Test users and evaluation approach

AccessGuide was developed further with the results from the user tests in October 2014 and another user test is planned for mid-September 2015. This test will be carried out on a bigger scale, with approximately 40 participants. Both 'thinking aloud' and 'eyetracking' techniques will be used and the results will be included in the final refinements of the public version. Beside testing the AccessGUIDE software the user tests will also focus on the integrated MOOC prototype including the vSubtitle rendering component that was described in section 2.3.6.

As soon as the service goes online, the final evaluation will be made using an online survey to gain user feedback.

4.7.4. Challenges, problems and potential solutions

No significant problems have been faced since the last version of the pilot report in HBB4ALL deliverable D3.1. The user tests for the AccessGUIDE application have been promising so far. Recommendations by users and partners have been taken into account and were integrated in the actual version. The work on the subtitle part of the MOOC sub-pilot is on time so the planned schedule can be met. One change has been made regarding the supported subtitle format. As the vPlayer is focusing mainly on PC and mobile platforms, the Vsonix *vSubtitle Component* natively supports the WebVTT subtitle format, which is widely available in the online world.

4.8. Overview of common challenges

No large problems are currently anticipated for Pilot-A. With respect to the operational phase of the sub-pilots, the partners have identified following common challenges:

- Recruiting and motivating test users:
  Open tests involving members of the general public always contain certain challenges and risks. Firstly there is the need to recruit users' representative of the intended end user group. To meet this

challenge partners are working closely with user associations in their test areas to raise awareness of the tests and reach the correct user groups. These testers not only need to be representative but also meet the technical and network requirements to participate in the tests especially if doing so from their own homes. In cases where testers do not have the required technology, partners will look into the possibility of providing the testers with the technology required for the test.

- Users understanding the testing method:
  This is an organisational issues, which will be tackled by cooperating as closely as possible with (local) user organisations during the final preparations of the sub-pilots and when carrying out the evaluations. In advance of any tests, partners will inform their user associations of their test methodology and provide any questionnaires to gather qualified feedback regarding the suitability of the methodology and accessibility of any material for the intended user group.
- Evaluation of sub-pilots:
  Using evaluation methods that differ too much may result in responses that are difficult to compare. To monitor user experience, the WP3-partners have agreed upon the System Usability Scale (SUS) as a standardized instrument to measure usability as a core part of user experience. This will be adapted to each pilot and translated into simple language versions for each single sub-pilot. The challenge here is not to change the nature and intent of the original questionnaire but at the same time make it relevant to the service and understandable for the testers. To this end each pilot will create their version of the questionnaire if a standard translation does not already exist. This version will be checked with the user associations and also UAB will cross-check and compare all SUS questionnaires.
- Delay in HbbTV 2.0:
  This is an organisational issue with a technical impact: the HBB4ALL partners need to rely on their own developments to support EBU-TT-D in the sub-pilots. Before the start of HBB4ALL, hardly any EBU-TT-D implementation existed. Integration of the current developments has the risk of incompatibility. To avoid this as far as possible, the partners using EBU-TT-D in their sub-pilots are actively cooperating to cross-test their implementations, e.g. by exchanging test content.

More general, and beyond Pilot-A, challenges with respect to the standard conformant usage of the novel EBU-TT-D subtitle format play an important role. These challenges are relevant for all companies who plan to migrate to using EBU-TT-D:

- Interoperability of EBU-TT-D implementations:
  HBB4ALL partners are working hard to achieve interoperable implementations, e.g. by exchanging EBU-TT-D files for mutual testing and by using the EBU-TT-D XSD for verification of the XML documents.
- Correct rendering of EBU-TT-D subtitles:
  The subtitle rendering can be verified as conforming to the standard (as is being done by HBB4ALL partners) by making use of reference material from the EBU-TT-D rendering test suite.

EBU-TT-D has been widely and collaboratively used by the partners in the project, as described in the previous paragraphs, since it satisfies very well the requirements of the new broadcast-broadband convergence multimedia paradigms. EBU-TT-D seeks harmonisation between different subtitle formats and currently gets a lot of market attraction (especially through the use of HbbTV 2.0).

One common challenge of HBB4ALL is the further promotion of EBU-TT-D as an open, standardised solution for the provision of subtitle services in connected TV environments and as a way to increase the future exploitation of the components implemented in Pilot-A.

5. Complementary user tests

The tests carried out by UAB mainly focused on how to improve subtitle content quality in the HBB4ALL environment, by analyzing different aspects about subtitle personalization and perception by end users. The data gathered from these tests, and the conclusions drawn from them, can be applied by the sub-pilots in their tests for the operational phase and further actions.

For the sub-pilot Automated HbbTV multilingual news broadcast subtitles – Spain, all user tests have been completed and the results are being reported in this chapter.

5.1. Preliminary user tests (until M20)

During the first phase of the project, UAB structured and adapted the target users and the evaluation methods according to testing goals of each experiment, in order to provide consistent results. The user tests carried out and their results are described in the following sections.

5.1.1. Subtitle usability across devices

In this experiment, the differences in the usability, comprehension and perception of subtitles were tested across three devices with different screens size: TV screens, smartphones and tablets. The aim was to identify if the screen size has an impact when watching a video with subtitles. Although the pilots mainly target Smart TVs, it is clear that the same content will be displayed on other devices with different screen size (such as tablets). As the production chain is the same, the presentation of the subtitles will not often be adapted to the specific devices. Nevertheless, the results from this experiment could be taken into account when publishing broadcast content on other platforms.

The experiment was developed in coordination with the University of Warsaw in Poland and the Universidad Autónoma de Barcelona in Spain. The tests were carried out as an eye-tracking study in which participants watched 6 clips: 3 in English and 3 in Norwegian. Each participant saw the clips with subtitles (in Polish or Spanish) on the three different devices. The comprehension and recollection of visual data was recorded as well as eye-tracking and preference information. The participants were students of both countries between the ages of 19 to 35 years old. Regarding the number of participants, the initial plan was to test 36 subjects in each country predicting the possibility of experimental death in 6 participants (18 participants for English and the other half in Norwegian, so that each country tested at least 15 participants in the appropriate language).

The results showed that there were minor differences when watching subtitled videos on smartphones in relation to TV screens and tablets. In fact, comprehension and visual information levels were better on TV screens and tablets. This finding leads to the conclusion that screen size is an element to take into account when subtitling videos specifically for smaller screens (that is, for smartphones). More participants should be tested to gather accurate data during the current year (2015), in order to understand if it is worth developing a different platform to implement improvements in the subtitle presentation for smaller screens (i.e. smartphones).

5.1.2. The influence of the original language of the film on subtitle reading

The aim of this study carried out by the University of Warsaw in Poland and the Universidad Autónoma de Barcelona in Spain was to identify differences in subtitle reading patterns depending on the language of the video for Norwegian and English films using a TV screen.

The tests were carried out as an eye-tracking study, as well as a comprehension questionnaire regarding the text in the subtitles and visual memory of the images. A 'within-subjects design' pattern was was applied to the tests by showing 6 videos in counterbalanced order in two languages (3 in English and 3 in Norwegian) to each participant. After each clip, they answered 10 multiple choice comprehension questions: 5 on text and 5 on visual aspects.

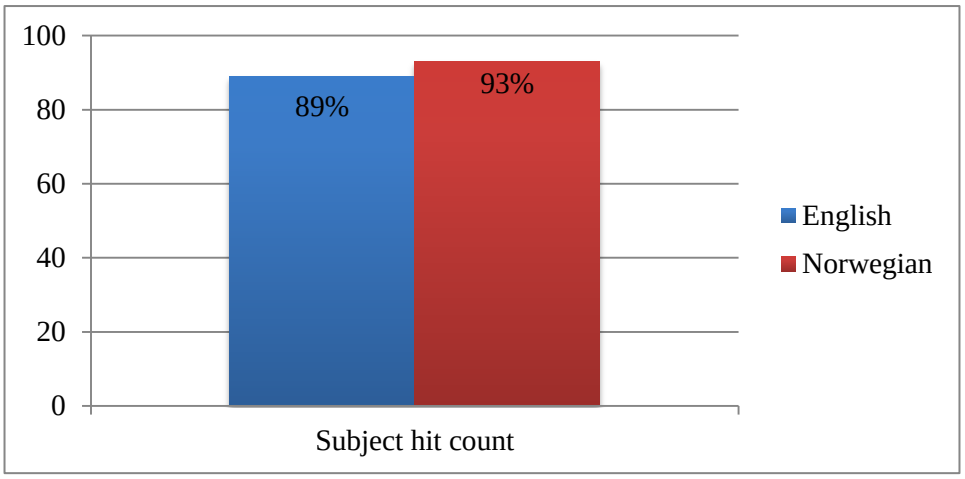

Figure 2. Percentage of subtitles looked at

The results showed that not knowing the original language of the video increases the probability of reading the subtitles. When knowing the language of the video, the probability of skipping the subtitles increases. Therefore, more time is spent on reading subtitles and less time is invested in watching the rest of the image when not knowing the original language of the video (see **Figure 2** and **Figure 3**).

Also, eye-tracking results showed that there were longer mean fixation duration on subtitles in the English videos and, in general, the mean fixation duration while reading subtitles is shorter than when reading printed text (see **Figure 4**). This finding leads to the conclusion that the audience only reads specific parts of the subtitles when knowing the language, as a support to understanding the audio. That is why the mean fixation duration is longer; the audience will focus their attention on specific words or segments.

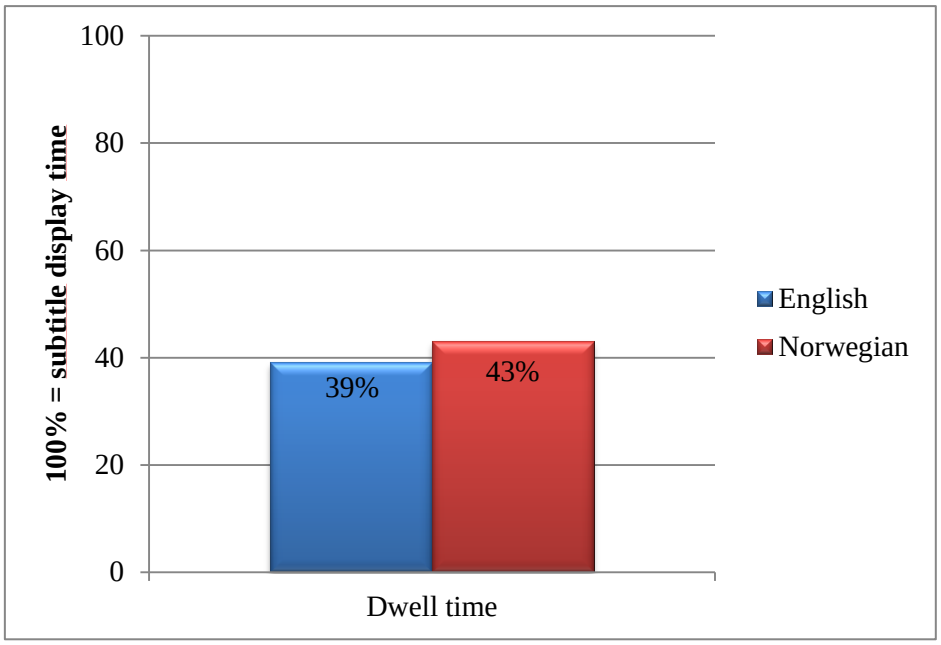

Figure 3. Percentage of looking at a subtitle and subtitle display time

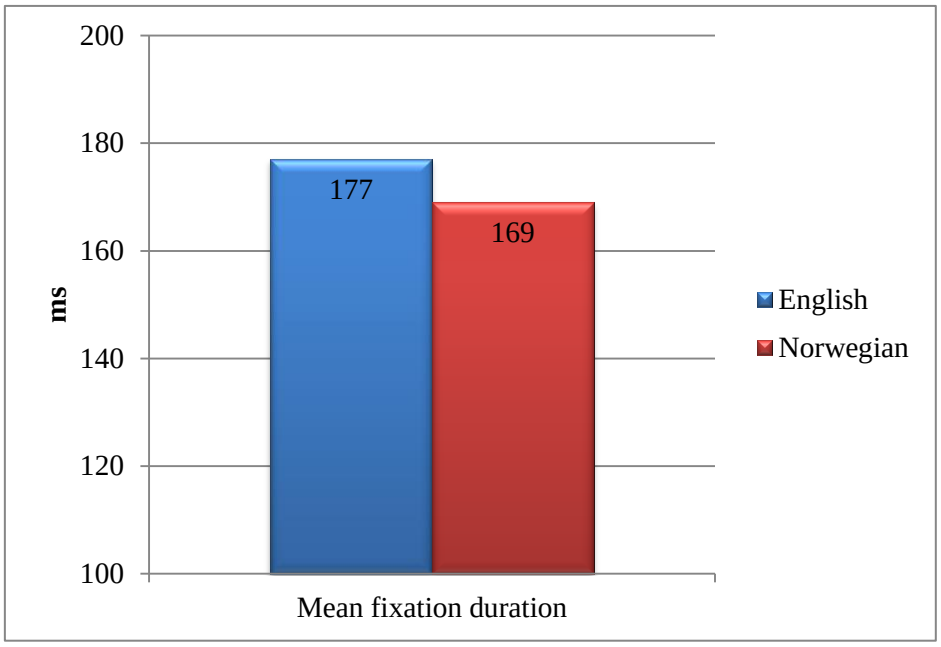

Figure 4. Mean fixation duration

These findings show that different strategies could be taken into account when creating subtitles. For instance, if the original language of the video is known or familiar to the target users, then the subtitles will be used as a support to better understand the audio of the video. As the probability of skipping the subtitles is higher, elements such as the number of characters could be reconsidered or a more literal translation might be chosen (as the audience will look for specific expressions, segments or words in the subtitles). When not knowing the language, the most important elements to consider are the readability of the subtitles and the

correct transfer of the sense or meaning of the dialogue. Strategies such as text reduction and maximum number of characters are more important for languages unknown by the target users, since they entirely rely on the subtitles to understand the video content.

5.1.3. Optimal subtitle synchronization

This study aimed at identifying the optimal subtitle synchronization before and after a shot change. Subtitle synchronization is a very important element that helps the audience to better follow the content. This investigation highlights how important frame accuracy is and provides indications for subtitle editors. Another possibility for future related experiments would be to carry out a test to analyze if the optimal frame accuracy can be reached.

The study was carried out by the University of Valladolid. Tests have been carried out in 25 hearing users studying at university (and it is still ongoing) who are between 19-45 years old (so far 15 females and 10 males) and whose working language is English. The test was carried out using eye tracking technology.

The variables tested were as follows :

- leaving a gap[6] after a [shot change] of +2/4/6/8 frames
- leaving a gap of +2/4/6/8 frames before a [shot change]
- leaving a gap of +2/4/6/8 frames before a [shot change] and then leaving a gap of +2/4/6/8 frames

The preliminary results showed significant differences between shorter (0/2/4 frames) manipulation times and longer (6/8 frames) manipulation times (similar results when leaving a gap before and after a shot change). It seems that leaving a gap of 8 frames after and before a shot change is the most effective option. There are also significant differences in reaction times (Time to First Fixation, First Fixation Duration shown in the eye tracking data) depending on the level of English of the participant: between B (higher level of English) and C (lower level of English), C speakers preferred shorter times. This might mean that speakers whose level of English is not that high are anticipating a subtitle more.

The University of Valladolid is currently undertaking another experiment regarding the optimal subtitle synchronization for inter-subtitle spacing (i.e. for sentences that have to be divided over multiple subtitles because they contain too many characters to be presented in one subtitle). For further details, see section 5.4.

5.1.4. Subtitle usability in different alphabets

The aim of this experiment was to determine user's preferences on subtitles in Chinese taking into account different variables, such as font, colour, and size, in the two kinds of Chinese orthography; traditional and simplified, to contribute to establishing standards in subtitles in Chinese. The objective is to study how other alphabets (such as Chinese) could be integrated in the system. Even though the HbbTV standard may not require support for Chinese fonts in devices, the system could reference Chinese fonts from the web.

[6] "gap" here means the number of video frames used to separate subtitles

Figure 5. Chinese subtitles, fontsize 12

Figure 6. Chinese subtitles, fontsize 14

This experiment was developed by Vienna University of Vienna in Austria, the University of Valladolid and the Universidad Autónoma de Barcelona in Spain. The design of the experiment was based on an online preference questionnaire designed with Google forms with no control over participant profile. It included demographic and control independent variables (font, size, and colour). The stimuli consisted in a snapshot taken from a video clip subtitled using the screen recorder Camtasia, changing subtitling parameters such as

font, size, and colour in the two Chinese orthographies (traditional and simplified). Participants watched 5 pairs of snapshots with different sizes (sizes according to Camtasia software) (12 and 14 in both orthographies, see **Figure 5** and **Figure 6** respectively) and 5 pairs of snapshots with different colour (white and yellow in both orthographies) using a device (computer, tablet or smartphone). The total number of participants was 112 and the average age was 29.94. The origin of the participants was distributed as follows: 55% Mainland China, 41.4% Taiwan, 1.8% Hong Kong, 1.8% other.

It is important to mention that sizes 12 and 14 are specifically related to the Camtasia software; the metrics are not "px". In our case the picture has a height of 362px (for the font with size 14) and the height of the font is approximately 18px. The relation between the rendering of the video and the font in this case is 4%.

The five fonts in two different orthographies tested were the following:

- Simplified:
  - Kaiti: 我叫夏巍，来自巴塞罗那
  - Simsun: 我叫夏巍，来自巴塞罗那
  - SimHei: **我叫夏巍，来自巴塞罗那**
  - HanWangYenLight: **我叫夏巍**，来**自巴塞**罗**那**
  - STXihei: **我叫夏巍，来自巴塞罗那**
- Traditional:
  - DFKai-SB: 我叫夏巍，來自巴賽隆
  - MingLiU: 我叫夏巍，來自巴賽隆納 **(新細明體)**
  - HanWangYenLight: **我叫夏巍，來自巴賽隆納**
  - FangSong: 我叫夏巍，來自巴賽隆納
  - SimHei: **我叫夏巍，來自巴賽隆納**

Results showed clear preferences for several variables: the preferred font type is SimHei, the participants preferred the larger font (14) over the smaller font (12) and the preferred colour is white. These preferences are valid for both orthographies and for both groups (Mainland China, and Taiwan and Hong Kong). Even though more than 80% of the sample used the same kind of device (computer) to answer the questionnaire, and therefore, no conclusions can be drawn as far as this variable is concerned, the results obtained show no major differences in this respect.

Based on the outcome, there seems to be a possibility to make a common policy on font selection for Chinese subtitles and to allow for bidirectional exchange flows.

5.1.5. Second screen subtitling utility

The aim of this experiment was to test subtitles displayed on a second screen device (a smartphone) using the application Artaccés (provided by the Department for Culture of Generalitat de Catalunya) to watch subtitles for a movie displayed in the cinema. The experiment was carried out with participants with hearing difficulties. The test took place during Sitges International Fantastic Film Festival of Catalonia.

The film was presented to the participants in a regular session during the festival. It was a Spanish film that had Russian as the main language and Japanese and German as secondary ones. Due to the fact that the film was spoken in foreign languages, subtitles were shown on screen in English, and in Catalan and Spanish on secondary screens under the main display screen. Participants were placed in the middle right section of the theatre. After the visualisation of the film with the subtitles activated on the device, seven participants were conducted to a separated room where a focus group session took place. The participants downloaded and manipulated the application on their own; no technical support was offered nor demanded by them. An interpreter assisted with the Focus Group discussion.

A series of guidelines was prepared for the focus session, in order to gather data. One of the main concerns raised by the participants during the focus group session was the position difficulty and the change of focus (between the two screens). The same issue would arise if the application was displayed in sign language. However, the focus session was also useful in suggesting different strategies to improve this application. It was suggested that the application could be used as a complement to the subtitles displayed in the main screen. If a system based on signaling were to be used (e.g. haptic feedback), the application could display specific notations such as music or context sounds. Thus, this feature could be useful especially when watching films in their original version at the theatre (with subtitles on the main screen). The vibrations (haptic feedback) would need to be placed carefully when the dialogs start in the first subtitle, not in each subtitle. Given all this feedback, UAB considers that the application should be tested properly in a more representative context. Thus, it should be tested with other films, perhaps with a tablet or different models of mobile phones in order to verify if there are any differences. More tests should be carried out taking into account different types of user profile. Some recommendations or different suggestions about how to hold the device should be also tested.

Finally, the possibility of using this technology for conferences came up from a couple of participants. Given the fact that there is no need to be looking at the speaker all the time, this application could be of much use in this scenario. Nevertheless, this concept would present other technical problems related to live subtitling that would have to be addressed for this idea to be implemented.

5.1.6. Subtitles and embodied cognition

This study was carried out by the Macquarie University, the University of New South Wales and the Universidad Autónoma de Barcelona. The aim was to investigate the cognitive processing of subtitled film in terms of subjective psychological measures and Electroencephalography (EEG), find out whether an audiovisual translation (AVT) product like subtitling alters a viewer's immersion into, and also enjoyment of the text and to offer a scientific methodology for investigating the cognitive processing of AVT products that will be beneficial in future studies on the process, product and reception of AVT.

The availability sample consisted of 88 university students aged 23 to 49 years. Participants were recruited from the population of Translation and Interpreting students at two universities in Sydney, Australia and had different first languages (21 English, 41 Mandarin Chinese and 22 Korean). Participants were assigned randomly to the subtitled and un-subtitled conditions. The group who saw the video without subtitles consisted of 10 English, 22 Chinese and 13 Korean participants, and the group who saw the video with subtitles consisted of 11 English, 19 Chinese and 9 Korean participants. The average age of the un-subtitled group was 27.5 years and of the subtitled group 29.5 years. The material used for the experiment was one episode from the eighth and final season of the American medical drama series *House, MD* with a duration of 44 minutes. This study made use of subjective post-report questions on 7-point Likert scales with participants circling the number representing their opinion on the film.

The results are in favour of the subtitle condition. In fact, the findings show higher levels of transportation, character identification, presence, perceived realism and enjoyment in the subtitled group than the un-subtitled group.

The conclusions indicated that subtitles would not decrease the immersion or enjoyment of the audience. The audience used the subtitles to support their comprehension of unfamiliar terms, giving them a stronger sense of understanding of the drama. Also, the audience was able to process the subtitles efficiently and that it facilitates embodied (and disembodied) cognition. Participants were more focused and had less time to become side-tracked by irrelevant visual elements or their surroundings. These findings are useful to confirm the value of subtitling AVT products and its benefits.

5.1.7. Caption utility

The aims of this study were to assess:

- the quality of the captions produced by CaptionCue (an automatic caption cueing system) as they are received by users with and without hearing loss through different devices
- the preferences regarding four different caption devices: an LED screen integrated in the stage, an LED screen on the side of the stage, tablets and smart glasses
- the users' views about the main features of the captions produced by CaptionCue
- the users' distribution of attention between the captions and the actors on stage, and
- the users' sense of presence and engagement with the performance.

The results contribute to understand the utility of this system and how to improve it. This study was carried out by the University of Roehampton.

The Temporary Theatre at the National Theatre prepared an especially-devised 30-minute performance for the experiment. The first performance was attended by hearing users watching the play with no captions. The second, third and fourth performances were attended by users with and without hearing loss watching the play with captions displayed on an LED screen integrated at the top centre of the stage, two LED screens on the side of the stage and tablets resting on holders in front of viewers sitting in the upper circle. The audience was seated on three-sides of the floor level acting area, with more seating in a gallery on three sides, above.

The participants were initially provided with a demographic questionnaire that they completed online before the experiment. After the performances, they completed a more detailed preference questionnaire and an ITC-Sense of Presence Inventory (ITC-Sopi) adapted to the context of the experiment. 4 cameras situated at the far end of the upper circle, 15 metres away from the audience, filmed the participants' eye movements as they watched the performances. These movements were analysed individually to determine if the participants' gaze was focused on the caption device, the actors or elsewhere around the stage.

166 people took part in the experiment: 103 (62%) had a hearing loss and 63 (38%) were hearing. Amongst those with hearing loss, 31 (30%) defined themselves as deaf, 12 (11.6%) as deafened and 57 (55.3%) as hard of hearing. 3 participants did not state their hearing characteristics or the data was missing.

The results showed that captions were accurate and readable, had the right length and display mode, were shown at the right speed (although some were too fast), were not fully synchronous and, in general, were at least as good as normal manual captions. 50% of the participants considered that it was possible to follow both the captions and the performance on the stage and 30% disagreed.

There were differences depending on the caption device used: highest percentage of time looking at the actors with LED screens (43% on the captions vs 56% on the actors), followed by the integrated LED screen (45% on the captions vs 51% on the actors) and the tablets (52% on the captions vs 43% on the actors). Also, differences were observed in distribution of attention, and particularly in the time spent elsewhere caused by the age of the participants: the younger they were when they lost their hearing, the more they looked elsewhere during the performance.

5.2. User tests regarding the Automated HbbTV multilingual news broadcast subtitles' sub-pilot - Spain

5.2.1. Aim and setup

The initial aim of this sub-pilot was to determine if automatic inter-lingual subtitling helps to better understand live news content broadcast in a newsroom environment. The hypothesis was that clips automatically subtitled would be better understood than clips without subtitling. This was to be tested with a proposed a sample of 30 students in their 4th year of Journalism Studies volunteering to participate in the study. Materials would include 2 clips, each about 2 minutes long (one with and one without subtitling) on a breaking news story.

After preparatory work in the experimental set-up, two changes were made to the initial plan. On the one hand, it was decided that three scenarios would be compared: automatic intra-lingual subtitles (English-English), automatic inter-lingual subtitles (English-Spanish), and no subtitles (original track in English). The rationale for this change is that the quality of intra-lingual automatic subtitles was expected to be higher than inter-lingual automatic subtitles, and significant differences were expected to be found. Therefore, the aims have been reformulated as follows:

- The aim of this research is to determine if automatic inter-lingual subtitling and automatic intra-lingual subtitling help to better understand news content broadcast.

- The hypothesis is that clips with intra-lingual automatic subtitles are better understood than inter-lingual automatic subtitles and original content without subtitles.

On the other hand, due to the difficulties of finding appropriate participants (students in their 4th year of Journalism Studies) to volunteer for our tests, it was decided that participants from other studies would be selected and classified according to their English level.

Materials for the new test scenario have been developed, and preliminary tests were carried out with two aims: (a) to identify initial trends, and (b) to test the proposed methodology.

The preliminary results showed almost no differences when comparing the different conditions tested. Moreover, the number of correct replies was very low. Due to these results, the methodology was revised to create a new experimental design.

The proposed new experimental set-up was as follows:

The aim of this experiment is to research whether automatic inter-lingual and intra-lingual subtitling helps to better understand live news content broadcast in a newsroom environment. This is why three scenarios are compared:

- automatic intra-lingual subtitles (English into English)
- automatic inter-lingual subtitles (English into Spanish)
- no subtitles

The hypotheses are:

- Hypothesis 1: intra-lingual automatic subtitling will increase comprehension as compared to clips with no subtitles.
- Hypothesis 2: inter-lingual automatic subtitling will increase comprehension as compared to clips with no subtitles.
- Hypothesis 3: inter-lingual automatic subtitling will not increase comprehension as compared to clips with intra-lingual subtitles.
- Hypothesis 4: the usefulness of subtitling will increase as the English proficiency of the participants decreases.

5.2.2. Timeframe

The preliminary user testing was finished at the end of January 2015. Then the new experimental design for the main test was created, tested and concluded in June 2015. At the beginning of July 2015, the analysis of the data was undertaken.

