Developing Accessibility Multimedia Services: The Case of EasyTV

Dimitrios Konstantinidis†
Centre for Research and Technology Hellas
Thessaloniki, Greece
dikonsta@iti.gr

Kosmas Dimitropoulos
Centre for Research and Technology Hellas
Thessaloniki, Greece
dimitrop@iti.gr

Kiriakos Stefanidis
Centre for Research and Technology Hellas
Thessaloniki, Greece
kystefan@iti.gr

Thanassis Kalvourtzis
Centre for Research and Technology Hellas
Thessaloniki, Greece
tkalv@iti.gr

Salim Gannoum
Centre for Research and Technology Hellas
Thessaloniki, Greece
salgan@iti.gr

Nikolaos Kaklanis
Centre for Research and Technology Hellas
Thessaloniki, Greece
nkak@iti.gr

Konstantinos Votis
Centre for Research and Technology Hellas
Thessaloniki, Greece
kvotis@iti.gr

Petros Daras
Centre for Research and Technology Hellas
Thessaloniki, Greece
daras@iti.gr

Sara Rovira-Esteva
Universitat Autònoma Barcelona
Barcelona, Spain
Sara.Rovira@uab.cat

Pilar Orero
Universitat Autònoma Barcelona
Barcelona, Spain
pilar.orero@uab.cat

Silvia Uribe
Visual Telecommunications Application Group
Universidad Politécnica de Madrid
Madrid, Spain
sum@gatv.ssr.upm.es

Francisco Moreno
Visual Telecommunications Application Group
Universidad Politécnica de Madrid
Madrid, Spain
fmg@gatv.ssr.upm.es

Alvaro Llorente
Visual Telecommunications Application Group
Universidad Politécnica de Madrid
Madrid, Spain
alg@gatv.ssr.upm.es

Pablo Calleja
Ontology Engineering Group
Universidad Politécnica de Madrid
Madrid, Spain
mpoveda@fi.upm.es

Maria Poveda-Villalón
Ontology Engineering Group
Universidad Politécnica de Madrid
Madrid, Spain
pcalleja@fi.upm.es

Pasquale Andriani
Engineering Ingegneria Informatica
Rome, Italy
pasquale.andriani@eng.it

Giuseppe Vitolo
Engineering Ingegneria Informatica
Rome, Italy
giuseppe.vitolo@eng.it

Giuseppa Caruso
Engineering Ingegneria Informatica
Rome, Italy
giuseppa.caruso@eng.it

Nicolamaria Manes
Mediavoice SRL
Rome, Italy
manes@mediavoice.it

Fabrizio Giacomelli
Mediavoice SRL
Rome, Italy
giacomelli@mediavoice.it

Jordi Fabregat
Corporació Catalana de Mitjans Audiovisuals SA
Barcelona, Spain
jfabregat.f@ccma.cat

Francesc Mas
Corporació Catalana de Mitjans Audiovisuals SA
Barcelona, Spain
fmas.z@ccma.cat

Jordi Mata
Corporació Catalana de Mitjans Audiovisuals SA
Barcelona, Spain
jmata.n@ccma.cat

Stavros Skourtis
ARX.NET S.A.
Thessaloniki, Greece
skourtis@arx.net

Chrysostomos Bourlis
ARX.NET S.A.
Thessaloniki, Greece
bourlis@arx.net

Giuliano Frittelli
Unione Italiana dei Ciechi e degli Ipovedenti
Rome, Italy
uicroma@uiciechi.it

Emilio Ferreiro Lago
Fundación CNSE para la Supresión de las Barreras de Comunicación
Madrid, Spain
formacion@fundacioncnse.org

Federico Alvarez
Visual Telecommunications Application Group
Universidad Politécnica de Madrid
Madrid, Spain
fag@gatv.ssr.upm.es

ABSTRACT

People with various degrees of disabilities (e.g., visually or hearing impaired) often find it difficult to access mainstream products and services and thus they are excluded from enjoying audio-visual services on an equal basis as people without disabilities. These people feel marginalized in today's digital society and they are unable to reach their maximum potential personally, professionally and socially. EasyTV aims to address these issues by providing numerous services that not only ease the access of people with disabilities to multimedia services, but also enhance interaction with the media through a multi-language approach that adapts content based on user's preferences. In this work, the main EasyTV services are described in detail. Furthermore, an evaluation of the main services by end users is presented and discussed in order to highlight the importance of offering accessibility services to people with disabilities.

