



Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 196 (2022) 647–654

Procedia
Computer Science

www.elsevier.com/locate/procedia

CENTERIS - International Conference on ENTERprise Information Systems / ProjMAN - International Conference on Project MANagement / HCist - International Conference on Health and Social Care Information Systems and Technologies 2021

Audiovisual translation models for visually impaired users of Interactive Television (iTV)

Rita Oliveira*, Jorge Ferraz de Abreu, Ana Margarida Almeida

DigiMedia, Department of Communication and Arts, University of Aveiro, 3810-193 Aveiro, Portugal

Abstract

Audiovisual translation techniques, such as audio description (AD), subtitling and dubbing, have a significant role in the overall television experience of people with visual or hearing impairments, illiterate or persons unfamiliar with foreign languages. These techniques are assured by qualified professionals using specific technical resources, resulting in expensive production processes. However, new models to support alternative and more affordable audiovisual translation techniques can be used in the current interactive television (iTV) ecosystem.

This paper aims to discuss two models to produce audiovisual translation (voluntary audio description and automatic reading of subtitles), which present the potential to increase the offer of TV contents adapted to the aforementioned persons.

An integrative literature review was carried out in order to identify and define the technical framework inherent to the proposed models. The main findings of the performed analysis show that the models have technical viability enabling its technical implementation.

© 2021 The Authors. Published by Elsevier B.V.

This is an open access article under the CC BY-NC-ND license (<https://creativecommons.org/licenses/by-nc-nd/4.0>)

Peer-review under responsibility of the scientific committee of the CENTERIS –International Conference on ENTERprise Information Systems / ProjMAN - International Conference on Project MANagement / HCist - International Conference on Health and Social Care Information Systems and Technologies 2021

Keywords: audiovisual translation; accessibility; audio description; subtitling; interactive television.

* Corresponding author.

E-mail address: ritaoliveira@ua.pt

1. Introduction

Audiovisual translation is commonly used in the television context, especially for foreign content interpretation, using techniques such as subtitling and dubbing. When visually impaired viewers are at stake, the most typically used techniques are voice-over and audio description in their native language, which enables them a richer understanding of Television programs. Audio description is, therefore, the most common current strategy that can be used to promote accessibility to television content, being an important vector in the creation of a more autonomous and inclusive TV experience, promoting the well-being of people with visual impairment.

In Portugal, the current situation concerning the relationship between users with visual impairments and television demands more effort in this research field since there are many people with this type of impairment who do not benefit from audio description. According to the 2011 Portuguese Census, 9.3 % of the population over 5 years old (about 921,000 people) is visually impaired, 97 % of them (about 893,000) have great difficulty in seeing and the remaining 3 % (about 28,000) are blind [1]. In addition, Portuguese viewers with visual impairment are currently limited when they access television content since they can only follow some TV series with audio description on a single Portuguese television channel: RTP1, the public TV station [2]. In the opposite direction, the international scenario of the audio description is very different. Countries like the United Kingdom and United States of America present an audio description television market much more regulated and expanded.

Regarding subtitling, it is the more used audiovisual translation technique in the Portuguese TV broadcasting [3]. Television viewers in Portugal can daily use subtitles in native language contents through Teletext to watch recorded (pre-prepared subtitles) and live (automatic subtitles) programs [4], [5]. In addition, in Portugal, like in several countries, all foreign language TV programs are subtitled instead of dubbed. Therefore, this translation technique helps many viewers who cannot have the sound turned on or that have hearing impairments or difficulties in understanding foreigner languages. However, in the case of visually impaired viewers, one major problem when they are watching TV and there are subtitles in foreign language TV programs is the inability to read them. In most of the cases, the solution is having a sighted person reading the subtitles, but this is not an effective solution. In this scenario, a strategy to automatically read the subtitles to viewers with visual impairments appears to be a worthy approach to meet the needs of this kind of users.