5.2.3. User tests and evaluation

This is the description of the evaluation and results from the new experimental user test finished and analysed in July 2015.

Participants

Tests were carried out with 30 students (13 male, 17 female, mean age: 25.2) from UAB, following ethical procedures approved by UAB ethical committee.

Materials

Testing materials included three breaking new stories in the three conditions indicated above: no subtitles, intra-lingual English subtitles, inter-lingual Spanish subtitles. The 2-minute clips chosen were from Reuters news service, and were selected after a thorough content analysis (number of speakers, topic, length). They dealt with financial issues.

Comprehension questionnaires were developed for each clip, prioritising multiple-choice questions for practical reasons. Comprehension questions included the key aspects of any news story (Who? What? When?

Where? How? Why?). An analysis of the clips allowed the control of the information provided visually. Preliminary testing allowed the fine-tuning of an initial version of the questionnaire. The final version included 9 questions per clip.

English proficiency test: experts in English language teaching suggested DIALANG but, after thorough examination, it was discarded due to the length of the test. Alternatively, an online system was proposed. It focuses exclusively on listening comprehension and lasted a maximum of 20 minutes.

Procedure

Participants were welcomed in a lab and signed informed consent sheets. They were instructed that they would be shown three clips in English on a news item (one without subtitles, one with English subtitles and one with Spanish subtitles) two times. After watching the first time they were allowed to read the questions. After the second time they had to reply to a questionnaire. Prior to the experiment, a demographic questionnaire and an English proficiency test were administered.

The English proficiency test allowed us to classify the participants into five levels: A1 (zero participants, A2 (two participants), B1 (8 participants), B2 (7 participants), C1 (8 participants), and C2 (four participants). The sample in the highest and lowest level is very low (2 and 4 participants), allowing only for qualitative interpretations. The sample between B1 and C2 is balanced, although limited in its number. Taking into account the low numbers per level, a qualitative descriptive analysis is provided instead of a statistical analysis.

Figure 7 summarises the results, showing for each condition (English subtitles, Spanish subtitles and no subtitles) the percentage of correct replies.

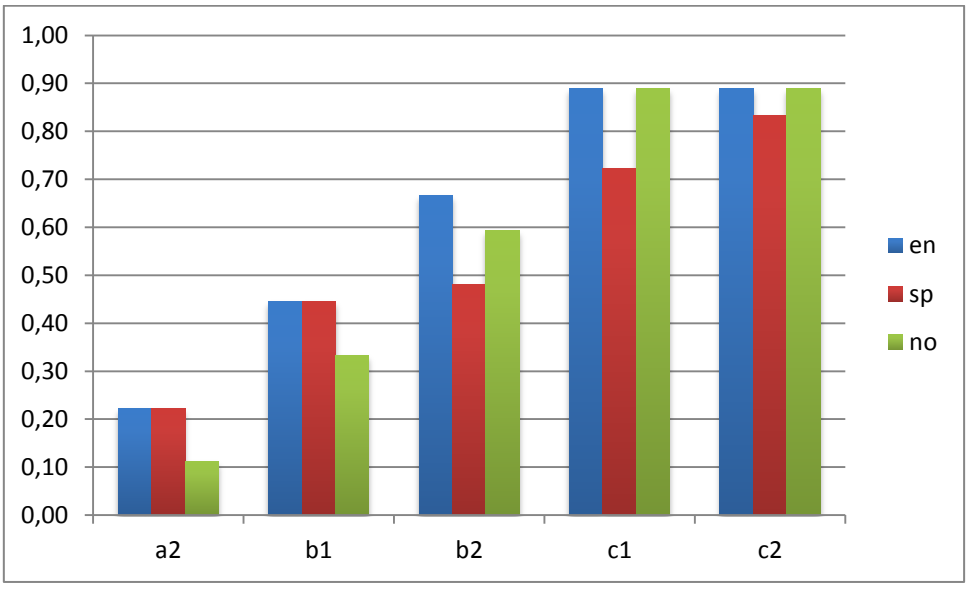

Figure 7. Summarised results for preliminary user tests on multilingual news broadcast subtitles

In the less proficient participants (A2), both automatic inter-lingual and intra-lingual subtitles help to increase the comprehension of the original content (from 11% to 22%), although comprehension is still very low (below 22%). This pattern is exactly the same for B1 participants, although the comprehension levels

increase up to a minimum of 33%. So, with both intra-lingual and inter-lingual subtitles comprehension of the original content is 44%, whilst without subtitles it is 33%.

In the most proficient participants (C2 and C1), on the contrary, an improvement is not observed. Comprehension without subtitles is 89% in both groups, the same value obtained for automatic intra-lingual subtitles in English. With the inter-lingual subtitling option comprehension levels decrease for both groups of participants, with the decrease more striking in C1 (83% for C2, 72% for C1). This may indicate, subject to further testing with wider samples, that automatic inter-lingual subtitles may detract the viewers' attention and affect comprehension negatively.

In the participants with a B2 level, a different trend is observed: comprehension with automatic intra-lingual subtitles (67%) is better than without subtitles (59%). However, comprehension decreases considerably in the automatic inter-lingual subtitles option (48%).

If 50% of correct replies is considered as a threshold to consider the news to have been understood, this is only achieved in the following conditions.

- B2 level: intra-lingual subtitles (67%)
- C1 level: intra-lingual subtitles (89%), no subtitles (59%)
- C1 level: intra-lingual subtitles (89%), inter-lingual subtitles (72%), no subtitles (89%)
- C2 level: intra-lingual subtitles (89%), inter-lingual subtitles (72%), no subtitles (89%).

Conclusions

The previous results allow the following conclusions, which should be verified with wider samples of participants.

1) Regarding to hypothesis 1 ("intra-lingual automatic subtitling will increase comprehension as compared to clips with no subtitles"), it has been confirmed for participants whose level of English is between A2 and B2, but levels of comprehension stay the same for intra-lingual automatic subtitling and no-subtitles for C1 and C2 participants.

2) Regarding hypothesis 2 ("inter-lingual automatic subtitling will increase comprehension as compared to clips with no subtitles"), it has been confirmed for the less proficient participants (A2, B1), although their levels of comprehension are very low. As for B2, C1 and C2 participants, comprehension is better without subtitles than with inter-lingual subtitles, demonstrating a distracting effect of automatic inter-lingual subtitles.

3) Regarding hypothesis 3 ("inter-lingual automatic subtitling will not increase comprehension as compared to clips with intra-lingual subtitles"), it has been confirmed for all participants. Comprehension stays the same (A2, B1) or improves (B2, C1, C2) with automatic intra-lingual subtitles in English as compared to inter-lingual subtitles.

4) Regarding hypothesis 4 ("the usefulness of subtitling will increase as the English proficiency of the participants decreases"), it has been confirmed. Automatic subtitles (both intra and inter-linguistic) increase comprehension for the lowest levels (A1, B1), but do not have a positive effect on

comprehension in the highest levels (C1, C2). In the middle range (B2), intra-lingual subtitles help but inter-lingual subtitles do not.

5.3. Summarised results and guidelines / recommendations

From all the studies carried out so far, following guidelines and recommendations can be developed.

- Screen size is an element to take into account when displaying subtitles on smaller screens, such as smartphones. Further tests for the experiment regarding reading across devices (TV screens, smartphones, tablets) should be carried out in order to understand to what extent the size of smartphones could have an impact on the comprehension and enjoyment of subtitled audiovisual products. It would be important to determine whether or not it is necessary to develop a different platform to implement improvements in the subtitle presentation. Also, more specific experiments on subtitles for smartphones could be undertaken on specific elements such as line-breaks, font size and maximum number of characters, to improve the quality of subtitles on smartphones.

- The language of the audio is also important when subtitling: subtitles for a very international language such as English (for the audio) are used by viewers as a support tool to understand the audiovisual material. Considerations such as the maximum number of characters could be taken into account when subtitles are created for known languages, since the subtitles are used as a complement to understand the audio of the video (in comparison with a more unknown language to the audience, for which the audience has to read all the text of the subtitles). Thus, different strategies could be considered depending on the language of the video (e.g. more literal subtitles for known languages and more reduced subtitles for unknown languages).

- The 'gap' of frames used to separate subtitles will have to be reconsidered because this could improve reading and comprehension for the participants. Broadcasters and subtitlers will have to take into account the results from the mentioned experiment to improve the quality of subtitles. More precisely, we found the optimum gap of frames that should be left before and after a shot change to be 8 frames.

- Also, the results for Chinese alphabets could be implemented to establish standards for subtitles and thus improve the enjoyment of audiovisual products in Chinese. Although smart TV sets may not have Chinese fonts, they could reference Chinese fonts from the web. As mentioned in the previous section, results showed clear preferences for the font type SimHei, size 14 and white color for the subtitles.

- The haptic effect could be integrated for smartphone applications that display subtitles on the screen (such as the application mentioned previously, Artaccés) to complement the subtitles on the main screen, so that hearing impaired viewers will be able to follow and appreciate better movies at the cinema. Systems such as Artaccés could also be used for conferences.

- It is always important to consider the fact that these recent findings indicate that subtitles do not decrease the immersion or enjoyment of the audience. Subtitles support comprehension of unfamiliar terms, giving viewers a stronger sense of understanding of the drama.

- Systems of automatic caption cueing, such as CaptionCue, prove to be efficient and could be useful for certain live TV programs that share elements with live performances and events.

- Regarding automatic subtitles, the test carried out shows that they are useful for participants with a middle-range level of English (B2) but only if intra-lingual. In participants with a low proficiency in English, both intra-lingual and inter-lingual automatic subtitling increase comprehension but levels remain very low, so no substantial effect is observed. In participants with a high proficiency in English, subtitles do not increase comprehension: on the contrary, inter-lingual subtitles may lower comprehension levels.

- Another recommendation is to use the System Usability Scale (SUS)[7] for the evaluation of HBB4ALL sub-pilots in the operational phase because it provides a global view of subjective assessments of usability. The SUS consists of a 10 item questionnaire with five response options for respondents, from "strongly agree" to "strongly disagree". This evaluation is useful for the user tests carried out in the project because it analyzes a variety of products and services, and gives criteria on how to improve these products from the perspective of the user (i.e. it is a valid measure of perceived usability). It has been proven to be easy to administer and can be used on small sample sizes with reliable results. This scale covers many aspects for measuring usability (need for support, training and complexity), so that it can be applied to different tests.

  Depending on the test carried out, the standard SUS template will need to be adapted. If the participant manipulates a device during the test, then the standard template works fine and does not need to be modified. However, when usability is tested in cases where the participant is doing the test without interacting with the device used to carry out the test, then the statements will have to be adapted to the context in order to identify the level of usability of the component or element that is planned to integrated in the system (e.g. the customization of the subtitle layout: no device will be used by the participant during the test, but the aim is to investigate the level of usability of the customization of the subtitle layout). In fact, many tests will most probably deal with subtitles and the parameter/component/element manipulated and tested for pilots related to subtitling and audiovisual accessibility.

5.4. User tests outlook

Subsequent to the research lines specified above, UAB is planning the following additional user tests during the remaining duration of HBB4ALL:

- Subtitle adaptation to mobile devices: to determine which factors need to be adapted in the subtitles for mobile devices in order to achieve better understanding of the content. Different formal aspects of subtitles will be tested individually to achieve our objective, such as subtitle segmentation.

- 6 seconds rule verification: the six seconds rule (i.e. maximum time a subtitle stays on screen) is an industry standard that has never been tested. UAB will be testing it in different universities across Europe to establish if it is a valid 'good practice' standard for the maximum time a subtitle can stay on screen to improve the overall audiovisual experience.

- Personalization of subtitle layout utility: the aim of the study is to test whether the personalization of subtitle layout has an impact on the comprehension, immersion and enjoyment of audiovisual material.

[7] Link to the standard SUS template: http://www.usability.gov/how-to-and-tools/methods/system-usability-scale.html

- The next experiment for subtitle usability in different alphabets (i.e. Chinese) is to focus on the relationship between different subtitle variables and comprehension tests, on the one hand, and with different devices, on the other hand. Another aim will be to study in depth font size in subtitling and its impact on quality with another subtitling software that offers more options to personalise the font and that can be measured consistently.

- The University of Valladolid is currently undertaking another experiment regarding the optimal subtitle synchronization for inter-subtitle spacing. When looking at past studies, the recommendation was to leave the minimum gap of 2 frames between subtitles, so that viewers could notice the change of subtitles. This study aims at identifying, by empirical research, the optimal gap of frames for these cases. The gap is "language dependent". Different time frames will be tested, without considering shot changes.

- The next experiment regarding the cognitive processing of subtitled films will be a study of the EEG beta coherence (to be conducted this year) to determine whether immersion correlates positively with reduced beta coherence.

6. Conclusions

Workpackage 3/Pilot-A has advanced well, and in general is performing in line with the project plan and current market and industry developments. Since the writing of D3.1, the remaining work for both tasks T3.1 and T3.2 has progressed well; the main objectives for Pilot-B can be achieved and no further specific measures, actions or contingency plans are considered necessary at this stage.

This document overviews the large number of service components that have been developed and adapted for use in Pilot-A, as well as their integration for the respective sub-pilots. Many implementation details are given in the Annexes.

For all service components, versions are available that can already be used in the respective sub-pilots. Updates and/or additional features and functionality are planned in some cases and will be added before the start of some of the sub-pilots.

The partners have also made good progress on the pilot integration. EBU-TT-D plays a central role as a common solution and is used in many service components and sub-pilots. However, the individual workflows are implemented quite differently because of the different starting points (e.g. workflows that already existed at broadcasters) and the sub-pilot goals. That said, the solutions developed are generic enough (as they are based on open standards) to allow relative easy porting onto other platforms (e.g. in addition to HbbTV).

The partners in WP3/Pilot-A who have implemented EBU-TT-D in their components / workflow have intensively cooperated and exchanged results to assure interoperability of the various solutions. In addition, subtitles generated for the respective sub-pilots in the EBU-TT-D format have been exchanged to allow cross-sub-pilot testing.

Complementary user testing has been carried out up until M20 of the project on various issues of subtitle presentation, use and comprehension from the end-user point of view. The resulting recommendations regarding subtitling services go well beyond Pilot-A. Further user tests are planned in parallel to the sub-pilots.

Workpackage 3 / Pilot-A is a wide ranging workpackage addressing many aspects of subtitles. The partners have worked together to develop technology offering customised subtitles, improve the standardized production and distribution of subtitles to multiple devices, and address issues of language, comprehension and presentation. The partners are well prepared to start the operational phase and are optimistic it will verify the work done so far and offer positive results that can be used by broadcasters and content providers to increase the availability and accessibility of subtitles for those who need them.

7. References

[1] D2.3.2 – *Common Technical Components (II)*, HBB4ALL deliverable, October 2015

[2] D3.1 – *Pilot-A Progress Report*, HBB4ALL deliverable, December 2014

[3] EBU Tech 3381: Carriage of EBU-TT-D in ISOBMFF. Version: 1.0. October 2014. https://tech.ebu.ch/docs/tech/tech3381.pdf (Last accessed 2015/07/27)

[4] ISO/IEC 14496-12:2012/Amd 2:2014. Carriage of timed text and other visual overlays. January 2015. http://www.iso.org/iso/home/store/catalogue_tc/catalogue_detail.htm?csnumber=63188 (Last accessed 2015/07/27)

[5] ISO/IEC 14496-30:2014. Information technology -- Coding of audio-visual objects -- Part 30: Timed text and other visual overlays in ISO base media file format. March 2014. http://www.iso.org/iso/home/store/catalogue_tc/catalogue_detail.htm?csnumber=63107 (Last accessed 2015/07/25)

[6] EBU Tech 3380: TECH3380 EBU-TT-D Subtitling Distribution Format. Version 1.0. March 2015. https://tech.ebu.ch/docs/tech/tech3380.pdf (Last accessed 2015/08/10)

[7] D5.2 – *Pilot-C Solution Integration and Trials*, HBB4ALL deliverable, scheduled for submission in September 2015

8. Annexes

8.1. Annex: Service components

This Annex provides detailed descriptions for service components which were outlined in chapter 2.

8.1.1. Automatic Subtitling Component (VIC)

The main objective of the automatic subtitling component is two-fold: (1) it allows the automatic generation of EBU-TT format intra-lingual live subtitles within the HbbTV workflow from broadcast English breaking-news, and (2) it performs continuous communication with *Machine Translation Component* (see section 8.1.2) with the aim of sharing the recognized English sentences for their automatic translation into Spanish.

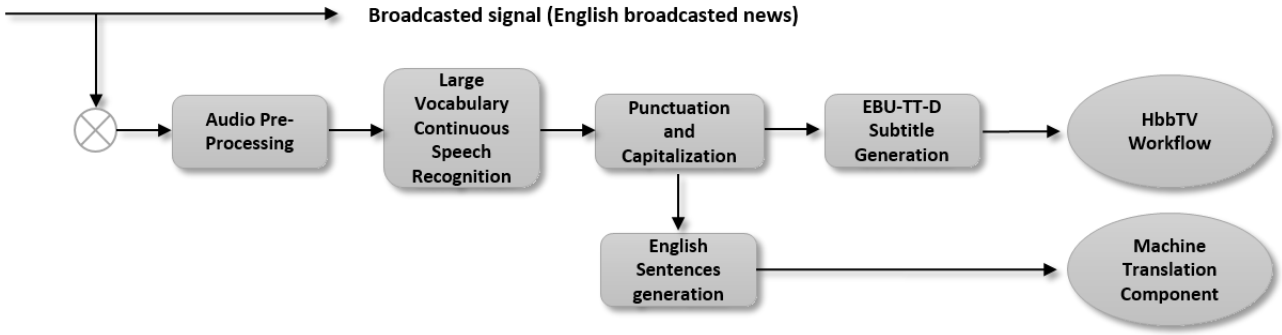

Figure 8. Automatic subtitling component system

As shown in **Figure 8**, this component is composed of a pipeline of processing modules, each providing a set of operational capabilities needed to automatically subtitle live audio-visual material. A description of each module is given below:

- The Audio Pre-Processing module discriminates between speech and non-speech segments and it is based on acoustic change detection information. Only the segments tagged as speech are consumed by the Large Vocabulary Continuous Speech Recognition module.

- The Large Vocabulary Continuous Speech Recognition (LVCSR) module transcribes the audio input stream according to an acoustic model, vocabulary and language model. The engine is based on a HMM-GMM (Hidden Markov Model – Gaussian Mixture Model) acoustic model with context-dependent phone states. It was trained using KALDI[8], a toolkit for speech recognition written in C++ and intended for use by speech recognition researchers under the Apache License. The language model (LM) is a trigram language model and it was trained using KenLM[9] toolkit. The final system composed by the acoustic model, the vocabulary and the LM was generated through KALDI toolkit.

  With regard to the corpora used to train the models and to define the final vocabulary, VIC used ~100 hours of planned speech from HUB97[10] and HUB98[11] corpora and 27 hours of spontaneous speech from

[8] http://kaldi.sourceforge.net/about.html
[9] https://kheafield.com/code/kenlm/
[10] https://catalog.ldc.upenn.edu/LDC98S71
[11] https://catalog.ldc.upenn.edu/LDC2002S10

the HUB97 corpus to train the acoustic model. The LM was trained over the transcriptions from the corpora described above, composed of ~1.5 million words, and texts crawled from the three data sources described below. The first two were collected from the OPUS repository (http://opus.lingfil.uu.se/) and using the English portion of the parallel corpora. The third corpus was distributed as part of the 9th Workshop on Statistical Machine Translation (http://www.statmt.org/wmt14/translation-task.html).

- News-commentary: A small parallel corpus of news commentary material. This corpus is the most representative of the news domain.
- MultiUN: A collection of official documents from the United Nations. The corpus is composed of professionally created translations and thus expected to be of a high quality.
- News-crawl 2012: A collection of texts extracted from news publication websites in the year 2012. The quality of the data in this corpus is not uniform and not all portions of it relate directly to the news domain.

To prepare a training dataset that could balance adaptation to the news domain and general language coverage, VIC performed data selection using a cross-entropy difference approach (Moore and Lewis 2010[12]). The news-commentary corpus was taken to represent in-domain data, as it is uniquely composed of news-related data. The news-crawl corpus, although also collected through crawling of online news publication sites, typically contains data either unrelated to the domain, due to crawling errors, or less directly related to the news domain due to the nature of the specific pages that were crawled. It was thus necessary to identify the portion of the corpus most related to the news domain in order to retain the most useful portions of the data for the task. Regarding the MultiUN corpus, the goal was also to identify subsets of the corpus most similar to the news domain data, although less for reasons of possible errors than for considerations of distance to the domain.

Individual language models were first created using the full monolingual data from each of the three corpora. Perplexity differences were then computed for each sentence of the News-commentary corpus on the one hand, and the MultiUN and News-crawl corpora on the other hand, using the corresponding language models. The sentences in the MultiUN and News-crawl corpora were then sorted according to perplexity differences, resulting in sorted corpora where sentences are distinguished by their closeness to the News-commentary corpus as well as their distance to the average of the MultiUN and News-crawl corpora, respectively. Finally, subsets of the sorted datasets were selected, to the tune of 5 million sentences for each of the MultiUN and News-Crawl corpora. These specific selection thresholds were empirically determined to balance general language coverage and adaptation to the news domain.

- The Punctuation and Capitalization module is in charge of automatically adding punctuation marks and capitalizing the words after a full stop. The Punctuation technology is based on acoustic information and statistical metrics to detect boundaries and automatically insert commas and full stop marks. It is capable of adapting its functionalities to each audio and speaker characteristics.

- Once the audio stream is recognized and the output text has been enriched with capitalization and punctuation marks, each of the sentences is converted to EBU-TT-D subtitle format. Each subtitle is then

[12] Moore, R. C. & Lewis, W. (2010). Intelligent selection of language model training data. In Proceedings of the ACL 2010 Conference Short Papers (pp. 220-224), Uppsala, Sweden.

inserted in the HbbTV workflow. In parallel, the sentences are shared with the *Machine Translation Component* in order to be translated to Spanish language.

Summarising, the automatic subtitling system was developed to work in real-time, generating live intra-lingual subtitles from English breaking-news content which is broadcast live. The subtitles are created once full sentences have been generated. To this end, sentences are split considering acoustic information related to silences and/or non-speech segments. When a short non-speech segment is detected, the recognized text is given by the LVCSR module. Once this output is enriched with capitalization and punctuation marks, the sentences are created. These sentences are inserted into the HbbTV workflow in order to provide intra-lingual subtitles. In addition, the sentences are sent to the Machine Translation Component using JSON (JavaScript Object Notation) objects as a data interchange format.

Software, Resources	
Name	*Further Info*
Kaldi	http://kaldi.sourceforge.net/about.html
KenLM	https://kheafield.com/code/kenlm/
OPUS	http://opus.lingfil.uu.se/
Spontaneous Speech from HUB97	http://www.statmt.org/wmt14/translation-task.html

Table 1. List of resources used for the Automatic Subtitling Component

8.1.2. Machine Translation Component (VIC)

This component allows the automatic translation of broadcast news subtitles, from English into Spanish, within the HbbTV workflow. EBU-TT-D is used as output subtitling format.

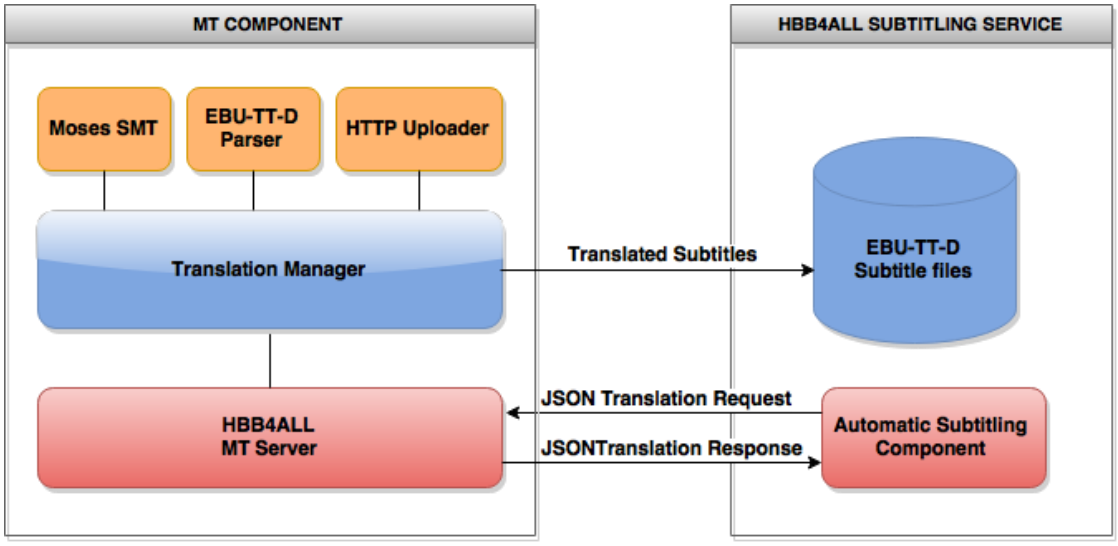

Figure 9. MT Component Architecture

The machine translation (MT) component interacts with the main HBB4ALL subtitling service (for The Automated HbbTV multilingual news broadcast subtitles sub-pilot, described in Section 3.6, see also **Figure 9**) through a service-oriented architecture whose main operations are described below:

1. Translation requests are generated by the *Automatic Subtitling Component* and sent to the RESTFUL API of the MT server. The server provides the appropriate response, which contains the status of the translation request, identifier and path of the translated files. The exchange format is JSON, with requests that are subtitle-oriented, as shown in **Table 2**.

TRANSLATION REQUEST		
FIELD	**DESCRIPTION**	**EXAMPLE**
audio_name	Output subtitle file name	*example_file-ES_0001.xml*
text	Subtitle to be translated	*This is an English subtitle.*
begin	Subtitle start timein hh:mm:ss.sss format.	*00:00:00.000*
end	Subtitle end time in hh:mm:ss.sss format.	*00:01:05.352*
client	Name of the request sender	*HBB4ALL*
model	Name of the translation model to be used	*hbb1_en_to_es*
hbb_time	Time already consumed for speech recognition (in milliseconds).	*0145*

Table 2. Translation requests for the HBB service

2. The Translation Manager module manages the translation process by executing automated translation of the subtitle text in the request. To do so, it communicates with a statistical MT decoder based on Moses, using the already trained translation model specified in the translation request. Once translated, the subtitle is automatically split if needed and an EBU-TT-D file containing the

translated subtitle is generated. This file is then sent to the HbbTV service through an HTTP upload and stored in the repository defined for machine translated subtitles.

The MT component has been adapted to the specific requirements of the HBB4ALL online processing scenario. Translation and casing models are for instance pre-loaded for higher processing efficiency, and the Automatic Speech Recognition (ASR) engine was set to deliver longer segments that would more closely match the training data used for the translation models. Additionally, an in-house algorithm was developed by VIC to split the subtitles generated by the MT component, using a combination of character count, segment length and ending constraints.

To train the machine translation model for the English to Spanish pair in the broadcast news domain, data were collected, selected and prepared using the two data sources described below, both being parallel corpora collected from the OPUS repository (http://opus.lingfil.uu.se/).

- News-commentary: A parallel corpus of news commentary material. Although small, with around 170K sentence pairs, this corpus is the most representative of the news domain.

- MultiUN: Composed of data from official United Nations documents, the corpus contains more than 11M sentence pairs and is expected to be of high quality, being based on professionally created translations.