CCS CONCEPTS

• Accessibility systems and tools • Accessibility technologies • Interactive systems and tools

KEYWORDS

Accessibility, Interaction, Sign language, Personalization, Artificial Intelligence, Semantics

1 INTRODUCTION

Even in the poorest countries and societies, media has great importance because it provides the channels, networks, formats, and languages through which much of life takes place and finds meaning. As a result, it is of real concern that people with disabilities may not have access to either the media or the information provided through specific channels (TV, movies, newspapers, magazines, the Internet, etc.). Ensuring a barrier-free life for people with disabilities and enabling access to the media and content for all is crucial for a full democratic participation. Accessibility prevents or removes barriers to the use of mainstream products and services. It allows the perception, operation and understanding of these products and services by people with disabilities, on an equal basis with others. The inclusion of people with disabilities is a key objective in any program of social and economic measures to create a real Information Society for all, where each citizen has the same opportunities to develop to its maximum potential his/her personal, professional and social aptitudes.

Achieving an inclusive digital society, in which people with disabilities have access to online services and information on an equal footing to people without disabilities is of paramount importance. Studies have shown that blind and visually impaired persons watch TV as much as normally-sighted persons as they consider TV watching to be an important family activity, while their desire is to improve the television experience [1]. On the other hand, the aged people link TV watching with their occupational well-being [2], highlighting the significance of having access to the media. Other studies have shown that being absent from social media can lead to social isolation and feelings of frustration and disappointment [3][4].

The need for social inclusion of people with disabilities has led to the funding of several EU projects that aimed to design and develop accessibility services targeting digital TV [5]. DTV4All[1] was responsible for providing accessible digital television programs for the analog to digital switch across the European Union, encompassing several mature access services, such as subtitling, signing and audio description [6]. Similarly, Cloud4All[2] was concerned with advancing the Global Public Inclusive Infrastructure (GPII) concept by researching, designing,

PETRA '20, June 30–July 3, 2020, Corfu, Greece
© 2020 Association for Computing Machinery.
ACM ISBN 978-1-4503-7773-7/20/06...$15.00
https://doi.org/10.1145/3389189.3397973

[1] https://cordis.europa.eu/project/id/224994
[2] https://cordis.europa.eu/project/id/289016

developing and testing key software infrastructure and piloting implementations to deliver accessibility to every individual where they need it, when they need it and in a way that matches their unique requirements [7]. To this end, Cloud4All created and/or refined, among several other things, user profiling standards and tools to capture the individual needs and preferences of users with disabilities, interface and material profiling standards and tools to characterize ICT material and the different techniques to access them, as well as mechanisms to allow users to locate solutions from many different sources in a single Internet search.

Prosperity4All[3] was a continuation and extension of the work performed in Cloud4All, focusing on developing the necessary infrastructure for a robust cross-platform spectrum of mainstream and assistive technology-based access solutions based on self-rewarding collaboration [8]. This infrastructure utilized cloud, crowd, game and smart technologies to bring new players with both low and high technical skills into the service development and delivery, thus managing to reduce redundant development, lower costs and increase market reach and penetration internationally. Finally, Hbb4All [4] focused on the hybrid broadcast-broadband TV (HbbTV) [9] concept to propose more and better TV access services tailored to the users' needs at lower cost [10]. To achieve its goals, Hbb4All designed and developed a multi-platform media environment for the delivery of audio-visual content on PC, tablets, smartphones and TVs using an array of communication solutions, such as subtitling, audio description, clean audio, and many customizable features.

EasyTV took into consideration the advantages and limitations of previous EU projects and designed and developed improved accessibility services for people with various degrees of disabilities, focusing on visual, hearing and mobility impairments, with the goal to provide equal access to audio-visual services for all users. More specifically, EasyTV designed and developed media improved access services and novel accessibility features with enhanced multimedia visual and sound experience in a way that they are more cost-efficient and yet more flexible and easier to use. EasyTV breaks the language barrier for all by developing technologies which can enhance the interaction and perform sign language translations towards an inclusive media interaction. The heart of EasyTV is an improved personalization of the content experiencing and interaction, towards a hyper-personalized experience to all. The main scientific and technological contributions of the EasyTV project are:

- **Improved access services** for enhanced multimedia visual and sound experience for people with disabilities, by adapting the content through image adaptation (e.g., face and text detection and magnification, color-blind subtitles, etc.) and content description improvement (e.g., adapted audio narratives, clean audio, etc.).
- **Improved personalization of the content and interaction**, towards a hyper-personalized experience to all, based on an auto-personalization profile tool and focusing on a) adaptation of the content provision using Dynamic Adaptive Streaming over HTTP (DASH) streaming services, b) personalized services that can learn from users' actions and c) 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 that can perform automatic translation in different languages with the aim to allow an inclusive media consumption. Such technologies include a realistic avatar available in different languages, a crowdsourcing platform for the collection of sign language data, as well as a definition of a multilingual ontology that map signs 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 a universal remote control.