In this context, the research presented in this paper proposes two new models (voluntary audio description and automatic reading of subtitles) for the creation of audiovisual translation techniques, addressed to television contents in a native language and non-native language, correspondingly. Besides that, the described technical framework, which is in the basis of these models, was structured through an integrative literature review.

The paper is organized as follows. In the next section, a state of the art concerning new approaches on audio description and automatic reading of subtitles is presented. In section 3, the two proposed translation models are described and the literature review that supported the definition of the technical framework inherent to them is explained. Also, in section 3 the systems architectures diagrams of the two proposed models are explained. The last section presents the most relevant conclusions and the work to be done in the future.

2. Related Work

The international audio description scenario is quite different from Portugal. Countries like the United Kingdom (UK) and United States of America (USA) have a much more structured and developed audio description market. In the UK, the broadcasters have the legal obligation of offering at least 10% of its programming through audio description [6] and in the USA, the local stations of the 60 zones of the country with greater use of Television are obliged to broadcast about 4 hours a week of programs with audio description [7].

In Portugal, the scenario is more unfavorable, although there has been some research towards the development of this field. Neves [8] is a pioneer in this field and has developed several projects that encourage the use of audio description in different domains. Oliveira, Ferraz de Abreu and Almeida [9] have been working in this field, particularly under the scope of the interactive television scenario, suggesting an adapted system to visually impaired users that supports the audio description access.

In Spain, the AudescMobile application was created allowing the access to AD of several types of audiovisual productions, using audio fingerprint (analysis of a sound excerpt recorded by the application) [10]. The WhatsCine

app [11], also created in Spain, allows the access of audio description, subtitling and sign language in television and cinemas. In Brazil, there are projects related to the promotion of audio description; for example, Campos, Araújo and Filho [12] suggests a mechanism based on a semantic web application for the automatic creation of audio description for movies; Domingues et al. [13] proposes an automatic system to create audio descriptions in Portuguese, based on the script, subtitles and/or other metadata related to the video. In the United States of America, there is a free and experimental tool developed as part of an academic research that allows adding audio descriptions to YouTube videos [14], although these audio descriptions can only be created and accessed through the ‘YouDescribe’ web tool.

In audiovisual translation, volunteering appears as a non-professional technique and is increasing in online communities and social networks [15]. Concerning the volunteer model to provide inclusive services, there are several European initiatives that take advantage of it for the creation of audiobooks [16], [17], [18], audio description book illustrations [18] or audio newspapers [19] and for the supply of geographical information [20] and social initiatives [21]. The mobile app ‘Be My Eyes’ is also supported by a volunteer model, connecting blind and visually impaired with sighted helpers via live video connection: the blind users request assistance and volunteers receive video and describe what the blind person want (for example to know the expiry date on the milk) [22].

In Europe, there are also more extensive audiovisual translation projects, such as the ‘HBB4ALL’ [23] that aims to promote a wider availability of accessible media aids to all viewers with special needs (like audio description, subtitles, and sign language). The project aims to develop a cross-platform production and to distribute accessibility features in a more cost-efficient and more flexible and uncomplicated way, helping users and also broadcasters and media producers. Similarly, the European research group ‘TransMedia’ has remarkable studies in audiovisual translation, including audio description [24].

In what concerns subtitling, it is worth mentioning the ‘Subtitle Quality’ project, of the British Broadcasting Corporation (BBC), which intends measuring and improving subtitle quality in UK [25] and the ‘Speech-to-Text’ project, which use the BBC subtitle archive to create more accurate speech-to-text [26].

Regarding the automatic reading of subtitles systems for television, between 2002 and 2005, the Germany [27] and Sweden [28] public television providers implemented a solution that automatically read the subtitles, but it had limitations inherent to the use of two set-top boxes by the user in order to receive the reading sound through a secondary audio channel. Latest studies were developed in this field, namely in Sweden and Denmark: in these countries, some prototypes were developed using complex solutions, which used Optical Character Recognition (OCR) [29] and text-to-speech (TTS) [30] systems. Also, in Sweden, Ljunglöf, Derbring and Olsson [31] developed a prototype that assures the automatic reading of subtitles of videos and television programs through a TTS engine using a computer. However, this solution was not specifically adapted to a television scenario nor allows the cohabitation of visually impaired users and sighted users. The media player ‘KMPlayer’ allows, since 2008, the automatic reading of subtitles that appear on the computer screen and are associated with a subtitle file [32].