To prepare a training dataset that could balance adaptation to the news domain and general language coverage, we performed data selection using a bilingual cross-entropy difference approach (Axelrod et al. 2011[13]). The news-commentary corpus was taken to represent in-domain data, as it is uniquely composed of news-related data. For the MultiUN corpus, the goal was also to identify subsets of the corpus most similar to the news domain data, for reasons of distance to the domain.

Individual language models were first created using the full monolingual data from each of the two corpora, for both the source and target languages, resulting in four language models. Bilingual perplexity differences were then computed for each sentence pair of the News-commentary corpus on the one hand, and for each sentence pair of the MultiUN corpus on the other hand, using the corresponding language models. The sentences in the MultiUN corpus were then sorted according to these perplexity differences, resulting in sorted corpora where sentences are distinguished by their closeness to the News-commentary corpus as well as their distance to the average of the MultiUN corpus.

Finally a subset of the sorted datasets was selected, using an empirically determined threshold to balance general language coverage, adaptation to the news domain and size of the translation model for online translation efficiency. Out of the 11+ million parallel sentences, only approximately 2 million sentence pairs were thus selected to train the translation model; the language model was trained on all data collected from the Spanish side of both corpora. A summary of the retained data can be found below:

- Parallel data:
    - In-domain (News-commentary):
        - 140,764 parallel sentences as training set;

[13]Axelrod, A., He, X. & Gao, J. (2011). Domain Adaptation Via Pseudo In-Domain Data Selection. In Proceedings of Empirical Methods in Natural Language Processing, Edinburgh, UK.

- 2000 sentence pairs as test set;
- 2000 as development set.

  - Out-of-domain (MultiUN):
    - 1,816,246 parallel sentences as training set;

- Monolingual data:
  - 11,555,134 Spanish sentences, from the News-commentary and MultiUN corpora.

The data were prepared using in-house tokenization and true casing models, and used to train two separate phrase-based models using the Moses toolkit. These two models were then combined through perplexity minimization on the News-Commentary development set, using the approach in Sennrich[14]. The final combined model was tuned using a 5-gram language model created from the entire selected monolingual data.

The results on standard automated SMT metrics are given below:

- BLEU: 37.3;
- METEOR: 62.1;
- TER: 44.3.

These results on the News-commentary test sets can be considered quite satisfactory, although the quality of translation is naturally lower when using the output of an ASR engine as input to the translation process.

Summarising, the MT component has been built to respond to the demands of the HBB4ALL scenarios, delivering on-line subtitle translation of the output of the automatic subtitling component in the required processing timeframe. An English to Spanish translation model was specifically created to cover the News domain as closely as possible given available corpora, taken into account quality goals as well as size of the model for efficient processing. The necessary tools for the entire processing chain were also created, from the development of an appropriate MT server and translation manager with a restful API, to the management of EBU-TT-D files containing automatically segmented subtitles.

Technology/Standards	
Technology used	*Task/Module*
RESTFUL API	Exchange of Translation Data between Machine Translation Server and Automatic subtitling

[14] Sennrich, R. (2012). Perplexity minimization for translation model domain adaptation in statistical machine translation. In Proceedings of the 13th Conference of the European Chapter of the Association for Computational Linguistics,Avignon, France.

	component
JSON	Exchange of Translation Data between Machine Translation Server and Automatic subtitling component.
Software, Resources	
Name	*Further Info*
OPUS	http://opus.lingfil.uu.se/

Table 3. List of technology, standards and resources used for the Machine Translation Component

8.1.3. Live Broadcast-Internet Subtitle Synchroniser (VIC)

This component will create specific timestamp mechanisms to match the broadcast video with Internet-based subtitle delivery and include a JavaScript library to be used in an HbbTV application to display subtitles for live TV.

The aim of this component is to create an HTTP media stream contributing a live broadcast content with subtitles generated automatically. This way, the generated stream can be then played on a player embedded in a web browser, such as an HbbTV application.

This component is responsible for creating a synchronised Internet multimedia stream, in MPEG-DASH format that comprises the broadcast video and the automatically generated subtitles, all aligned to a common time basis. The automatic subtitling component is formed by the integration of the *Automatic Subtitling Component* (see section 8.1.1) and the *Machine Translation Component* (see section 8.1.2) providing real-time automatically generated subtitles in different languages. On the other hand, the *Live Hybrid Broadcast-Internet Subtitle Synchroniser and Viewer* (see section 8.1.4) will enable the client player to consume an MPEG-DASH stream that includes the aggregated subtitles.

Regarding the formats, the data source will contain an MPEG-2 TS with H.264 encoded video and MPEG AAC audio. The generated MPEG-DASH MPD will manifest synchronised ISOBMFF fragments containing MP4 fragments that include H.264 encoded video, MPEG AAC audio and EBU-TT-D subtitles (considering EBU tech 3381 [3], ISO/IEC 14496-12 [4] and ISO/IEC14496-30 [5]). This way, no video transcoding or quality loss is introduced in the dataflow. However, the option of video transcoding has not been discarded, and the component design could include this feature with minor modifications.

This component does not consider the processing of the subtitles potentially present in the data source MPEG-2 TS stream. They will not be employed, and are excluded from the final MPEG-DASH stream.

Figure 10 summarises the component functionality and formats.

It is important to highlight the delay introduced by the intrinsic processing tasks for creation of the fragments of the MPEG-DASH stream and the necessary buffering of the audio to generate the subtitles. In case of the Internet stream generation, the lag compared with the broadcast stream is closely related to the size of the fragments and the window size, as setup in the manifest MPD, to enable the player buffering that mitigates networking fluctuations. For the automatic subtitle generation, it is also required to have a buffer that contains the past and future audio speech context to provide a consistent transcription, meaning a real-time processing that pulls bursts of subtitles with a variable delay.

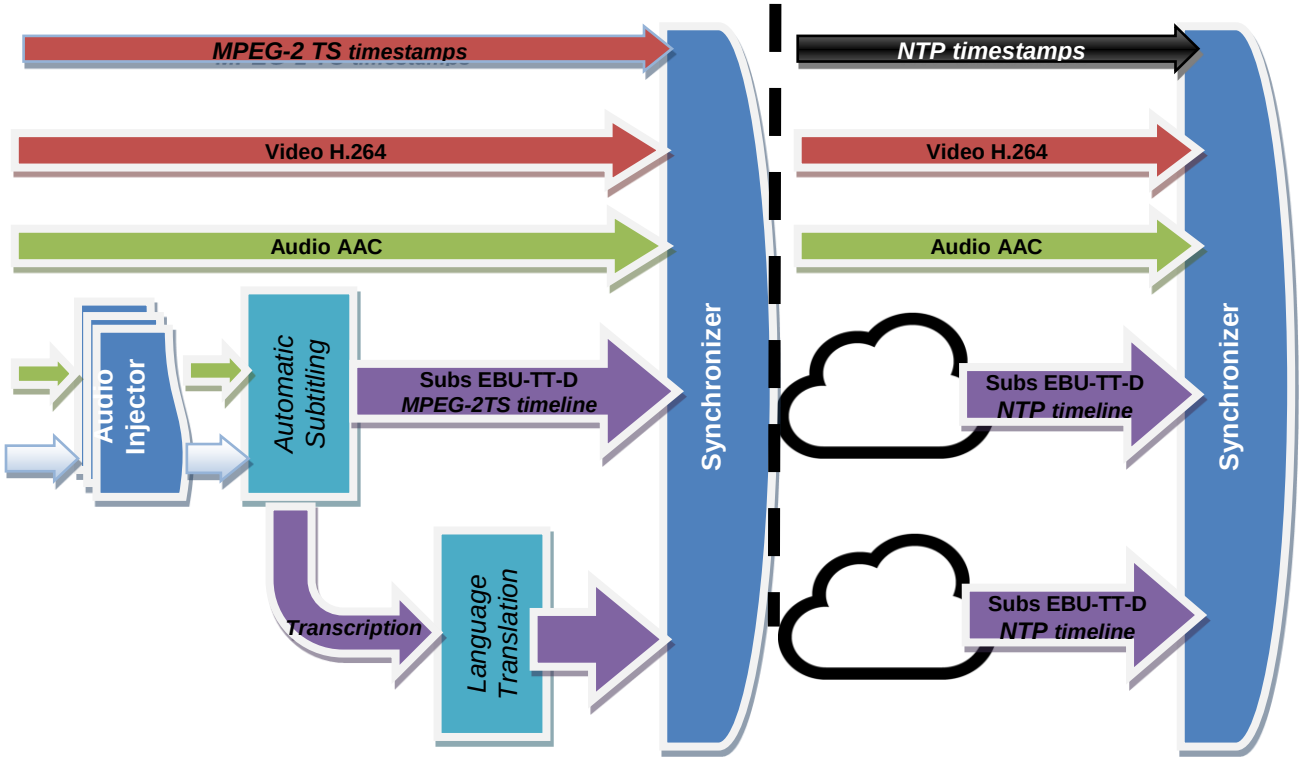

Figure 10. Synchronisation of video / audio with subtitles scenarios

Concerning the synchronisation of the video and audio with the subtitles, two different scenarios are being considered:

a) Subtitles generated automatically by this component invoking the *Automatic Subtitling Component*. Thus, the timestamps embedded in the MPEG-2 TS are used to signal the audio buffers injected to the automatic subtitling generator to have a common timeline with the video and audio streams and synchronise them properly.

b) Subtitles provided by a third party Internet service. In this case, the different data sources, the live broadcasted content and the live subtitles contributed from the Internet, do not share a time basis, thus making necessary the use of a NTP protocol to create a common timeline.

To overcome these scenarios the component design would comprise:

1) a parser of the broadcast video format to extract the media timestamps;

2) a demuxer of the live broadcast videos, splitting the encapsulated video and audio and dropping the embedded DVB subtitles;

3) a timestamped audio injector to the automatic subtitling component;

4) a parser of the automatically generated EBU-TT-D formatted subtitles;

5) a synchroniser that aligns the timestamps of the different media streams;

6) an MPEG-DASH video and audio fragment generator in ISOBMFF format;

7) an MPEG-DASH subtitles formatter and fragment generator in EBU-TT-D ISOBMFF format;

8) an MPEG-DASH live manifest generator and fragment manager;

and an MPEG-DASH player integrated in a Web browser.

The general scheme and workflow of this component is depicted in **Figure 11**.

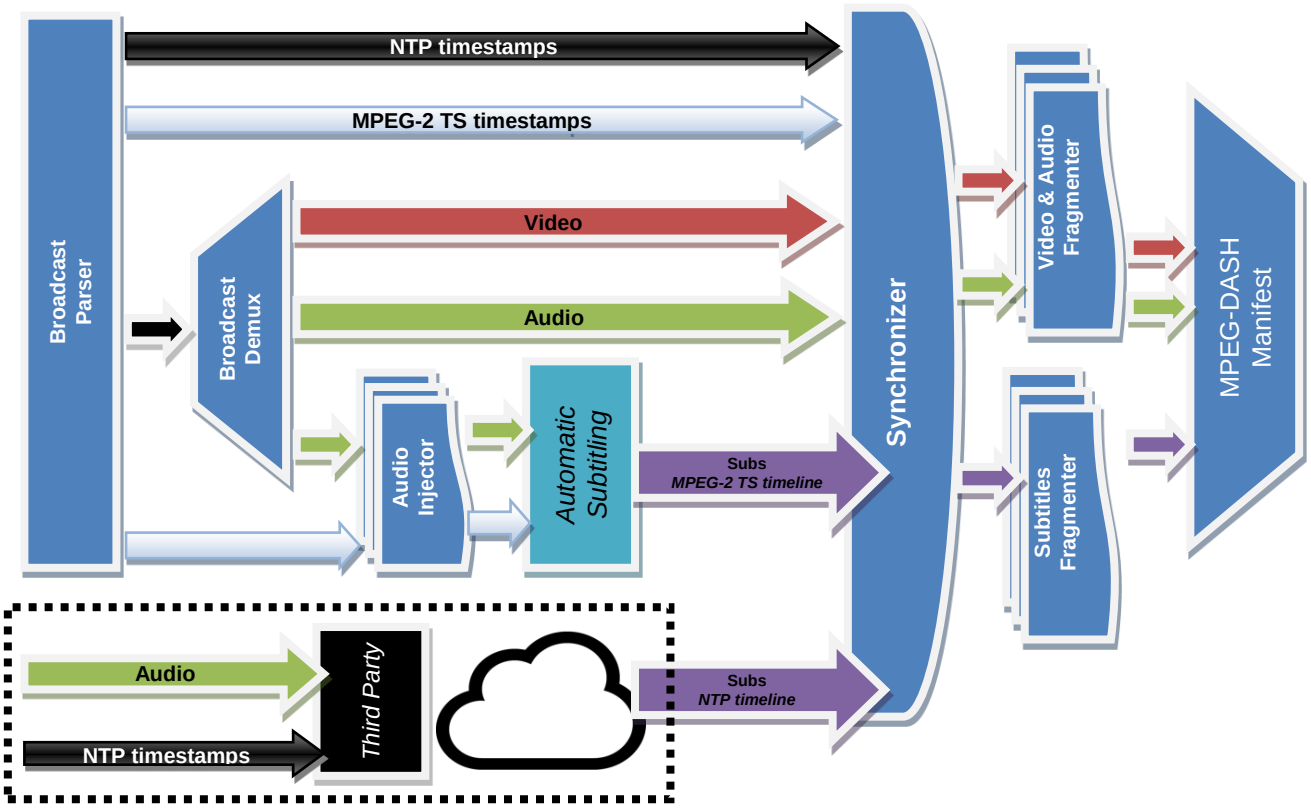

Figure 11. Live Hybrid Broadcast-Internet subtitle viewer and synchroniser

The broadcast parser will retrieve the MPEG-TS timestamps and the NTP clock to build a time basis that can be attached to the incoming video and audio signals. Thereby, this block gets the timestamps from the live

broadcast content or from a NTP server and provides it to the synchroniser in order to signal each media with a common timeline.

The demuxer of the broadcast signal MPEG-2 TS takes the embedded video (H.264) and audio (MPEG AAC audio) streams, dropping the embedded DVB subtitles, if any.

The audio injector ingests the audio samples with the MPEG-2 TS timestamps to assure a correct alignment of the EBU-TT-D subtitles with the timeline of the video and audio streams.

A parser of the subtitles will capture the live generated results coming from the automatic subtitling component or from a third party Internet provider and it will deliver them to the synchronization component. The format accepted by the subtitles parser is EBU-TT-D for live subtitling.

The synchronisation of the different media is needed because of the latency added by the automatic generation of subtitles which requires a sound buffer containing enough data, and time, to create a transcription. Therefore, the synchroniser is responsible for creating the buffers to queue the different incoming signals, the live broadcast video in MPEG-DASH format and the subtitle stream from the Internet subtitling service. Then, the synchroniser has to absorb the latency and the jitter from the subtitles employing the different timestamps to match the timing delays.

The different 'fragment generators' ('Fragmenter' in **Figure 11**) must include the proper time segments that create a synchronised subtitled experience. The fragmentation mechanism must be triggered by the business models and under terms of Quality of Experience according to some previously defined patterns.

The manifest generator must manage all the data and the fragments to create a fully standard compliant manifest MPD ISOBMFF.

Last but not least, a web browser player will request the MPD manifest and the corresponding fragments and updates in order to visualise synchronous subtitles, video and audio Internet stream.

The implementation has been done using existing Open Source technologies such as:

- Gstreamer: eases the live broadcast signals management, the parsing and de-multiplexing tasks, the buffer timing and synchronisation.
- DASH IF player: due to its support by the DASH community. It has an up-to-date player that supports different configurations to validate generated streams.

These technologies comprise the implemented pipeline that pivots around an MPEG-DASH server purely based on Gstreamer. This server accepts as input:

- a live source DVB with a MPEG-TS multiplexed stream containing a Video H.264 and Audio AAC
- a live HTTP URL folder with EBU-TT-D subtitles multiplexed in ISOBMFF

This server processes the inputs to:

- create the fragments and store them inside an Apache server
- aggregate subtitles from different languages from third parties with EBU-TT-D subtitles multiplexed in ISOBMFF

The timing of the DASH server (the MPD and the published fragments) can be set:

- according to the Timestamps from the MPEG-TS of the DVB signal
- according to a NTP clock to synchronize the DVB signal reception and the remote subtitles generation from a third party

This server produces as output:

- an MPD with a on demand profile

The main actors of the server are:

- MP4DASHMUX Gstreamer plugin. ISOBMFF(mp4) multiplexer fragments (M4S) for Video (H.264), Audio (AAC) and Subtitles (EBU-TT-D), where each fragment only contains video or audio or subtitles. This plugin integrates the modifications reflected in EBU tech 3381 [3], ISO/IEC 14496-12:2012/Amd.2:2014 to carry timed text streams [4] and ISO/IEC14496-30 for the carriage of EBU-TT-D in ISOBMFF [5].
- DASHSINK Gstreamer plugin. DASH server to publish a manifest (MPD). The MPD publishes all of the video, the audio, and the subtitles in the setup languages.

In **Figure 12** the pipeline and key parameters are depicted:

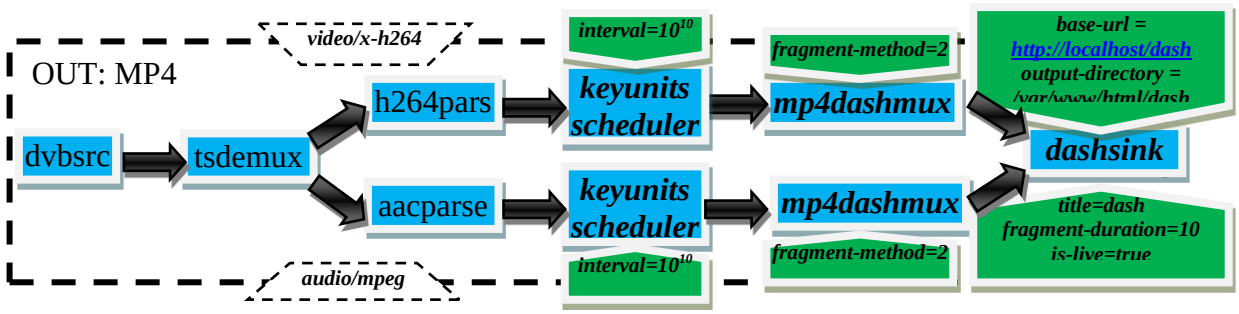

Figure 12. Server Pipeline

Therefore the steps for the subtitles would be:

The *Automatic Subtitle Generator* directly generates texts and timestamps.

1) It publishes EBU-TT-D files corresponding to some seconds of the content and signalled according to a common time basis shared between the *Automatic Subtitle Generator* and the rest of the DASH pipeline.

2) It writes the files with a correlative pattern shared by a HTTP server onto a unique URL folder.

3) Then, the DASH pipeline reads them and integrates them into the DASH pipeline, to convert them into ISOBMFF + EBU-TT-D fragments according to the video and audio fragments and timestamps (in this case all the participants have the same clock).

4) Last the DASH server publishes the fragments into a MPD to provide DASH subs.

The solution is fully modular to enable flexibility.

Technology/Standards	
Technology used	*Task/Module*
MPEG-DASH, H264, AAC	Contribution and Distribution
MPEG-2 TS	Source Signal
EBU-TT-D	Subtitle Format
Software, Resources	
Name	*Further Info*
Gstreamer	http://gstreamer.freedesktop.org/
DASH IF player	http://dashif.org/software/

Table 4. List of open technology and resources used for the Live Broadcast-Internet subtitle synchroniser

8.1.4. Live Broadcast-Internet subtitle viewer (VIC)

This player is a modified version of the open source DASH IF client. It accepts as an input:

- an MPD with an on demand profile

This player processes the inputs to perform:

- fragments of video, audio and subtitles for downloading
- basic subtitle processing to turn (parsing just the timestamps) EBU-TT-D subtitles, multiplexed in ISOBMFF, into HTML5 browser standard subtitles

Technology/Standards	
Technology used	*Task/Module*
MPEG-DASH, H264, AAC	Distribution formats
HTML 5, CSS, JavaScript	Client side player Components
EBU-TT-D	Subtitle Formats
Software, Resources	
Name	*Further Info*
DASH IF player	http://dashif.org/software/

Table 5. List of open technology and resources used for the Live Broadcast-Internet subtitle viewer

8.1.5. Subtitle interpretation tool for browsers (UPM)

The functionality of this service component consists of the interpretation of EBU-TT-D subtitles and the depiction of these subtitles on the PC screen. A significant change has happened in the development of this component since previous deliverables. The previous work included an embedded specific player to integrate this development in a webpage. However, the current version is built on JW Player as a plugin. JW Player is the web player currently deployed on the RTP website. This component is designed for the Portuguese subtitling service pilot (as described in section 4.4). It was decided to move to using JW Player, as this is a common and widely deployed player, and thus this service component could be used by other HBB4ALL partners and other broadcasters and content providers worldwide. In this way, the service pilot will be open and available for any user who visits the RTP website.

Among the development methods supported by JW Player, this service component is based on JavaScript to achieve a larger compatibility with HTML5.

This component is designed to depict subtitles in PC and other IT devices, e.g. smartphones, taking advantage of the support of JW Player in the browsers that those devices embed. The subtitles are received as an XML file (created by using the *Subtitling Format Conversion Framework* described in section 2.1.3) in EBU-TT-D format and this tool is able to interpret the subtitles and the styles.

Technology/Standards	
Technology used	*Task/Module*
HTML5, CSS, JavaScript	Videoplayer Components for Subtitle Rendering
EBU-TT-D	Subtitle Format
Software, Resources	
Name	*Further Info*
JW Player	http://www.jwplayer.com/

Table 6. List of open technology and resources used for Subtitle interpretation tool for browsers

8.1.6. HbbTV-based Customisable Subtitles for VoD Playback (RBB)

The HbbTV-based customisable subtitle for VoD playback component allows for an interpretation of EBU-TT-D subtitles and provides synchronised subtitles for VoD and live video in HbbTV devices. In addition to end-user customisation of the subtitle representation on the screen, fast forward/rewind is also supported.

The service component HbbTV-based customisable subtitle for VOD playback consists mainly of an EBU-TT parsing module, the enhancement of the player and the HbbTV Mediathek application's GUI for the HBB4ALL German sub-pilot. The client-side part of the system is able to receive and parse subtitle data and split and transform it into JSON packets. These packets can then be handed over to the actual player module, which can easily add the text bits from the JSON packets to the correct video fragment.

Technology/Standards	
Technology used	*Task/Module*
EBU-TT-D	Subtitle parser
JSON	communication between subtitle parser and player module

Table 7. List of open technology and resources used for the HbbTV-based customisable subtitles for VoD playback

8.1.7. IPTV Subtitle Publishing System (TVC)

For providing multiplatform subtitling services a subtitle publishing system is needed in order to expose the subtitles, using the EBU-TT-D format, to IPTV systems. This component usually consists of an HTML engine with a playback mechanism, but in some cases, like mobile environments, more native components can be used additionally. TVC's platform uses the HTTP protocol to make available the published subtitles to the client side application that includes among others the playback component and the synchronization mechanism, commonly JavaScript. A key advantage of this subtitle publishing system is the interoperability between platforms, as the published subtitles should be valid for all platforms without having to replicate infrastructure components for each case.

Technology/Standards	
Technology used	*Task/Module*
EBU-TT-D	Subtitling Format
JavaScript	playback component and synchronization system

Table 8. List of open technology used for the IPTV subtitle publishing system

8.1.8. HbbTV Device and Version Detection (TVC)

Detecting the specific device and the supported HbbTV version is necessary to ensure that the features used by the running HbbTV application are compatible with this specific device and will work correctly. This detection is done by retrieving the device's user agent string.

If some functionality of the HbbTV application requires a higher HbbTV version than the one implemented on the device these functionalities will not work and using that functionality could make the application crash. Ideally, the application should determine if this is likely and adapt to use alternative ways to execute the same operations or try to avoid using them, if possible. And, if that's not possible, the HbbTV version detection could be useful in determining the cause of problems or bad behaviour in devices.

In the same way, the device detection can determine a specific TV brand, or device vendor, which might have incorrect behaviour or known open issues that affect all devices of this brand similarly. With knowledge of which device which is running the HbbTV application, some vendor-specific corrections or adaptations of its functions could be added to the application, to avoid problems if possible.

Technology/Standards	
Technology used	*Task/Module*
JavaScript	HbbTV device and version detection

Table 9. List of open technology used for the HbbTV device and version detection

8.1.9. EBU-TT-D HbbTV Personalised Subtitle Playback (TVC)

To fulfil the personal requirements of the users who are the target audience for an accessibility subtitle service, the subtitle playback system on HbbTV devices should support subtitle personalisation. TVC has developed a component to fulfil that requirement. By default, the subtitles initially presented to the viewers are the non-personalised "basic" subtitles.

Currently transport is done via HTTP; no MPEG-DASH is used. The EBU-TT-D subtitle and the MPEG-4 AV file are both transported separately. In order to solve the synchronization playback problems a JavaScript time-based synching mechanism has been developed and adapted.

Technology/Standards	
Technology used	*Task/Module*
MPEG-DASH	Distribution Format
EBU-TT-D	Subtitle Format

Table 10. List of open technology used for the EBU-TT-D HbbTV personalised subtitle playback

8.1.10. vSubtitle: Adaptive Subtitle Rendering (VSX)

The vSubtitle component is available as a standard add-on for Vsonix' vPlayer, a webcast player that was designed for the playback of video lectures. E.g. for the realization of MOOCs or to be used corporate learning platforms. It was developed in the course of the HBB4ALL project. The vSubtitle component can be used as a standard subtitle renderer with no adaptation or personalisation features as well as with personalised subtitle rendering. For this purpose we have successfully added support for the UI and subtitle adaptation framework that was developed by Vsonix in WP5. The framework provides the user profiling

application called AccessGUIDE (see section 8.2.5 and D5.2 [7]), that could be integrated by developers into their services to determine e.g. the user's subtitle rendering preferences.

The vSubtitle component as part of vPlayer handles synchronization, rendering and adaptation of subtitles. vSubtitle was designed as an addon for vPlayer since it requires extensive communication in order to keep in sync with the media that is displayed by vPlayer.

