This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement n: 761999

EasyTV: Easing the access of Europeans with disabilities to converging media and content.

D1.4 - Final release of the EasyTV system architecture

EasyTV Project

H2020. ICT-19-2017 Media and content convergence. – IA Innovation action.

Grant Agreement n°: 761999

Start date of project: 1 Oct. 2017

Duration: 30 months

Document. ref.: Deliverable 1.4

Disclaimer

The EasyTV Consortium consists of the following partners:

	Partner Name	**Short name**	**Country**
1	Universidad Politécnica de Madrid	UPM	ES
2	Engineering Ingegneria Informatica S.P.A.	ENG	IT
3	Centre for Research and Technology Hellas/Information Technologies Institute	CERTH	GR
4	Mediavoice SRL	MV	IT
5	Universitat Autònoma Barcelona	UAB	ES
6	Corporació Catalana de Mitjans Audiovisuals SA	CCMA	ES
7	ARX.NET SA	ARX	GR
8	Fundación Confederación Nacional Sordos España para la supresión de barreras de comunicación	FCNSE	ES

PROGRAMME NAME:	H2020. ICT-19-2017 Media and content convergence - IA Innovation action
PROJECT NUMBER:	761999
PROJECT TITLE:	EASYTV
RESPONSIBLE UNIT:	CCMA
INVOLVED UNITS:	UPM, ENG, MV, ARX, CERTH, CCMA
DOCUMENT NUMBER:	D1.4
DOCUMENT TITLE:	Final release of the EasyTV system architecture
WORK-PACKAGE:	WP1
DELIVERABLE TYPE:	Report
CONTRACTUAL DATE OF DELIVERY:	30-09-2018
LAST UPDATE:	28-09-2018
DISTRIBUTION LEVEL:	PU

Distribution level:

PU = *Public,*

RE = *Restricted to a group of the specified Consortium,*

PP = *Restricted to other program participants (including Commission Services),*

CO = *Confidential, only for members of the LASIE Consortium (including the Commission Services)*

Document History

VERSION	DATE	STATUS	AUTHORS, REVIEWER	DESCRIPTION
V0.10	31/07/2018	Draft	CCMA	First draft Table of Contents, document structure and first contributions.
V0.20	07/09/2018	Draft	CCMA, MV, UPM, ENG, & ARX	Contributions
V0.30	19/09/2018	Draft	CCMA, MV, UPM, ENG, & CERTH	Contributions
V0.40	24/09/2018	Draft	ENG, CCMA	First release version ready for review
V0.50	24/09/2018	Draft	UAB	Review
V0.60	25/09/2018	Draft	CERTH	Review
V0.70	28-09-2018	Final Candidate	CCMA	Final candidate once reviewed by CCMA, UAB & CERTH

Definitions, Acronyms and Abbreviations

ACRONYMS / ABBREVIATIONS	DESCRIPTION
AD	Audio description
API	Application programming interface
ASR	Automatic Speech Recognition
C4A	Cloud for All project
CS	Companion Screen
CSS	Cascading Style Sheet
DASH	Dynamic Adaptive Streaming over HTTP
DRM	Digital Right Management
DoW	Description of Work
FAQ	Frequently Asked Questions
GPII	Global Public Inclusive Infrastructure
GPL	General Public License
GUI	Graphical User Interface
HbbTV	Hybrid Broadcast Broadband TV
HTML5	HyperText Markup Language v5
ID	Identifier
IT	Information Technology
MGMT	Management
MPD	Media Presentation Description
NLP	Natural Language Processing
OCR	Optical Character Recognition
REST	REpresentational State Transfer
SDK	Software Developer Kit
SL	Sign Language

SOAP	Simple Object Access Protocol
ST	Subtitles
TTS	Text To Speech
UCD	User Centred Design
UI	User Interface
UIDLs	User Interface Definition Languages
VUI	Voice User Interface
WSDL	Web Service Description Language

Table of Contents

List of Figures

Executive Summary

This document describes the needed technical system specification of the EasyTV platform. It provides the specification for the functionality and basic architecture based on the results of WP1's user cantered design and the resulting requirements established in T1.1 and T1.2. This includes an initial functional description, leading to a detailed platform related specification for each EasyTV services to be implemented, including production modules, delivery and end-user equipment.

EasyTV follows a user centered design, where project developments are driven by real user needs. This is achieved by involving these users in each step of the design and implementation of the EasyTV project. A group of end consumers and professional users were consulted in order to define the accessibility requirements needed with the aim to provide an equal access to TV and audio-visual services to all citizens, including those with different sensorial issues, like for example low vision or deaf persons.

These consultations allowed collecting a detailed list of requirements which were considered to define the EasyTV platform satisfying the required functionalities and defining a basic architecture.

The document has been structured in 4 chapters:

Chapter 1 provides an introduction and overview of this document and the project and describes the interconnections of WP1 with the other packages within the project.

Chapter 2 defines the methodological approach adopted to obtain the architecture of the platform.

Chapter 3 defines 'How to do it'. This defines the basic architecture of the EasyTV platform, divided into different functional modules, and involves not only the definition but the technologies that will be used for the development.

Chapter 4 analyses the deployment and test phases, which are vital in the project life cycle, since they represent the main tools for detecting issues and planning improvements that have to be applied to the system in order to achieve an optimal performance.

The information in this document will be used as the basis for WP2, WP3, WP4 and WP5, where the technical architecture of the EasyTV platform will be studied in a much deeper technical way to develop the multiple platform-based services classified in:

- Improved Access Services
- Novel Technologies for Interaction
- Improved Personalization
- Crowdsourcing platform for accessibility production

Closed and open pilots in WP6 will allow testing and validation of the results from WP3 to WP5 and further refine these requirements, functionalities and architecture in a second iteration of the EasyTV platform specification.

1. INTRODUCTION

This deliverable describes the final version of the EasyTV architecture according to the project objectives comprised in the Description of Work (DOW) [1] of the project and it starts from the users' requirements gathering in document D1.1 [2] (User scenario and requirements definition) as well as from the scenarios and system specifications reported in D1.2 [3] (System requirements specifications).

The document, which is an evolution from previous Deliverable D1.3 [4], is divided into four main sections:

- Introduction, which describes the overall document and involves the project overview and WP1 interconnections with other packages.
- The methodological approach adopted to obtain the architecture of the platform.
- The EasyTV architecture definition, which involves how it is defined and how it will be tested.
- The EasyTV deployment & test plan

1.1. Project Overview

The EasyTV Project aims to provide an equal access to television and audio-visual services to ensure all users, including persons with various degrees of disabilities regarding sight and hearing, are able to participate in the multimedia environment on equal basis, removing the actual barriers and improving the user experience for this kind of mainstream products and services in terms of choice and quality.

EasyTV will focus on developing multimedia and sound services with novel accessibility features that will enhance multimedia experience by making it more cost-efficient and yet more flexible to use, which is directly related to its ease of use. Specifically, it is defined to break the language barrier for all by developing technologies which will improve the interaction by providing a wide set of innovative tools for accessing the visual content which may lead to the inclusive media scenario proposed in the environment of the Information Society for all.

The project considers past developments in accessibility services for digital TV as starting points (such as DTV4All[1], Cloud4All[2], Prosperity4All[3], Hbb4All[4]), as well as multiple devices and platforms in the definition, design and test specific interaction tools. In this way, it's focused in three main objectives:

- To **provide well-proven access services** to be adapted and improved in new media delivery mechanisms.
- To **improve the interaction** by following a multi-language approach, including sign language & subtitling, which vary from one country to another.
- To **ensure a user-centered approach** since the end users will be at the centre of the development, giving the guidance through a close cooperation that will facilitate a more cohesive and integrated media content.

That said, the outcomes of EasyTV will be delivered "as-a-service" that will be joined in a platform which will be further refined to reach the required maturity levels through a detailed testing and the

[1] http://www.psp-dtv4all.org/

[2] http://www.cloud4all.info/

[3] http://www.prosperity4all.eu/

[4] http://www.prosperity4all.eu/

feedback analysis. Regarding this, the main scientifically and technologically challenging contributions that the EasyTV project will offer are:

- **Improved access services** for enhanced multimedia visual and sound experience for people with specific disabilities, by adapting the content through different aspects:
  - Image adaptation: focusing on improving accessibility for people with different visual impairments.
  - Content description improvement by adapted audio narratives, clean audio, etc. that will help people with hearing disabilities.
- **Improved personalisation of the content and interaction**, towards a hyper-personalised experience to all. This will be based on an auto-personalisation from profile tool and it will be focused on:
  - Adaptation of the content provision using DASH streaming services.
  - Personalised services, which may include self-adaptive and tailored services which can learn from users' actions.
  - Recommendation of available access and interaction services according to the user's profile.
- **Novel technologies to break the sign language & subtitles barrier** by the development of solutions which can perform automatic translation in different languages with the aim to allow an inclusive media consumption. This will provide, for example, a realistic sign avatar that will be available in different languages, which can be widened in the future by means of a crowdsourcing platform, as well as a definition of a multilingual ontology that will map sign to concepts.
- **Innovative solutions for voice and gesture/gaze recognition to control the TV set** and applications that will be delivered as a part of an universal remote control for easing the interaction.

1.2. WP1 positioning and interdependencies

This document summarizes the tasks carried out within the WP1 of the project, which started with the identification of user scenarios and the users' needs as an essential task for the definition of an optimal and clear system specification which, at the same time, will allow to design and develop the EasyTV architecture with a core set of components for the entire EasyTV system.

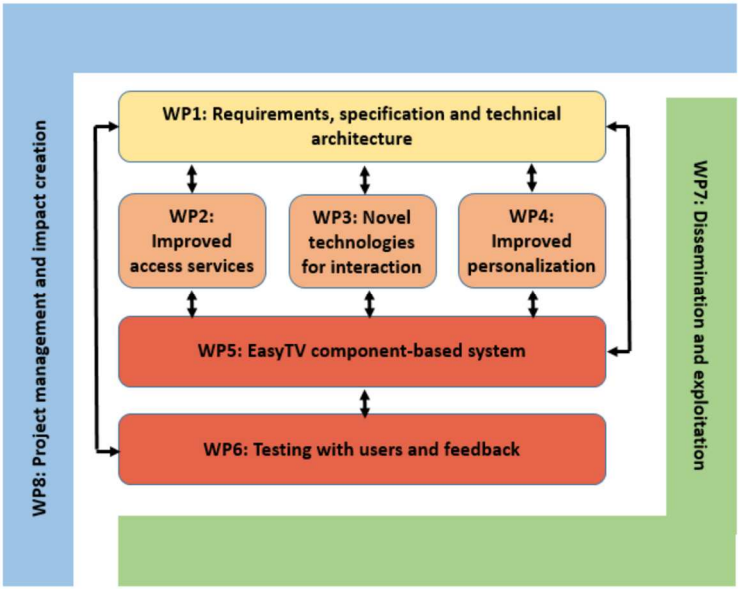

Figure 1: EasyTV Pert Diagram

More specifically, WP1 provides technical requirements from the innovation needs, gathering them from the end consumers and professional users. These requirements guide the system specification and the design of the technical platform which delivers different services that are implemented along the different technical work packages (WP2, WP3 & WP4).

The platform will be assembled in WP5 (together with a service catalogue, a service registry and a service development kit) and the whole EasyTV system will be completely tested with end users in WP6. Feedback gathered from the users within WP6 will help to update the system by providing relevant guidelines for its improvement in the second iteration of the platform.

2. METHODOLOGY FOR REQUIREMENT IDENTIFICATION AND ARCHITECTURE DESIGN

The design process that led to Requirement Identification and Architecture Design started from the evidences emerged in the research activities carried out in Task 1.1 – End user requirements gathering.

The outcomes of the focus groups with the super end users present in D1.1 [2] were translated into a system of functional specifications through a series of steps, based on the methodology of the User Centered Design (UCD) in the D1.2 [3].

2.1. The methodology

The design activities of the EasyTV architecture were based on the focus group carried out with the end users on the D1.1 [2].

These system requirements were achieved with a series of steps based on the methodology of the User Centered Design elaborating the needs gathered by the users and applying the indications provided in the methodology of the *User Sensitive Inclusive Design* [1].

This approach has led designers to elaborate extremely holistic scenarios, finalizing each story in a framework that integrates different solutions from the beginning in a native way, without incurring successive overlapping operations that could then be addressed only with the logic of additional components.

The requirements emerged from the needs of end users involved in the focus groups. These were collected and managed in order to be used to create and define the different profiles of **personas**[5]**,** strongly characterized in the normal course of everyone's daily life.