In Section 2, the architecture of the platform proposed by the EasyTV project is described, as well as the main components that were developed in the framework of the project. Section 3 presents and analyzes the user feedback during the evaluation of the EasyTV services, while Section 4 draws some conclusions.

2 THE EASYTV ECOSYSTEM

2.1 The EasyTV architecture

The EasyTV architecture describes the various components and services developed in the framework of the project, as well as the interconnections between them. The basic architecture is presented in Figure 1 and it consists of the following three blocks with distinctive functionalities:

- **Broadcaster premises / Content Owner** that englobes the main workflows of the broadcaster or a content owner related to the management, storage, broadcast and publication of audio-visual contents.
- **EasyTV platform** that contains several modules grouped in service components that aim to provide equal access of people with disabilities to media.
- **Consumer platform**, in which end-users will consume the curated audio-visual content through multi-platform devices, such as smartphones, desktops or SmartTVs, interacting with their devices through improved accessible interfaces that facilitate content access and consumption.

2.2 The EasyTV main components

In this section, the main components of the EasyTV ecosystem that allow users with different disabilities to access broadcaster content in an easy and flexible way are described. In this regard, each component represents an innovative service, which is easy to use, low cost and useful for improving the interaction with terminals and to access multimedia content. For a more detailed description of the EasyTV project and all its components, the interested reader is urged to visit the EasyTV website[5].

[3] https://cordis.europa.eu/project/id/610510
[4] https://cordis.europa.eu/project/id/621014
[5] https://easytvproject.eu/

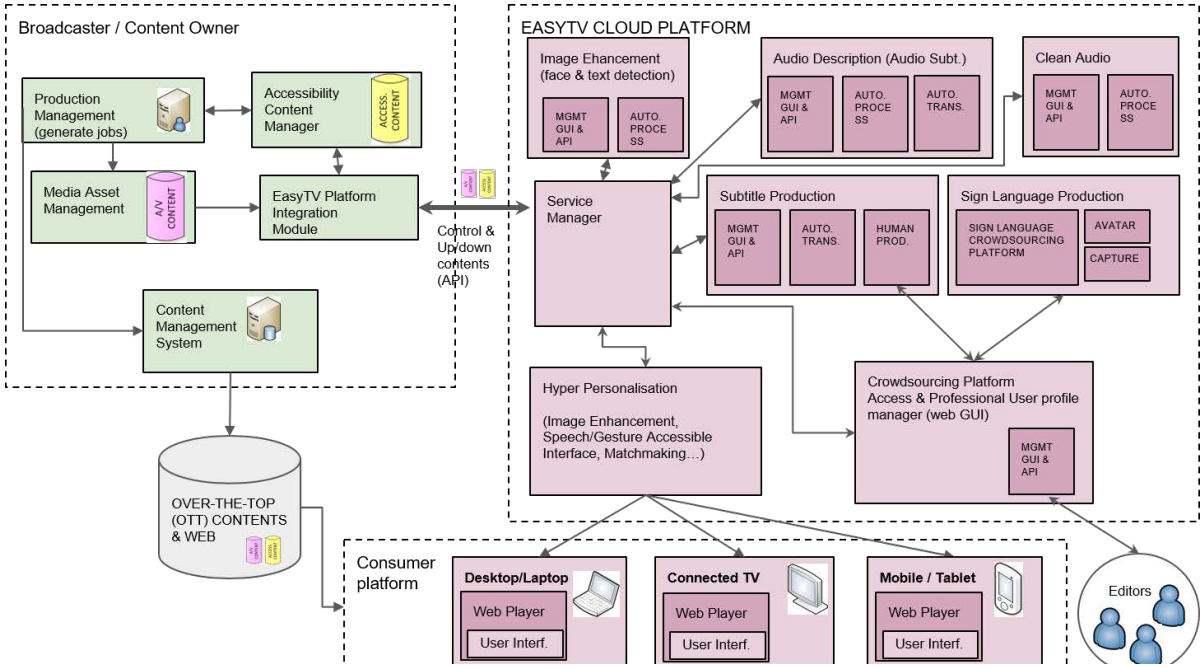

Figure 1: The basic EasyTV architecture

2.2.1 Broadcaster Premises

A broadcaster or content owner is considered the professional client of the EasyTV cloud platform. This client can be the owner of the audio-visual content, and/or has the rights for broadcasting them. This client needs the outsourcing of accessibility content production that the EasyTV platform offers through an interface for requesting production tasks that can involve one or more services.