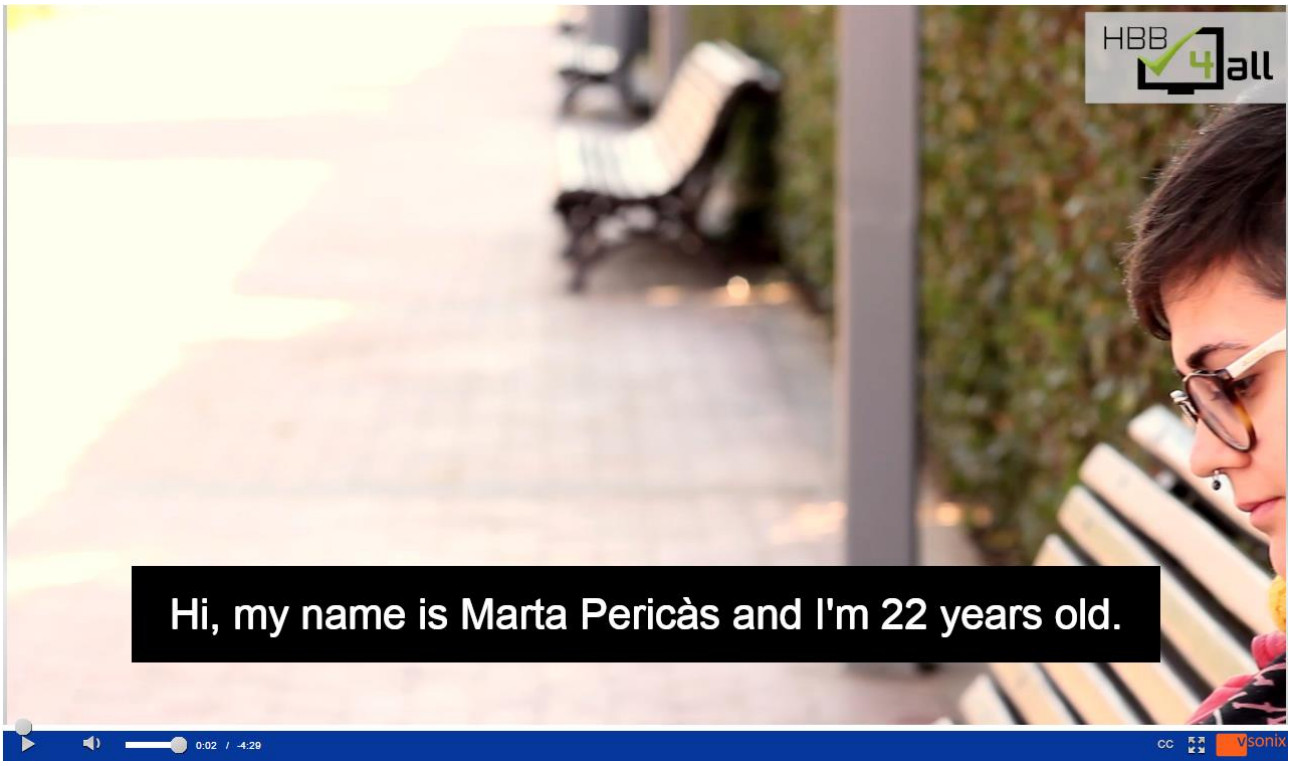

Figure 13. vSubtitle component as part of the vPlayer

vSubtitle allows users to adapt the subtitle display according to their user profile, which has to be established using the AccessGuide application beforehand. It can be enabled and disabled manually, and will automatically establish the correct adaptation values by querying the UI adaptation service upon initialisation. The communication to the UI adaptation service as well as the adaptation of the subtitle display take place using asynchronous communication, so that no noticeable lag or drop in application responsiveness can be noticed. Of course, vSubtitle can also be used without the enhancements provided by the UI adaptation framework or in the case where no user profile has been created by the user. It will then fall back to well-defined default values for all adaptation parameters.

vSubtitle provides support for multiple subtitle tracks (commonly one track per language), but allows for at most one track to be displayed at any one time. The user may switch subtitle tracks at any time or switch them on or off altogether with the operation effective immediately, so that the user does not experience any delay and media playback is not affected by the vSubtitle component. The subtitles are loaded upon initialisation so that the component is able to provide instant feedback to the user whenever a switch in subtitle track has been requested and no delay in loading, parsing and queueing can be observed.

A first version of the vSubtitle component has been designed, implemented and tested with several videos. Furthermore, the vSubtitle component has been integrated with the UI adaptation service. Applicable adaptation criteria such as minimum legible font size, subtitle position and subtitle font family can be queried, retrieved and applied to the subtitle display if the user created a respective user adaptation profile with the adaptation framework prior to using the vSubtitle component.

Currently, vSubtitle supports subtitle metadata in the WebVTT file format only and does not provide direct support for other formats, such as the concurrent SRT subtitle file format or EBU-TT-D. SRT is only indirectly supported by using one of the available third-party tools to convert SRT into the WebVTT file format. Support for EBU-TT-D is not planned since it incorporates semantics and style, which interferes with the styles that are applied to the subtitle display based on the profile received by the UI adaptation service.

It is planned for the future that the vSubtitle component automatically shows the subtitle track that corresponds to the user's language as set in the adaptation profile, as opposed to making the user select the track they can understand best manually.

Technology/Standards	
Technology used	*Task/Module*
WebVTT	Subtitle format

Table 11. List of open technology used for vSubtitle

8.1.11. HbbTV MPEG-DASH/ISOBMFF Segmenter and Packager (IRT)

The EBU cooperated with GPAC to create an MPEG-DASH based workflow which is based purely on open source software (aka MP4box). A respective showcase was presented at IBC 2014. However, this development was not focused on the MPEG-DASH profiles defined in HbbTV 1.5 and 2.0. As a result, when creating content for HbbTV, manual modification of the MP4box output was required. Also at a very base level MP4box was able to include TTML files in the ISO Base Media File Format (ISOBMFF) which is required by MPEG-DASH. This was a first step to be able to generate video content with embedded subtitles.

IRT, as part of its HBB4ALL activities, subcontracted GPAC to make MP4box compliant to directly support HbbTV formats. GPAC extended MP4box software, while IRT defined requirements and tested the developments. As a result, there is now a new profile switch for HbbTV 1.5 in MP4box when generating MPEG-DASH streams. A currently ongoing task is to check that MP4box is fragmenting and packaging EBU-TT-D documents correctly, according to the specification referenced by HbbTV.

The following command line examples showing the use of ffmpeg and MP4box show how a compliant MPEG-DASH stream can be generated with those tools:

1. *Encoding video with ffmpeg*

```
ffmpeg -i input.avi -pix_fmt yuv420p -aspect 16:9 -c:v libx264  -b:v 1536k -profile:v high -r 50 -force_key_frames "expr:gte(t,n_forced*2)" -keyint_min 50 -g 50 -maxrate 1536k -bufsize 2M -preset slow -x264opts nal-hrd=vbr:sps-id=70450:scenecut=-1  -an -sn output_pass1.264
```

This encodes a video clip to just one video elementary stream of approximately 1.5 Mbit/s. The codec used is H264. The GOP size is 1 second and it is enforced that an I-frame is generated every second by using multiple parameters (-force_key_frames, -keyint_min, scenecut). For more bitrates this call should be repeated with different bitrates. The resolution might be changed between those supported by HbbTV.

2. *Embedding EBU-TT-D document in ISOBMFF*

```
mp4box.exe -add subtitles.xml:ext=ttml  subtitles.mp4
```

It is as simple as this, just use the –add option to define the source file including a hint that the format is TTML and the name of the output file.

3. *Generating MPEG-DASH stream*

```
mp4box.exe -dash 6000 -profile hbbtv1.5:live -profile-ext urn:hbbtv:dash:profile:isoff-live:2012 -segment-ext m4s -out manifest.mpd video.m4v:role=main audio.m4a:role=main subtitles.mp4:role=subtitle
```

This will generate an MPEG-DASH stream compliant with HbbTV 1.5 including the HbbTV 2.0 feature EBU-TT-D (if the TTML file is compliant). The segment length will be 6 seconds, in case multiple representations are needed, the video tracks with different bitrates should be added. Of course the encoding must be done in the proper way to allow for seamless switching by the client.

The profile switch ensures that all settings are set as required for HbbTV 1.5. The additional profile also signals compliance with DVB DASH which is used in HbbTV 2.0.

Technology/Standards	
Technology used	*Task/Module*
MPEG-DASH	Contribution and Distribution
EBU-TT-D	Subtitle Format

Software, Resources	
Name	*Further Info*
MP4box	https://gpac.wp.mines-telecom.fr/mp4box/
ffmpeg	https://www.ffmpeg.org/

Table 12. List of open technology and resources used for the HbbTV MPEG-DASH/ISOBMFF Segmenter and Packager

8.1.12. SCF-Subtitling Conversion Framework (IRT)

The Subtitling Conversion Framework (SCF) supports the conversion between broadcast subtitle file formats and the subtitle file formats for broadband distribution that are used in the HBB4ALL pilots and beyond. The main focus is the EBU-TT-D format, one of the main technical components in HBB4ALL.

The start format for the conversion is an EBU STL file. EBU STL is possibly the most widely used subtitle archive and exchange format in Europe. In addition to the conversion to EBU-TT-D, the translation into the subset EBU-TT-D-Basic-DE (used by e.g. RBB) and to the archive and exchange document profile EBU-TT Part 1 is also supported.

The SCF is structured as different modules where each module implements one conversion step. The conversion modules that are currently supported are:

- EBU STL to STLXML (STLXML is an XML representation of EBU-STL[15])
- STLXML to EBU-TT Part 1
- EBU-TT Part 1 to EBU-TT-D
- EBU-TT-D to EBU-TT-D-Basic-DE

[15] The structure of STLXML is fixed in a W3C XML schema

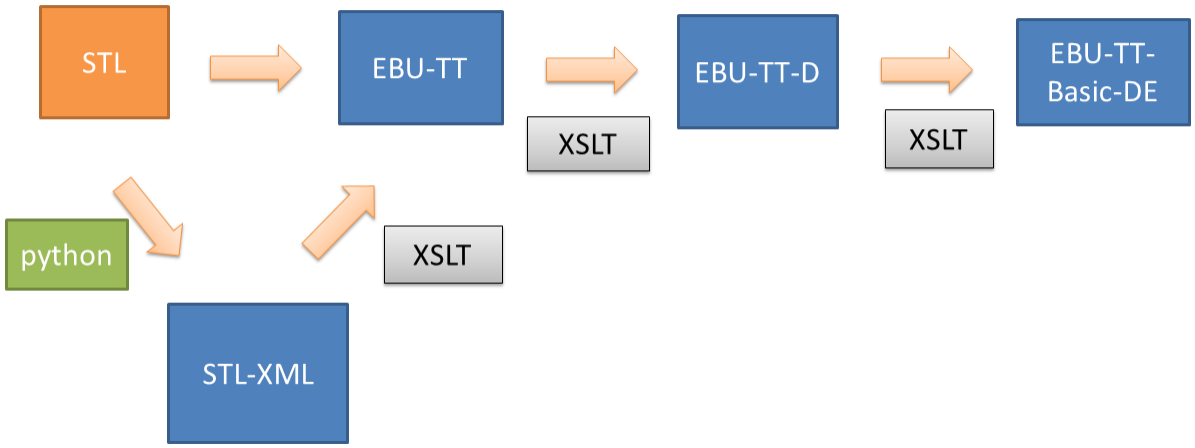

Figure 14. Subtitling Conversion Framework – implemented transformations

For the transformations, the Python scripting language and Extensible Stylesheet Language Transformations (XSLT 1) are used as technology; this is illustrated in **Figure 14**. With the use of these widespread and open technologies, adoption on a wide range of different platforms is possible.

For the validation tests the technologies used are W3C XML Schema and Schematron. To allow wide adoption the framework is published under a free open source license (Apache 2) on github.com.

The SCF is being used by these partners for their sub-pilots:

- RBB (see section 4.1)
- TVC (see section 4.2)
- RTP /UPM (see section 4.4)

The SCF is also proposed as an EBU-TT reference implementation. Therefore not only the source code but also the test files and requirements have been published as open source.

Technology/Standards	
Technology used	*Task/Module*
EBU-TT Part 1	Subtitle Format
EBU-TT-D	Subtitle Format

EBU-TT-D-Basic-De	Subtitle Format
EBU-STL	Subtitling Format
Python	Format Transformation
XSLT	Format Transformation
DITA	Documentation
Schematron	Unit Tests
Software, Resources	
Name	*Further Info*
SCF	https://github.com/IRT-Open-Source/scf
to_srt.py	https://github.com/yanncoupin/stl2srt

Table 13. List of open technology and resources used for the Subtitling Conversion Framework

8.1.13. Lightweight Subtitle Editor (IRT)

The production of subtitles for broadband distribution involves many people and often only small changes need to be made to the subtitle files that are pre-produced for broadcast. Furthermore, the possibilities to improve subtitle presentation for broadband distribution are greater than for broadcast production.

With EBU-TT-D and XML as format and technology for subtitle distribution the editing possibilities open up.

IRT has therefore started collaborating with the software development company SyncROsoft SRL from Romania which publishes the popular XML editor oXygen. Based on information provided about the EBU-TT format, SyncROsoft already has published a first version of their editor that allows the re-purposing of EBU-TT Part 1 subtitle files, see **Figure 15**.

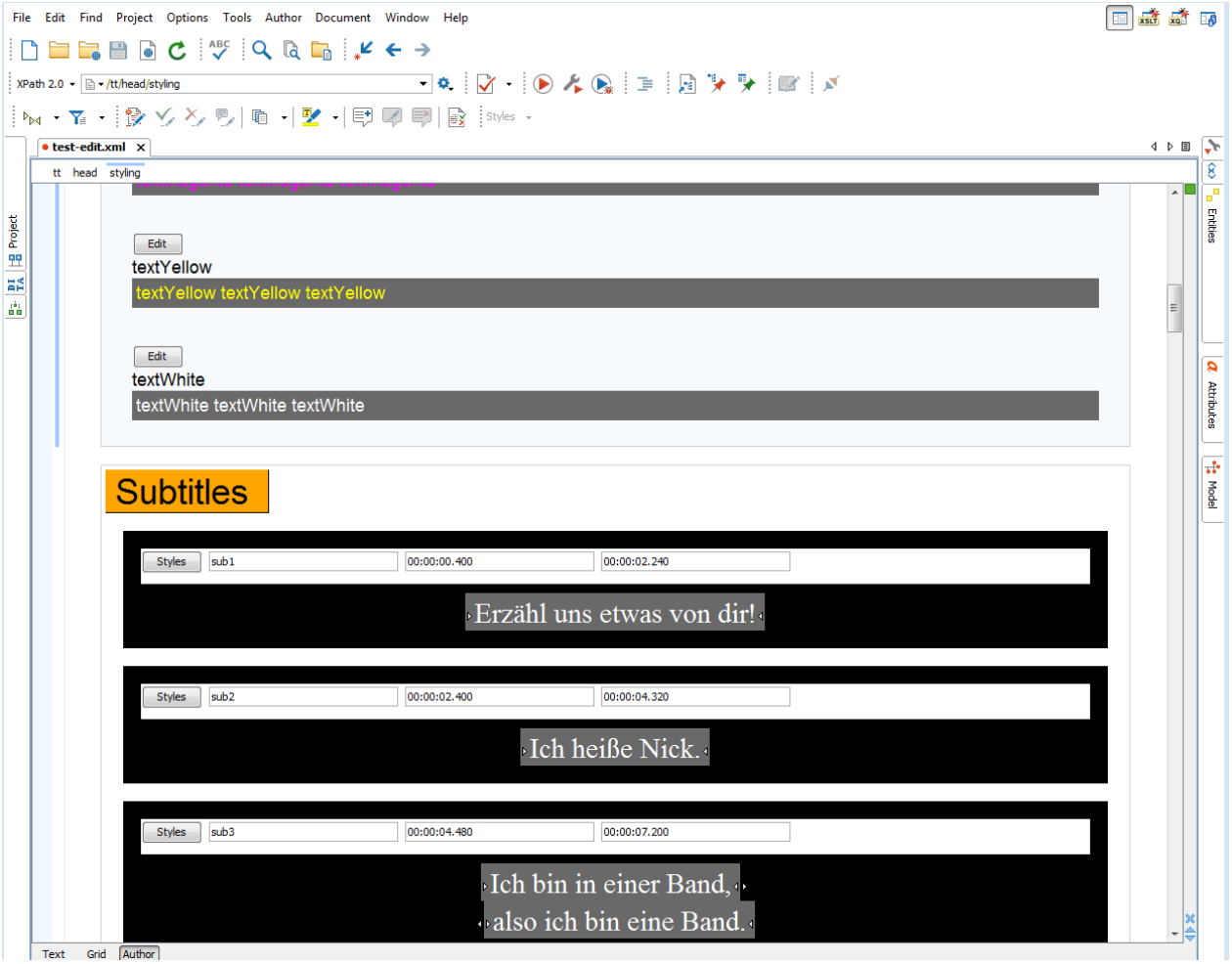

Figure 15. oXygen XML editor with support for EBU-TT Part 1

IRT started to build on this work to make this also accessible for EBU-TT-D production. It is planned to test this editor with the partners in HBB4ALL.

The oXygen version with EBU-TT Part 1 support was published in April 2015, and the implementation work for an EBU-TT-D version has started; a first prototype should be ready end of 2015. The final release of the editor is planned for 2016.

Technology/Standards	
Technology used	*Task/Module*
CSS	Customizing of Interface

Software, Resources	
Name	*Further Info*
oXygen Editor	http://www.oxygenxml.com/
to_srt.py	https://github.com/yanncoupin/stl2srt

Table 14. List of open technology and resources used for the Lightweight Subtitle Editor

8.1.14. EBU-TT-D XML Schema (IRT)

For the introduction of new standard technologies in electronic data exchange, a standard conformant implementation is essential. Only if software systems can safely trust that the data received conforms to the standard, do the possibilities for interoperability open up.

One main technology used in all pilots and activities of WP3 is the XML Subtitle format EBU-TT-D. The key technology that is used to verify that an XML Document conforms against a specific standard is W3C XML Schema. At the time of the publication of EBU-TT-D no W3C XML Schema was available. IRT therefore implemented an XML Schema and provided this to the HBB4ALL partners to test their implementations. Because W3C XML Schema is supported in a wide range of software environments the validation could be easily implemented by the partners.

IRT contributed this work to the EBU and the XML Schema is now published as draft under a free open source license on github[16]. It can be used by anybody working with EBU-TT-D. In their role as an EBU member, IRT is still the editor of this XML Schema.

[16] https://github.com/ebu/ebu-tt-d-xsd/tree/editors-draft

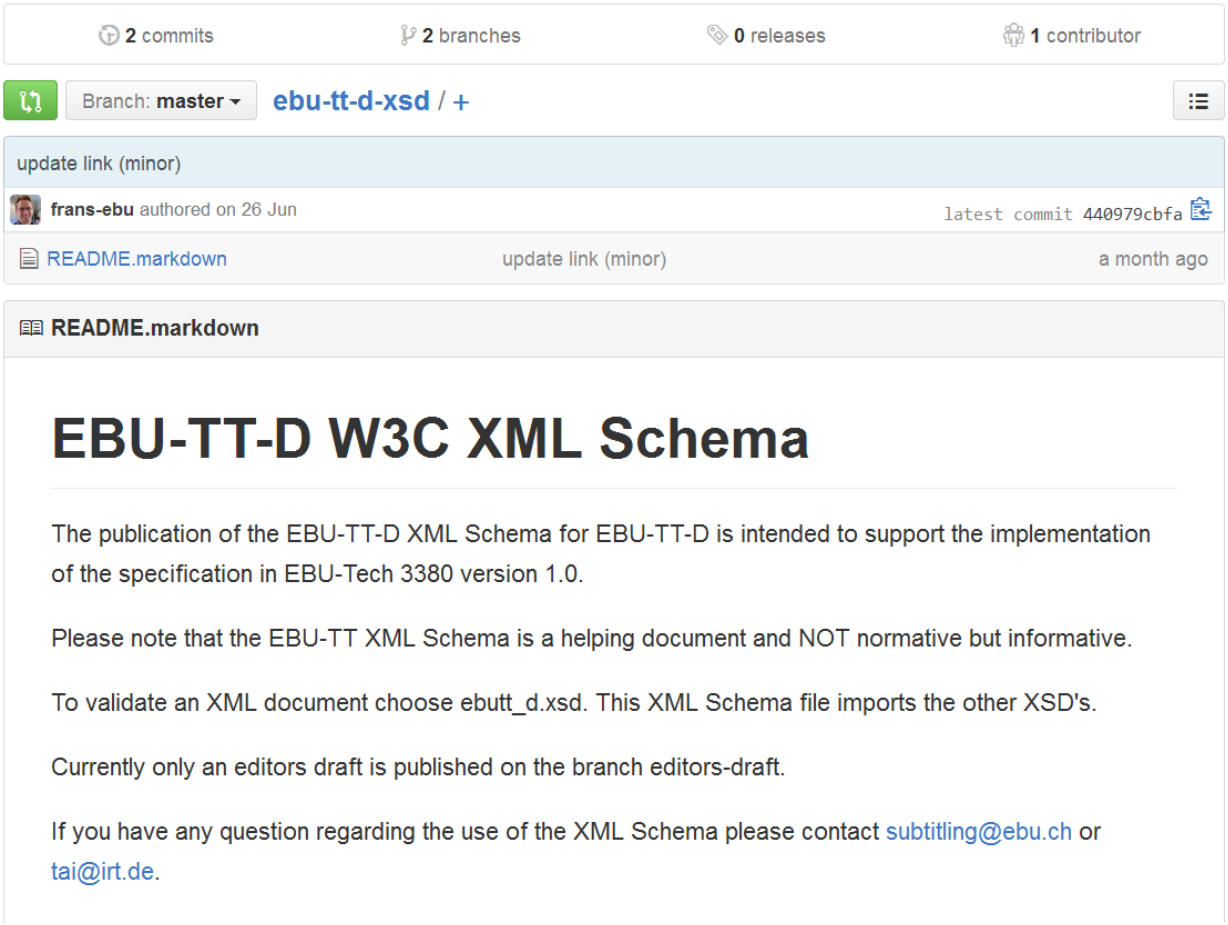

Figure 16. EBU-TT-D W3C XML Schema for validation of EBU-TT-D XML documents

The XML Schema has still to be reviewed by the Standards Working group and it is currently published as draft. It is expected to be finished in 2016.

Apart from the XML Schema, IRT further investigated the possibilities for the validation of EBU-TT-D. Amongst others, a master thesis was written on this theme[17]. Building on this work, a module in the context of the *Subtitling Conversion Framework* was published by IRT. This module implements further checks on an EBU-TT-D document instance that cannot be checked by an XML Schema.

[17] "Strategies for User-Oriented Conformance Testing of XML Documents", Barbara Fichte, IRT 2014, see https://www.irt.de/de/irt-studenten/abschlussarbeiten.html

Technology/Standards	
Technology used	*Task/Module*
W3C XML Schema	EBU-TT-D Validation
Schematron	EBU-TT-D Validation
Software, Resources	
Name	*Further Info*
EBU-TT-D XML Schema	https://github.com/ebu/ebu-tt-d-xsd
SCF	Modul

Table 15. List of open technology and resources used for the EBU-TT-D XML Schema

8.1.15. EBU-TT-D Rendering Tests (IRT)

While the published *EBU-TT-D XML Schema* (see section 8.1.14) guarantees that the produced XML document conforms to the standard, it is not guaranteed that a presentation processor will make a conformant rendering of the EBU-TT-D document. This check has to be done manually. For the manual review, reference material is necessary to compare to the presented material.

IRT has built a set of reference tests to accomplish this. The test suite has the following structure:

- EBU-TT-D sample files in XML.
- PNG format images that show the expected rendering of the EBU-TT-D sample.
- Video that contain a 30s long 1280x720px mp4 video with the rendered PNG as still image.
- A test video with following structure
    - Title page
    - Short info about the text context
    - Display of the image how the subtitles should be rendered

- Blank screen (only the rendered EBU-TT-D subtitles should be shown)
- 2x changing blank screen and pre-rendered subtitles to double check possible observations during testing.

Some key EBU-TT-D features are already covered. The final test set is planned for release in 2016.

Technology/Standards	
Technology used	*Task/Module*
XSL-FO	Rendering of Test Images
EBU-TT-D	Subtitle Format

Table 16. List of open technology used for the Live Broadcast-Internet subtitle viewer

8.1.16. HbbTV 2.0 Demo App (IRT)

As documented in HBB4ALL deliverable D2.3.2 ([1] chapter 6.1), IRT intensively contributed to the HbbTV 2.0 as well as the EBU-TT-D specification, which enables applications to include subtitling for live streaming and video on demand via IP. After the specifications were published early 2015 and the HbbTV 2.0 test assertion definition was also completed, IRT focused on further cooperation with the industry to encourage a wide adaptation of the new specifications. For this purpose a showcase has been planned that includes live streaming on the Internet, using MPEG-DASH streaming and EBU-TT-D subtitles. At the time of writing, a presentation of the demo is planned at both IFA 2015 (Berlin) and IBC 2015 (Amsterdam). As partners from the industry, Samsung prepared an HbbTV based prototype integrated digital television (IDTV) with support for MPEG-DASH and EBU-TT-D, and from the encoder manufacturer side three different manufacturers have implemented HbbTV 2.0 support. RBB supplied sample A/V and subtitle material for the demonstrations.

For interoperability testing as well as for presentation of the showcase, IRT prepared an HbbTV application, which allows a user to navigate through available on demand content with so-called ‘out-of-band’ subtitles and MPEG-DASH streams with embedded subtitles. The application provides easy access for TV manufacturers to the available demo content.

The on demand content available for testing was generated by IRT, mainly small test cases checking for a particular rendering feature of EBU-TT-D. A test case consists of a video with a TTML file, where the video shows how a subtitle should look like when rendered by the TV. For further details see the previous chapter.

Streams containing subtitles embedded in MPEG-DASH were provided by various encoder manufacturers, thus supporting interoperability testing as well.

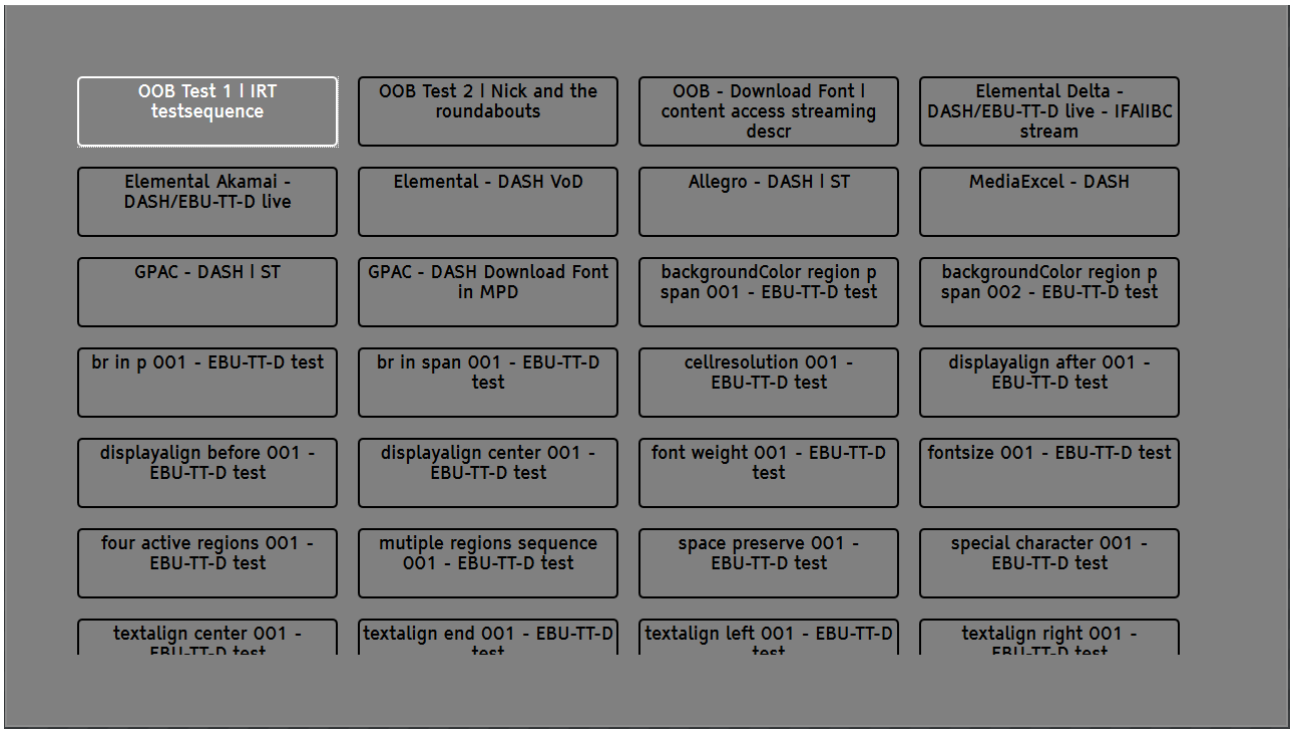

Figure 17. HbbTV EBU-TT-D Test App for evaluation of various DASH streams and EBU-TT-D features

While the application for testing mainly focusses on quick access to the test content (see **Figure 17**), the design for the presentation at IFA and IBC has been adapted to the layout of the RBB catch up service "RBB Mediathek" (see **Figure 18**).