The personas were then used as protagonists in the preparation of **scenarios**[6], representations in which specific users pursue a certain objective in a specific context of use. The scenarios have been elaborated starting from short stories in which personal experiences by users regarding specific contexts of use are narrated. The scenarios provided contextualization of different moments in different environments.

Afterwards, the **User journey**[7] of each scenario was elaborated in order to visualize in a schematic form the description of the scenario. This is a map to show the main dimensions in which every single scenario operates (daily routine, context of use, interaction channel), articulating them for each set

[5] **Persona** is an archetype or character that represents a potential user of the system that will be developed. In a narrative way, personas express and focus on the major needs and expectations of the user groups.

[6] **Scenario** describes a basic story of an action or goal that a user wants to accomplish in a specified context of use.

[7] **User journey** describes the journey of a user by representing the different touchpoints that characterize his interaction with the service.

of expected tasks.

The subsequent activity concerned the generation of the **User tasks**[8] related to each analyzed persona, to define the objectives that the user is trying to achieve with the use of the EasyTV platform, the ways in which it can reach them, but also how it is influenced by the environment in which it is located and by the available technologies.

A further step in defining the users´ needs was elaborating the **User stories**[9], continuing to work on the analysis of the results of previous analysis activities.

Each single defined task represents a high-level transition of the user experience. This level of definition has been further analyzed and decomposed into a series of sub-tasks, to specify in more detail all the micro-objectives of the user that will have to be satisfied through one or more functional features of the system.

The description of each sub-task was then elaborated with the perspective of the end user, using a format organized on a construct articulated as follows:

> As a [type of user] I want to [do something] So that I can [get some benefit]

2.2. End user requirements gathering

In document D1.1 [2] users' needs, requirements and specification were gathered in order to feed the technical architecture. To collect them a specific methodology was defined.

Firstly, "super end users" were involved. This term is used to define end users who, besides being regular users, also have some knowledge on the technologies foreseen for the EasyTV system. The preliminary consideration was that it would make no sense to consult end users with no knowledge or experience with neither functional diversity nor technological background since the issue was related to technology development and user expectations in order to match the innovation.

The criteria adopted to select "super end users" was based on the following points:

- Users who are experts regarding the world of sight and hearing loss people, which is the target of EasyTV project.
- They don't necessarily have to be people with sight or hearing loss, but of course is better if they are because they are also an everyday user of the technology. The important thing is that they know the technology and their needs since they deal with them every day.
- Users who are everyday users of the technology because they teach the technology to other users, so they should be able to tell the researchers, better than others, what the final users really need.

Regarding methodologies identified for collecting super end users' needs and expectations, the Focus Group option was selected between different research techniques, in order to have a more interactive interview and to *share and exchange* opinions across experts during discussions on user needs. Focus groups are typically useful for exploring arguments, especially when little is known about the question of interest. Moreover, focus groups are most commonly used at the beginning stages of a research project and they are an interesting way to share ideas, and express opinions and attitudes. Focus groups also generate ideas or gather feedback about what final users think

[8] **User Tasks** are steps that users currently take in order to achieve their goals.

[9] **User Stories** are short, specific and goal-oriented descriptions of what a user will do with a part of a system. User Stories have the following structure: "As a , I want so that ".

about the discussion topic (products and/or services).

Two Focus Groups were conducted, one for each type of disabilities: people with visual and hearing loss, according to the target of EasyTV project. For these two Focus Groups two groups were defined: one of five super end users for people with hearing loss and one of six super end users for people with visual loss who were requested feedback.

The process to gather information from super end user expectation during the Focus Group consisted in:

- Stage 1: Preparation and first interaction
  - Ethical permission
  - Questionnaire for data gathering
  - Drafting/Translating Consent/Information forms

- Stage 2: User scenarios
  - Asking for the main component of the system (and user scenarios related to services and their functionalities)
    - Matchmaking
    - Avatar
    - Crowdsourcing platform
    - Image magnification
    - Audio Narratives
    - Speech Interface
    - Gesture/Gaze Interface
    - Sign Language capturing technology

- Stage 3: User expectations
  - Asking about interoperability, manageability, scalability and exploitation

The Focus Groups outcomes gathered in D1.1 [2] have been translated into users' requirements and system specifications collected in D1.2 [3].

2.3. EasyTV system requirements specification

The system requirements were presented in D1.2 [3], subdividing them into three sections: **Functional requirements**, which described the design of functionality, **Component specification**, which organized the functional macro components previously described, and **Non-functional requirements**, which illustrated the non-functional constraints of the system.

The **Component specification** section was organized to group functional requirements into logical clusters. This organization does not aim to identify the individual software components of the architecture that are instead described appropriately in the following chapter.

The requirements identified guided the system specification and the design of the EasyTV platform architecture which will deliver different services that will be further implemented in different Work Packages (WPs).

3. EASYTV ARCHITECTURE

3.1. Introduction

User requirements are essential to define the system requirements specifications.

In the same way, these user requirements and the resulting system requirements are the basis of a successful design of the EasyTV architecture, which will guide the development process in work

packages WP2, WP3, WP4.

The following sections will list the system functionalities and break out the different service components, each one formed from several functional modules that define the EasyTV architecture.

3.2. System Functionalities

The main aspect of the EasyTV platform is the multiple functionalities it provides with the aim to allow users with different disabilities accessing broadcaster content in an easy way. In this regard, each platform-based service module that is going to be developed represents an innovative service which can be considered easy to use, low cost and useful for improving the interaction with terminals and to access multimedia content.

That said, the **system generic services** that have been defined together with their main functionalities are the following:

- **Audio description:** this service will provide different solutions for helping people with visual loss for accessing the multimedia content. In particular, it will comprise two different services:
    - **Automatic descriptive narratives:** this service will be able to provide additional information about a content derived from the related metadata. Moreover, this service will be also in charge of analyzing the video content in order to detect textual information and to extract it in order to give the users some contextual data that may help them to have a more complete experience. This will be done by a OCR service.
    - **Automatic voice synthesis of subtitles:** also known as audio subtitles, this service will be in charge of providing audio information obtained from the subtitles files. As in the previous one, this service will make use of a OCR solution to enable spoken subtitles for burned-in video subtitles, or simply by parsing them if they are provided in a standard protocol.
- **Clean audio:** this service will help to improve the intelligibility of access services by providing a smart equalization service according to different profiles. In this regard, and advance process of the audio will be done, providing a tool for emphasizing the proper audio frequencies for improving their access in terms of dialog intelligibility by user with different hearing losses.
- **Universal remote control:** this service will enable a universal and accessible interaction between users and TV through the recognition of different users' gestures, gaze estimation and speech recognition (voice control).
- **Image enhancement:** this service will be focused on the processing of an image for obtaining a more suitable bespoke service, adapting it to the requirements of the user, especially when this user has impaired vision. This will include different functionalities such as:
    - Improving the presentation of subtitles and sign language video, in a customizable or an automated manner.
    - Magnifying specific elements in the image, such as faces and text, through custom or automated selection.
- **Crowdsourcing platform:** this service will allow the creation and management of a new sign language & subtitling framework in which, for example, users will be able to upload and collect sign language content, as well as to check their correctness.
- **Realistic avatar for sign language presentation:** this service will provide different SL solutions for people with hearing loss that will allow:
    - The generation of automated realistic avatar for sign language content, combining head and hands movements for a closer experience.
    - To add new sign language representations through crowdsourcing by making use of a real-time hand and face motion capture solution.
- **Automated translations among different sign languages:** this service will help the

annotation of sign language captures and to translate original sentences in other languages by means of a multilingual ontology that can be populated through the crowdsourcing platform.

- **Subtitling translation:** this service will allow the production of subtitles with tools to automate translation to different languages and allow human monitoring and improving through a crowdsourcing platform.
- **Hyper personalised access:** this service will be focused on enabling automated services' personalisation and interface adaptation according to users' profile and context information. This will also include matchmaking for personalised DASH streaming services.

3.3. Basic Architecture

The basic architecture is divided in three blocks:

- **Broadcaster premises / Content Owner** – this block englobes the main workflows of the broadcaster or a content owner related to the management, storage, broadcast and publication of audiovisual contents.
- **EasyTV platform** – within this block several modules are grouped in service components that will be defined in the next sections.
- **Consumer platform** – end-users will consume the contents with accessible services through multi-platform devices like smartphones, desktops or SmartTV, interacting with their devices through improved accessible interfaces that will ease the access and consumption.

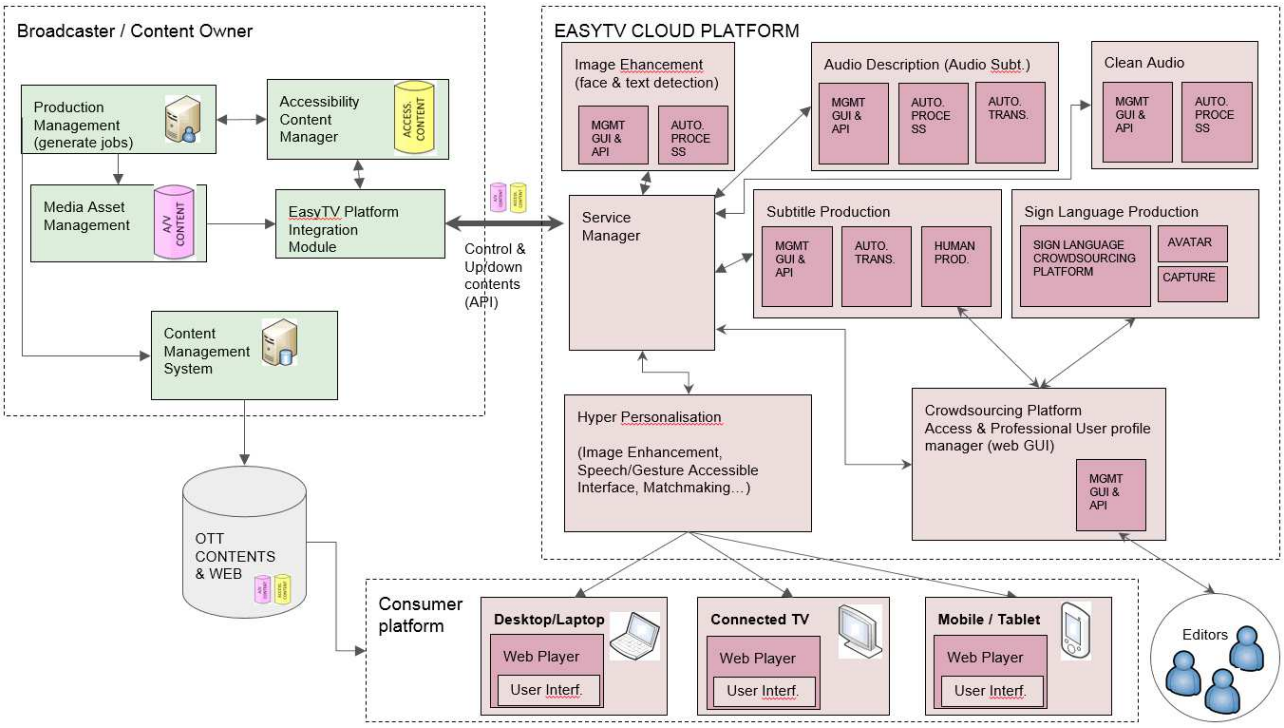

Figure 2: EasyTV Basic Architecture

3.3.1. Broadcast Premises

In the EasyTV architecture presented in Figure 2, a Broadcaster or content owner block has been considered as the professional client of the EasyTV cloud platform. This client can be the owner of the audiovisual contents, and/or has the rights for the broadcasting of them.

This client needs the outsourcing of accessibility contents production, and EasyTV platform offers this outsourcing with a range of tools available through its Service Manager which will offer an

interface for the requesting of production tasks that can involve one or more tool modules.

Several modules also exist within the internal premises of the broadcaster or content owner (Figure 3). This is a general overview of possible modules involved in the workflow from the content management to the final delivering through digital platforms.