The described state of the art highlights that the research area in audio description is growing and the international Television market is giving more attention to viewers with visual impairment increasing the AD offer. Usually, audio description for television is performed by qualified professionals using specific technical features, which makes it financially expensive. Thus, the creation of a volunteer model to perform audio descriptions becomes clearly helpful. Actually, from some of the examples mentioned above, it can be said that volunteering is usually applied to help visually impaired viewers. So, the idea of producing voluntary audio descriptions seems sustained and appropriated.

Regarding the automatic reading of subtitles in television contents, two methods are commonly used in order to read the subtitles: OCR and TTS. The TTS technology appears to be easier to implement considering the OCR calls for a high degree of efficiency and effectiveness in order to assure the alphabet recognition in the TV screen.

3. Proposed Models

As mentioned before, this paper proposes two new models for the creation of audiovisual translation techniques: voluntary audio description and automatic reading of subtitles.

The voluntary audio description technique addresses native language content and is based on a new model for the creation of TV audio descriptions, taking advantage from volunteering. This model is supported in the recording of

real-time audio descriptions, rather than in the common process where the AD is processed in a professional way. This new approach relies on volunteers (e.g. relatives/caregivers that usually help these viewers during television viewing) that do not necessarily have specific audio description professional skills. The rationale behind this technique is based on the enhancement of the current poor offer of AD in the Portuguese scenario, in which there is a weak financial investment in hiring professionals for AD production and, also, aiming to take benefit from the familiar dynamics between people with visual impairment and their relatives who can support the creation of informal AD. A preliminary study using this technique was already conducted and the results show that not only volunteers were satisfied with the procedure used to create AD but also visually impaired users felt that this type of AD, despite its non-professional quality, assisted them [33].

In the second scenario (automatic reading of subtitles technique addressed to non-native language content) a dedicated TTS process is needed. The purpose of this technique is to perform the subtitles reading of foreign TV contents in order to visually impaired viewers who do not understand a second language can have access to such contents. This approach appeared to be relevant since most of the television contents broadcasted in foreign languages in Portugal (such as movies and series) are subtitled and the offer of dubbed TV contents is very limited. In this scope, the access to such contents by visually impaired persons who are not fluent in foreign languages is clearly compromised.

Section 3.1.3 describe these two models, being the first (voluntary audio description) supported by an approach that mainly takes advantage of the family relationships between viewers with and without visual impairment for the creation of audio descriptions; whereas the second (automatic reading of subtitles) is based on a mechanical and automated approach with no human intervention.

3.1. Technical Framework of the Proposed Models: An Integrative Literature Review

After identifying the audiovisual translation models, it was necessary to understand its technical viability. An integrative literature review was, therefore, conducted in order to validate the technical framework of the two proposed models aiming to support the requirements analysis, the identification of features and the design of its system architectures diagrams, as described in the next sections.

The integrative literature review is a refined form of research that requires a commitment of research skill and insight. With this research method, it is possible to identify a topic or problem with appropriate literature resources, analyze and review the literature and create new understandings of the topic through one or more forms of synthesis. Moreover, most integrative literature reviews consider two kinds of topics: mature and new or emerging topics [34], [35]. In the case of this study, the review led to a preliminary conceptualization of the topic (a new model or framework), benefiting from a holistic conceptualization and from the previous synthesis of the literature to date.