The application is available on http://hbbtv-live.irt.de/ebuttd/showcase

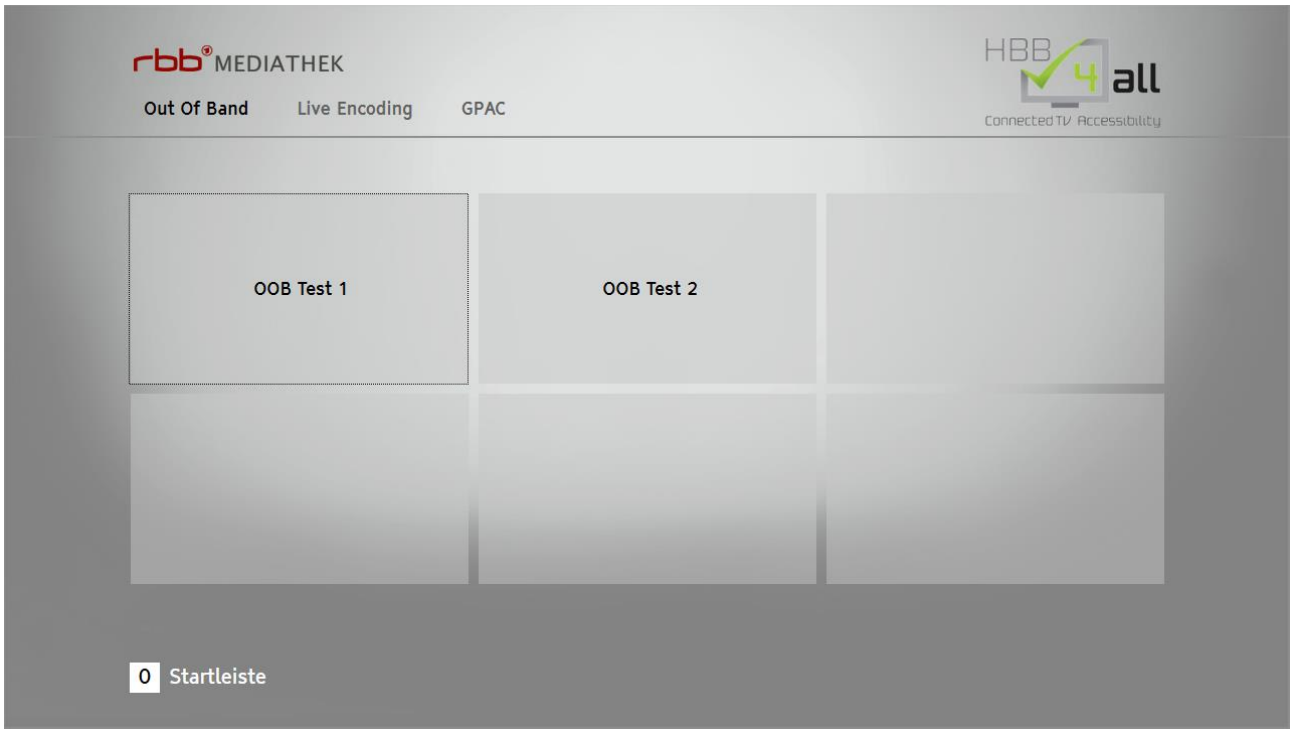

Figure 18. HbbTV demo for IFA and IBC 2015; adapted UI layout

Technology/Standards	
Technology used	*Task/Module*
MPEG-DASH	Distribution format
EBU-TT-D	Subtitle Format
Software, Resources	
Name	*Further Info*
HbbTV 2.0	https://www.hbbtv.org/pages/about_hbbtv/HbbTV_specification_2_0.pdf

Table 17. List of open technology and resources used for the HbbTV2.0 demo app

8.1.17. Subtitle Authoring Component (SCREEN)

In support of the HBB4ALL project, and to investigate the implications of the requirement to generate EBU-TT-D subtitle files for broadband delivery, Screen has added an EBU-TT –D export function to its standard commercial subtitle creation tool (WINCAPS QU4NTUM). This tool allows a subtitler to use the conventional methodology for subtitle authoring, and to export the EBU-TT-D file as an additional export format. See **Figure 19** and **Figure 20**.

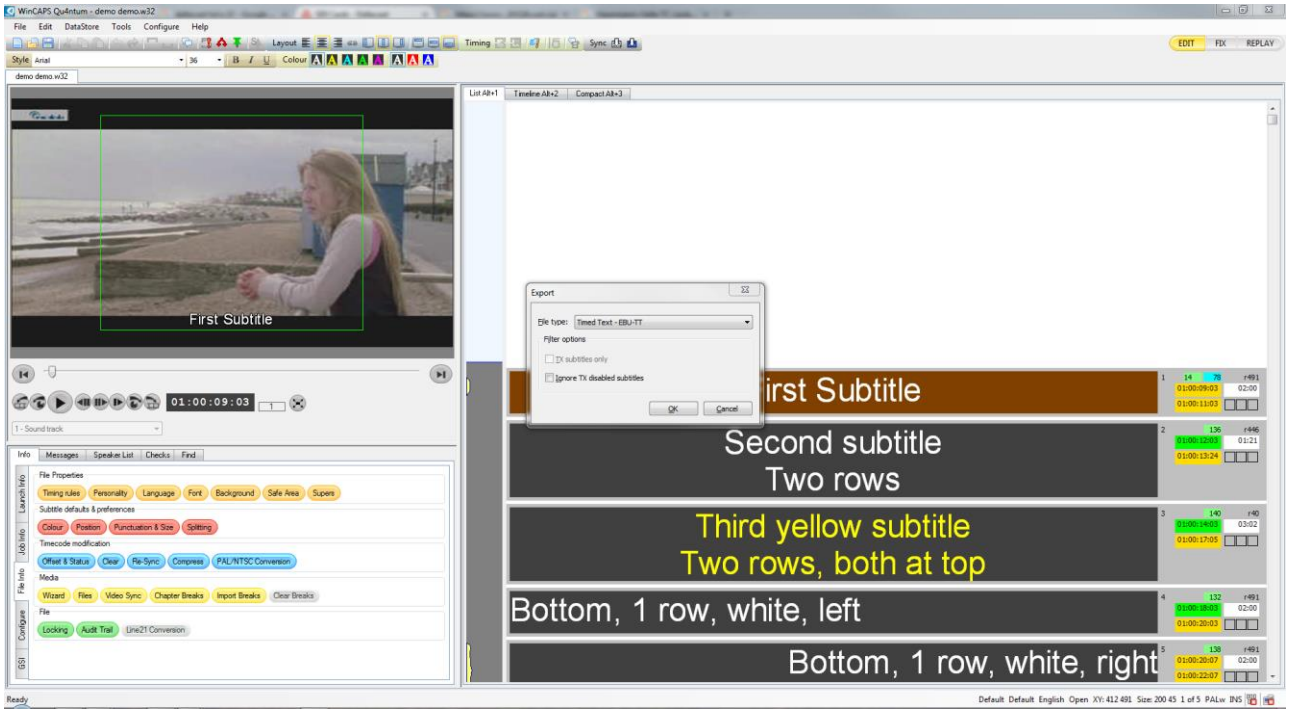

Figure 19. Quantum EBU-TT Export Option

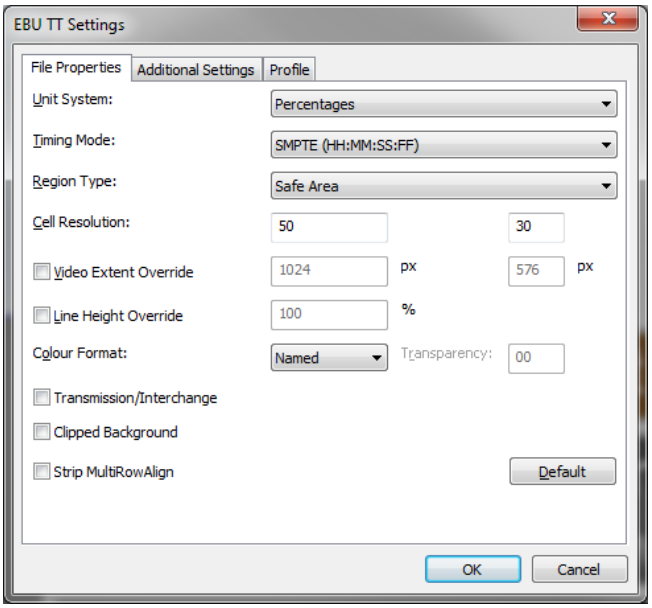

Figure 20. Quantum EBU-TT Export settings

It is apparent that using EBU-TT based subtitle files will have a potential impact on the subtitling authoring process, since EBU-TT format files assume a different rendering philosophy to existing 'legacy' subtitle formats (e.g. EBU STL). In particular the use of custom browser / player rendering, where individual players and scripts may result in substantial differences between the end presentation of the subtitles, significantly alters the current perspective of how the viewer sees what the subtitler authored.

In support of investigating custom browser / player rendering, Screen have also developed a prototype renderer within a pure browser environment (HTML 5 and JavaScript). The prototype renderer includes a module that can break a single EBU-TT format file into a series of Intermediate Synchronic Documents (ISD). Each ISD contains only the information from the source EBU-TT file that is active during the presentation of subtitles at a specific time coordinate on the timeline of the document. For typical EBU-TT format files, each ISD will contain a single 'pop on' subtitle, the series of ISDs generated for a 'snake' presentation of live subtitles is more complicated, since an ISD would also be created for each new word or phrase added during a subtitle display. This process of deconstructing an EBU-TT file into a series of subtitle ISD documents is also relevant to the conversion of EBU-TT format content into legacy format content such as EBU-STL or proprietary formats, since the majority of legacy formats use 'inline' styling. I.e. the style applied to the text is inserted between the text characters as style control codes, in contrast to the EBU-TT format where style information is shared between subtitles by reference to a common section at the head of the file. See **Figure 21**.

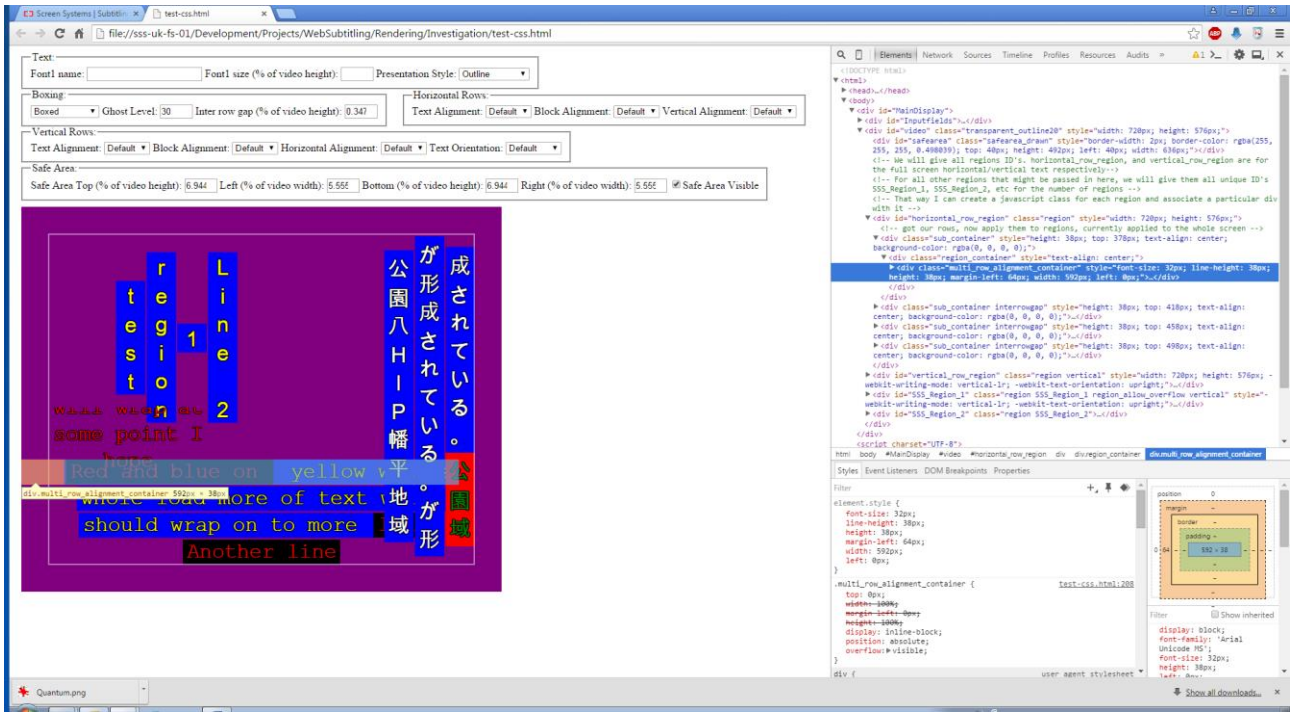

Figure 21. Rendering tests for EBU-TT content using HTML5 and CSS

8.1.18. Subtitle Contribution Components (SCREEN)

Screen is investigating the transport of EBU-TT data in the VBI / VANC spaces of an SDI video signal by extending and adding components within a commercially available Screen subtitle transmission system (Polistream).

The VANC/VBI insertion component under development receives EBU-TT-D documents over an IP connection and converts the EBU-TT data into a form that can be opaquely tunneled within VBI in SDTV or VANC in HDTV.

The second component under development converts any supported subtitle data into EBU-TT data format and can be used within the framework for example for the real-time immediate conversion of (live) subtitle content for onward delivery to HbbTV distribution systems.

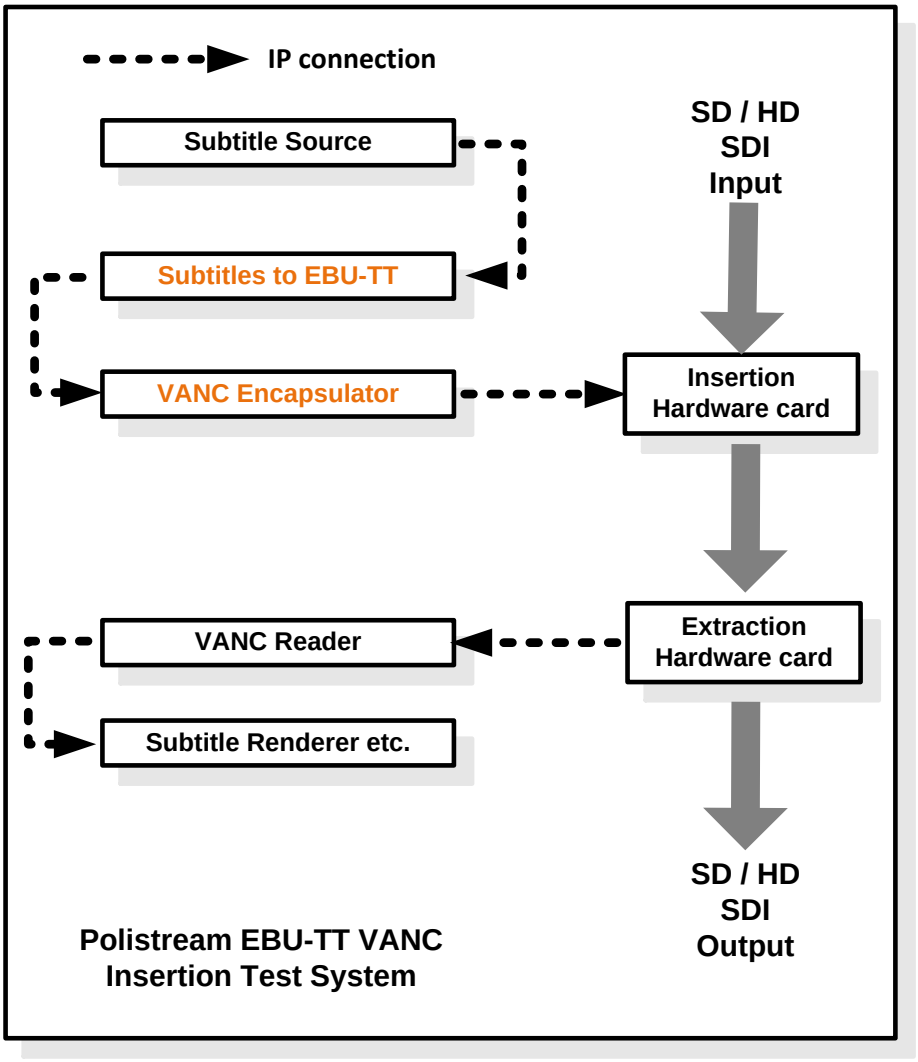

Figure 22. EBU-TT VANC Insertion Test System Block Diagram

Figure 23. VANC Insertion Test Modules

Using this Test System Screen are investigating options for using ANC packets for the distribution of EBU-TT data over SDI, which would be intended for downstream extraction in real time. ANC packets are limited to containing a maximum of 255 8 bit values per packet. The EBU-TT data for a single subtitle is likely to exceed this, so a mechanism for multiple packets to be used to represent a single subtitle is under investigation. In addition, Screen are trialling some simple fixed text substitution mechanisms to reduce the data size, as well as comparing typical file compression techniques such as 'ZIP' and XML specific compression strategies. Early trials suggest that XML specific compression techniques are of very little benefit for the relatively small EBU-TT XML files, so the current focus is on text substitution and 'ZIP' types of compression.

Another aspect which is being explored is the insertion strategy of the EBU-TT data, either repeating the data for the current subtitle in each frame (regardless of the number of packets required), spreading the data for an EBU-TT document over a number of frames (i.e. using a limited number of packets in each frame), or simply sending the ANC packets when required; i.e. on a subtitle change, with frames that contain as many ANC packets as needed.

8.1.19. Subtitle Distribution Component (SCREEN)

Screen is developing functionality within a commercially available data conversion tool (MediaMate) that will support the conversion of subtitle files into EBU-TT-D files. The design for this conversion capability is split into two components, one component that can read EBU-TT format data and a second component that can write EBU-TT-D data. As this functionality has been split into two parts, and these parts are separately usable in workflows designed within the tool, it is possible to convert any supported format of subtitle file into an EBU-TT format file and vice versa.

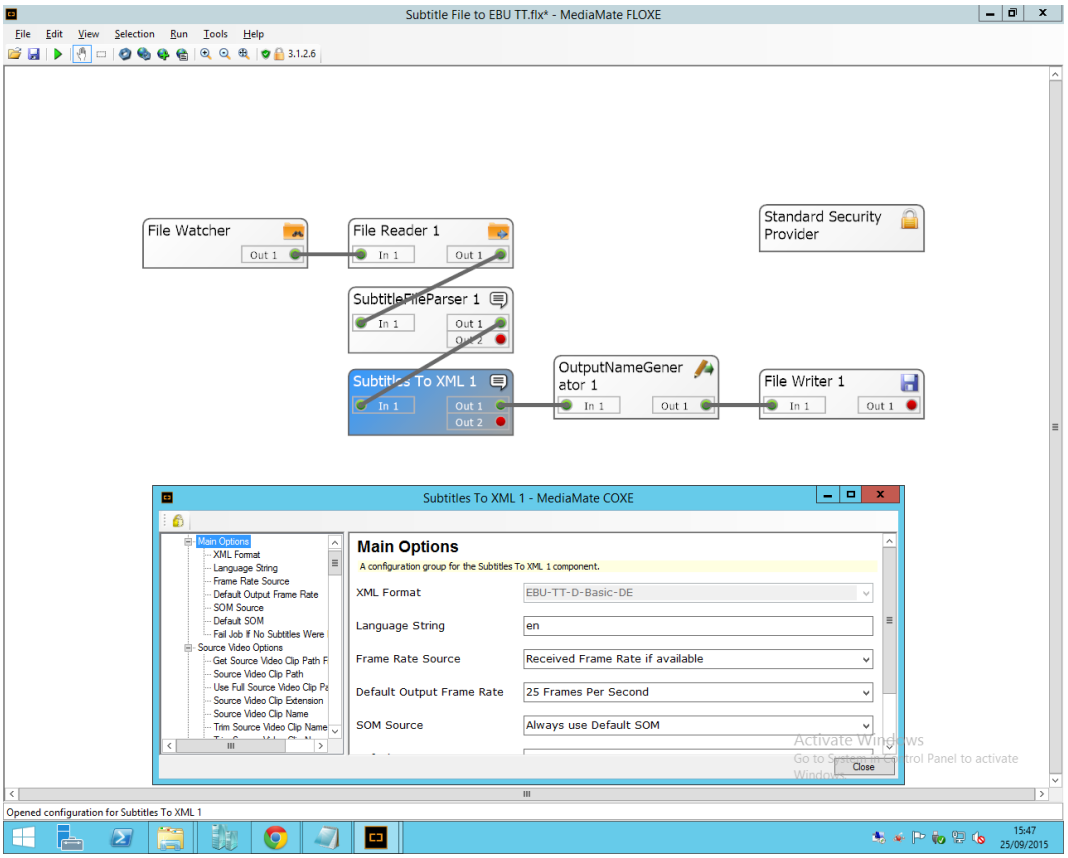

Figure 24. EBU-TT-D Output Option in Data Conversion Tool

8.2. Annex: Integration sub-pilots

This Annex provides detailed descriptions of the workflows planned for the sub-pilots in Pilot-A, in addition to the overviews given in chapter 3.

8.2.1. Customised HbbTV Subtitles for VOD Portal - Germany (RBB)

In summary, RBB coordinated in this pilot the addition of the service components HbbTV-based customisable subtitles for VoD playback and the integration of the *Subtitling Format Conversion Framework*.

8.2.1.1. Production

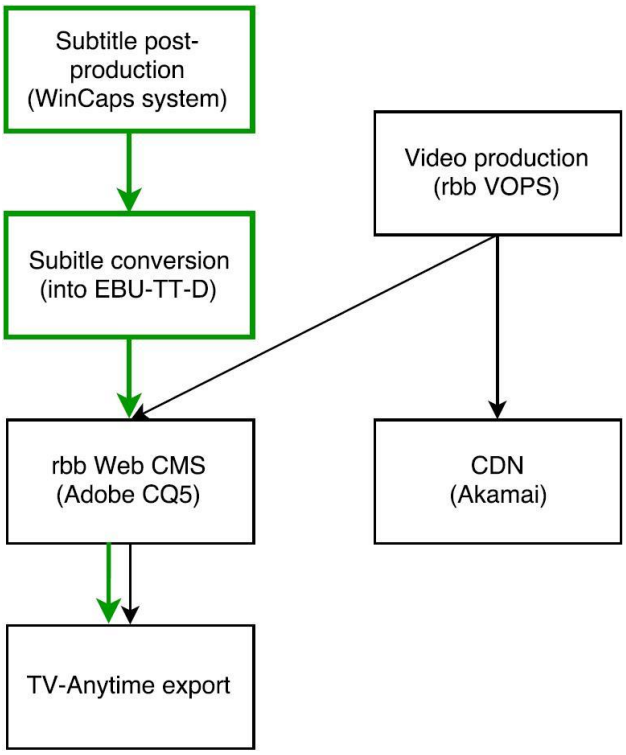

Figure 25. German sub-pilot workflow for subtitle production, new implementations in GREEN

The production of the required VoD RBB subtitle data started in May 2015. At RBB all subtitle data for non-live programmes are pre-produced and temporarily stored on network shares as EBU STL files (for the actual use in the TV playout) and later in the WinCaps proprietary format (.w32) for backup and possible future re-purposing.

Re-purposing of the produced subtitle data for the multiple variants of RBB's on-demand service ("rbb Mediathek") was introduced. Here, RBB is targeting desktop services (browsers on PCs), mobile website and especially the HbbTV-based VoD client. To enable the provision of subtitles for these services, a new workflow was established, allowing for the transfer of subtitle data to publishing and delivery facilities. As defined in HBB4ALL the new EBU-TT format is desirable for an implementation, the German sub-pilot has

added an instance of HBB4ALL's component *Subtitling Format Conversion Framework* (SCF) to the new workflow (see section 2.1.3) and now produces this format.

To produce subtitles in the EBU-TT format, professional users need to load a WinCaps format subtitle file into the subtitle authoring system, adapt and check timebases, check text alignment and save the file in EBU STL format. Next they need to start the SCF tool, open the EBU STL file and export it as EBU-TT-D. This workflow is illustrated in **Figure 25**.

8.2.1.2. Publishing & Delivery

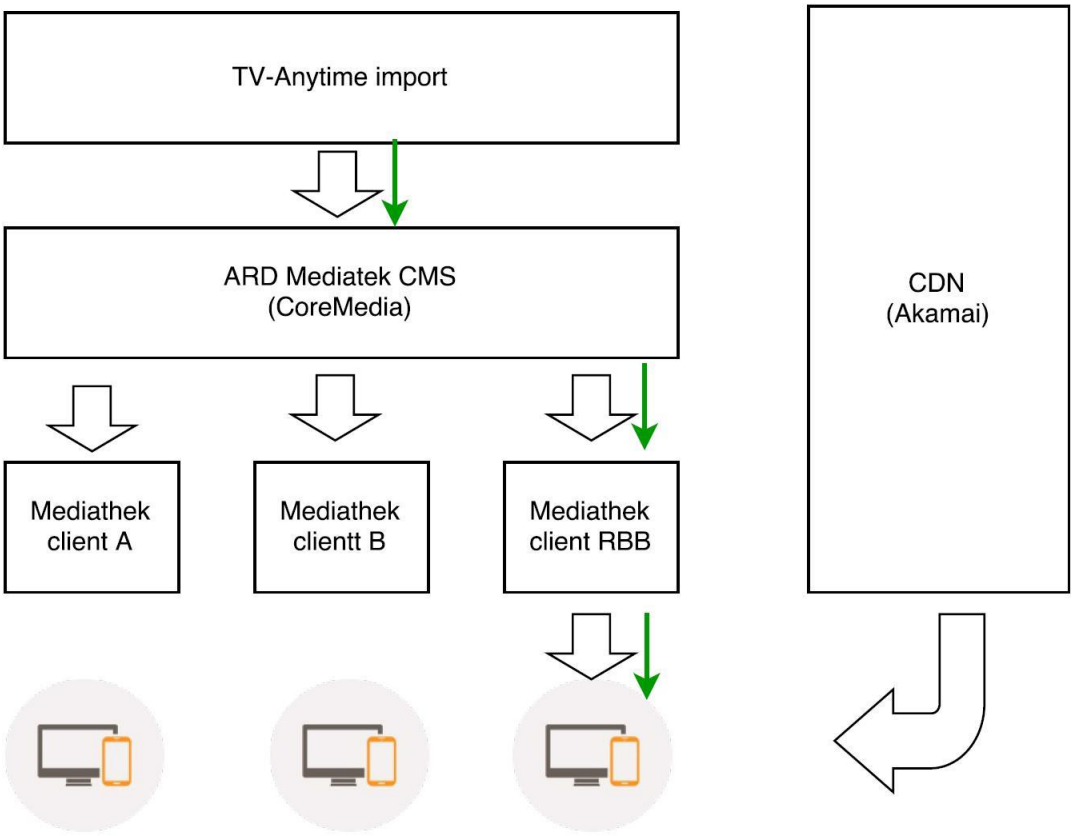

Figure 26. German sub-pilot workflow continues through to end user terminals, new implementations in GREEN

In parallel RBB implemented an upload functionality for its web content management system (CMS) (see **Figure 26**), allowing editors to upload the pre-produced and converted subtitle files to the CMS and assign these subtitle items to video items, coming from the video production system and then pushed to the CDN. The automatic ARD specific TV-Anytime interface is then able to export video and subtitle references to 2nd party systems, especially the actual Mediathek system ("CoreMedia"). CoreMedia is then able to automatically look up the referenced subtitle resources, import and process them for delivery to end user terminals (after going through local broadcaster-specific Mediathek clients (PC, mobile and HbbTV), e.g. RBB Mediathek client). For the internet delivered video content, CMS professional users need to upload the desired EBU-TT-D file to the container assigned to the video item.