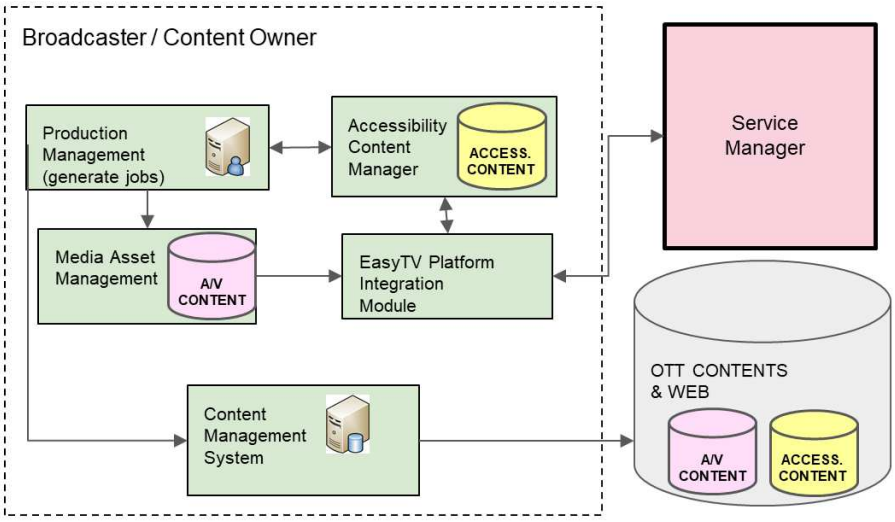

Figure 3: Broadcaster/Content Owner general architecture

Description of modules:

- **Production Management** – allows to manage the contents and rights, scheduling for the broadcasting, publishing to digital platforms, and requesting for accessibility contents.
- **Accessibility Content Manager** – Manages the database and archive of accessibility contents, with direct connection to the production management. Keeps the archive of accessibility contents identified with a unique ID related with the original Audiovisual Content to which it belongs. If the requested accessibility content is not available, it will request the job of production of this accessibility content to the EasyTV platform through its EasyTV Platform Integration Module.
- **Media Asset Management** - A MAM system provides a single repository for storing and managing video and audio files.
- **EasyTV platform integration module** – will allow the translation of broadcaster or content owner requested jobs to be sent to the EasyTV platform through the Service Manager API. Therefore, this EasyTV platform integration module must be developed by the client to allow the adaptation of internal workflows with the platform.
- **Content Management System –** A CMS is a software application or set of related programs that are used to create and manage digital content allowing, for example, the publishing of the broadcaster contents to digital platforms involving not only the audiovisual and accessibility contents but also the web page graphical and textual contents.

The broadcaster generates the video, audio & metadata files required for the publication of the audiovisual content to internet, using the adaptive bitrate streaming technique MPEG-DASH which includes the MPD (Media Presentation Description) file with all the metadata required for playing the video and related resources.

Required files are uploaded to the EasyTV platform through its Service Manager Module which will manage the workflows to author new accessibility contents and modify the MPD file including the

required metadata to signalize the new content.

The client application (Player App) accesses the MPD file and, depending on the user´s preferences and the Hyper Personalisation module in the EasyTV Cloud Platform, it selects from the MPD those contents that best satisfies its needs.

3.3.2. EasyTV Service Manager

The Service manager will be the main communication hub between the broadcasters' premises, and the different components, modules, and services at the internal of the EasyTV platform. It will act as a gateway and orchestrator of the full platform and it will allow an abstraction of the work and processes that can involve multiple modules.

The Service Manager will have two main components:

- A **web graphical user interface** (GUI), which will allow the broadcaster to request content from the EasyTV platform in a centralized and unified way. This web GUI will keep track of their request, status, and serve the contents to the broadcaster premises in a user-friendly way without the need of knowing the underlying structure of the EasyTV modules, services, and components.
- A **REST API**, acting as a middleware between the user interface (and other possible services that the broadcasters may deploy in the future) and the EasyTV services.

The Service Manager will be able to address the specific content (accessibility and audiovisual contents) request to the appropriate EasyTV component where the repository for that content resides. If the requested content is available, it will be directly served to the broadcaster. If the requested content is not available (or doesn't exists yet), the Service Manager will generate the needed tasks in the EasyTV modules, allowing the broadcaster to know in every moment the status of the request.

As an example to better illustrate the process, let's take in consideration the following example:

The broadcaster is interested in audio subtitles for an audiovisual production in original and English languages which are not currently available. The audiovisual production has its original subtitles file. The process would implicate multiple modules from the EasyTV platform, and would be handled through the web GUI of the Service Manager in the following way:

1) The broadcaster will request audio subtitles and, as they are not available, the request will be divided in several tasks:
    - Task1: Upload of the audiovisual and original subtitles files.
    - Task2: Automatic translation from original to English subtitles.
    - Task3: Human improvement and production of English subtitles through the crowdsourcing platform.
    - Task4: Generation of audio subtitling in both original language and English.
2) Once all the tasks are completed, the new accessibility contents would be available for downloading to the broadcaster premises.

As some of the tasks could be dependant of previous tasks, the Service Manager will take care of the correct workflow progress and monitoring.

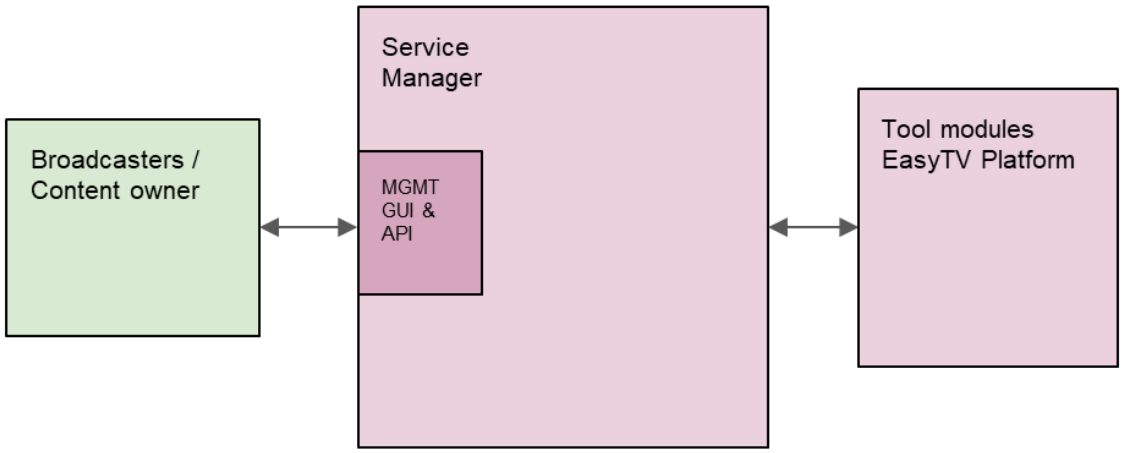

Figure 4: Service Manager.

3.3.3. EasyTV Audio Description

3.3.3.1. Automatic descriptive narrative

Audio Description (AD) allows persons with visual loss to hear a verbal version of the visual information, but it has been shown to be useful for anyone who wants to truly notice and appreciate a complete perspective on any visual event. Audio describers provide services for different kinds of contents and events, such as news, documentaries sports events, and on Internet web sites, but this process is costly and a high percentage of contents broadcasted lack of this type of service, having as a result that the targeted audience of these materials obtain reduced user experience. For this reason, it is necessary to automatize the process for obtaining high quality contents available for all users independently of their physical condition.

The descriptive narratives are generated in a technologically advanced module found in the EasyTV cloud server. The Audio narratives module generates a textual file similar to a subtitles file that contains valuable automatically generated data obtained from the intrinsic information of the multimedia contents and from different external information sources, such as the metadata associated to the file and the television guides. This information is related to the preferences of the user and it is suitable for obtaining real-time data about the contents that the user is watching. Moreover, this process will be completed with the textual information extraction from the video in order to provide more data that can improve the experience of the users, especially if they are visually impaired.

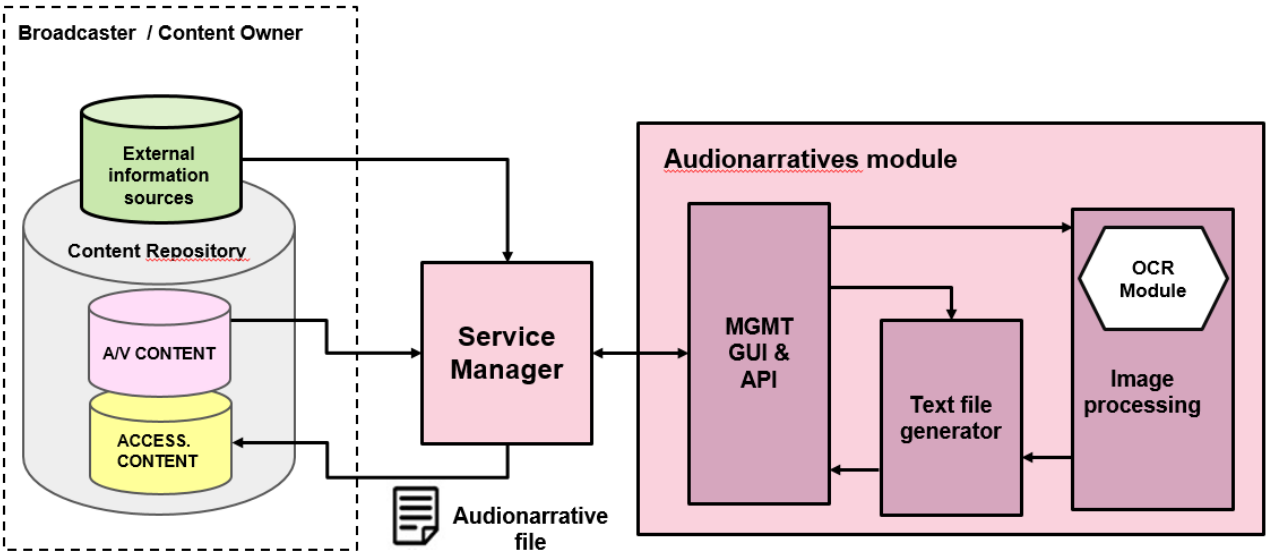

Figure 5:Audionarratives generation module in EasyTV Cloud Platform

The Image Processing module will use also a Server Based OCR provided on the "Automatic Voice Synthesis of subtitles architecture" that will be able to translate the burned text (burned subtitles or any other text) contained in the video into a text content and will generate the text file accordingly (i.e. EBU-TT-D or narrative file).

Once the audio narrative file is generated, the file is stored in the accessibility contents database through the service manager.

3.3.3.2. Automatic Voice Synthesis of subtitles

EasyTV project will enable spoken subtitles by making use of two different approaches: using a Text to Speech Service (TTS) available on the server-side architecture and another Text-to-Speech Service on the cloud for a more flexible and effective solution. Subtitles will be provided by the subtitling production platform available in EasyTV or directly through a Web Based application that can be used by professional users. The Automatic voice synthesis of subtitles will include a set of functionalities available through two different interfaces: the Web API interface and the Web GUI interface.

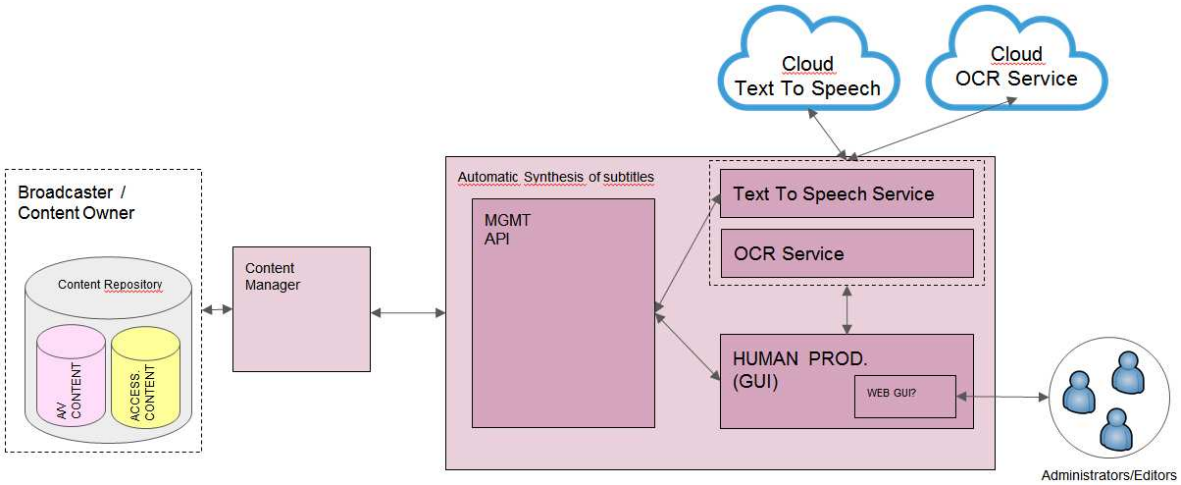

Figure 6: Automatic Synthesis of Subtitles

The Text To Speech Service (TTS) (local or on the cloud) will be also used for translating any text

in an audio track to be included or mixed to an audio video streaming content (for example for audio narratives or audio subtitling). The final user will be able to activate/deactivate the audio subtitling service and using both the client-side Text To Speech Engine to play the subtitles or the prerecorded audio subtitles stored in one of the audio track of the streamed content.