3.1.1. Selection of the sources

Advanced searches were carried out in ‘Scopus’ and ‘Web of Science’ databases. The searches in these databases included the at least two combinations of the following keywords in English: i) audio description, Television, TV, voluntary, volunteer; ii) subtitling, Television, TV, automatic, reading. The inclusion criteria defined for the literature selection were articles or conference papers published since the year 2000 to present (2021) and the presence of the keywords in the title. On the other hand, the exclusion criteria defined for the literature selection were articles that had been mentioned in the related work and were not linked to the television context and/or the target audience of the study: visually impaired persons. At the end of the process, the literature review was based on the resources identified below. The selected resources related to voluntary AD included four references (two journal articles [36], [37] and two conference papers [38], [39]) between 2010 and 2016. Regarding the automatic reading of subtitles, the selected resources included four references (one journal article [40] and three conference papers [41], [42], [43]) between 2007 and 2014.

3.1.2. Critical Analysis

Despite this study is supported on an extensive theoretical framework, the conducted integrative review enables the identification and definition of the technical challenges of the proposed techniques.

First, to clearly understand the proposed models, it is important to divide its components into production and consumption side. Regarding voluntary audio description, on both sides, there will be human intervention. In the production side, the volunteer will record audio description supported by a web/mobile application and, in the consumption side, the visually impaired viewers will have access to the created AD through an iTV application. In Table 1 are identified the technical challenges inherent to voluntary audio description.

Table 1. Technical challenges of voluntary audio description

Voluntary AD	
Production	AD recording Web/Mobile App development
Consumption	TV App development

Concerning automatic reading of subtitles, in the production side there will be no human involvement because the process will be completely automatic, but in the consumption side, like in the previous audiovisual approach, visually impaired viewers will be engaged. In the production part, the main work will be made by a TTS engine, which will synchronously read the subtitles, and, in the consumption side, the visually impaired viewers will have access to the automatic reading through an iTV application. In Table 2 are identified the technical challenges of the automatic reading of subtitles approach.

Table 2. Technical challenges of voluntary audio description

Automatic Reading of Subtitles	
Production	Synchronization
Consumption	TV App development

3.2. Features and System Architectures Diagrams of the Proposed Models

The system architectures diagrams of each proposed audiovisual translation techniques, supported on the critical analysis of the literature review, are following described.

- Voluntary audio description

The concept for the voluntary approach that can support the development of a cross-platform solution for iTV should be sustained in the following features: i) recording of the audio description to be carried out by the volunteers synchronously with the television content; ii) uploading of the resulting audio file into a repository of audio descriptions; and iii) distribution of the audio description using the application associated to the corresponding television program (usually a catch-up TV program or other non-linear one). It is also foreseen to provide the possibility of selecting the audio description by various criteria (e.g. name of the author or rating).

The workflow of this model is depicted in Figure 1 and is following described (the technical challenges are identified with pink colored text and followed by an asterisk.). The voluntary AD process begins on the Audio Description Production side, where the volunteer has access to the television content to be described through an application available on a mobile device or computer. The volunteer writes the AD segments, uploads them to the server and after that, they are stored in a database. Finally, on the side of the Audio Description Consumption, the visually impaired user through a TV application available in the set-up box requests the AD for the TV content that wants to view.

This model could allow that the audio descriptive content is accessible not only to people with visual impairment, but also to sighted people that are temporarily unable to see (e.g.: after eye surgery) or, for other reasons, are not able to view the TV content (e.g.: when performing other tasks while watching TV).