The premises of the broadcaster or content owner are composed of several internal modules that are responsible for creating, managing and storing the content (i.e., video, audio and metadata files) and rights in a single repository, scheduling for the broadcasting and publishing of content to digital platforms using the adaptive bitrate streaming technique MPEG-DASH [11] and requesting and managing the database of accessibility content.

2.2.2 EasyTV Audio Description

Audio description allows people with visual impairments to hear a verbal version of the visual information, but it has also shown its usefulness to anyone who wants to truly notice and appreciate a complete perspective on any visual event. However, audio description is costly and thus a high percentage of broadcasted content lack this type of service, meaning that the targeted audience of this content obtain reduced user experience. EasyTV proposes the automation of the audio description process for obtaining high quality content for all users independently of their physical condition through the automatic descriptive narrative and automatic voice synthesis of subtitles services.

The automatic descriptive narrative service generates a textual file similar to a subtitle file that contains automatically generated data obtained from the intrinsic information of the multimedia content and from different external information sources, such as the metadata associated to the file and the television guides. This service is further enhanced by an optical character recognition algorithm that translates the burned text (burned subtitles or any other text) contained in the video into a text content. The audio narrative file (i.e., EBU-TT-D [12] or narrative file) that is generated by this service is finally stored in the accessibility content database and used to improve the experience of the users.

The automatic voice synthesis of subtitles enables spoken subtitles by making use of a Text to Speech (TTS) algorithm that is available both on the server-side and on the client-side for a more flexible and effective solution in the case no connection with the server can be established. Subtitles are initially provided by the EasyTV subtitle production service or directly through a web-based application. Then, this service translates the given subtitles to an audio track to be included or mixed to an audio video streaming content (e.g., for audio narratives or audio subtitling). Finally, the end user is able to activate/deactivate the audio subtitling service and play the subtitles or the prerecorded audio subtitles stored in one of the audio track of the streamed content.

2.2.3 EasyTV Clean Audio

Dialogue intelligibility of television audio content is an important issue especially for people with hearing impairments. EasyTV presents an innovative solution based on Background Audio Enhancement technologies to separately adjust 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 EasyTV clean audio component allows the audience to change the balance of the audio mix according to the listening environment or

personal needs, obtaining important benefits for the hearing-impaired people. This component interacts with the sound from the source multimedia content and develops a process for the equalization of frequencies, while advanced filtering techniques enhance the intelligibility of the audio by identifying, processing and reinforcing the bandwidth where the most important audio information is located, especially the voices. The output of this component is two different audio files, one containing the source audio file and the other the clean audio approach. The users can select one of these two files depending on their preferences and in the case of the clean audio file, they have the possibility of dynamically adapting the conditions of the sound by configuring the volume of the voices with respect to the background.

2.2.4 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 extended to allow access into the content to people who do not understand its spoken language, regardless of whether they have any impairment. For this reason, EasyTV proposes a subtitle production component that constitutes a powerful and cost-effective tool to achieve the desired feature without having to assume the extensive bill of the traditional subtitling production. The EasyTV subtitle production component receives as input the content, along with the subtitles in its original language and the required metadata for the translation process management (i.e., original and desired languages, minimum trust level required for editor users, etc.) from the broadcasters. Then, an automatic subtitle translation tool initially provides a first translated version of the subtitles in the languages of interest indicated by the broadcaster/content owners before collaborative crowd-workers are tasked with reviewing the automatic translations and make corrections. Finally, users that assume the role of reviewers revise and accept or reject the translations made by the crowd-workers prior to the publication of the translated subtitles by the broadcaster/content owner over its own publication platform.

2.2.5 EasyTV Sign Language Production

The EasyTV Sign Language Production component is responsible for the creation, validation and exploitation of the sign language content in different languages for easing the access of hearing impaired individuals to media content and services. This component consists of the sign language capturing, the sign language crowdsourcing and the realistic avatar services [13].

The sign language capturing service is a software module that utilizes a RGB-D sensor that is connected to a second screen (i.e., desktop PC/laptop) to capture videos and process them to extract valuable motion information (i.e., body, hand and face motions) that can be used for the correct differentiation among different signs. The recorded data are also accompanied by a text describing the signing sequence (word/sentence) that can be inserted manually by users or computed automatically by deep learning algorithms developed in the project [14][15]. The recorded data and their annotations are finally uploaded to the EasyTV sign language crowdsourcing platform as a response to a sign language task (SLT) assigned to a user.