3.3.4. EasyTV Clean Audio

Dialogue intelligibility of television audio content is an important issue especially for people with hearing impairments. EasyTV platform presents an innovative solution based on Background Audio Enhancement technologies to separately adjusting the levels of the voices and the background sound in order to satisfy the requirements of the user.

Finding the right balance between ambient sound and dialogue is a key challenge for audio engineers and an often common source of complaints by the audience.

The technology of EasyTV Clean-Audio component allows the audience to change the balance of the audio mix according to listening environment or personal needs, obtaining important benefits for the hearing-impaired, including the people affected by hearing loss that increases gradually with age.

The clean-audio component interacts with the sound from the source multimedia content and develops a process for equalization of frequencies and advanced techniques of filtering for enhancing the intelligibility of the audio by identifying, processing and reinforcing the bandwidth where the most important audio information is located, especially the voices.

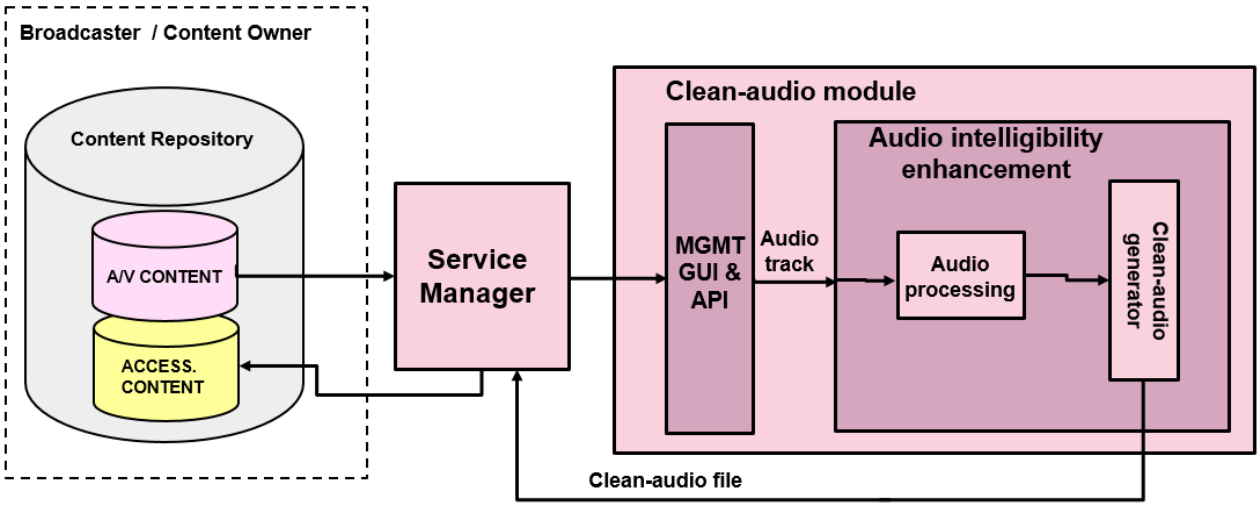

Figure 7: Clean-audio generation module in EasyTV Cloud Platform

On the client-side, the user can receive two different audio files, one containing the source audio file and the other the clean audio approach, which will improve the intelligibility of the contents for users with hearing impairments. The users can select one of these two files depending on their preferences. If a user selects the clean-audio file, he/she will have the possibility of dynamically adapting the conditions of the sound by configuring the volume of the voices with respect to the background.

3.3.5. EasyTV Subtitle Production

The use of subtitles on audio visual contents for facilitating the accessibility of hearing impaired people is a well-known issue, but, with the translation of subtitles into different languages, this kind of service can be also extended to bring access into the contents to those people that do not understand its spoken language, regardless of whether they have any impairment. Aligned with this concept, the broadcasters have also shown a keen interest in subtitling the content in alternative

languages due to the fact that this feature brings the opportunity to engage audience who normally do not understand the original language of the contents, and also to promote actions targeted to integrate immigrants through information technologies.

The crowdsourcing multi-language production service offered by the EasyTV platform constitutes a powerful and cost-effective tool to achieve the desired feature without having to assume the extensive bill of the traditional subtitling production and, at the same time, it looks for the voluntary collaboration of persons from those collectives that are going to take advantage of the produced material. The design of this module has taken into account the user requirements and the broadcaster constraints gathered in the deliverable D1.2 [3].

The technology of the EasyTV subtitle production module takes profit of some components in the cloud to guarantee a high quality of the produced contents such as an automatic translation service and the collaborative crowdsourcing validation workflows. The architecture of this modules is shown in Figure 8.

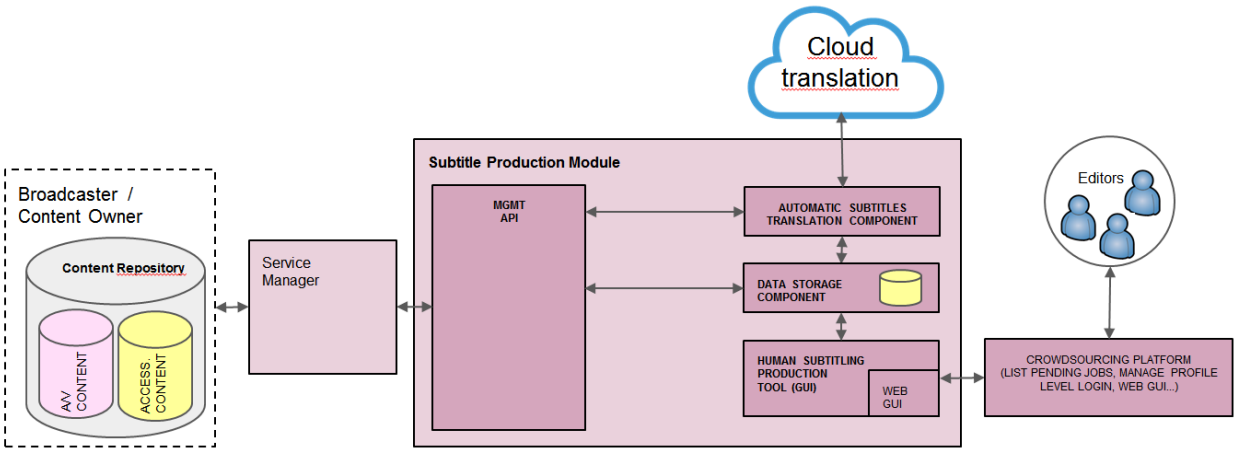

Figure 8: EasyTV Subtitle Production architecture

The workflow followed by the broadcaster/content owner starts with the uploading of contents into the EasyTV platform by connecting to the Service Manager, including the subtitles in its original language together with the required metadata for the translation process management: original language, desired translation languages, minimum trust level required for editor users, identification of the program, etc. A Data Storage Component is needed to temporally keep the uploaded content and its metadata, but also for storing the state of the jobs and the crowdsourcing users' activities performed.

The aim of the Subtitle Production Module is to facilitate the translation jobs in order to achieve a good quality of translations in a reasonable time. With this objective in mind, two components have been included in the workflow of subtitle translation: the **Automatic Subtitles Translation Componen**t and the **Human Subtitling Production Tool**.

The Automatic Subtitles Translation Component provides a first translated version of the subtitles in the languages of interest indicated by the broadcaster/content owner. This is achieved by using free translation services available in the cloud. The translation process takes place as soon as a new job is mandated to the platform, and its results are also stored in the Data Storage Component and used as a starting point by the editorial users of the crowdsourcing platform in the Human Subtitling Production Tool.

The Human Subtitling Production Tool consists of a web-based interface that allows the collaborative users to perform the requested translation tasks. The users will be granted to the tool by means of the crowdsourcing platform. There is expected to have different profiles of users in the platform that can be grouped as follows:

- **Editorial Users:** each broadcaster/content owner shall be able to manage different user's levels of trust in order to control content rights and quality. In this sense, collaborative editorial users can be divided in four profiles: **low, medium, high and professional**. Each user profile has a degree of revision before it's work can be published, so a new user is always profiled as 'low'. The professional user profile is created to manage those contents with strong right limitations, so this kind of users has a professional relation with the broadcaster/content owner.
- **Reviewer Users:** this profile is needed for revision and acceptance of the work in a particular language done by the users before its returning to its original content owner. This kind of users can also promote and reject editorial users basing his/her decision on its activity history.
- **Administrative users:** each broadcaster/content owner shall have at least one administrative user for the assignment of reviewer and professional profiles to the corresponding users.

The tool will assist each user in order to facilitate its tasks of subtitle translation on the platform. In this sense, editorial users will be assisted by showing each subtitle phrase in the original language with the automatic translation as a first proposal. On the other hand, reviewer users will perform a guided revision according to the trust level of the author of the translation ranging from complete revision for 'low' users, partial revision for 'medium' and 'high' users, and direct acceptance for professional users. Reviewer users will also take profit of automated mark-up processes to ease its work such us the highlighting of differences with the automatic translated version of the subtitles, the detection of forbidden words, spotting words that are not in the dictionary and so on.

Once the translated material is validated by a reviewer user and a final version is available in the Data Storage Component, the broadcaster/content owner can download it by connecting to the Service Manager. In the case of contents with rights management, neither the original nor the translated content can be made publicly available directly from the EasyTV platform, the publication of these content will be performed exclusively by the broadcaster/content owner over its own publication platform.

3.3.6. EasyTV Sign Language Production

The EasyTV Sign Language Production is responsible for the creation, validation and exploitation of the sign language content in different languages for easing the access of people with hearing loss to media content and services. The architecture of the Sign Language Production module, along with its interactions with other components, is shown below:

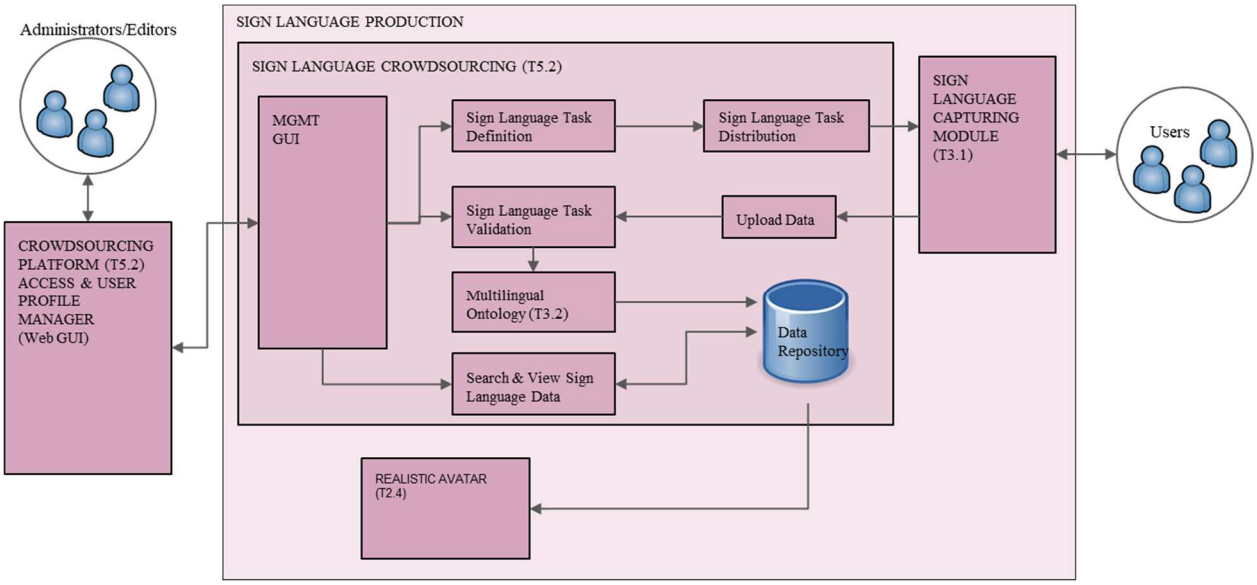

Figure 9: EasyTV Sign Language Production architecture

From Figure 9, it can be observed that there are several components that make up the Sign Language Production Module. More specifically, this procedure consists of the sign language crowdsourcing, the sign language capturing and the realistic avatar. These components are further analysed below:

- **Sign Language Crowdsourcing:** The purpose of this procedure is to allow users to contribute with sign languages and translations of available signs, creating a multilingual sign language repository. The administrators of the EasyTV Crowdsourcing platform will be responsible for defining Sign Language Tasks (SLTs) by proposing words/sentences that need to be translated in sign language and distribute these SLTs to expert sign language users. The users will then use a capturing setup in order to perform the requested signs, and by using a capturing component, record and upload the data to the Crowdsourcing Platform. The administrators will validate the users' contributions by stating an SLT as completed. Afterwards, the multilingual ontology will be responsible for annotating the recorded signs, the text corresponding to the signing with the concepts in the multilingual ontology. This annotations will be later used for identifying 1-to-1 mappings from the recorded videos signs in different languages. The enriched multilingual ontology can also be linked with existing repositories that contain signs in different languages (e.g. Spreadthesing), laying the foundations for the creation of a sub-cloud of resources that contain sign language data. Finally, the motion data along with the corresponding text and concept translation will be stored in a sign language repository for use by the other EasyTV services/modules. The Sign Language Crowdsourcing will also allow a user to search and view already recorded signs.
- **Sign Language Capturing component**: The purpose of this component is to allow users to record signs and upload them to the EasyTV Crowdsourcing Platform. The architecture of the Sign Language Capturing component Architecture is shown in Figure 10.