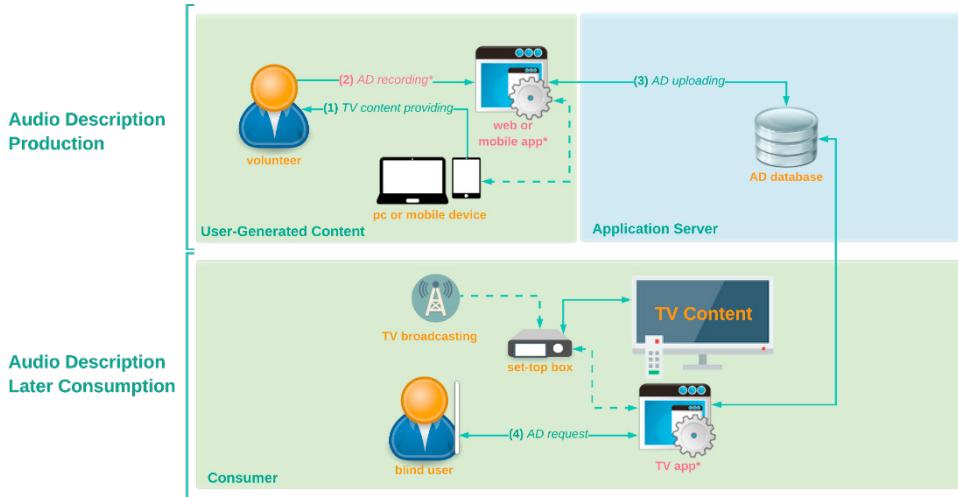


Fig. 1. System architecture diagram of voluntary AD.

- Automatic Reading of Subtitles

The approach that supports the proposed automatic reading of subtitles system can be sustained in the following features: i) identification of the subtitles for each TV program using its subtitle file (e.g.: via a free online repository); ii) TTS reading of the subtitles to produce an audio stream; iii) connection to the TTS engine with the corresponding audio stream file; iv) creation of a cross-platform application that allows the audio stream synchronization with the television program that is being broadcasted (which can be performed through an interconnection with an existing TV infrastructure that must enable the identification of the cue points required for synchronization). The workflow of this model is depicted in Figure 2 and is following described. The mechanism of the automatic reading of subtitles production is processed from srt files that are stored in a database and are read through a text-to-speech engine accessible in a web service. The process of audio synchronization of the subtitles allows the visually impaired user to be able to perform automatic reading requests to a TV application in order to watch TV contents that are broadcasting with audible subtitles. Like audio description, subtitled content can be accessible not only to visually impaired persons who do not have familiarity with a second language, but also sighted people that are temporarily unable to see or for other reasons are not able to read the subtitles (e.g.: illiterate people).

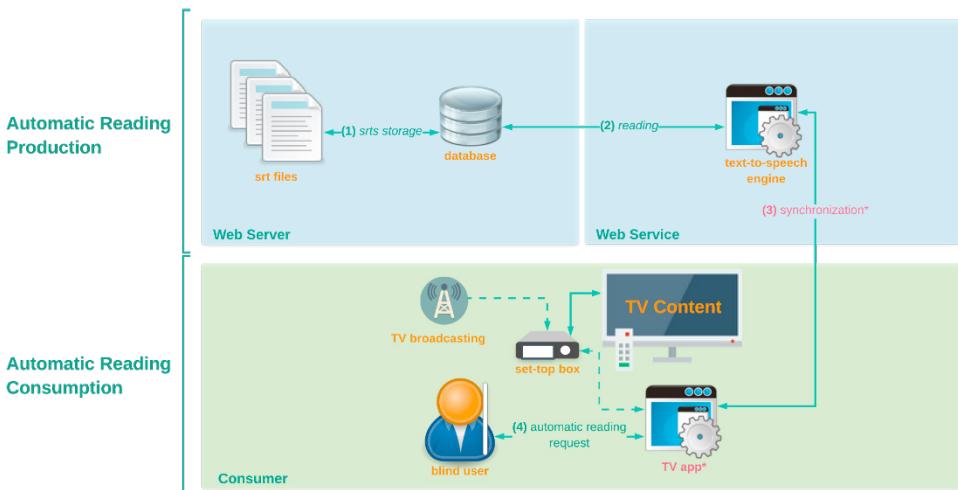


Fig. 2. System architecture diagram of automatic reading of subtitles.