The Mediathek system in the ARD is based on a master-client system, see **Figure 26**. Each regional broadcaster has a client version of the "master" ARD Mediathek. The client versions have the same or a

subsection of the functions available in the master and each client has its own corporate branding and offers its own content.

8.2.1.3. Reception & User Application

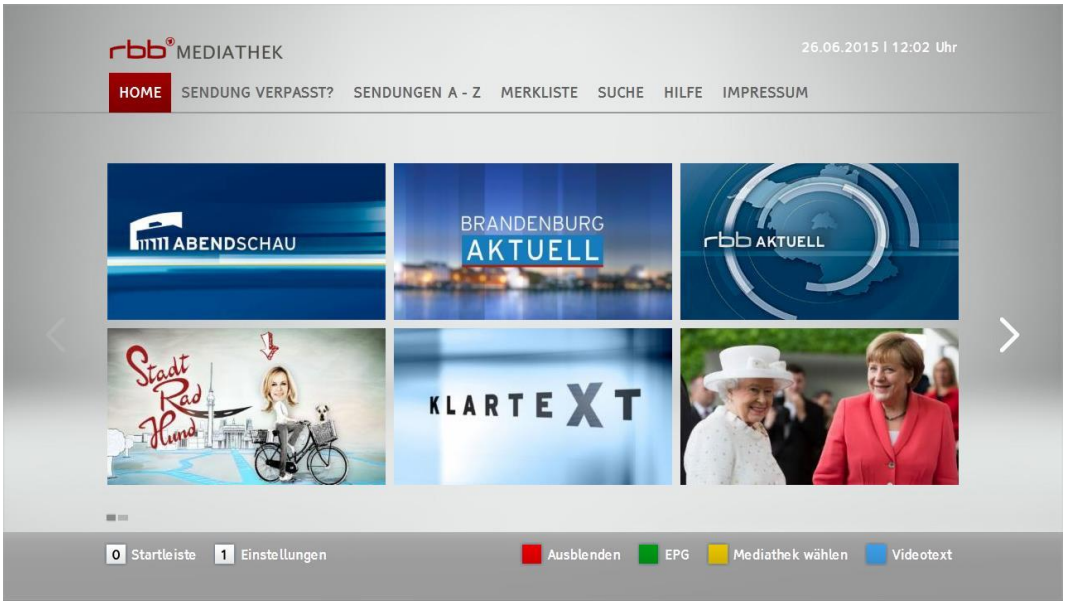

Figure 27. Landing page of the RBB Mediathek client, http://hbbtv.ardmediathek.de/hbbtv-rbb/mediathek/

To access the RBB Mediathek HbbTV client application a viewer must start the HbbTV launcher bar on the TV, by pressing the red button on the remote control. The launcher bar appears on the bottom of the TV screen and contains links to any HbbTV application provided by RBB TV, see **Figure 27**. The same procedure applies to all TV channels provided by ARD network broadcasters. In each ARD network, Mediathek client content items (videos) are categorised and ordered following specific rules, like programme schemes, topics etc.

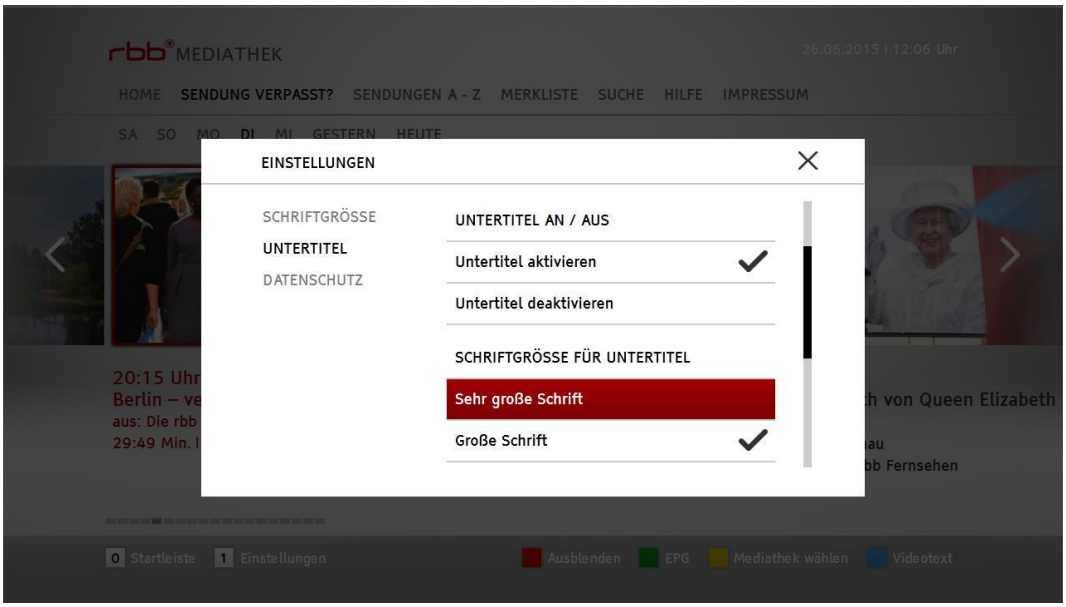

Figure 28. Subtitle settings GUI

RBB has added a module to the Mediathek implementation, which offers a GUI for switching on/off and adapting the graphic representation of the subtitles. It can be accessed by users by pressing the [1] button on their remote control (see **Figure 27**), and navigating to the subtitles category ("Untertitel"). Here font size (in four different levels) and position (top, bottom and dynamic) has been made adaptable (see **Figure 28**).

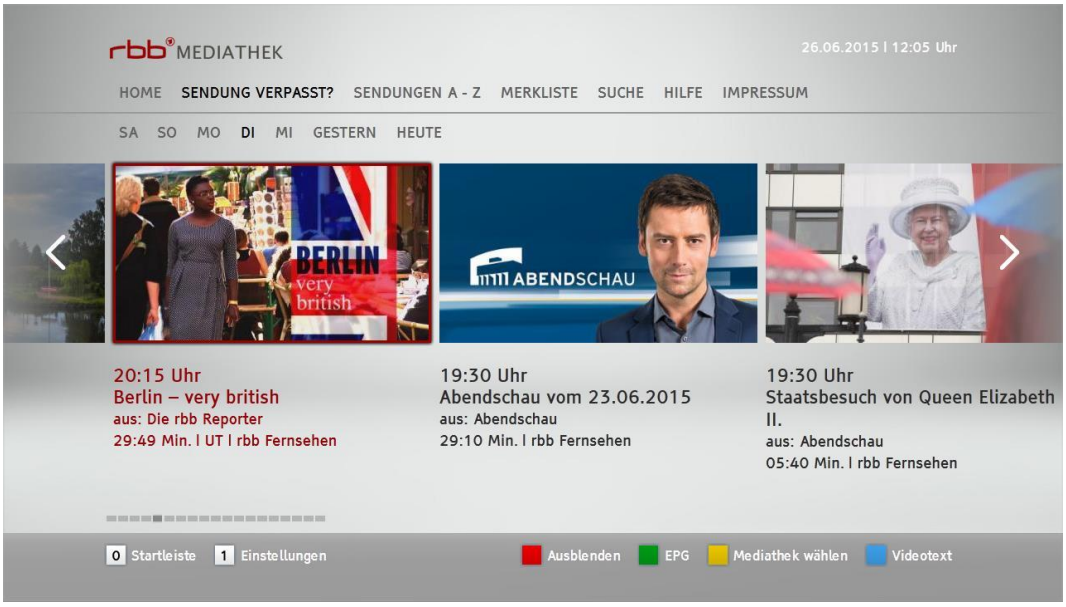

Figure 29. Selected video with HBB4ALL- enabled subtitles

Users can generally browse through Mediathek content by using the arrow keys of their remote control. Newly provided videos with HBB4ALL-enabled subtitles have a textual annotation ("UT", abbreviation for subtitles) right behind the playing time, see **Figure 29**.

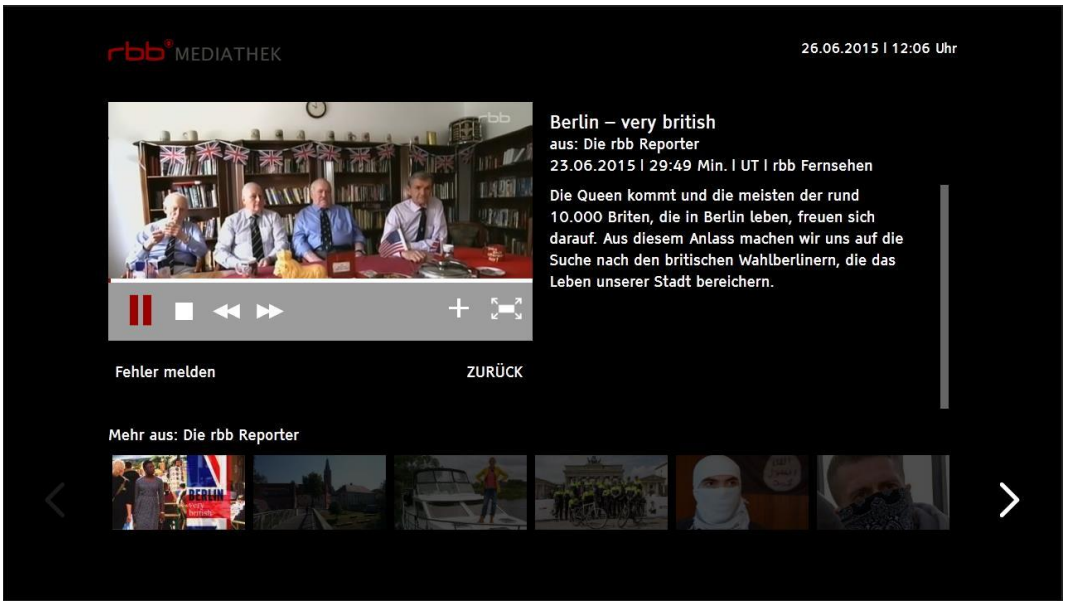

Figure 30. Detail view of video item

When a video is selected, the detail view of that video item is called up and the video starts playing (see **Figure 30**). When the full screen mode is activated, a subtitle switch is added to the video controls, see **Figure 31**. If the video controls are visible, the subtitles are pushed to the top of the screen regardless of which setting the user has chosen, in order to remain visible.

Figure 31. Full screen video player with controls

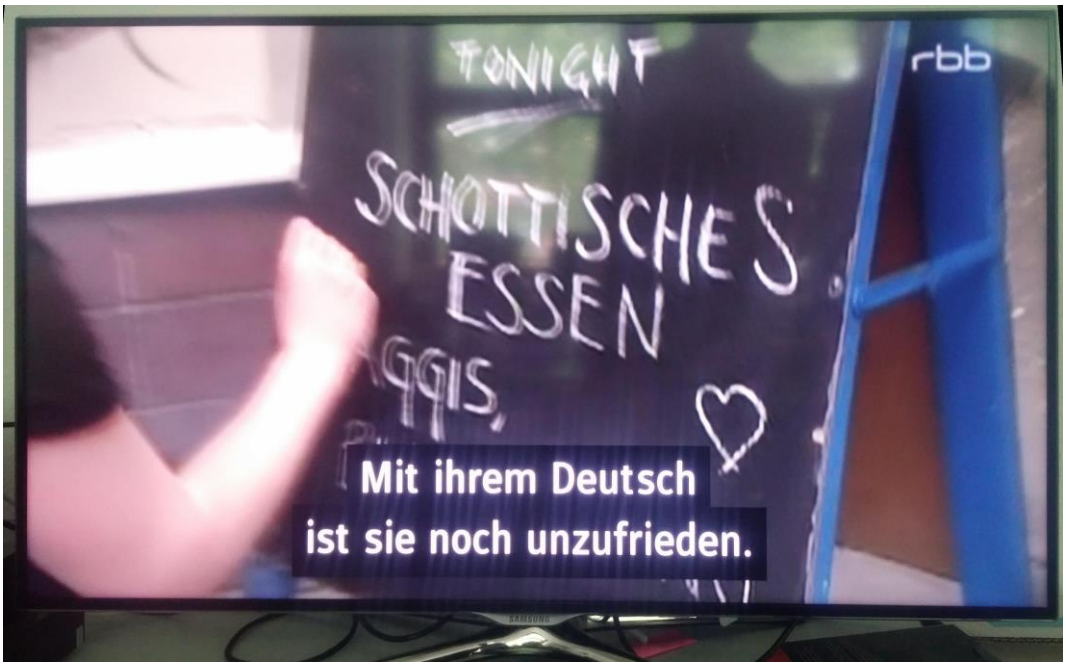

Figure 32. Full screen video with subtitles

8.2.2. Customised HbbTV Subtitles for VOD Portal - Spain (TVC)

8.2.2.1. Production

Figure 33 shows the workflow involved in the production of subtitles for the different media over which the CCMA[18] distributes its on-demand contents:

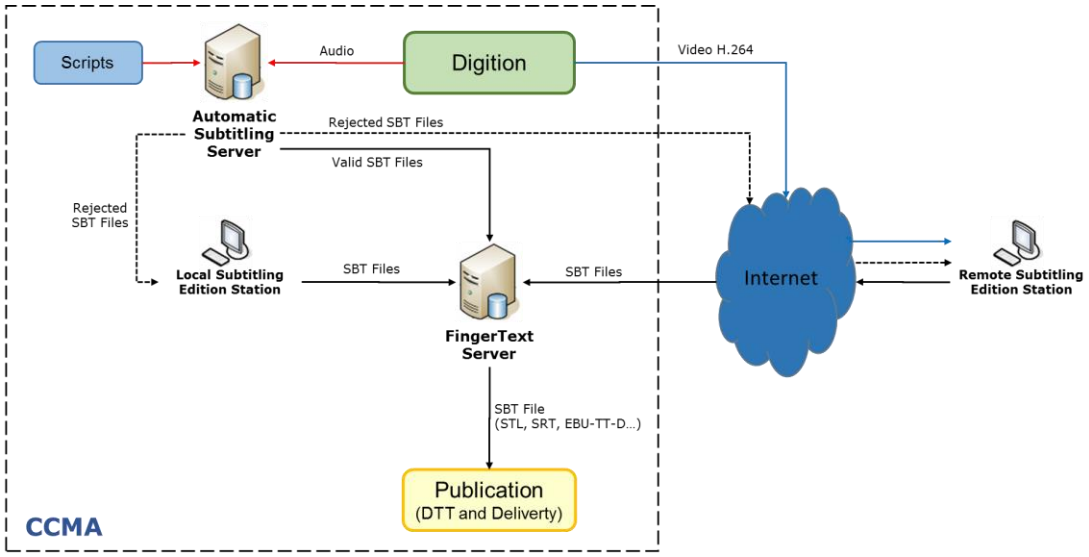

Figure 33. TVC Subtitle Production Workflow

The central element in the subtitle production chain is the FingerText Server, which stores the subtitle files generated by two different feeds:

- **Manual Subtitling Edition:** the editor staff uses the "FingerText Offline Subtitling Editor" software in order to generate and synchronise the subtitle content over an imported digital video file and, after that, they store the output subtitling file into the FingerText Server with its metadata. This tool can be installed in two types of stations:
    - Local Subtitling Edition Station: the editor tool is installed at a station that is placed inside the CCMA's Intranet and it is able to directly access the resource servers. The digital video files are imported from Digition, the local digital video library of the CCMA, and the output subtitling files are stored directly into the FingerText Server.

[18] Corporació catalana de Mitjans Audiovisuals, the public radio and television corporation in Catalonia

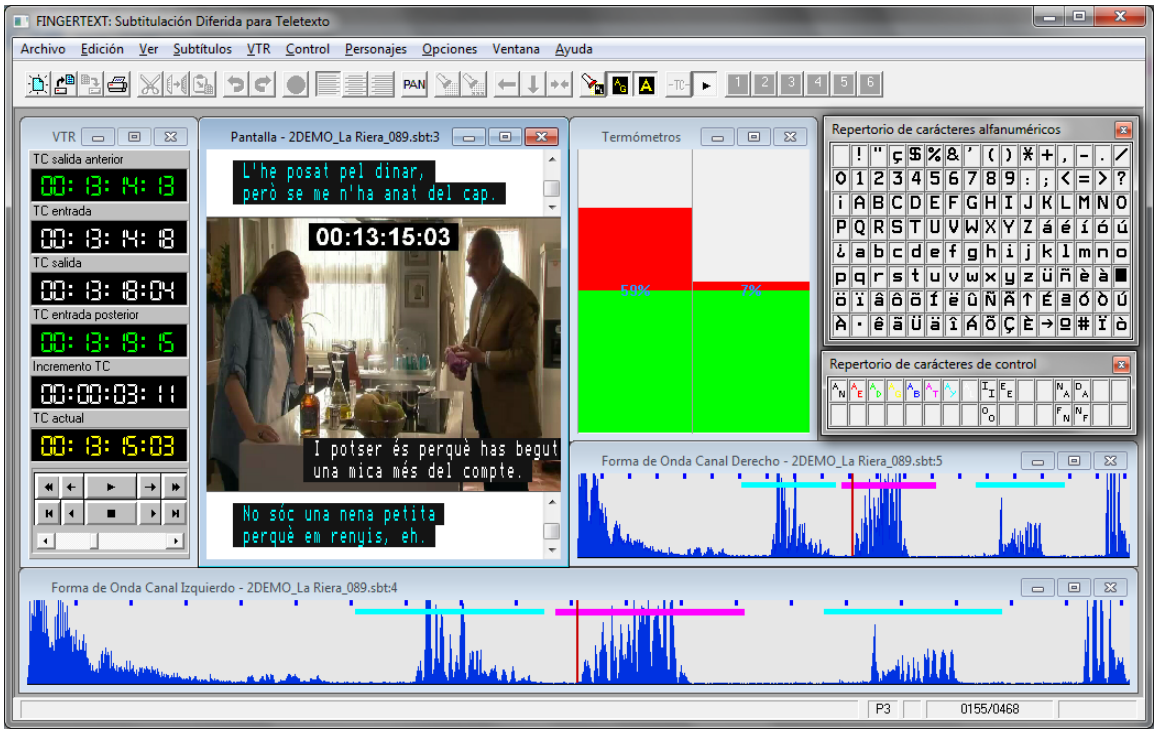

Figure 34. FingerText Offline Subtitling Editor sample

- Remote Subtitling Edition Station: the editor tool is installed at a station that is placed outside the CCMA's Intranet allowing the work of external editor staff. The digital video files are sent to the station using iClip, a system for cataloging and sharing videos through Internet in a compressed version. The output subtitling files are also introduced into the FingerText Server through Internet using secure access.

- **Automatic Subtitling Edition:** the CCMA operates the FingerText Automatic Subtitling System that allows CCMA to meet the subtitling ratio required by the Spanish legislation. The system takes the audio extracted from the original video file and the script of the event in order to automatically generate the output subtitling file. This file is checked against a quality test and is uploaded into FingerText server when passed, or sent to the subtitling editor staff when rejected for deeper revision.

The FingerText Server is integrated with the publishing workflows of the CCMA for different media such as broadcast channels (DVB-T subtitling) and web portals. The system enables the export (with the required transformation) of a subtitle file in a suitable format for each target media: STL, SRT, EBU-TT-D, etc.

In the scope of HBB4ALL, the FingerText system has incorporated the export of subtitle files that conform to the standard "tech3380 - EBU-TT-D Subtitling Distribution Format" [6], and the CCMA has integrated its publishing into the web portal though the Delivery platform and its use in the video-on-demand content inside the HbbTV "3alacarta" application.

8.2.2.2. Publishing & Delivery

This part of the workflow takes the VoD content from TVCs broadcasting Media Asset Management (MAM) system and its subtitles from Fingertext and prepares them for delivery at the IP publishing MAM TVC

system. The programming system of TVC, GREC, is responsible for initiating and completing the publishing and delivery workflow, when it is necessary, for each media asset scheduled for the next seven days.

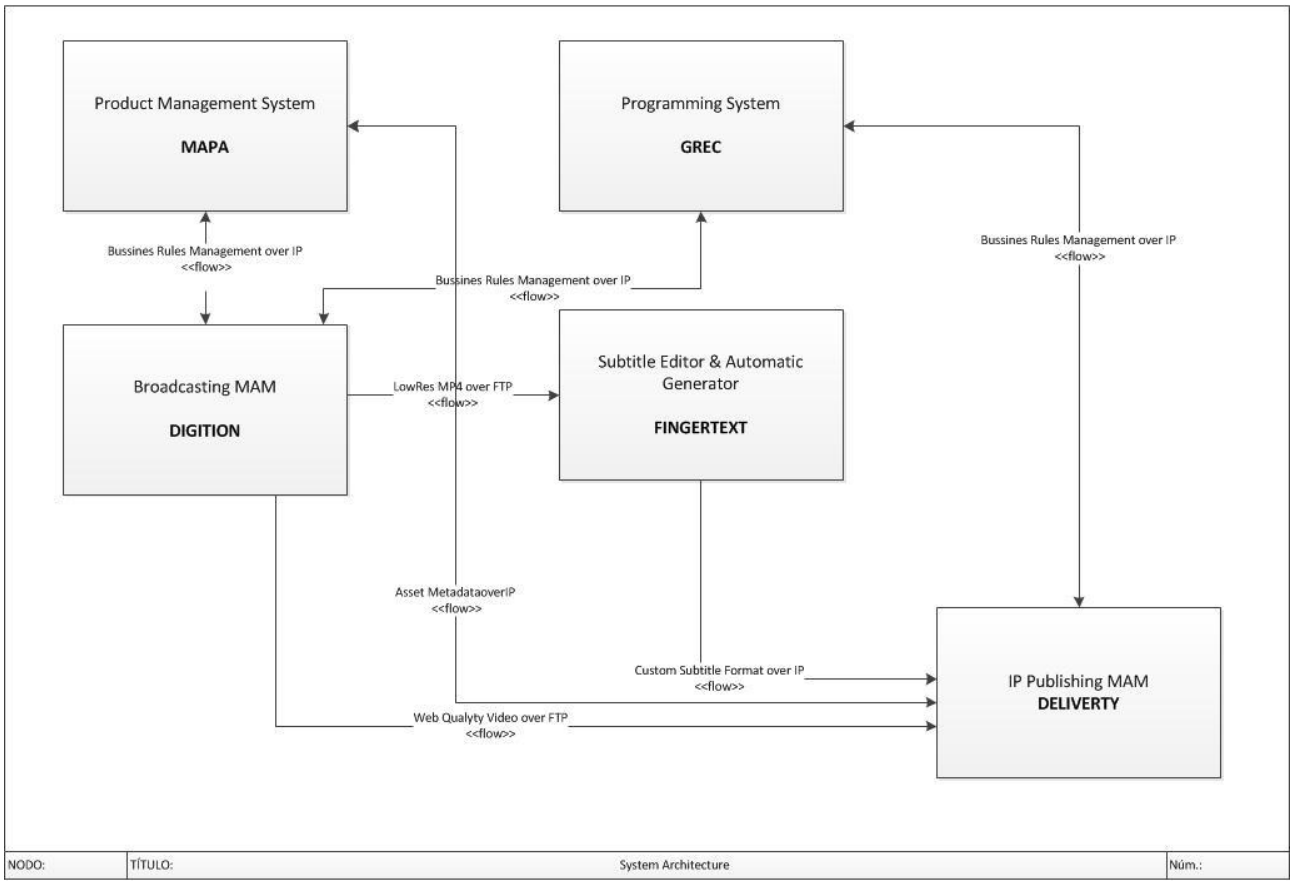

Figure 35. TVC Publishing & Delivery Workflow

This workflow basically consists of the following procedure for each content item to be published (see **Figure 35**):

1. Start of workflow. It starts some days before scheduled transmission date, for each program.

2. Wait until media asset is available on DIGITION, which is TVCs Broadcasting MAM system. In some cases the media asset is not available until few minutes before scheduled time.

3. Transcoding of the media assets present on DIGITION to web quality video. There may be multiple video representations in various qualities for the web, depending on the user application.

4. Delivery of the web quality video to DELIVERTY, the IP Publishing MAM of TVC, over an FTP connection.

5. From DELIVERTY, request transfer of the metadata for the video content to MAPA, over IP.

6. From DELIVERTY, request transfer of custom subtitles file to FINGERTEXT, over IP. The custom subtitles are not always available at same time as the media asset. The subtitles request process is repeated until the subtitle file is available or until 10 days after emission of the content.

7. End of workflow.

Once the video content, the associated EBU-TT-D subtitle file and other metadata are ready on the DELIVERTY platform, which is part of TVCs Content Management System, they are also automatically available at TVC's server file system (VoD as well as web servers). The users can request the VoD content and, if they want to display the subtitles, the subtitle file is requested too and it is synchronised with the current playback of the video, all through the TVC web portal or the "3alacarta" HbbTV application.

8.2.2.3. Reception & User Application

For a broadcaster a multiplatform player is a key tool at the heart of their business, containing a lot of intelligence for reporting audiences and handling advertising.