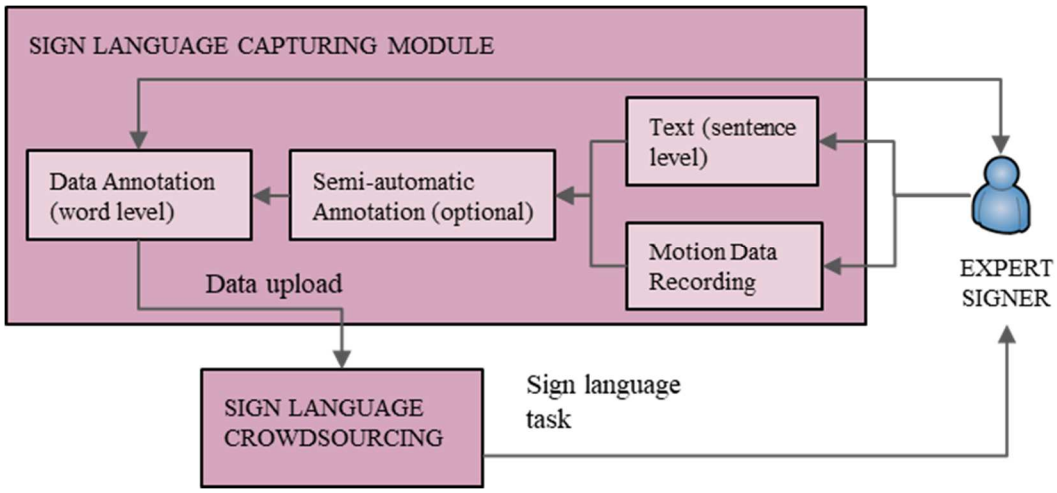

Figure10: Sign Language Capturing Module Architecture

Initially, the user will connect to the crowdsourcing platform and through his/her account, accept the assigned SLTs which recommend the signing of specific words or sentences. However, the user is not limited to completing the SLTs but he/she can perform his/her own signs given that he/she annotates them correctly. The Sign Language Capturing component will be responsible for processing the recorded data in order to extract valuable motion information that will be used for the correct differentiation among different signs. The recorded data will also accompany a text describing the signing sequence (word/sentence). At the next phase, there is an optional possibility of building a semi-automatic annotation algorithm that will accept the recorded data and the text describing them and output a data annotation that breaks down the video in frames and provides starting and ending frame numbers for each recorded sign. The semi-automatic annotation tool will be based on accurate and robust machine learning techniques applied on already learned signs. Finally, the expert user/signer will be able to view data annotation and perform corrections before uploading the data on the Sign Language Crowdsourcing component.

- **Realistic Avatar:** The purpose of this component is to allow the visualization of the recorded signs using an avatar. This component will communicate directly with the data repository of the Sign Language Crowdsourcing component and will acquire and play the sign representations using 3D motion data of face and hands.

3.3.7. EasyTV Crowdsourcing Platform

The Crowdsourcing Platform provides an infrastructure for implementing the sign language production and subtitle production procedures. More specifically, it includes functionality for task definition, distribution and validation, and its purpose is to allow the creation and management of professional user profiles and the access of users in the sign language and subtitle production procedures through web Graphical User Interfaces (GUIs). The architecture of the Crowdsourcing Platform, along with its interactions with other modules is shown in Figure 11.

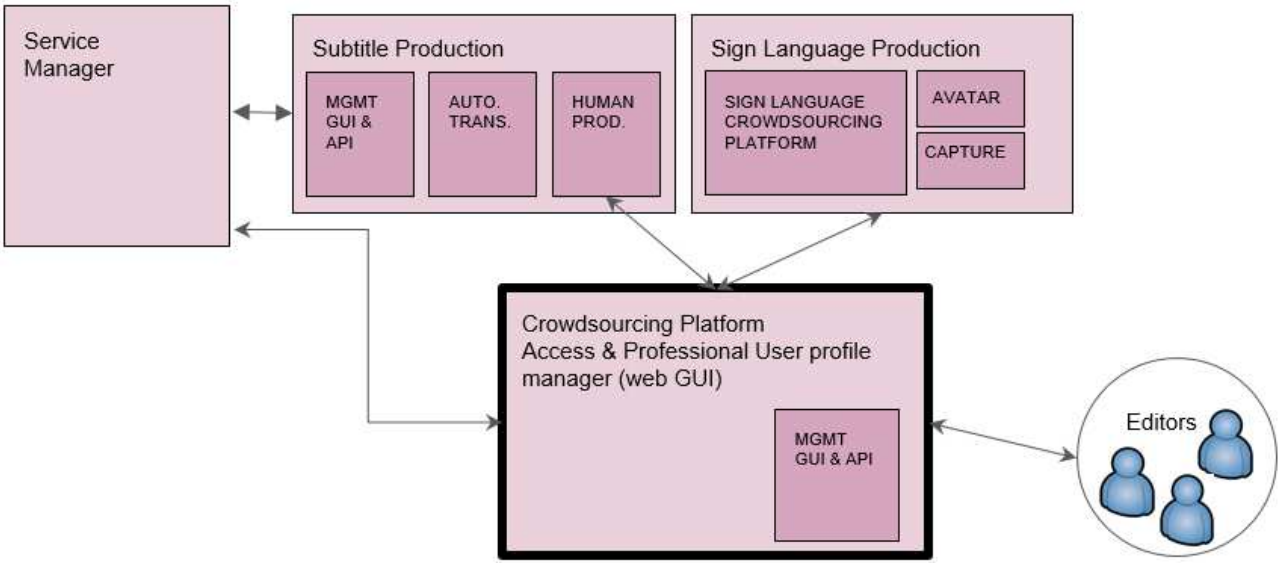

Figure 11: Crowdsourcing Platform Architecture

3.3.8. EasyTV HbbTV Multi Terminal Application

Connected TV allows users to combine the spirit of consuming conventional television in big screens with the advantages of the Internet through different applications and services. Previous HbbTV specifications 1.0 and 1.5 presented a series of limitations, which the new version HbbTV 2.0.1 is trying to overcome. Among the novel specificities, the terminal connection with Companion Screens (CS) (second-screen devices, such as smartphones and tablets) is highlighted, because it allows the synchronization and the exchange of information between both devices. This connection between devices plays a vital role in improving the accessibility, being beneficial for users with specific demands, such as customized subtitles, image enhancement, audio description or clean-audio. The user can access accessibility tools through his/her individual device, but without missing the chance of sharing the experience with the people in the same room. For EasyTV, different interfaces based on multimedia players and tests will be developed taking advantage of this new specification of HbbTV 2.0.1 in order to show the benefits of Connected TV (see Figure 12).

The new standard of HbbTV includes Advanced User Experience functionalities. HbbTV 2.0 uses HTML5 as the basis of the interface structure of applications and a set of related web technologies including modules of CSS3 for defining styles, DOM3, WOFF, Canvas, 2D, Web Messaging, Web Sockets, Web Workers, Server-Sent Events and Web Storage for the exchange of data. In contrast, HbbTV 1.0/1.5 used HTML4, and CSS Level 2. DOM Level 2 and a set of related web technologies have evolved over time.

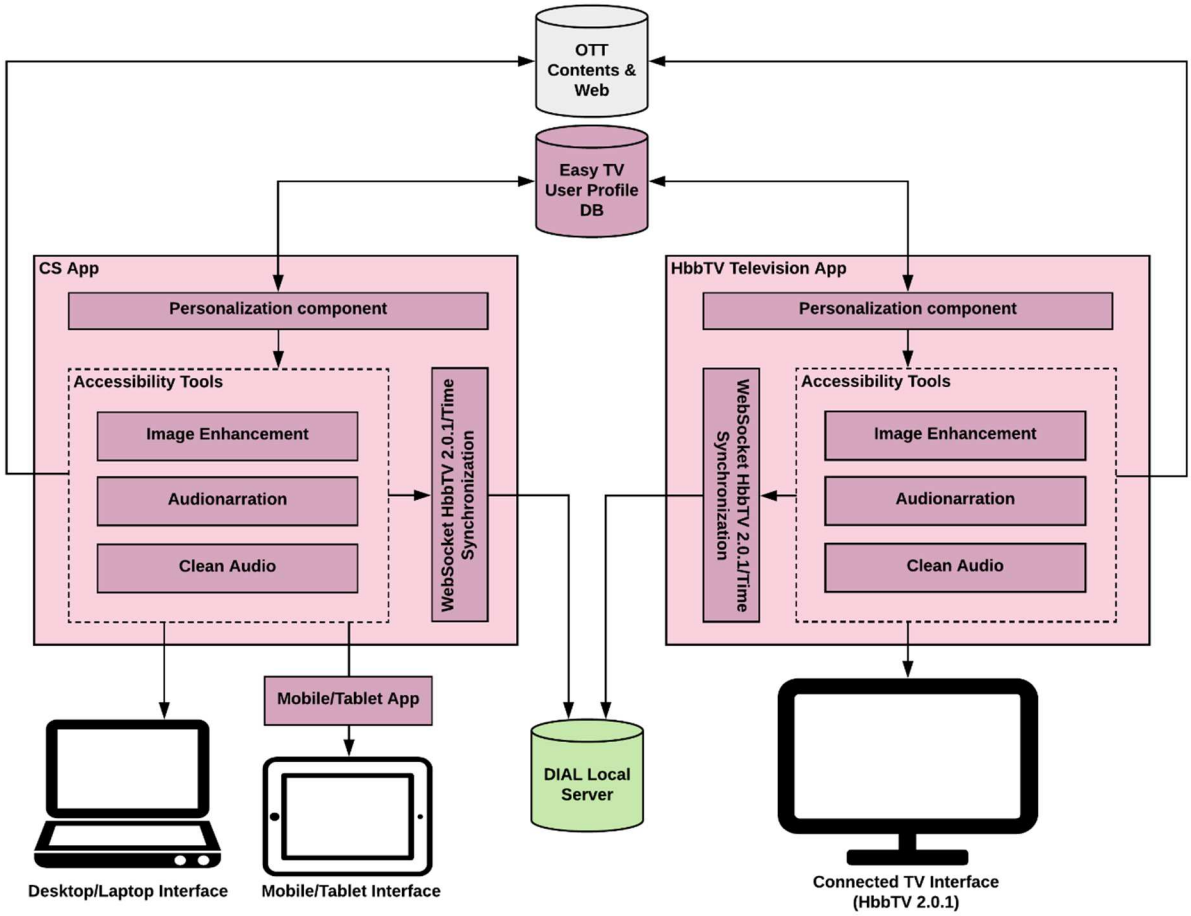

Figure 82: Client interfaces for HbbTV 2.0.1: Terminal and CS Apps

The architecture is based on two interactive applications, one for terminal (television) and one for CS (Companion Screen, such as smartphones or tablets) connected with WebSocket calls through a DIAL server (**Error! No s'ha trobat l'origen de la referència.**). The applications are allowed to discover other devices in their local network and stablishing connection among them. The CS App is able to present accessibility tools, such as audio description and subtitles, synchronized with the contents appearing on the Terminal App, which are contents broadcasted on television. Both applications will be developed in HTML and JavaScript languages, but in the case of the mobile application, an Android application in Java will also be developed in the EasyTV environment to facilitate the interaction of the user. This application includes an embedded WebView to include the facilities of HTML application. Additionally, a web player will be responsible for presenting the contents through different plugins and inner tools for adapting the development to the preferences of the end-users. The web player is based on HTML5 universal player, but the possibility of using the open-source player VideoJS in order to fulfil the requirements of usability is also taken into account, because it presents a collection of libraries including menus of interaction, graphical user interface (GUI) and tools for distributing MPEG-DASH contents.

The application contains a module for personalization management to configure the applications adapted to users' preferences. The connection established between the EasyTV User Profile Database will produce an exchange of information that feeds the system to improve the prediction and recommendations, as a consequence of user's interaction and stored personal profile.