4. Conclusions and Future Work

The proposed models prove to have technical viability enabling its implementation and the future development of two products/services that can represent new forms of inclusion in the TV eco-system. These new models have an interesting potential to modify the television paradigm inherent to the experience of Portuguese viewers with visual impairment who are currently limited in the accessibility to television content as, nowadays, they can only follow some TV series of RTP1 channel through audio description and are unable to read subtitles in foreign TV programs. Accordingly, this study will contribute to the increase of the level of participation of visually impaired viewers in society, expanding also their levels of well-being and literacy, since television, as a means of social communication, has an interesting potential at this level. From the presented results, we can extend this research to a more practical level, namely in what concerns the conceptualization and specification of the more challenging features of the proposed models. Thus, low-fidelity prototypes are being developed and will support solutions that will reproduce the technical challenges identified in this study. After that, these prototypes will be validated by volunteers (in the case of AD), low vision and blind viewers (in both techniques) through accessibility and usability tests.

Acknowledgements

The authors would like to thank FCT and FSE for the financial support to this research (ref. SFRH/BPD/112129/2015) under the 3rd European community support framework. The authors would also like to thank FCT for the financial support to DIGIMEDIA (UIDP/05460/2020+ UIBD/05460/2020) through national funds.

References

- [1] Instituto Nacional de Estatística, Dificuldades (N.º) da população residente com dificuldades por Local de residência (à data dos Censos 2011), Sexo, Grupo etário, Tipo de dificuldade, Grau de dificuldade e Dimensão (indivíduos com dificuldades), (2012). http://censos.ine.pt/xportal/xmain?xpid=CENSOS&xpgid=ine_censos_indicador&contexto=ind&indOcorrCod=0006369&selTab=tab10 (accessed 5 May 2021).
- [2] Rádio Televisão Portuguesa, RTP Acessibilidades – Audiodescrição, (2021). <http://www.rtp.pt/wportal/acessibilidades/audiodescrecao.php> (accessed 5 May 2021).
- [3] Neves, J. (2006). Audição pública sobre o acesso de pessoas com necessidades especiais às emissões de televisão em Portugal, Assembleia da República: Lisboa.
- [4] Rádio Televisão Portuguesa, RTP Acessibilidades – Legendagem em Teletexto, (2021). http://www.rtp.pt/wportal/acessibilidades/legendagem_teletexto.php (accessed 5 May 2021).
- [5] Voice Interaction, Automatic Closed Captioning, (2019). <https://www.voice-interaction.com/us/solutions/automatic-closed-captioning/> (accessed 5 May 2021).
- [6] Office of Communications, Ofcom's Code on Television Access Services, (2021). https://www.ofcom.org.uk/_data/assets/pdf_file/0035/179954/tv-access-services-code.pdf (accessed 5 May 2021).
- [7] The Audio Description Project, Top 60 Markets Which Must Provide Audio/Video Description, (2021). <http://www.acb.org/adptvcctop60.html> (accessed 5 May 2021).
- [8] Neves, J. (2011). Imagens que se Ouvem – Guia de Audiodescrição, Instituto Politécnico de Leiria: Leiria.
- [9] Oliveira, R., de Abreu, J.F. & Almeida, A.M. (2017). “Promoting interactive television (iTV) accessibility: an adapted service for users with visual impairments”. Univ Access Inf Soc, 16: 533. Springer Berlin Heidelberg. doi: 10.1007/s10209-016-0482-z.
- [10] Organización Nacional de Ciegos Españoles, Audesc Mobile: la audiodescripción en nuestra mano, (2014). <https://cti.once.es/noticias/audesc-mobile-la-audiodescripcion-en-nuestra-mano> (accessed 5 May 2021).
- [11] Whatscine, Accesibilidad audiovisual en cines y televisión, (2016). <http://www.whatscine.es/> (accessed 5 May 2021).
- [12] Campos V, Araújo, T. and Filho, G. (2014). CineAD: “Um Sistema de Geração Automática de Roteiros de Audiodescrição”. Chapter 1 in Simpósio Brasileiro de Sistemas Multimídia e Web (WebMedia) 2014 - IX Workshop de Teses e Dissertações (WTD). <http://www.lbd.dcc.ufmg.br/colecoes/wtdwebmedia/2014/001.pdf>.
- [13] Domingues, L., Campos, V., Santos, D., Araújo, T. and Filho, G. (2016). “GT-AaaS 2.0 - Acessibilidade como um Serviço com Foco em Pessoas com Deficiência Visual”. Poster presented at the 17th Workshop RNP. http://wrnp.rnp.br/sites/wrnp2016/files/wrnp16_lamina_gttaas_edit_v3.pdf.
- [14] YouDescribe, A free accessibility tool for adding audio description to YouTube videos, (2021). <http://youdescribe.org/> (accessed 5 May 2021).