To serve all possible players available on its portfolio, TVC centralizes information and offers an API to publish information to all available players via AJAX metadata requests. Initially, TVC generated and offered subtitles in EBU-TT-D format. In HBB4ALL deliverable D3.1, section 3.3.1 it was described how TVC offers EBU-TT-D compatible subtitles for all available content (see [2]). The following shows a simplified XML snapshot:

```
<tt:tt xmlns:tt="http://www.w3.org/ns/ttml" xmlns:ebuttExt="urn:ebu:tt:extension"
xmlns:ttp="http://www.w3.org/ns/ttml#parameter" xmlns:tts="http://www.w3.org/ns/ttml#styling"
xmlns:ebuttm="urn:ebu:tt:metadata" ttp:timeBase="media" xml:lang="ca_ES" ttp:cellResolution="40 25">
   <tt:head>
            <tt:metadada>
                     <ebuttm:documentMetadata>
                              <ebuttm:documentEbuttVersion>v1.0</ebuttm:documentEbuttVersion>
                              <ebuttm:documentCountryOfOrigin>es</ebuttm:documentCountryOfOrigin>
                     </ebuttm:documentMetadata>
            </tt:metadada>
            <tt:styling>
                     <tt:style xml:id="s3" tts:fontFamily="Tiresias" tts:fontSize="85%" tts:lineHeight="normal"
tts:textAlign="center" tts:color="aqua" tts:backgroundColor="black" tts:fontStyle="normal" tts:fontWeight="normal"
tts:textDecoration="none" tts:wrapOption="wrap"/>
                     <tt:style xml:id="s4" tts:fontFamily="Tiresias" tts:fontSize="85%" tts:lineHeight="normal"
tts:textAlign="center" tts:color="white" tts:backgroundColor="black" tts:fontStyle="normal" tts:fontWeight="normal"
tts:textDecoration="none" tts:wrapOption="wrap"/>
            </tt:styling>
            <tt:layout>...</tt:layout>
   </tt:head>

   <tt:body>
            <tt:div>
                     <tt:p xml:id="1" style="s4" region="r12" begin="00:00:00:00" end="00:00:02:13">
                     <tt:span>La carta funciona, Mercè.</tt:span>
                     <tt:br/>
                     <tt:span>Als clients els hi agrada.</tt:span>
                     </tt:p>
                     <tt:p xml:id="2" style="s3" region="r12" begin="00:00:02:17" end="00:00:04:17">
                     <tt:span>Sí, sí, sí, no ho poso en dubte,</tt:span>
```

```
                    </tt:p>
                    <tt:p xml:id="3" style="s3" region="r12" begin="00:00:04:21" end="00:00:07:06">
                    <tt:span>però el que jo proposaré</tt:span>
                    <tt:br/>
                    <tt:span>també agrada a molta gent.</tt:span>
                    </tt:p>
                    <tt:p xml:id="4" style="s4" region="r12" begin="00:00:07:10" end="00:00:09:24">
                    <tt:span>Tu saps el temps que vaig dedicar</tt:span>
                    <tt:br/>
                    <tt:span>a crear aquesta carta?</tt:span>
                    </tt:p>
                    <tt:p xml:id="5" style="s3" region="r12" begin="00:00:10:03" end="00:00:13:20">
                    <tt:span>
                    Ja m'ho imagino, molt, Maribel. Però n'aprofitarem moltes coses.
                    </tt:span>
                    </tt:p>
                  ...
            </tt:div>
    </tt:body>
</tt:tt>
```

Once a user decides to playback an asset from TVC's VoD service, the player (software in the device) requests metadata according to its profile, and receives a response containing the basic VoD information (duration, title, synopsis, mp4 url, key frames and so forth) and also audience, advertising and geolocation rights information. Not all the content has subtitles available, but those assets that have subtitles receive information about the subtitles to be played back, the following shows an example.

```
"subtitols": [{
    "text": "Català",
    "iso": "ca",
    "url": "http://statics.ccma.cat/multimedia/xml/2/8/1434086536782.xml",
    "format": "ttaf"
}, {
    "text": "Català",
    "iso": "ca",
    "url": "http://statics.ccma.cat/multimedia/xml/0/8/1433337779480.xml",
    "format": "ebuttd"
}],
```

TVCs approach for the implementation of HbbTV subtitles is based on a detailed analysis of available subtitling delivery and fulfils following requirements:

- uses EBU-TT-D subtitle data format;
- realises subtitles for Access Services in a cost-efficient manner;
- optimises integration in the existing workflows.

As result, the TVC implementation reuses existing API call requests, also opening the possibility to reuse EBU-TT-D subtitle format for other connected devices, such as PCs and mobile devices, thus increasing interoperability. EBU-TT-D has been adopted in 1.5 HbbTV specification release. Due to the lack of subtitling support within the early HbbTV specifications, the EBU-TT-D format is not supported by current software and hardware solutions for subtitle production and distribution as a native format. TV models that

support this standard are rare, almost non-existent, and market penetration is still incipient. TVC decided to make its implementation compatible for all HbbTV versions, which means that this approach is able to cover all available HbbTV devices on the market, instead of adding more fragmentation to the existing offer.

The adopted solution works thanks to a new Javascript plugin developed for that purpose. Once the user enables subtitles during playback, the plugin parses the EBU-TT-D file and translates the TTML tags to common HTML tags ready to be rendered in the browser and schedules phrases according to the defined timing information about when each subtitle must appear. The regions and positions for rendering on the screen must be defined (in the EBU-TT-D document) using percentages in order to draw it properly in full screen or half screen video. Also, the translation to common HTML tags helps to not replicate infrastructure components across platforms.

When a user accesses the HbbTV player's playback screen, a playback control menu is available as usual. In cases where there are no subtitles available, this feature is not presented to the user (see **Figure 36**).

Figure 36. TVC HbbTV player without subtitles option

Otherwise, when this functionality is available, a button appears on the playback menu. Then, users may activate subtitles, and they are informed about the present status (ON or OFF). **Figure 37** shows the example when the subtitles are "OFF".

Figure 37. TVC HbbTV player with subtitles option

When subtitles are activated, they are superimposed on the video and the state is shown as ON, see **Figure 38**.

Figure 38. TVC HbbTV player rendering subtitles (phase 1)

8.2.3. Customised Subtitles for Wide Focus Multi-Platform – Portugal (RTP/UPM)

The currently available subtitle service for broadcast at RTP is the teletext subtitle service (for prepared and automatic subtitles), for which the workflow is depicted in **Figure 39**.

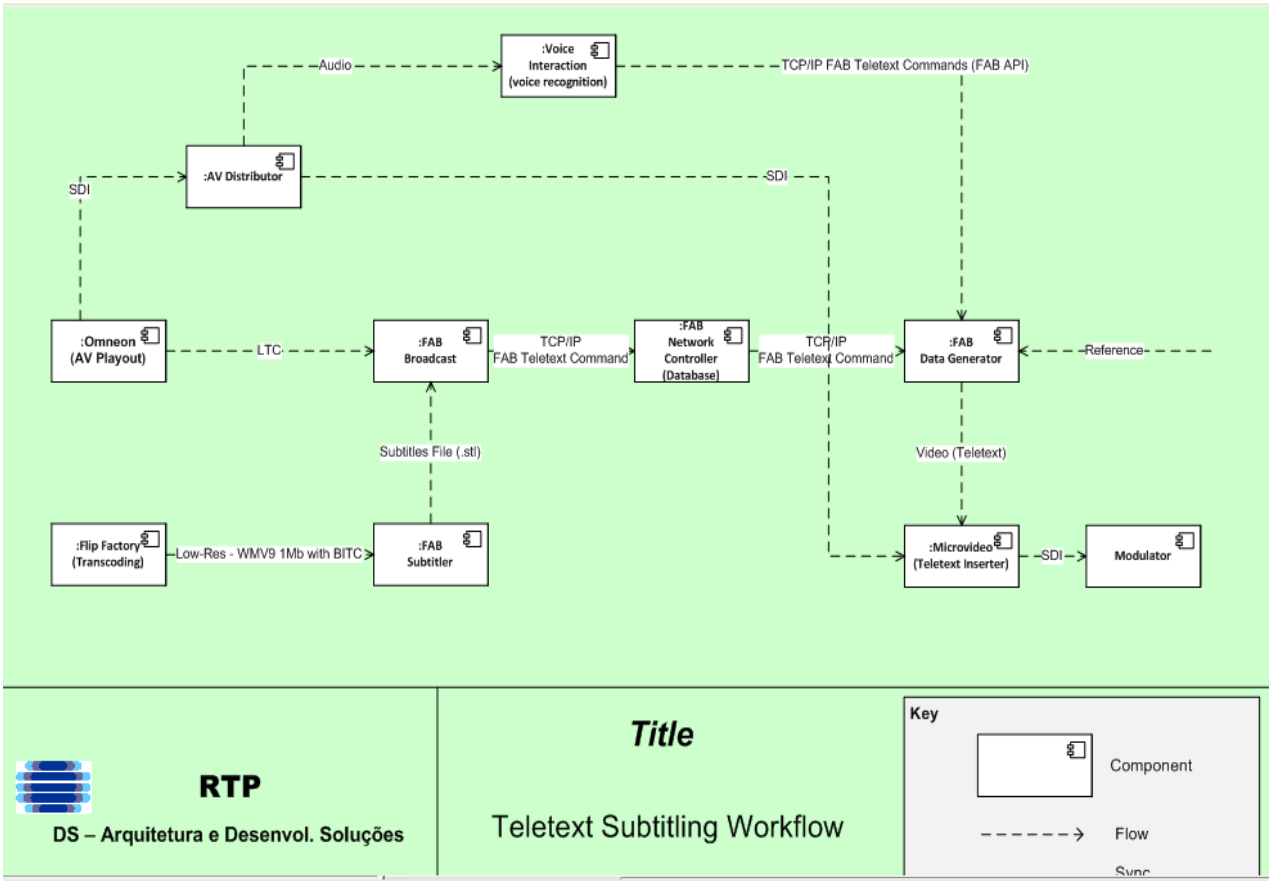

Figure 39. RTP's Teletext Subtitle Workflow

The Flipfactory video transcoder system indicated in **Figure 39** will be replaced by the WFS system and some channels will be sourced from the iTX system (currently the playout is made by the Omneon servers).

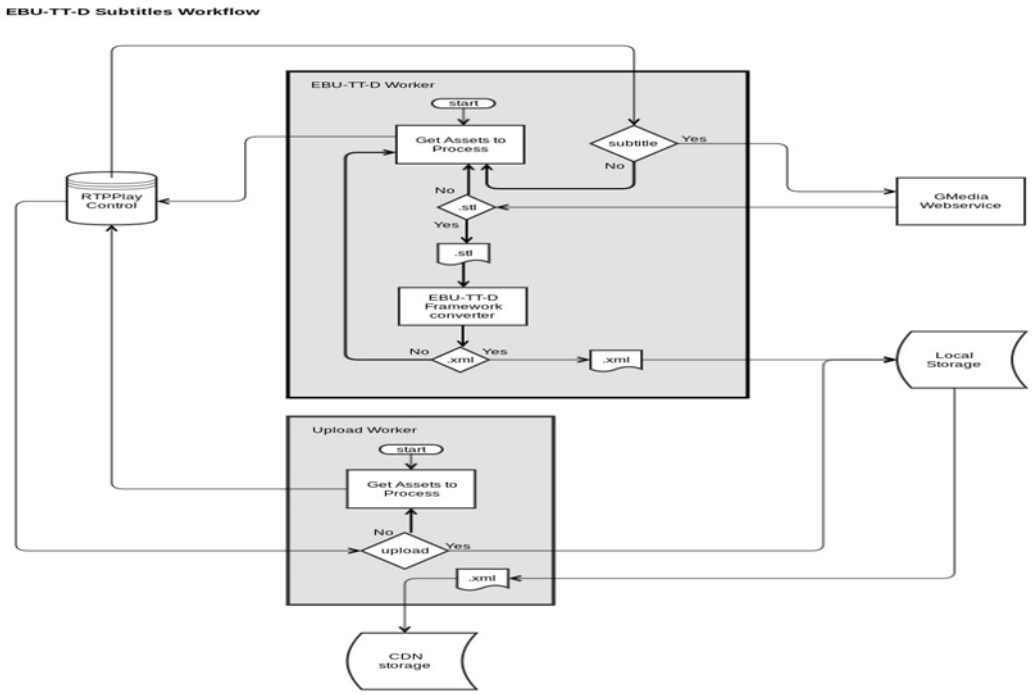

Figure 40. RTP's EBU-TT-D subtitle workflow

8.2.3.1. Production

The basic workflow used in this process starts inside the RTP's Accessibility Unit, with the production of an EBU STL format subtitle file for broadcasting the teletext accessibility subtitle service. Then, the file is uploaded into GMediaTX, in the episode program database, with the same media id as the program. The publishing system, based on AsRun log file detects the programs to publish and records them in a database.

8.2.3.2. Publishing & Delivery

As EBU STL is the subtitle format traditionally used by the industry, a tool is needed to generate the EBU-TT-D file, which will be interpreted by the web-based application implemented in the sub-pilot. For this purpose, this sub-pilot uses the technical component provided by IRT named "*Subtitling Format Conversion Framework*" (described in section 2.1.3). To manage all format conversions, several worker tasks were defined for the different formats (jpg, mp4, mp3...). Each format corresponds to one worker task, which works independently and searches for a new job each minute. For the subtitle conversion of the EBU-TT-D format, a new worker task was defined. The EBU-TT-D worker task searches for the EBU STL file in GMedia and starts the workflow described in the framework conversion, converting EBU STL → XML → EBU-TT → EBU-TT-D.

After all steps have concluded, the EBU-TT-D worker task signals the format is converted and ready for upload. The upload worker task finishes the process, uploading and recording the subtitle file in the RTPPlay database. After this, the subtitle file, video and other metadata are available via the RTPPlay API. This process is represented in **Figure 40**.

8.2.3.3. Reception & User Application

The reception and presentation of the EBU-TT-D subtitles in this sub-pilot is based on JW Player, a player widely used by broadcasters and content providers to distribute their contents by means of the Web. This is, in fact, the player that supports RTPPlay, which is embedded in the RTP website. For this reason, the availability of the subtitles in the Portuguese service sub-pilot requires a plugin that is able to interpret the EBU-TT-D format and to render the subtitles on the video content. This plugin is being developed by UPM in the project and will available in September 2015.

On the other hand, as the JW Player is embedded in a web page, this solution also requires a web browser as user application. The objective of the implementation of a plugin for JW Player aims to provide subtitles over broadband to be presented on a variety of devices, including PC, tablets and smartphones. The purpose of creating a plugin for JW Player is enabling the user to configure and embed subtitles in video content.

The plugin for JWPlayer developed for HBB4ALL is based on the EBU-TT-D specification. Although JWPlayer is capable of presenting captions in different formats, such as WebVTT (Web Video Text Tracks[19]) or WebSRT (Web Subtitle Resource Tracks), including other advanced caption standards was necessary for offering an improved experience to the user.

Figure 41. Example of usage of plugin for JW Player in RTPPlay platform.

EBU-TT-D subtitle files: Reading the XML-based file for obtaining information of timing, text and styling

The behaviour of navigators to manipulate the EBU-TT-D file contents was implemented in two different ways:

[19] http://dev.w3.org/html5/webvtt/

- Google Chrome and Safari use jQuery (a cross-platform JavaScript library designed to simplify the client-side scripting of HTML).

- Mozilla Firefox and Internet Explorer (v.11) required the usage of XML DOM (a standard for accessing and manipulating XML documents). The DOM presents an XML document as a tree-structure.

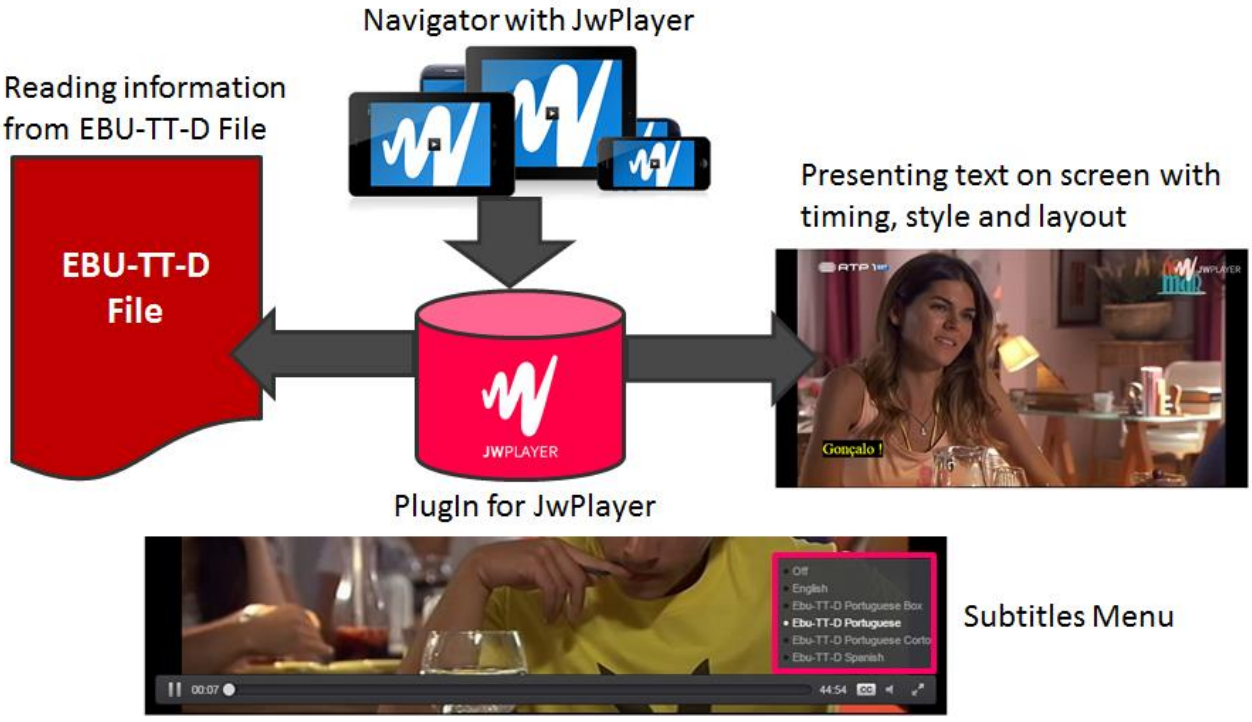

Figure 42. Architecture for the implementation of the plugin

Work Process

The design process of the implementation is based on the schema shown in **Figure 43**.

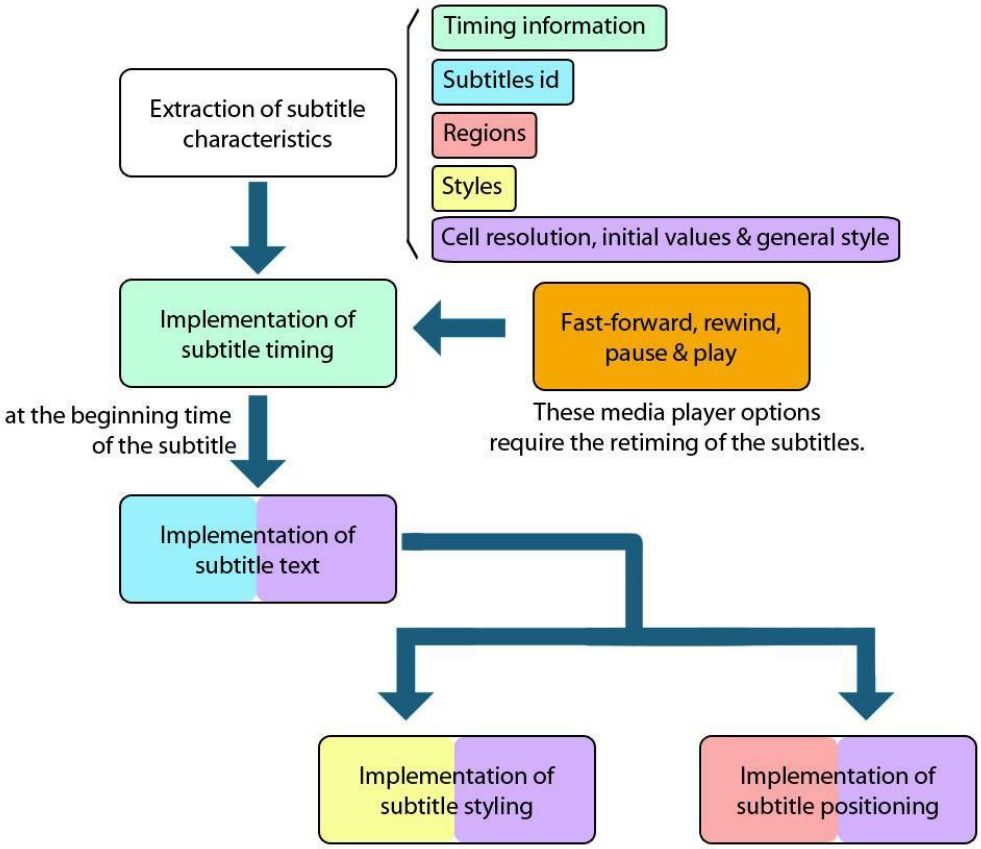

Figure 43. Schema of the different stages in the implementation of the subtitles

The phases for the development and their status are as follows:

- Extraction of subtitle characteristics. This phase is completed and tested for four different navigators. Information about regions, styles and texts is extracted from the EBU-TT-D file.

- Implementation of subtitle timing. This phase is completed using the jwplayer.onTime()function, which controls the timing by clock ticks. The timing information included in EBU-TT-D file is transform to milliseconds, which is the type of content that this tool understands. Completed with the information extracted from <p> tags.

- Implementation of subtitle text. Completed with the information extracted from tags.

- Fast-forward, rewind, pause & play. This phase requires an improvement for cases in which an interaction with video by the user is necessary.

- Implementation of subtitle styling and positioning. This phase requires an improvement for some subtitle characteristics which differ CSS from style description.

Development

The plugin requires no previous installation and, once the subtitles are included in the HTML file, are presented in the official closed captions menu (CC menu). Subtitles in EBU-TT-D format are selectable as shown in **Figure 44**.

Figure 44. Accessible subtitles through conventional closed captions menu

Each EBU-TT-D file includes a header and a body with all the subtitling information. The header includes three different types of data:

- Metadata
- Styling (with the <tt:style> labels). This connects different characteristics of the subtitles presented on screen with style attributes such as font colour or size, bold or italic text or background colour.
- Layout (with the <tt:region> labels), which specifies the different types of regions or positions where subtitles may be presented on screen.

The body is formed by different types of grouped components

- The tag defines a division or a section in an HTML document. The tag is used to group block-elements to format them with CSS styles. This type of tag includes a collection of <p> tags.
- The <p> tag defines a paragraph. CSS (with the margin properties). The <p> tag in EBU-TT-D files includes information about timing and also styling and layout characteristics. This type of tag includes a collection of tags.
- The tag is used to group inline-elements in a document. The tag provides no visual change by itself. The tag contains information about styles and the text information of the subtitle itself.

The distribution of this three types of components are included in the tag of JW Player whose main function is showing captions, with the “id” of “myElement_caption”. , <p> and tags are nested in this specific part of the code. The element “visibility” of the element style is set to “visible” or “hidden” defined by timing for each subtitle in the XML-based file. An example can be seen in **Figure 45**.

```
▼<div id="myElement_caption" class="myClass" style="visibility: visible;">
  ▼<div id="auxdiv0" style="font-family: monospaceSansSerif; font-size: 80%; line-height: 125%;
  text-align: center; color: rgb(255, 255, 255); font-style: normal; font-weight: normal; text-
  decoration: none; background-color: rgb(0, 0, 0);">
    ▼<p id="sub6" style="font-size: 100%; text-align: center; position: absolute; left: 80px;
    bottom: 40px; max-width: 640px; max-height: 320px; padding: 0%;">
       <span id="auxSpan0_6" style="color: rgb(255, 255, 0); font-size: 200%; text-align: center;
       background-color: rgb(0, 0, 0);">que neles procura esclarecer dúvidas</span>
       <br>
       <span id="auxSpan1_6" style="color: rgb(255, 255, 0); font-size: 200%; text-align: center;
       background-color: rgb(0, 0, 0);">ou confirmar certezas.</span>
     </p>
    </div>
  </div>
```

Figure 45. Example on HTML JavaScript interaction with subtitles

The presentation of subtitles with most of their corresponding characteristics is completed as shown in next example (**Figure 46**).

Figure 46. Example of subtitles presented embed in the video content

Tests

The plugin has been tested with:

- Navigators on PC / Mac:
  - Google Chrome
  - Mozilla Firefox
  - Safari
  - Internet Explorer (v.11), (with implementation also valid for IE v.6-10)
- Smartphones
  - Google Chrome

- Tablets
    - Ipad with Safari
    - Samsung Galaxy Tab with Google Chrome

Future Work

As future work, there is a chance of adding customization by user to the subtitling process, as shown on the example in **Figure 47**, e.g. by allowing the user to change:

- Font Size
- Font Color / Background Color
- Layout (Regions)

Figure 47. Customized font size in plugin

8.2.4. Automated HbbTV Multilingual News Broadcast Subtitles- Spain (VIC/UAB)

8.2.4.1. Production, Publishing & Delivery

This sub-pilot service focuses on the "News" domain, being broadcast at international level but only produced in the original language. The real-time transcription service will be used to generate subtitles of the broadcast event while the machine translation service will make those comprehensible in other languages. Those modules will be integrated in an HbbTV workflow, providing real-time automatically subtitled content accessible from a TV.

For this purpose, this sub-pilot integrates three different components, described previously:

- *Automatic Subtitling Component* (see section 2.1.1)
- *Machine Translation Component* (see section 2.1.2)
- *Live Broadcast-Internet Subtitle Viewer and Synchroniser Component* (see sections 2.3.1 and 2.2.3)

The *Automatic Subtitling Component* creates subtitles based on an audio signal. It applies a Large Vocabulary Continuous Speech Recognition technology from an audio in English and creates EBU-TT-D subtitles in English. The English transcription engine was trained for the "news" domain within the SAVAS European project.

The *Machine Translation Component* works on top of Statistical Machine Translation technology. It takes the English transcription created by the Automatic Subtitling Component and translates it to Spanish, creating EBU-TT-D Spanish subtitles. This component is also trained for the broadcast news domain.

The *Live Broadcast-Internet subtitle viewer and synchroniser component* creates an HTTP media stream from a broadcast video and audio source, and injects the automatically generated EBU-TT-D subtitles. The output of this component is an MPEG-DASH stream with the video, audio and the automatically generated subtitles in English and Spanish, all synchronised.

Figure 48 depicts a high-level integration of the components.

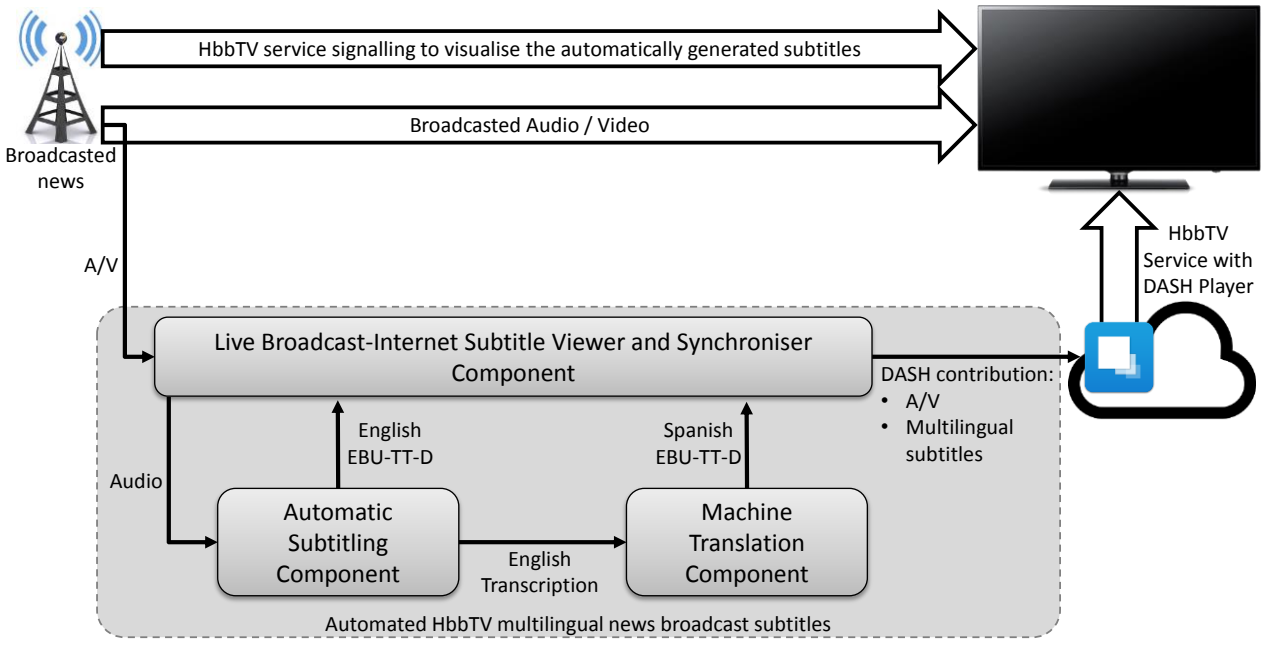

Figure 48. High-level integration overview of Automatic HbbTV multilingual news broadcast subtitles sub-pilot

The DASH stream will be delayed in comparison with the broadcasted audio and video. Part of the delay is created by the *Automatic Subtitling Component* and the *Machine Translation Component*. The *Live Broadcast-Internet Subtitle Viewer and Synchroniser Component* also introduce a delay, which is common for DASH servers, working with segments of around 30 seconds. As a result the user will be able to watch the broadcast audio and video in real-time, or launch the HbbTV application to consume the DASH stream and watch the audio, video with the automatically generated multilingual subtitles. In the second option, all the sources will be synchronised but delayed in comparison with the original broadcast source.