3.3.8.1. Hyper-Personalisation

The EasyTV hyper-personalisation module will be built on top of profiling- user experience mining techniques; extraction, abstraction, homogenisation of dynamic user profiles; and innovative interaction techniques for interface adaptation. This module will be able to handle dynamic and continuous changes in user experience in a user-transparent way in order to recommend new personalised services, dynamically adapted to the current context and device of the user. The Figure 13 presents an overview of the architecture of the EasyTV hyper-personalisation module.

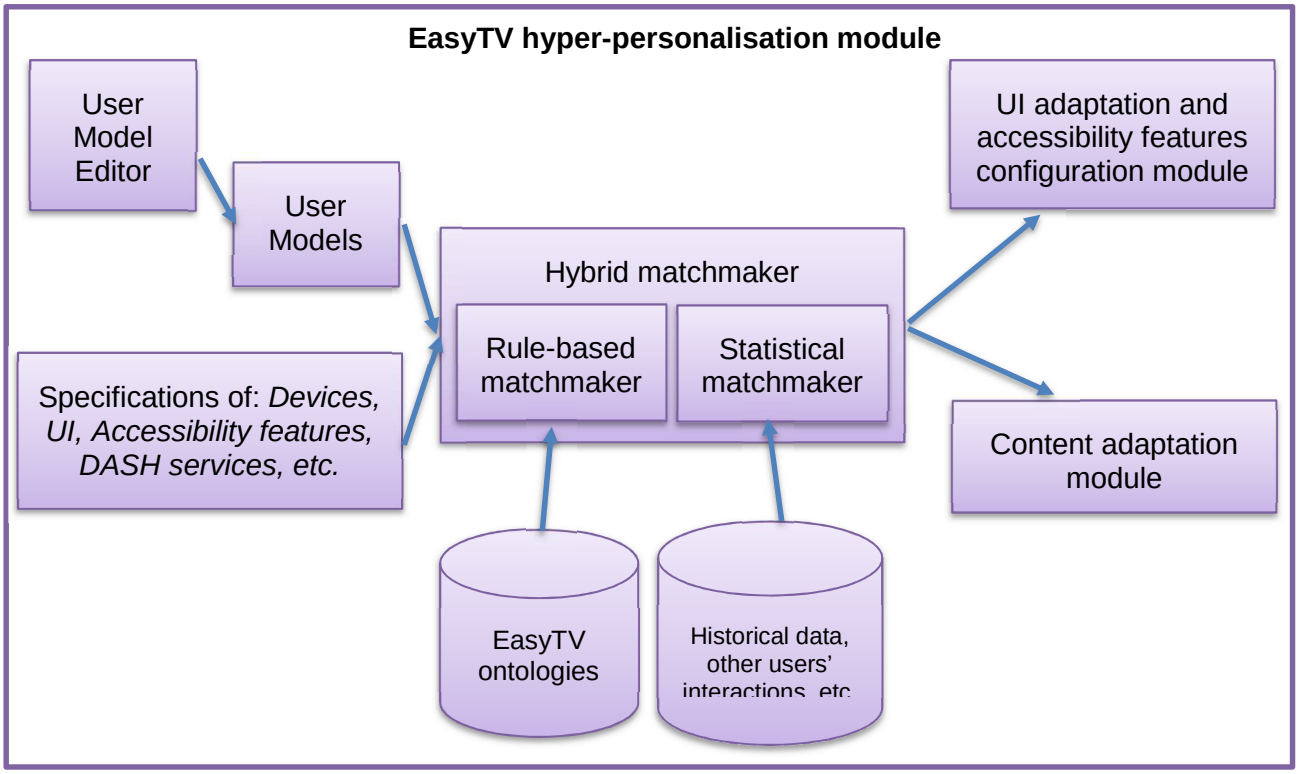

Figure 93:EasyTV hyper-personalisation module

The EasyTV hyper-personalisation module will consist of the following sub-components:

- **EasyTV user models:** The EasyTV user models will define user needs and preferences, including also the disabilities and functional limitations of the user. The structure of the EasyTV user models will be based on previous work conducted in previous projects, such as Cloud4All and VERITAS, as well as on the user models proposed by the VUMS cluster. Further extensions of the aforementioned user models will be developed to cover the EasyTV specific needs.
- **User model editor:** The EasyTV User Model Editor will be a web-based tool that will enable the easy creation and editing of EasyTV end user models through intuitive web forms.
- **Hybrid matchmaker:** The EasyTV hybrid matchmaker will perform matchmaking between user needs and preferences defined in user profiles and UI capabilities, accessibility features specifications and DASH streaming services specifications. The hybrid matchmaker will consist of the following two sub-components:
    - **Rule-based matchmaker:** The rule-based matchmaker will perform matchmaking on the content and metadata stored in the EasyTV ontologies by applying semantic rules. The implementation of the EasyTV rule-based matchmaker will be mainly based on the Cloud4All rule-based matchmaker.

- **Statistical matchmaker:** The statistical matchmaker aims at improving the accuracy of the matchmaking results provided by the rule-based matchmaker by supporting self-adaptive and tailored services, which can learn from users' actions. Statistical methods that will take into account not only the history of actions of the specific user, but also previous corresponding interactions of other users will be applied for this purpose. This will eventually lead to a self-learning system that evolves and fine-tunes its personalisation capabilities over time by taking hundreds of thousands of individual user experiences into account (i.e. tapping into the "wisdom of the crowd"). The crowdsourcing approach of the project will be combined with the EasyTV personalisation framework in order to automatically adapt user interfaces to the language preferences of the user by also providing reliable sign language functionalities. The implementation of the statistical matchmaker will be based on previous results mainly coming from the Cloud4All, PROSPERITY4ALL and Amara project.

- **UI adaptation and accessibility features configuration module:** This module will take as input the results of the hybrid matchmaker and will perform automatic turn on and configuration of accessibility features (e.g. volume, rate, pitch, colour preferences, etc.) that are built into different TV operating systems, applications and embedded devices that will be supported. This mechanism will be mainly based on the Cloud4All auto-personalisation framework, which supports recommendation and automatic launching of assistive technologies (e.g. screen readers, magnifiers, etc.) along with automatic adjustment of corresponding settings (e.g. speech rate, magnification factor, etc.) by applying rule-based and statistical matchmaking on user profile, application capabilities/available settings and current context (e.g. device/OS used, environmental factors, etc.). For the development of the UI adaptation mechanism, the use of well-known User Interface Definition Languages (UIDLs) like UsiXML, which will enable the formal and standardised description of user interfaces, will be also considered.
- **Content adaptation module:** The implementation of the EasyTV content adaptation module will be based on standardised DASH streaming services. In a simple streaming scenario between an HTTP server and a DASH client, the content exists on the server in two parts: a) Media Presentation Description (MPD), which describes a manifest of the available content, its various alternatives, their URL addresses and other characteristics, and b) Segments which contain the actual multimedia bitstreams in the form of chunks, in single or multiple files. In order to play the content, the DASH client first obtains the MPD and by parsing it the client learns about the timing of the program, the availability of media content, the media types, resolutions, minimum and maximum bandwidths and the existence of various encoded alternatives of multimedia components, the accessibility features and the required digital right management (DRM), the location of each media component on the network and other characteristic of the content. The EasyTV content adaptation module will perform content adaptation based on the results of the hybrid matchmaker in order to offer streaming content in the best possible form for a specific user (e.g. by selecting the audio that corresponds to user's language). This matchmaking process will be based on previous experience of CERTH and results from other EU projects, mainly the matchmaking approach followed by the IN LIFE H2020 project, where semantic rules are applied to enable automatic service selection and service composition.

3.3.8.2. Image Enhancement

The main objective of image enhancement is the processing of an image for obtaining a more suitable result that the original image/video, which is adapted to the requirements of the user, especially when this user has impaired vision. Digital image enhancement techniques provide a multitude of choices for improving the visual quality of images. Appropriate choice of such techniques

is greatly influenced by the imaging modality, task at hand and viewing conditions. The variety of algorithms commonly used for image enhancement is large, depending on the available processing resources and the real-time limitation. There are multiple techniques based on spatial domain techniques, with special reference to histogram processing and point processing methods.

In EasyTV platform, a set of different solutions will be involved in the image enhancement service, which will help blind or visually-impaired users to access multimedia content in an easier way, improving their user's experience.

The different solutions presented in the EasyTV cloud platform are:

- A tool for improving the presentation of subtitles and sign language video, allowing two kinds of approaches:
    - **Custom definition:** it is related to the user's control, making possible that users decide how they want to see this content in relation to the position, colour and size.
    - **Automated definition:** related to the auto-personalisation of the settings, based on the information given by the hyper-personalisation module.

  This solution also includes a text to speech tool for reading the subtitles.

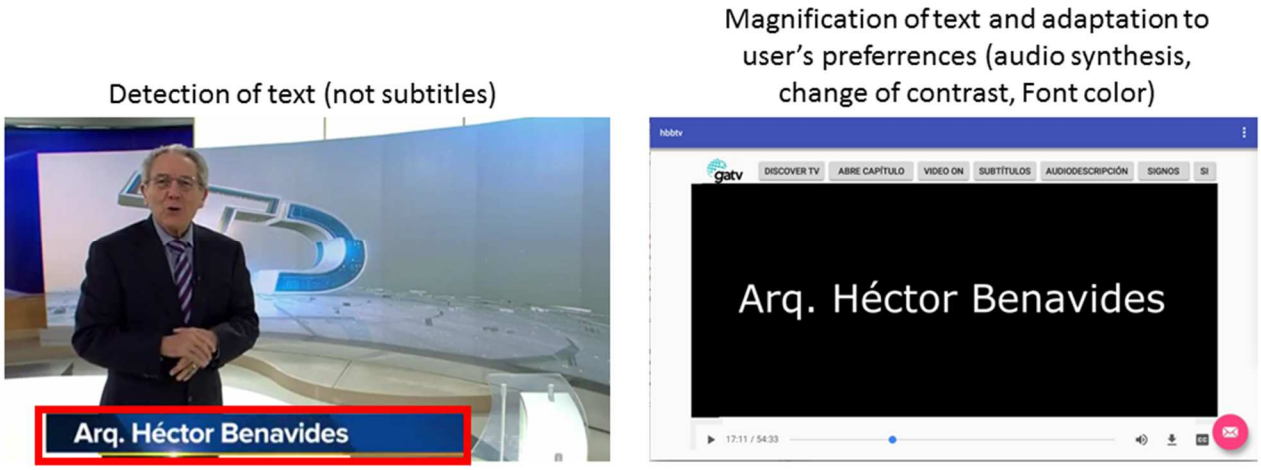

Figure 104: Example of image magnification of text and customizable subtitles

- **Image magnification:** this tool will allow to scale up texts and images in order to facilitate watching them in more detail. As in the previous case, it one also involves two different approaches:
    - **Custom magnification:** like in a magnifying glass, the system will be able to surround and magnify a specific point clicked by the user. This type of functionality is performed on the client side (application).

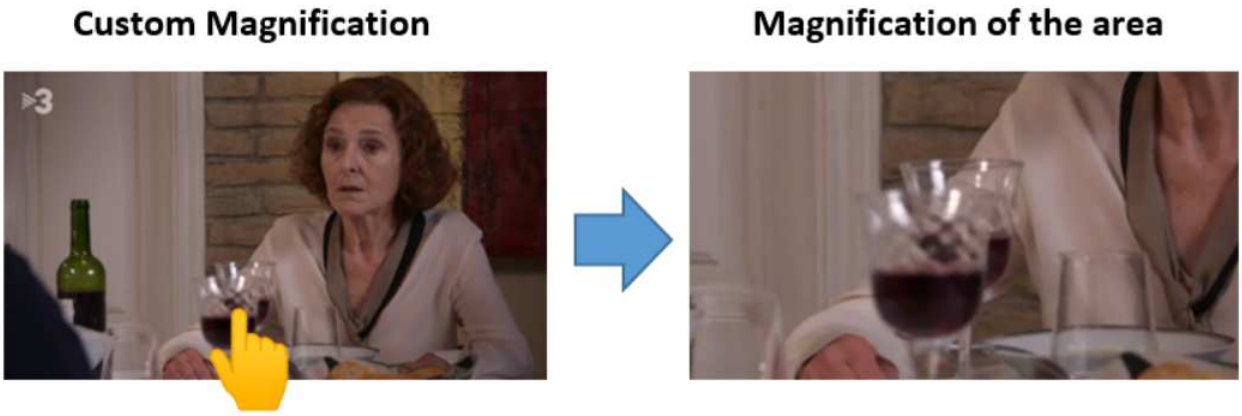

Figure 115: Example of custom magnification

- **Automated magnification:** in order to facilitate access to the content, the image magnification tool will be able to detect and magnify faces and text in specific contents and contexts. On the EasyTV Cloud Platform there is a module that automatically detects texts and faces of an audiovisual content. This module generates a text file with the position where the content to be magnified is located. Once the text file is generated, the file is stored in the accessibility content database through the service manager. Then, the client application can obtain the text file and magnifies the content with the information obtained.