- [15] Olohan, M. (2013). "Why do you translate? Motivation to volunteer and TED translation". *Translation Studies*, 7:1, 17-33. doi: 10.1080/14781700.2013.781952
- [16] Biblioteca Nacional de Portugal, Leitura para Deficientes Visuais – Recursos áudio, (2021). http://www.bnportugal.gov.pt/index.php?option=com_content&view=article&id=121&Itemid=155&lang=pt (accessed 5 May 2021).
- [17] Librivox, Acoustical liberation of books in the public domain, (2021). <https://librivox.org/> (accessed 5 May 2021).
- [18] Projeto de Leitura Inclusiva Partilhada, KITS PLIP, (2021). <http://plip.ipleiria.pt/dinamizacao/> (accessed 5 May 2021).
- [19] National Talking Newspapers and Magazines, About Our Service, (2021). <http://www.tnauk.org.uk/index.html> (accessed 5 May 2021).
- [20] Parker, C., May, A., Mitchell, V., Burrows, A. (2013). "Capturing Volunteered Information for Inclusive Service Design: Potential Benefits and Challenges". *The Design Journal*, 16(2), pp. 198-218. doi: 10.2752/175630613X13584367984947
- [21] Paredes, H., Fernandes, H., Sousa, A., Fortes, R., Koch, F., Filipe, V., Barroso, J. (2015). "CanIHelp - A platform for inclusive collaboration". *Universal Access in Human-Computer Interaction - Access to Interaction*. Vol. 9176 of the series LNCS pp. 474-483. Springer Verlag. doi: 10.1007/978-3-319-20681-3_45
- [22] Be My Eyes, Bringing Sight to the Blind and Visually Impaired, (2021). <http://bemyeyes.com/> (accessed 5 May 2021).
- [23] HBB4ALL, Hybrid Broadcast Broadband For All - Connected TV Accessibility, (2018). <http://www.hbb4all.eu/> (accessed 5 May 2021).
- [24] Matamala, A., Remael, A. (2014). "Audio-description reloaded: An analysis of visual scenes in 2012 and Hero". *Translation Studies*, 8:1, 63-81. doi: 10.1080/14781700.2014.943678
- [25] British Broadcasting Corporation (BBC), Research and Development - Subtitle Quality, (2016). <http://www.bbc.co.uk/rd/projects/live-subtitle-quality> (accessed 5 May 2021).
- [26] British Broadcasting Corporation (BBC), Research and Development - Speech-to-Text, (2018). <http://www.bbc.co.uk/rd/projects/speech-to-text> (accessed 5 May 2021).
- [27] Verboom, M., Crombie, D., Dijk, E. and Theunisz, M. (2002). "Spoken subtitles: Making subtitled TV programmes accessible". *Computers Helping People with Special Needs*. Vol. 2398 of the series Lecture Notes in Computer Science pp 295-302. doi:10.1007/3-540-45491-8_62
- [28] A-focus, Utredning avseende TV-tillgänglighet för personer med funktionsnedsättning, (2010). <https://docplayer.se/18064-Utredning-avseende-tv-tillganglighet-for-personer-med-funktionsnedsattning-for-myndigheten-for-radio-och-tv.html> (accessed 5 May 2021).
- [29] Folke Eliasson, IT i praktiken. Slutrapport, (2005). <http://docplayer.se/994886-It-i-praktiken-75-projekt-om-it-fragor-for-olika-typer-av-funktionshinder-slutrapport.html> (accessed 5 May 2021).
- [30] Nielson, S. and Bothe, H. (2007). "SubPal: A device for reading aloud subtitles from television and cinema". 5th Conference and Workshop on Assistive Technology for People with Vision and Hearing Impairments. <https://www.semanticscholar.org/paper/SUBPAL-A-Device-for-Reading-Aloud-Subtitles-from-Nielsen-Bothe/c2fda12b0f3ce9a6ca8c025640ec0014a48f82a9.pdf>