From the deployment point of view, the sub-pilot has been implemented in the Digital TV Laboratory in Vicomtech-IK4. There is a dedicated server for this sub-pilot (server in **Figure 49**). It is a Vibox PC VBX-

PC-4746. It has an Intel Core i7 processor with 32 GB of RAM and a NVIDIA GTX 960 graphic card. The OS is a version of Linux (Debian 8 64 bits).

The server has a USB tuner. It is a Hauppage WinTV-NOVA-T-Stick for DVB-T. One of the HD broadcast channels is the input for the sub-pilot, with H264 video codec and AAC audio codec (the Spanish HD DVB-T channels). For testing purposes, we use i24 News TV[20] that streams English news the whole day. This channel is not included in Spain, but VIC used this for the laboratory testing activities.

When the sub-pilot system starts to run, the launcher component is the *Live Broadcast-Internet Subtitle Viewer* and Synchroniser Component ("sub viewer and sync." in **Figure 49**). This module is responsible for capturing the TV channel. This component launches the *Automatic Subtitling Component*, which is located on the same server. The "sub viewer and sync" sends to the *Automatic Subtitling Component* the audio signal (coming from the TV), an offset value (to reference the subtitling timestamps in accordance to the "sub viewer and sync" and a name to set a namespace for the incoming subtitles. In addition, the "sub viewer and sync" module also implements a Gstreamer-based DASH server, with the broadcast audio and video, plus the automatically generated subtitles. The "server" has an APACHE server and both the *Automatic Subtitling Component* and the *Machine Translation Component* publish the EBU-TT-D files there. The "sub viewer and sync" component creates the DASH stream and publishes an MPD file which identifies the locations of the different sources.

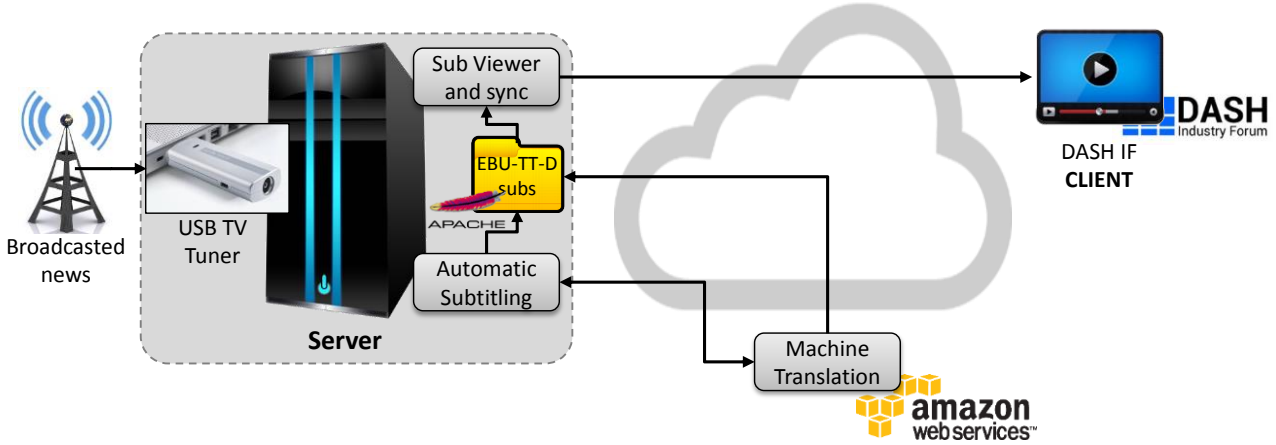

Figure 49. Deployment diagram of the Automatic HbbTV multilingual news broadcast subtitles sub-pilot service in VIC's Digital TV lab

The *Automatic Subtitling Component* is launched by the "sub viewer and sync" component. It creates an English transcription from the English audio and provides two outputs. On the one hand, it creates the EBU-TT-D files to feed the DASH server. On the other hand, it launches the *Machine Translation Component*, which is located in the cloud as an Amazon web service. The *Automatic Subtitling Component* communicates with the *Machine Translation Component* through a web service, exchanging a JSON file with the English transcription to be translated to Spanish, as well as the timing information, namespace information, etc.

The *Machine Translation Component*, when it is launched by the *Automatic Subtitling Component*, creates Spanish EBU-TT-D files, which are pushed to the same folder in the "server".

[20] http://www.i24news.tv/en/

8.2.4.2. Reception & User Application

The main output of the Automated HbbTV multilingual news broadcast subtitles sub-pilot is a DASH stream with the broadcast audio and video, and the automatically generated multilingual subtitles. A DASH streaming protocol is being used because of the interoperability it provides in comparison with other streaming protocols. DASH is being included and developed widely and it has been included in the HbbTV standard since the 1.5 version. HbbTV 1.5 includes a simple profile of MPEG-DASH which is being used successfully today and HbbTV 2.0 extends this to add extra features.

The deployment of the client-side application has been used with a modified DASH IF client. VIC has extended the DASH IF implementation to be able to show multilingual EBU-TT-D subtitles. At the current stage, the client implementation is only supported by a PC-based browser, such as Google Chrome of Firefox.

The next step, during the last months of 2015, will be to test an HbbTV 1.5 device and evaluate if the MPEG-DASH profile is sufficient to render the DASH stream and modify it to be able to show EBU-TT-D subtitles. Some difficulties are expected to provide an HbbTV-based client that fulfils all the functionalities already performed by a desktop browser. However, with the extra features provided by the HbbTV 2.0 version, the followed approach anticipates interoperability with the final HbbTV user application.

The use case will be to have broadcast English news content, without subtitles, and a Multilingual Service signalled in HbbTV. The users will be able to choose to consume the live broadcasted content without subtitles or to launch the HbbTV service. When the HbbTV service is launched, all content will come from the Internet, including the video and the audio. It will still be live content but delayed in comparison to the broadcast video due to the automatic generation of subtitles and to the delays imposed by the DASH server. The user will be able to switch between the subtitles generated in different languages.

8.2.5. Customised Subtitles for Online learning (MOOC) - Germany (VSX)

From the application point of view, the adaptive subtitle component that was developed by Vsonix and described in 3.13 was integrated in the online learning platform that was extended by Vsonix during the course of the project. The overall approach also includes the usage of the AccessGUIDE, which is part of the UI adaption service provided by Vsonix in WP5. In WP3, Vsonix are concentrating on the adaptive subtitling part of the application and its underlying service.

8.2.5.1. Production

Vsonix has defined a workflow for its subtitle production which is in line with the company's webcast production and post-production process. Based on the *vSubtitle* component (see section 2.3.6), the company has already extended its service offering for corporate learning customers, offering subtitle production and rendering for content localisation. This is often needed by multinational industry customers that need to translate their training materials (e.g. video lectures and webcasts) to be used in international subsidiaries.

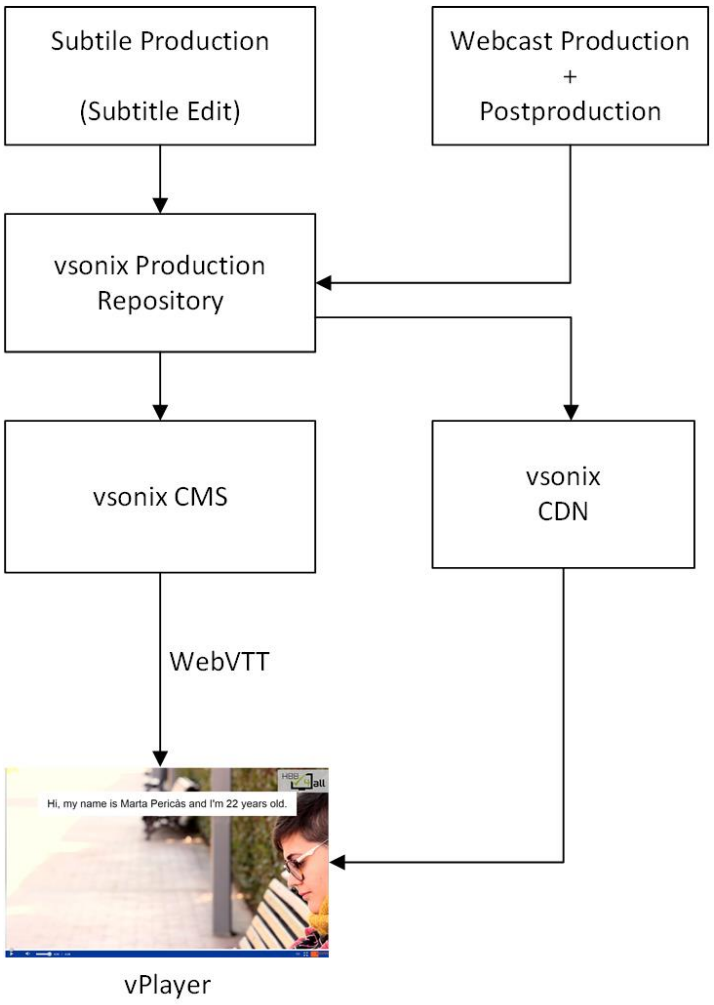

Figure 50. Subtitle production workflow at vsonix

For the production of the subtitles, Vsonix utilizes an open source subtitle editor called Subtitle Edit (http://www.nikse.dk/subtitleedit/), which has all the features for subtitle editing that Vsonix requires. The editor supports a number of formats and advanced editing functions. As an open source solution it has the potential to also be used in a database driven production environment as Vsonix is using. The subtitle production occurs in parallel with the webcast production and post production process. The content Vsonix is producing is mainly on demand webcast content for corporate learning. Both (subtitles and content) are managed and stored in the company's production repository during the editing process before they are published by the vPlayer based using the Vsonix CMS and CDN for content delivery, see **Figure 50**.

8.2.5.2. Publishing & Delivery

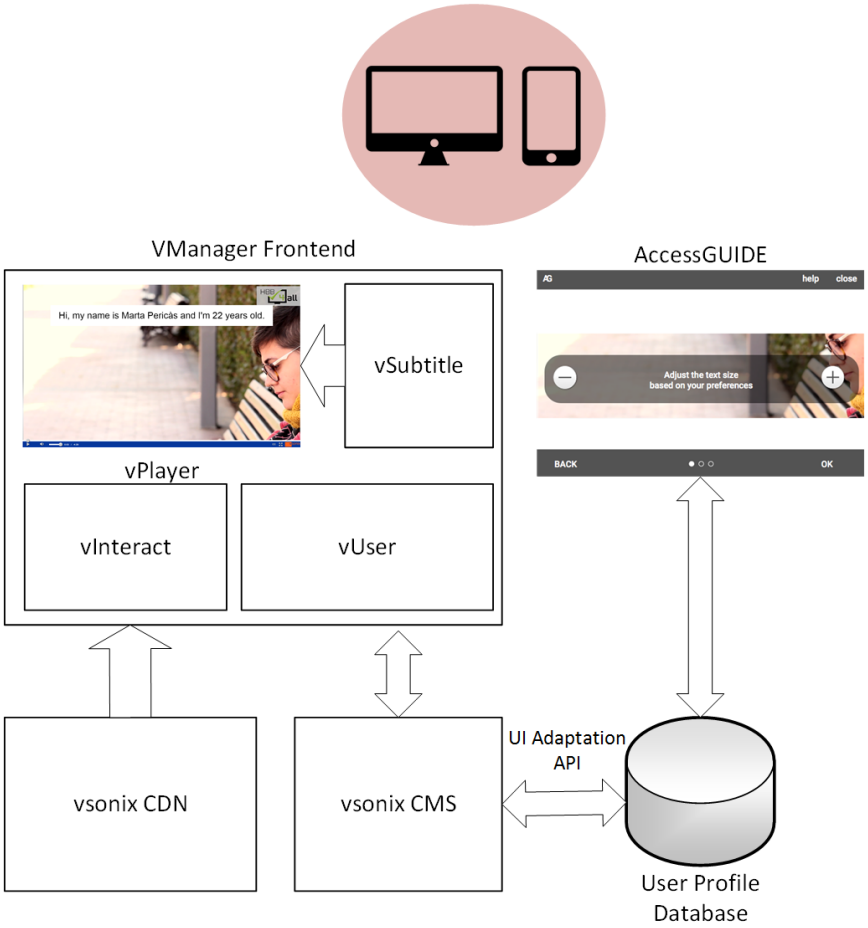

Figure 51. Delivery of adaptive subtitles at vsonix

As described in the chapter before, the edited subtitles are published via the Vsonix CMS, which as well as the video content that is published by the Vsonix CDN, stores all the data and metadata needed for the MOOC as well as the MOOC application itself.

The MOOC application is integrated and published using the Vsonix VManager service component. The vManager was designed to integrate entire webcast solutions including different functionality for interactive content access, user management, content search and retrieval as well as interactive user communication as a "micro-site". This micro-site can be integrated on a customer's corporate learning platform via <iframe> or <script> tag while it is hosted and managed by Vsonix via the Vsonix CMS and the Vsonix CDN.

In case of the MOOC, the integrated solution provides a number of functions for video lecture access via playlists and content catalogues, full text search on the lecture content, user access management, content related comments (in sync with video content) as well as a timely synchronized interactive question and answer module, which give learners the possibility to rehearse the content learned, while consuming it.

The playback of the content itself is done by the vPlayer component. For this we have developed an add-on for personalized subtitle rendering (*vSubtitle*, see section 2.3.6 and Annex 8.1.10). The subtitles are published via the Vsonix CMS. The adaptation parameters for the individual users are derived from the user profile databases provided by the UI adaptation service. It is accessed by the CMS using the UI adaptation API provided by Vsonix. The user profiles are generated using the AccessGUIDE application, which is a white label service that can be used also by other services to derive user preferences for UIs and subtitles (see **Figure 51**).

8.2.5.3. Reception & User Application

In the following section we will describe the end user perception of the personalized subtitle rendering based on the vSubtitle component as well as for the AccessGUIDE application.

Figure 52. Adaptive subtitle rendering in the vPlayer Component of the MOOC

As it can be seen in **Figure 52**, the *vSubtitle* rendering component is capable of rendering the subtitles in different variations, based on the user profile that was derived from the user profile database of the UI adaptation service (see above). The supported parameters for adaptation include font size and color, transparency level of the background as well as background color. Moreover, the user can choose the font type that is most suitable for him.

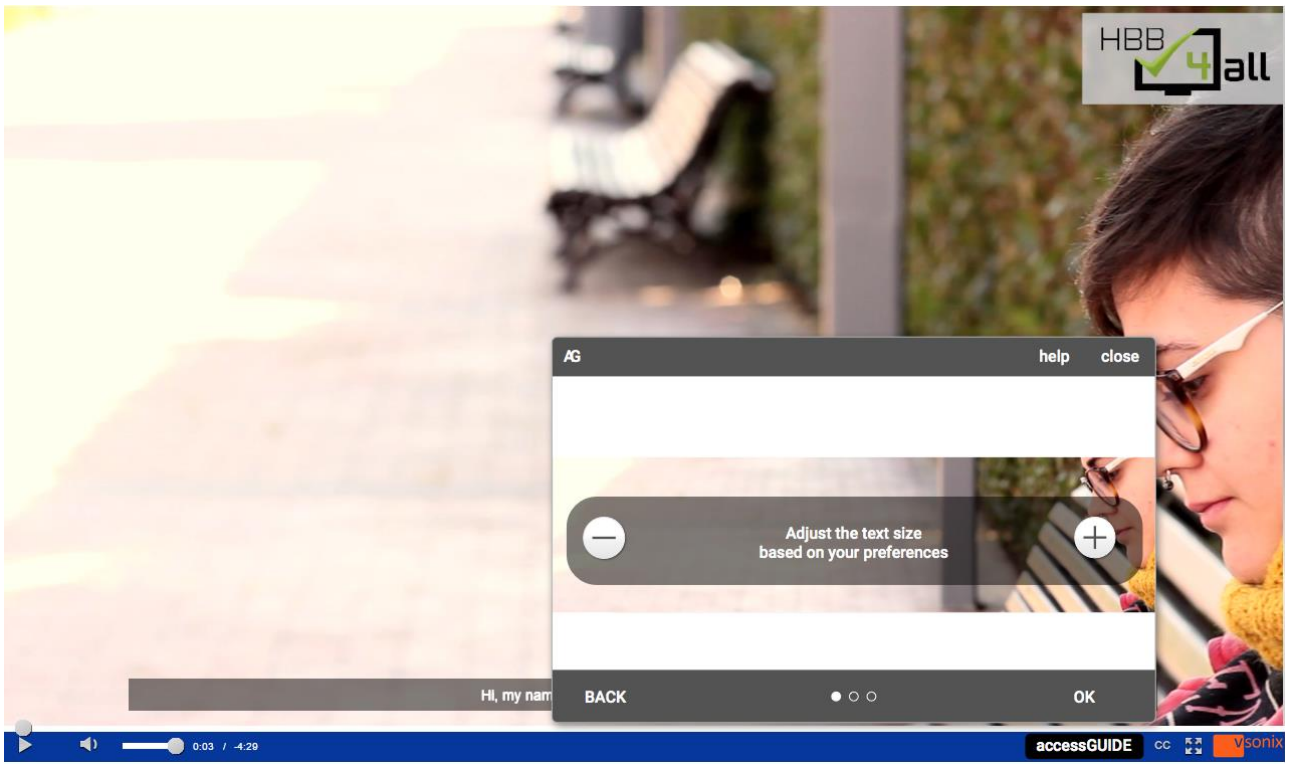

Figure 53. AccessGUIDE integrated in MOOC application

The subtitle parameters are defined by the user using the AccessGUIDE application as part of the UI adaptation service. This is already integrated in the MOOC to provide the necessary personalization features for the UI (see HBB4ALL deliverable D.5.2 [7]) as well as for the subtitle adaption. **Figure 53** shows the AccessGUIDE integrated in the MOOC application.

8.2.6. Testing Environment for End-To-End Subtitle Workflows (IRT)

8.2.6.1. Production

The Testing Environment is a rebuild of a typical broadcaster's production system. The system is used to test the potential integration of new services into an existing subtitle workflow. One of the main requirements in this scenario is that existing services continue to be delivered with their respective reliability.

The central component of the subtitle workflow typically is called the Subtitle Server. It manages all subtitle sources and the synchronisation to one or more given Timecodes. Typically it supports different output formats, including DVB subtitles and different carriage formats for Teletext. The management of subtitle sources mostly relies on an automation system that triggers activation of e.g. pre-prepared subtitle files. Traditional distribution channels will very likely depend on such a device in the system chain and thus the supported interfaces play a key role when integrating new subtitle services on the production side. Currently EBU-TT is not supported by such devices. In the HBB4ALL Test Environment a subtitle server from Screen Systems is implemented. It has been configured in such a way that it supports EBU STL and live input, DVB rendering, and Teletext generation. The Subtitle Server will be operated manually since an automation system is not used in the Test Environment (see **Figure 54**).

On the production side the broadcast signal is typically carried in HD-SDI, coming either from a live source or from a playout server. The HD-SDI signal also contains timecode information that is needed to synchronize pre-prepared subtitles to the video. Information for Teletext subtitles are inserted into HD-SDI by the Subtitle Server, where the timecode from the HD-SDI is used to trigger the subtitle insertion to ensure a synchronized playout. Two formats are common: for carrying Teletext in HD-SDI, OP47 and SMPTE 2031. Both are supported by the Subtitle Server used.

Generating DVB subtitles requires additional timecode information. As the Teletext subtitles they are generated in the Subtitle Server and triggered by the HD-SDI timecode (in pre-prepared scenarios). The Subtitle Server creates a DVB subtitle PID. For a correct synchronization this PID needs to be synchronised with the DVB timeline that is generated by the DVB encoder. Therefore, the encoder's output is routed to the Subtitle Server that then retrieves the DVB timecode information from the stream (see **Figure 55**).

The DVB encoder is fed with HD-SDI (including Teletext) and transmits a DVB stream comprising audio, video and Teletext (subtitles). A DVB multiplexer receives the DVB stream from the encoder as well as the DVB subtitle stream and merges them to one TV program. The merged DVB stream is then modulated and sent to the receiver via DVB-T.

There are two subtitle sources: 1) subtitles that are transmitted live to the subtitle server for immediate playout and 2) pre-prepared subtitle files that are copied to the subtitle server in advance. The latter use the EBU STL file format. For live transmissions there are a number of formats used, mostly proprietary ones. The Nufor format is a simple but commonly known format that is not standardised, however. Both EBU STL and Nufor are formats that comply with the Teletext format and thus limit the carried information to what is needed for teletext subtitles. For subtitle creation various editor applications are available on the market, supporting EBU STL as well as Nufor. WinCaps software is used in the Testing Environment both for editing live and pre-prepared content.

For pre-prepared content and for integration of web-based subtitle services, exported EBU STL files will additionally be transformed into EBU-TT using the *Subtitle Format Conversion Framework* (SCF, see sections 2.1.3 and 8.1.12). As a basic solution subtitles can be used for web playout that way. In a long term this solution is not ideal, since the EBU STL format misses much of the style information that is needed for web playout. The aim is to author subtitles in EBU-TT format directly. Thus the *SCF* will be extended to support a transformation from EBU-TT to EBU STL.

The management of live content is slightly more complex. Therefore, a different approach is used to provide subtitle playout for web in live scenarios. The uncompressed HD-SDI signal including Teletext subtitles will be supplied to the web encoder that creates an MPEG-DASH stream including subtitles in EBU-TT. The encoder needs to support that functionality.

In the following, a description of two scenarios further illustrates the setup of the Test Environment. For simplicity the generation of Teletext and DVB subtitles are described separately. However both services can be delivered simultaneously.

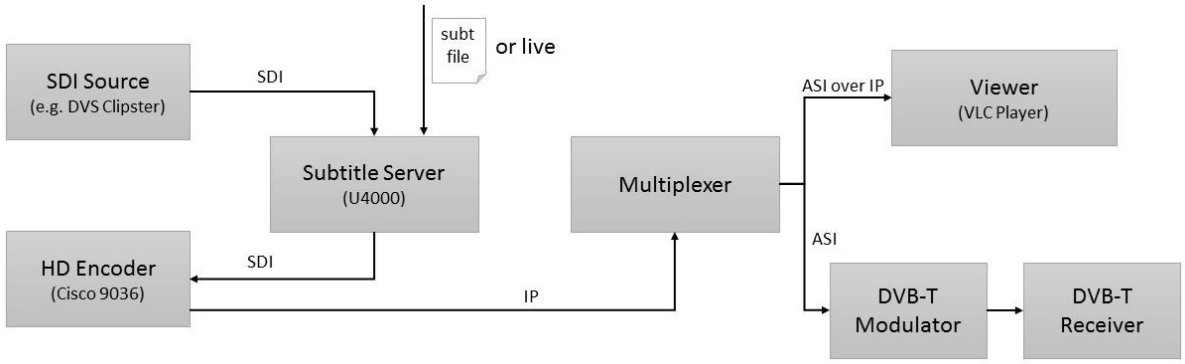

Figure 54. Overview of the test workflow for Teletext Subtitles

Signal chain for Teletext subtitling:

As shown in **Figure 54,** the SDI source sends an SDI signal to the Subtitle Server. In the Subtitle Server the subtitles are either played from file, based on the time code in the SDI Signal, or they are received live. The subtitles are then embedded into the SDI signal (as SMPTE2031 or OP47) and the SDI signal is routed to the HD encoder. The encoder creates an MPEG2 Transport Stream (TS) from audio and video and also encodes the embedded subtitles into a Teletext PID. The TS is then transmitted to the multiplexer (via IP). In this case the multiplexer merely needs to create the necessary DVB tables so that the signal can be processed by a DVB modulator. It is also possible to send the TS via IP e.g. to a VLC player, which is an easy way to monitor the multiplexer output.

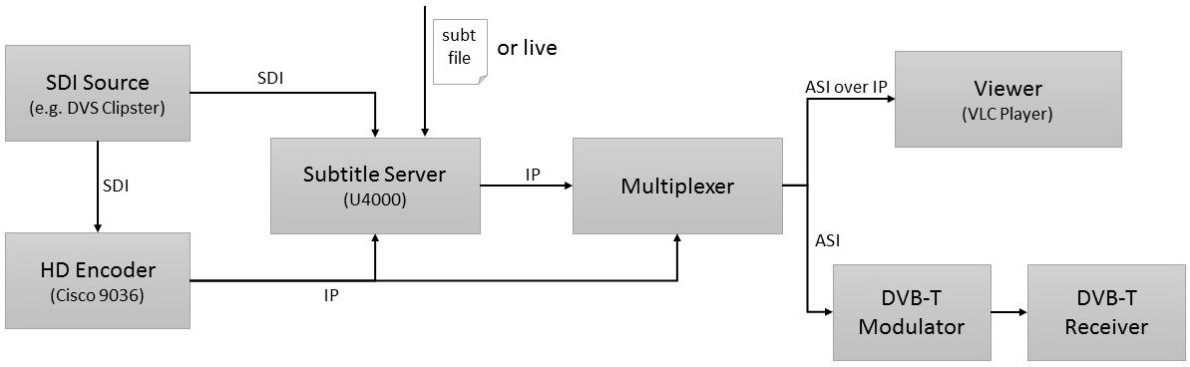

Figure 55. Overview of the test workflow for DVB Subtitles[21]

Signal chain for DVB subtitles:

As shown in **Figure 55**, the SDI source sends an SDI signal to both the Subtitle Server and the HD encoder. The encoder creates an MPEG2 Transport Stream (TS) from audio and video. The TS is sent to the Subtitle Server as well as to the multiplexer via IP. The Subtitle Server acquires the PTS timestamps from the TS, which are needed to create a DVB subtitle PID that runs synchronous with the TS. In the Subtitle Server the subtitles are either played from file, based on the Time Code in the SDI Signal, or they are received live. The subtitles are then encoded and put into the DVB subtitle PID that is sent to the multiplexer via IP. In the

[21] Please note that the operational use of the components / products may differ from their use in our test workflow

multiplexer all PIDs are combined into a single Program Stream. The output of the multiplexer is a TS that can be sent via ASI to a DVB Modulator. It is also possible to send the TS via IP e.g. to a VLC player, which is an easy was to monitor the multiplexer output.

8.2.6.2. Publishing and delivery

Focus of the testing environment is production. For publishing and delivery traditional DVB transmission systems (e.g. DVB-T) are used.

8.2.6.3. Reception & User Application

Focus on this component is to ensure the backward compatibility of the modified playout system with existing receivers. Thus, standard Set-top Boxes and TV sets are used to test the broadcast signal. In this Testing Environment the DVB-T modulation is used to transport the MPEG2 Transport Stream to the end user device. Additionally, the Teletext component within the TS can be received and tested with a compatible software player (like the VLC Player) via RTP/IP.

For the development of the *HbbTV 2.0 Demo App* (see section 8.1.16) the Test Environment was used to generate test Transport Streams.