The sign language crowdsourcing service provides a platform that allows users to contribute with sign languages and translations of available signs, creating a multilingual sign language repository. Currently, Greek, Spanish, Catalan, Italian and English sign languages are supported, although the platform can be easily expanded to accommodate additional sign languages. The administrators of this platform are responsible for defining SLTs by proposing words/sentences that need to be translated in sign language and distribute these SLTs to expert sign language users. Then, the users record and upload motion data to the platform, along with the corresponding text annotations. The administrators can finally revise the users' contributions and accept or reject them. The ultimate purpose of the crowdsourcing platform is to create a large vocabulary of signed videos, along with their translations in all supported languages that can be used to assist researchers towards developing accurate sign language recognition and translation algorithms.

The motion data and the corresponding text is semantically annotated using the EasyTV ontology. Then, an annotator is executed to identify the meaning of the terms that appear in the natural language representation of the signs and to link the terms to their corresponding concept in the multilingual knowledge base BabelNet [16], which can represent words and their senses in different languages.

The realistic avatar service is concerned with the development of a 3D virtual character, as well as the visualization of the recorded signs. This service communicates directly with the sign language data repository in order to acquire the motion data for a specific gloss/word that needs to be signed by the avatar. Then, the service is responsible for accurately and robustly mapping 3D motion data of body, hands and face to the avatar for realistic playback.

2.2.6 Hyper-Personalization

The EasyTV hyper-personalization component utilizes user experience mining and innovative interaction techniques to handle dynamic and continuous changes in user experience in a user-transparent way in order to recommend new personalized services, dynamically adapted to the current context and device of the user.

The EasyTV hyper-personalization component consists of several modules that define user needs and preferences, as well as disabilities and functional limitations, perform automatic turn on and configuration of accessibility features (e.g. volume, rate, pitch, color preferences, etc.) that are built into different TV operating systems, applications and embedded devices and perform content adaptation based on standardized DASH streaming services in order to offer streaming content in the best possible form for a specific user.

2.2.7 Image Enhancement

Image enhancement concerns the processing of an image to obtain a more suitable result that the original image/video, adapted to the needs and requirements of users. Digital image enhancement techniques provide a multitude of choices for

improving the visual quality of images. The EasyTV image enhancement component offers a set of different solutions to assist blind or visually impaired users in accessing multimedia content in an easier way, improving their experience. These solutions include a tool for improving the presentation of subtitles and sign language video manually where the users decide how they want to see the content in relation to the position, color and size or automatically based on the information given by the hyper-personalization module. Furthermore, a magnification tool is proposed that allows the scaling up of texts and images in order to facilitate watching them in more detail. The magnification can be performed either manually, where the tool is able to surround and magnify a specific point clicked by the user like a magnifying glass or automatically where the tool is able to detect and magnify faces and text in specific contents and contexts. To achieve automatic image magnification, an image analysis algorithm was developed to automatically detect texts and faces of an audio-visual content. This algorithm generates a text file with the position where the content to be magnified is located. Once the text file is generated, it is stored in the accessibility content database, from which it can be obtained and used for image magnification in the main screen (e.g., TV) or in the second screen of an external device (e.g., tablet, PC).

2.2.8 Speech Accessible Interface

The Speech Accessible Interface is available on a tablet or a smartphone device when users consume content on a second screen application. Blind and visually-impaired users interact with this service 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 speech accessible interface consists of several components that allow speech recognition of the user via well-known speech recognition engines, TTS synthesis to manage the voice, language, volume and speed, natural language processing for querying the user voice and understanding the user intent and processing of the dialog flow between the user and the application through pre-defined voice dialog templates, including voice prompt templates and semantic annotations for user actions. 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.

2.2.9 Gesture/Gaze Accessible Interface

The gesture/gaze accessible interface service involves the communication between a software module and its corresponding interface (GUI) with the TV set through users' hand and eye movements. More specifically, two well-known sensors (i.e., a RGB-D sensor for capturing hand movements and an eye tracker sensor for capturing eye motions) are utilized. These sensors are connected to a second screen (i.e., desktop PC/laptop) and gather hand and eye movements from users. By employing AI algorithms, hand movements are recognized and gaze data are collected and transformed to specific commands for TV. An intuitive set of hand movements (e.g., index finger in front of mouth for mute) was selected for matching with the movements of the users. Then, the corresponding commands are transmitted to the connected TV set using node.js, WebSockets and the HbbTV protocol.