- [31] Ljunglöf, P., Derbring, S. and Olsson, M. (2012). "A free and open-source tool that reads movie subtitles aloud". Third Workshop on Speech and Language Processing for Assistive Technologies (SLPAT '12). Association for Computational Linguistics, Stroudsburg, PA, USA, 1-4. www.aclweb.org/anthology/W12-2901.
- [32] KMP, KMPlayer | Multimedia Player, (2018). <http://www.kmplayer.com/> (accessed 5 May 2021).
- [33] Removed due to blind review
- [34] Torraco, R. (2016). "Writing Integrative Literature Reviews: Using the Past and Present to Explore the Future". *Human Resource Development Review*. Vol. 15(4) 404–428. doi: 10.1177/1534484316671606.
- [35] Souza, M., Silva, M. and Carvalho, R. (2010). "Integrative review: what is it? How to do it?". *Einstein (São Paulo)*, 8(1), 102-106. doi: 10.1590/s1679-45082010rw1134
- [36] Gorlova, N., Troska, Z., Starovojtova, I. and Kononova, T. (2016). "Volunteers training for audio description service within the Sochi 2014 project as exemplified by activity of the Russian state social university volunteer center: research and findings". *Indian Journal of Science and Technology*. 9(14): 91096. doi: 10.17485/ijst/2016/v9i14/91096
- [37] Lopez, M., Kearney, G. and Hofstädter, K. (2016). "Enhancing audio description: Sound design, spatialisation and accessibility in film and television". *Institute of Acoustics*. Vol. 38. Pt. 2, 38, pp. 174-182. <http://enhancingaudiodescription.com/wp-content/uploads/2016/11/RS2016-paper-Lopez-et-al.pdf>
- [38] Álvarez, M. (2014). "Subtitle, video marked and audio description at the Spanish current television". *Historia y Comunicación Social* 19, pp. 161-172. <http://revistas.ucm.es/index.php/HICS/article/view/45018>
- [39] Caldera-Serrano, J. (2010). "Thematic description of audiovisual information on television". *Aslib Proceedings*, Vol. 62 Issue: 2, pp.202-209. doi: 10.1108/00012531011034991
- [40] Mihkla, M., Hein, I., Kiissel, I., Räpp, A., Sirts, R.; Valdna, T. (2014). "A System of Spoken Subtitles for Estonian Television". *Frontiers in Artificial Intelligence and Applications*, Volume: 268, 19 - 26. doi: 10.3233/978-1-61499-442-8-19
- [41] Pražák, A., Psutk, J.V., Psutka, J., Loose, Z. (2013). "Towards live subtitling of TV ice-hockey commentary". 10th ICETE 2013; SIGMAP 2013 and WINSYS 2013, pp. 151-155. ieeexplore.ieee.org/document/7184690/
- [42] Garcia, J.E., Ortega, A., Lleida, E., Lozano, T., Bernues, E., and Sanchez, D. (2009). "Audio and text synchronization for TV news subtitling based on automatic speech recognition". IEEE International Symposium on Broadband Multimedia Systems and Broadcasting. doi: 10.1109/ISBMSB.2009.5133758
- [43] Obach, M., Lehr, M., Arruti and A. (2007). "Automatic speech recognition for live TV subtitling for hearing-impaired people". *Assistive Technology Research Series*, Vol. 20, pp. 286-291. url: <http://ebooks.iospress.nl/publication/641>