Figure 126: Example of faces magnification

In the case of face detection, the module is in charge of selecting all the frames and analyse it to find faces. In case one is detected, it returns in the text file the centre of the face and the size of the bounding box in order to make the magnification.

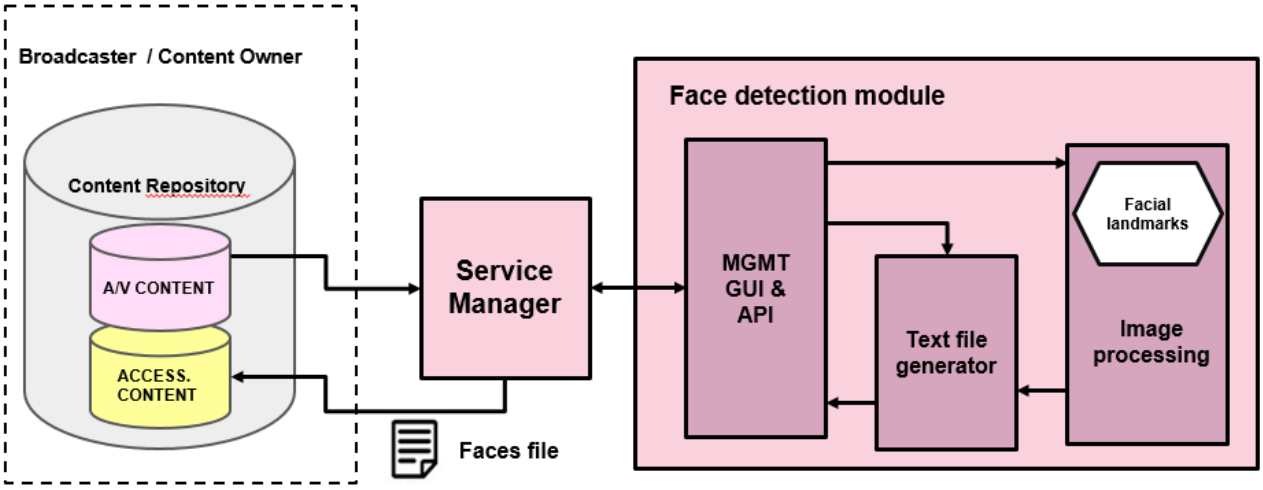

Figure 137: Face detection module in EasyTV Cloud Platform

The magnification of the content can be done in the main screen or in the second screen of an external device. For that, the applications exchange information using Websocket technology.

3.3.8.3. Speech Accessible Interface

The Speech Accessible Interface is part of the Universal Remote component of EasyTV project. This component will be available using a tablet or a smartphone device when users consume content on a Second Screen application. Blind and visually-impaired users will interact with the platform and services using a special voice enabled remote control (equipped with a microphone and a Push To Talk button) or directly using audio microphone and speakers available on the client device. The remote control will communicate with the client device through a Radio Frequency connection that allows the user to control the EasyTV applications remotely, even from a different room.

The Speech Interface Architecture will be based on existing technology and enhanced by exploiting the latest advances in Natural Language Processing as described later in this document. Figure 18 shows the overall architecture of the speech interface.

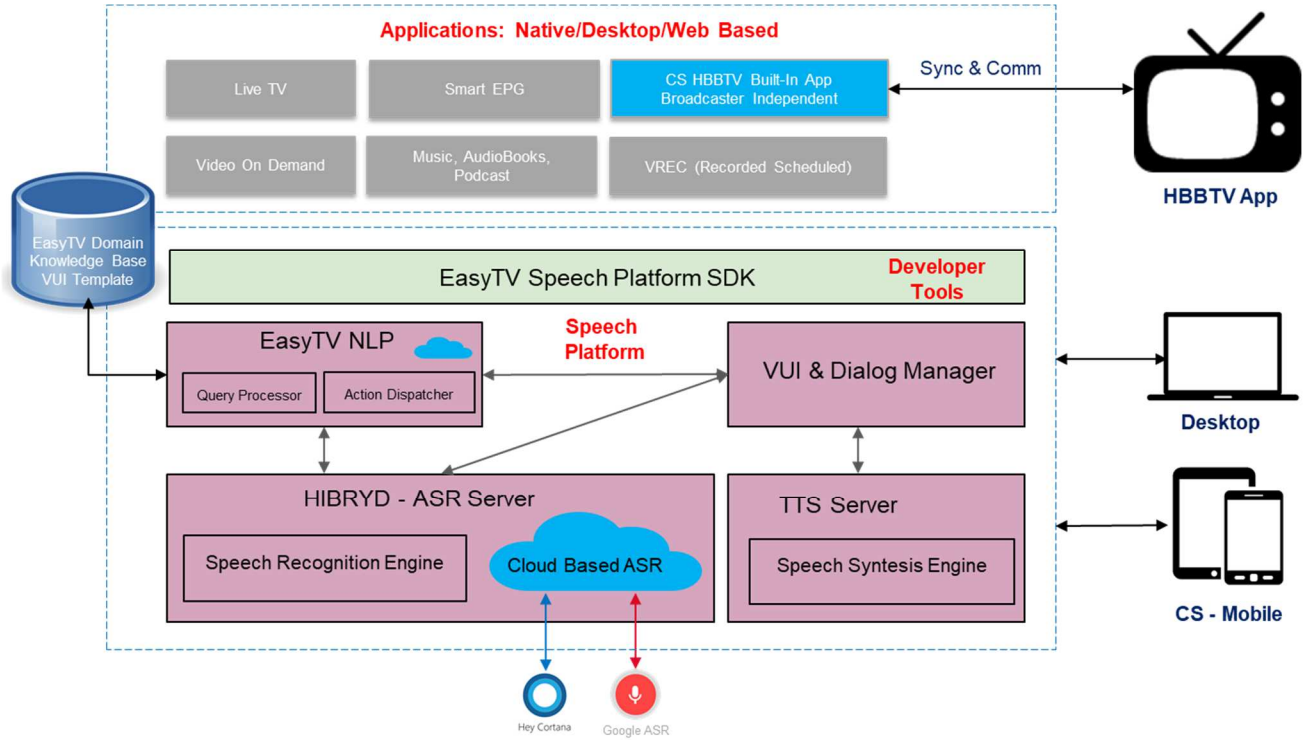

Figure 14: Speech Platform architecture for EasyTV voice control

In this architecture we have the Speech Platform components which include the core components that manage the voice user interactions:

- Hybrid ASR Server: this component manages the Local Speech Recognition Engine on the client device and the remote Speech Recognition Engine on the cloud. The speech platform will be able to manage both local and remote ASR depending on the functionality that is going to be used by the application and the network availability. Regarding the remote speech recognition engine both Cortana and Google ASR can be used based on the client device operating system.
- TTS (Text to Speech) Server: this component manages the speech synthesis engine which will be included in the client device. It will be able to manage the voice to use, the language and the volume and speed.
- EasyTV NLP: This will be the Natural Language Processing Component of the speech interface. It will process the voice query of the user and understand the user intent. It will also manage the subsequent communication with the Dialog Manager based on a reasoning and planning process. It will also dispatch the actions and receive events from the EasyTV SDK components to interact with the TV based application.
- VUI and Dialog Manager: this component is responsible to process the dialog flow between the user and the application through pre-defined voice dialog templates, which include voice prompt templates, semantic annotations for user actions and dialog flow templates.

The EasyTV Domain Knowledge Base and VUI Template is the repository of the EasyTV ontology and data available for the NLP component as described above.

On top of the Speech platform there are the Developer tools that will be based on a specific SDK (Voice Software Developer Kit) which will enable developers to add voice interface to HBBTV/HTML5 based applications other than native applications developed for EasyTV users. Both Speech Platform and SDKs will be available for developing any voice enabled application and will run on the

client device.

The application level module is the high level part of the architecture and includes all the EasyTV application set. EasyTV applications can be implemented in any native device language and of course in HBBTV/HTML5 technology. Applications will interact with the speech platform using a common communication protocol and technology that is native code for native applications, the WebView API interface for embedded HTML5 applications or WebSocket Technology for any other HTML5 application running on a Web Browser (Chrome, Edge, Mozilla, etc.) on any terminal, including HBBTV.

The Speech Interface component will be able to control and access all the main functionalities of the Universal HTML5 Player (as described above), for both audiovisual and accessible content.

3.3.8.4. Gesture Accessible Interface

The purpose of the Gesture Accessible Interface is to enable the communication of users operating on a second screen (i.e. desktop PC/laptop) with the smart TV by means of gesture or gaze information. The Gesture Accessible Interface can be employed in both use cases, meaning that a user can either remotely control the TV using gestures/gaze information from his/her second screen or watch TV on his/her second screen by directly controlling the TV program and user interface with gestures/gaze information.
The Gesture Accessible Interface pipeline is described in detail in Figure 19. It will employ a depth sensor to capture a RGB-D video of a user performing gestures or eye movements. Afterwards, the video will be processed, and motion data will be extracted allowing the differentiation among gestures and the identification of eye locations/movements. The classification of motion data will be performed by employing accurate and robust action recognition classification techniques. The detected gesture or eye movement will be detected and translated to a corresponding TV command for the remote TV control operation. The command will then be transmitted to the EasyTV platform using HTML5/HbbTV technology through web sockets.

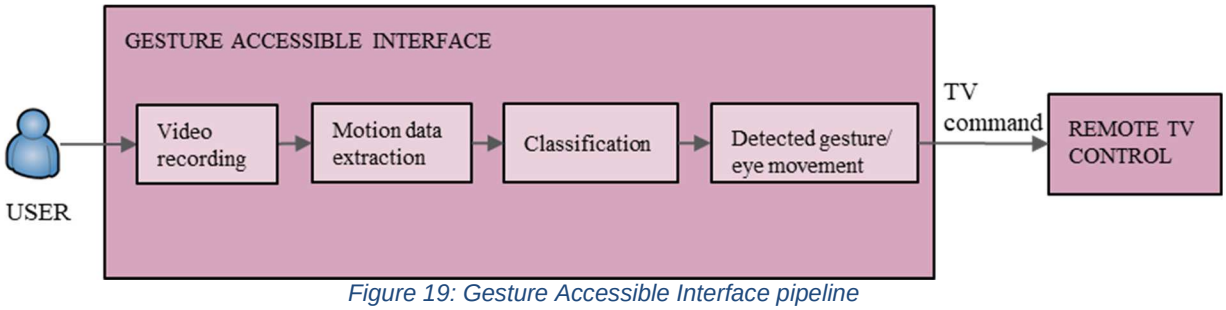

Figure 19: Gesture Accessible Interface pipeline

3.3.8.5. HbbTV SmartTV remote control through CS

The purpose of remote control through Companion Screen (mobile device, pc, tablets, etc.) is to give the user the ability to attain complete access to the HbbTV application content and to provide him with the full experience of interacting with it. The communication between the CS and the HbbTV application will be bidirectional and established in such a manner that the user will be able to remotely control his/her TV screen through an application installed in his CS.

The CS will provide the available actions to the user through a simple UI and will use a web socket to communicate with the HbbTV application. After the HbbTV application receives the user's action from the CS, it will use the same web socket connection to respond to the user's choice. When the CS receives the response from the HbbTV application, it will inform the user in various ways,

depending on his/her disability, such as vibrations, sounds, etc.