3 USER EVALUATION AND DISCUSSION

In this section, the user evaluation procedure is initially described and then the evaluation of several EasyTV services is presented and analyzed.

3.1 Evaluation procedure

Two test modalities were chosen to be followed during EasyTV's intermediate testing: live face to face and online. In live face to face tests, the users were initially introduced to the evaluation procedure and then paired with a computer or tablet connected to the EasyTV testing platform. The users were given a consent form in order for the evaluation procedure to be compliant to the GDPR guidelines and ethical requirements. Then, they were asked to interact with the specific EasyTV service and fill an online questionnaire. On the other hand, in online tests, the users were contacted via email by user associations and briefed on the evaluation procedure. Then, the users entered the EasyTV testing platform, read the information and consent form and, if willing to take part in the test, proceeded with watching a video of how the specific EasyTV service functions and answered a questionnaire.

As far as the questionnaire is concerned, it was decided by all EasyTV partners that the users participating in the evaluation tests should share the same questionnaire towards comparing results.

The questionnaire was translated in 5 different languages (i.e., Greek, Spanish, Italian, Catalan and English) so that they could be filled by users in any testing site. The questionnaire initially contained two demographics questions, before moving on service evaluation questions. For the evaluation of the EasyTV services, the System Usability Scale (SUS) was adopted. The SUS is amongst the most popular usability testing tools due to its many advantages, such as its brevity and robustness, as well as it being free of charge [17][18]. Despite its simplicity, Tullis and Stetson in [19] noted that the SUS yielded very reliable results across sample sizes. It has also been successfully applied to a wide range of devices and systems (e.g., learning management systems, landline telephones, web-based interfaces, etc.), which proves its flexibility and lack of dependence towards the system under study. It was designed by John Brooke and it consists of just ten questions, half of which are positive, while the rest are negative statements. These questions are alternated and presented in a fixed standardized order. Informants need to express how much they agree with the proposed statements selecting one of the five options available, ranging from "strongly disagree" to "strongly agree". Final scores for the SUS can range from 0 to 100, where higher scores indicate better usability. Based on research, a SUS score above a 68 is considered above average and anything below 68 is considered below average.

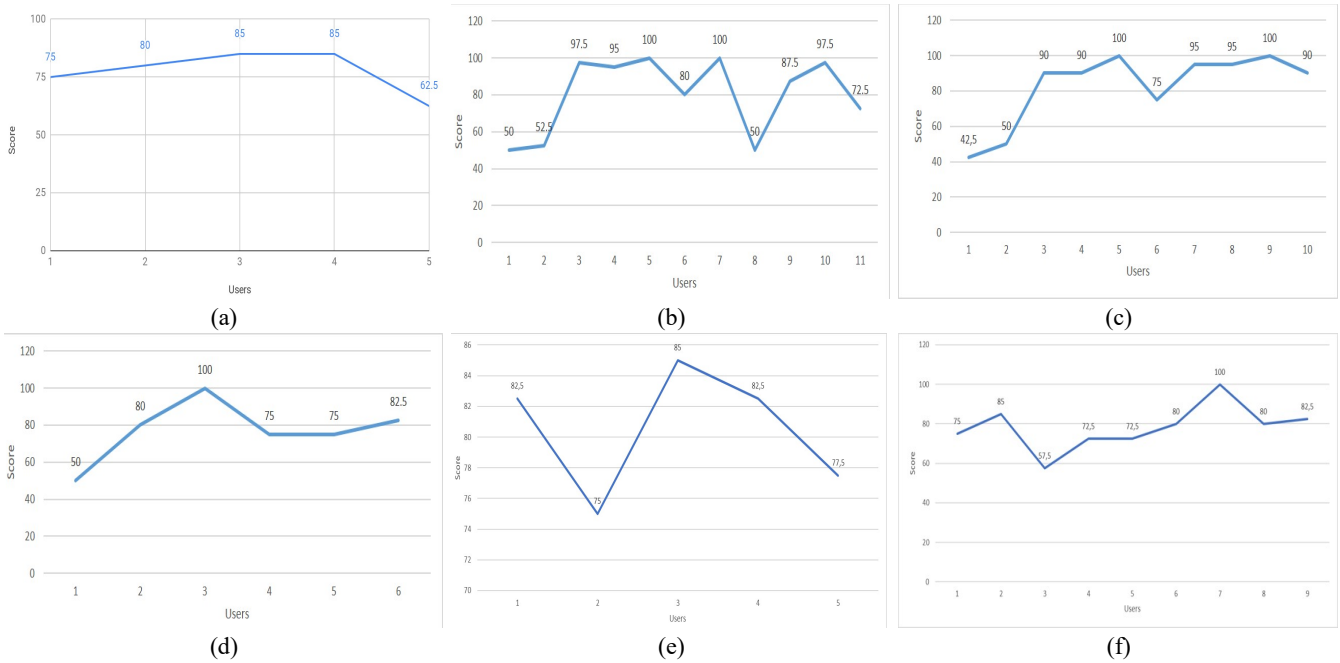

Figure 2: Individual SUS scores for the services that received the most positive user feedback: (a) gesture/gaze remote control, (b) custom magnification, (c) subtitles customization, (d) character detection, (e) screen reader and (f) color-blind subtitles.

3.2 Results and discussion

The intermediate evaluation of the EasyTV services by the users showed promising results, as most services have been positively evaluated by users. More specifically, an average SUS score of 71.1 was achieved for all EasyTV services, after weighing their individual SUS scores by the number of users that evaluated them. The fact that the average SUS score surpasses the value of 68 means that users grade the functionalities of the EasyTV services above average.

Figure 2 depicts the individual SUS scores of the EasyTV services that received the most positive user feedback out of those tested. Some observations can be made by looking at Figure 2, regarding the number of users that evaluated the services and the scores that assigned to them. The custom magnification service was tested by 11 users and received an average SUS score of 80.23. Three users evaluated this service with a score lower than average (< 55), while the rest of the users evaluated it with a higher than average SUS score (> 72). Similarly, the subtitle customization service was evaluated by 10 users and received an average SUS score of 82.75. Two users evaluated this service with a low SUS score (< 55), while the rest of them showed a very positive attitude towards it (SUS score > 75).

Except from the quantitative feedback (i.e., SUS scores), the services received also qualitative comments from users on actions that should be taken into consideration to improve the EasyTV services. Both qualitative and quantitative feedback were taken into account by the partners of the EasyTV project so that all services and especially the ones that performed mediocre were improved in a way to meet user needs and requirements before the final testing of the EasyTV services by users.

Most user comments on the EasyTV services concern the design of more intuitive interfaces and the implementation of additional functionalities or the improvements of existing ones. More specifically, users asked for reduced complexity in the text detection service, as well as the possibility for addition of voice control. Moreover, users wanted more polished navigation for the screen reader service and more user friendly and intuitive interfaces for the sign language capturing and gesture/gaze remote control services. Regarding the speech platform service, users needed simplified voice commands, as well as improved interpretation of voice utterances. Finally, as far as sign language crowdsourcing platform is concerned, users opted for simpler workflows, while regarding the 3D avatar service, users commented on the lack of realism of hand movements that is the result of inaccuracies in the collected data. Even though the results of these intermediate tests are positive in general, necessary steps were taken for the improvement of the EasyTV services, so that they were well received in the final testing.

Although the SUS can provide accurate scores with a relatively small sample, the results obtained in these tests are not statistically representative. This means that the information elicited does not necessarily apply to the general European population. However, the informants already shed some light on what can be improved with the services developed so far, which was the objective of this phase. Furthermore, some improvements could be made in the online testing platform so that it is fully accessible to all kinds of users for the next phases of the project. The project partners that are user associations themselves were

consulted during the intermediate testing process and will carry on piloting the platform until it is ready for the final testing. In the final testing, a much larger number of users is expected to test and validate the EasyTV services.

4 CONCLUSIONS

In conclusion, EasyTV is designed to provide equal access to audio-visual services for all users, especially those with hearing and visual impairments. Several services were designed and developed to facilitate the interaction of users with audio-visual content based on sophisticated image and video processing techniques, new communication protocols and crowdsourcing platforms. An intermediate testing phase was performed so that users assessed whether the developed EasyTV services meet the user needs and preferences. The evaluation results revealed that most EasyTV services were positively received by users, while user qualitative and quantitative feedback will be used to improve all services so that they can better accommodate user needs and requirements and achieve better SUS scores in the final testing. As a future work, EasyTV aims improve its services, as well as analyze and exploit the business potential of its services to provide enhanced accessibility to people with disabilities either freely or under a paid subscription scheme.

ACKNOWLEDGMENTS

This work has been supported from European Research Council under Grant agreement no. H2020-ICT-19-2016-2 "EasyTV: Easing the access of Europeans with disabilities to converging media and content".