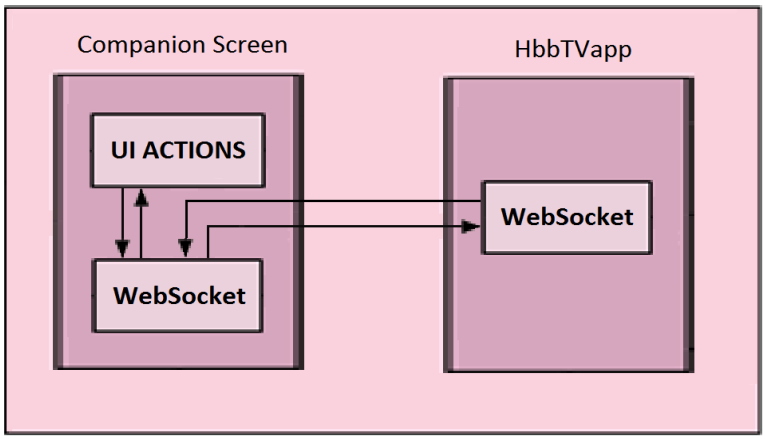

Figure 20: Remote Control through CS Interface

3.3.9. EasyTV Technical Integration Guidelines

The EasyTV architecture has a heterogeneous set of components and applications that must be placed in the cloud and must be able to communicate with each other, ensuring service continuity, ease of updating, expandability, portability, interoperability and the possibility to perform changes without necessarily interrupting the entire service. For this reason, two solutions have been identified to allow communication between the components:

- REST Api
- WebSocket

3.3.9.1. Rest API

Representational State Transfer (REST) is a type of "architectural style for distributed hypermedia systems" [8]. The REST architecture is based on HTTP; the operation provides a well-defined URL structure (designed to uniquely identify a resource or a set of resources) and the use of specific HTTP verbs for retrieving information (GET), for editing (POST, PUT, PATCH, DELETE) and for other purposes (OPTIONS, etc.). By using a stateless protocol and standard operations, REST systems aim for fast performance, reliability, and the ability to grow, by re-using components that can be managed and updated without affecting the whole system, even while it is running. REST provides that the web scalability and growth are direct results of a few key design principles, such as:

- application status and features are divided into web resources
- each resource is unique and addressable using universal syntax for use in hypertext links
- all resources are shared as a uniform interface for the transfer of status between clients and resources

The REST client-server separation of services simplifies component implementation, reduces the complexity of connector, improves the efficiency of performance tuning, and increases the scalability of server components. Layered system constraints allow intermediaries – proxies, gateways, and firewalls – to be introduced at various points in the communication without changing the interfaces between the components, allowing them to assist in translating communication or improving performance through large-scale shared caches.

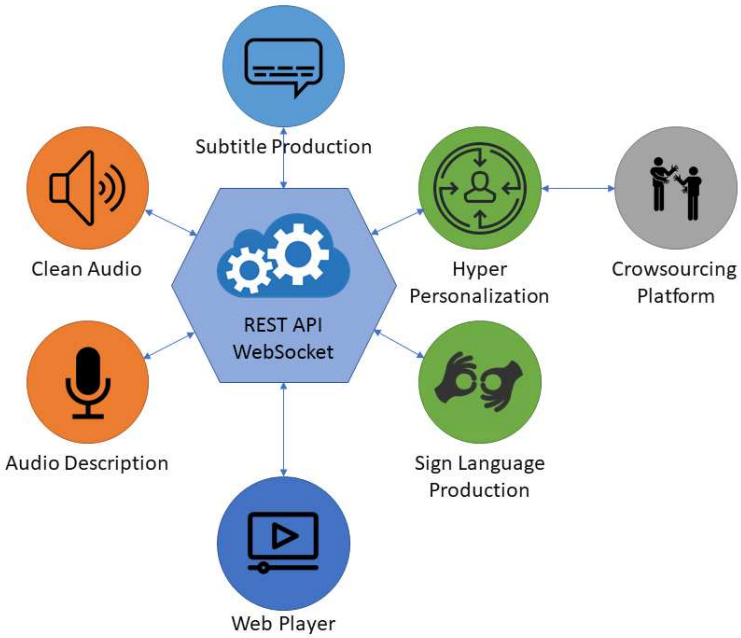

Figure 21: REST API and services

A syntax for exchange data format must be defined, preferably in JSON format, a very common and universally supported format by many programming languages.

The APIs of each component can communicate with other systems through a token based authentication mechanism (or, alternatively, a form of authentication such as Basic Authentication or authorization protocol like OAuth), in order to guarantee the necessary security and the recognition of the authorized user. However, for greater security, it is recommended that all systems are protected by TLS protocol, typical of HTTPS connections.

3.3.9.2. Web Socket

Another method with which some of the components can communicate is using WebSockets [9]. This is a computer communications protocol, providing full-duplex communication channels over a single TCP connection. This system has the advantage of being able to keep open a constant and continuous connection between the clients and server's components, allowing a stream of data in real time. Moreover, it is supported by all major browsers, and it also has the advantage of using port 80 (HTTP standard port) or 443 (HTTPS standard port) that are usually open and available even in environments with particularly restrictive proxies or firewalls.

As for REST API, Web Socket can also be used in many programming languages, through different development environments and on different devices (Java, Android, iOS, and so forth).

WebView or browser-based applications in EasyTV can take advantage of low-latency and always-on connections for fast transmission of user's status or data and full applications information in a way previously only emulated by methods such as Ajax and Comet polling.

Security is guaranteed by the native protocol system through a first phase of client-side handshaking and an original-based security model. The system also masks data transmissions to avoid sniffing of plain text packets.

3.3.9.3. HTML5 web player

The HTML5 web player is a video element for playing videos in web based applications or in mobile applications utilizing webview technology. It can support multiple audios and subtitles and give the user the ability to personalize his/her visual experience, in full accordance with the EasyTV project specifications and functionalities. Furthermore, it can support multiple video formats (such as mp4, MPEG-DASH, webm, and ogv). There is also a possibility to use other libraries in order to improve

the functionalities of the player. The most common library, which is also open sourced, is Video.js, which provides many different menus of interaction, GUI and tools for distributing MPEG-DASH.

4. DEPLOYMENT AND TEST

Deployment and test phases are vital in the project life cycle, since they represent the main tools for detecting issues and planning improvements that have to be applied to the system in order to achieve an optimal performance. This will assure the achievement of the results that were defined at the very beginning of the project, providing complete and usable results, which will be presented in the following sections.

4.1. Platform deployment

This phase comprises all the activities that allow a solution to move from the development/test environment to the production environment in order to make it operational and available to stakeholders. It depends on the characteristics of the project, so it should be performed after a careful analysis of the involved aspects, such as the number of users of the system, the scalability, the storage size for data system if needed, etc. Furthermore, in real cases, when talking about deployment, we usually distinguish between two kinds of environments:

- Pre-production environment, focused on experimentation and testing.
- Production environment, after obtaining a stable and complete version of the system in the previous phase.

In this case, along this section we are going to present a detailed plan for the platform implementation that may guide the process on the pre-production phase based on a state of work analysis. Following, we are going to propose a thoughtful testing plan for assuring the correct performance of the platform's components that may lead to the final result.

The EasyTV deployment plan is defined as user case-centred, since the solution is going to be deployed to different user case scenarios to, not only ensure that the solution will perform all of its function to meet full scale operational requirements, but also to verify that it is acceptable for the users. In this regard, the EasyTV deployment scheduling pattern will follow a functional approach, that is, it will be based on the deployment of the different applications or services. Nevertheless, before this stage can begin there are some previous steps that need to be done. The first one is related to the testing phase, since a solution that has not been tested cannot be deployed. The second phase is related to the correct documentation of the system, given the availability of user manuals that help the stakeholders know what they are using is essential. Once this is accomplished, the deployment phase is ready to start.

The plan comprises the following steps:

1. Defining the characteristics of the user case environment.
2. Checking the user case environment is correctly established, which includes checking the correct status of the involved services.
3. Checking the correct integration of the involved components to verify its performance.
4. Providing access to the test environment, that is, to let testers enter the system in order to check its availability and correct performance.
5. Conducting a training session, if necessary.
6. Operating the system in the user case environment.
7. Documenting the results and making recommendations, since the final report should include the following information concerning the errors detected:
    a. Function being performed and segment of the system in use at the time of the error.
    b. Detailed description of the problem.
    c. List of the messages that have been received.
    d. Time and date of the occurrence.

4.2. Platform testing

A complete test plan is crucial in the EasyTV environment, due to the complexity of the project. In this regard, the test plan will be made up of a regular flow of sub tests that may help to limit the maintenance process by resolving bugs since the very beginning of the project, when component dimensions are still not relevant for the system. The execution of successive steps will be carried out only when the results of the previous one are positive. According to this idea, the main steps of the test plan flow are presented in the Figure 22.

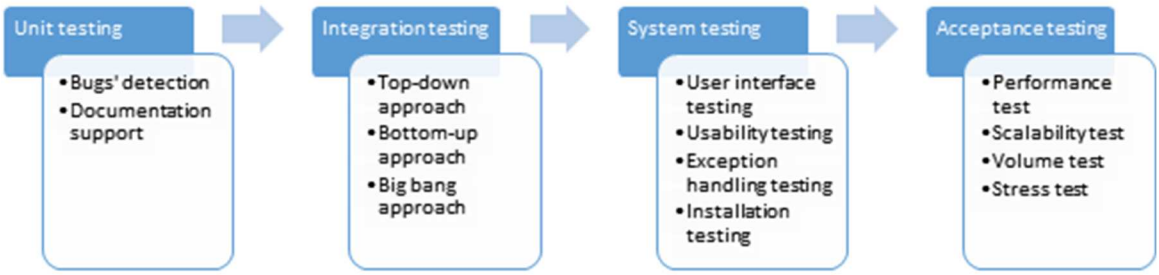

Figure 22:Test plan

4.2.1. Unit Testing

This first step of the plan consists in checking the activity of every single module that is part of the EasyTV platform as an autonomous component, with the aim of detecting possible bugs and prevent the system from accumulating too many errors that could not be managed in successive steps. Therefore, the main objectives of this phase are:

- To simplify the management of bugs.
- To simplify the integration testing step.
- To support documentation.

Unit testing can be performed in manual or automatic manner (please, review the functional validation tools list with some examples of automated tools for this purpose in D1.2 [3]). Nevertheless, we aim at adopting a waterfall approach, where testing for each component will be done in a sequential way, and the execution time and resource will be fixed. Besides, all components will be tested in parallel to reduce time consumption and to start the following stage as earlier as possible. Finally, it is important to point out that this first stage will help to reveal which further functionalities can be implemented in a component to achieve prefixed requirements.

4.2.2. Integration testing

This stage will provide information about the integration and the cooperation of all the modules in the EasyTV Platform. It will be performed with the use of appropriate test cases, which are finalized to the validation of the interconnection, and can be performed in three different approaches:

- Top-down approach that consists of starting from an empty system and aggregating every module with continuous tests step by step.
- Bottom-up approach that involves the test of lower components and their aggregation until complete integrated system is reached.
- Big bang approach, when all modules in the platform are first integrated and then tested.

Regardless the adopted option, this type of test will be performed iteratively, with the aim of helping the redefinition and re-evaluation if there are some critical issues.

4.2.3. System testing

The next stage in the testing phase consists of the complete evaluation of the system according to its functional requirements, and it can only be performed if the previous phase has returned successful results. This evaluation is usually made of a set of different tests. For EasyTV, the tests to be considered are:

- User interface testing, which consists of testing graphical interfaces to evaluate their approach to the desired goals.
- Usability testing that must be performed to detect if the system guarantees the user-friendly system usability, that is, if the objectives of the system can be achieved in an efficient and effectively way.
- Functional testing that is in charge of evaluating whether the system allows to achieve the desired results given in the form of functional requirements.

The tests to be performed in this phase will be planned along WP6, which focuses on defining a detailed evaluation methodology for testing and results gathering. Moreover, these results will be provided in the form of feedback reports in order to give some guidance for the system enhancement.

4.2.4. Acceptance testing

Acceptance testing involves some tests aimed to performance measurement and tuning, and it can be applied only after system evaluation returns positive results. Besides, it is important to previously define which kind of results are expected.

There are many types of acceptance testing, such as:

- Load test: to evaluate behaviour of the system under normal circumstances, that is, with an expected number of users.
- Stress test: to evaluate the critical usage of the system, typically with a high number of users executing many different operations.
- Volume test: performed with a certain amount of data.
- Scalability: to detect if the system is capable of scaling up or out in a desired manner.

5. REFERENCES

[1]. EasyTV – Annex I “Description of Work”

[2]. D1.1 User scenario and requirements definition. (http://easytvproject.eu)

[3]. D1.2 EasyTV system requirements specification (http://easytvproject.eu)

[4]. D1.3 First release of the EasyTV system architecture (http://easytvproject.eu)

[5]. Newell, A. F., Gregor, P., Morgan, M., Pullin, G., & Macaulay, C. (2011). User-sensitive inclusive design. *Universal Access in the Information Society*, *10*(3), 235-243.

[6]. Clarkson, J., Coleman R., Hosking I., Waller S. (2007). Inclusive design toolkit. Engineering Design Centre. Cambridge

[7]. Rosson, M. B., Carrol J. M. (2002). Scenario-Based Design. *The Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies and Emerging Applications.* Lawrence Erlbaum Associates, pp. 1032-1050

[8]. REST: https://www.ics.uci.edu/~fielding/pubs/dissertation/rest_arch_style.htm

[9]. WebSocket protocol RFC: https://tools.ietf.org/html/rfc6455