REFERENCES

[1] Woods, R. L., & Satgunam, P. (2011). Television, computer and portable display device use by people with central vision impairment. *Ophthalmic and Physiological Optics*, *31*(3), 258-274.

[2] Nilsson, I., Lundgren, A. S., & Liliequist, M. (2012). Occupational well-being among the very old. *Journal of Occupational Science*, *19*(2), 115-126.

[3] Ahn, D., & Shin, D. H. (2013). Is the social use of media for seeking connectedness or for avoiding social isolation? Mechanisms underlying media use and subjective well-being. *Computers in Human Behavior, 29*(6), 2453-2462.

[4] Coelho, J., Rito, F., & Duarte, C. (2017). "You, me & TV"—Fighting social isolation of older adults with Facebook, TV and multimodality. *International Journal of Human-Computer Studies, 98*, 38-50.

[5] Orero, P. (2016). From DTV4ALL to HBB4ALL: Accessibility in European Broadcasting. In *Researching Audio Description* (pp. 249-267). Palgrave Macmillan, London.

[6] Itagaki, T., Owens, T., & Orero, P. (2016, May). Digital TV accessibility—Analogue switch off in Europe and Africa. In *2016 IST-Africa Week Conference* (pp. 1-8). IEEE.

[7] Clark, C., Basman, A., Markus, K. G., & Zenevich, Y. (2013). Cloud-scale architecture for inclusion: Cloud4all and GPII. *Assistive Technology: From Research to Practice–Proceedings of AAATE*.

[8] Peissner, M., Vanderheiden, G. C., Treviranus, J., & Tsakou, G. (2014, June). Prosperity4All–Setting the Stage for a Paradigm Shift in eInclusion. In *International conference on universal access in human-computer interaction* (pp. 443-452). Springer, Cham.

[9] Van Deventer, M. O., de Wit, J. J., Guelbahar, M., Cheng, B., Marmol, F. G., Köbel, C., ... & Stockleben, B. (2013). Towards next generation Hybrid broadcast broadband, results from FP7 and HbbTV 2.0, *IET Conference Proceedings*, pp. 12.3-12.3(1).

[10] Orero, P., Martín, C. A., & Zorrilla, M. (2015, June). HBB4ALL: Deployment of HbbTV services for all. In *2015 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting* (pp. 1-4). IEEE.

[11] Sodagar, I. (2011). The mpeg-dash standard for multimedia streaming over the internet. *IEEE multimedia*, 18(4), 62-67.

[12] Hrbak, M., Herceg, M., Stefanovic, D., & Fimić, N. (2017, September). Integration of EBU-TT-D subtitles in HbbTV surrounding. In *2017 IEEE 7th International Conference on Consumer Electronics-Berlin (ICCE-Berlin)* (pp. 173-174). IEEE.

[13] Stefanidis, K., Konstantinidis, D., Kalvourtzis, T., Dimitropoulos, K., & Daras, P. (2020). 3D technologies and applications in sign language. *Recent Advances in 3D Imaging, Modeling and Reconstruction*, IGI Global, Pennsylvania, US.

[14] Konstantinidis, D., Dimitropoulos, K., & Daras, P. (2018, June). Sign Language Recognition Based on Hand and Body Skeletal Data. In *3DTV Conference*, IEEE, Stockholm, Helsinki.

[15] Konstantinidis, D., Dimitropoulos, K., & Daras, P. (2018, Oct). A Deep Learning Approach for Analyzing Video and Skeletal Features in Sign Language Recognition. In *IEEE International Conference on Imaging Systems and Techniques*, Krakow, Poland.

[16] Navigli, R., & Ponzetto, S. P. (2010, July). BabelNet: Building a very large multilingual semantic network. In *Proceedings of the 48th annual meeting of the association for computational linguistics* (pp. 216-225). Association for Computational Linguistics.

[17] Katsanos, Christos; Tselios, Nikolaos; and Xenos, Michalis (2012). Perceived usability evaluation of learning management systems: a first step towards standardization of the system usability scale in Greek. *Proceedings of the 2012 16th Panhellenic Conference on Informatics*.

[18] Bangor, Aaron; Kortum, Philip T; & Miller, James T. (2008). An Empirical Evaluation of the System Usability Scale. *International Journal of Human-Computer Interaction, 24*(6): 574-594.

[19] Tullis, T. S., & Stetson, J. N. (2004, June). A comparison of questionnaires for assessing website usability. In *Usability professional association conference* (Vol. 1).