



# Accessibility in 360° video players

Chris J. Hughes<sup>1</sup>  · Mario Montagud<sup>2,3</sup>

Received: 1 May 2020 / Revised: 18 August 2020 / Accepted: 15 October 2020

Published online: 22 October 2020

© The Author(s) 2020

## Abstract

Accessibility is a key requirement for any multimedia tool and application. With the current trend towards immersive experiences, such as Virtual Reality (VR) and 360° video, it becomes key that these environments are adapted to be fully accessible. However, until recently the focus has been mostly on adapting the existing techniques to fit immersive displays, rather than considering new approaches for accessibility designed specifically for these increasingly relevant media experiences.

This paper surveys a wide range of 360° video players and examines the features they include for dealing with accessibility, such as Subtitles, Audio Description, Sign Language, User Interfaces and other interaction features, like voice control and support for multi-screen scenarios. These features have been chosen based on guidelines from standardization contributions, like in the World Wide Web Consortium (W3C) and the International Communication Union (ITU), and from research contributions for making 360° video consumption experiences accessible. The in-depth analysis has been part of a research effort towards the development of a fully inclusive and accessible 360° video player. The paper concludes by discussing how the newly developed player has gone above and beyond the existing solutions and guidelines, by providing accessibility features that meet the expectations for a widely used immersive medium, like 360° video.

**Keywords** 360° video · Accessibility · Audio description · Immersive video · Sign language · Subtitling

---

✉ Chris J. Hughes  
c.j.hughes@salford.ac.uk

Mario Montagud  
mario.montagud@i2cat.net

<sup>1</sup> School of Science, Engineering & Environment, University of Salford, Salford, UK

<sup>2</sup> i2CAT Foundation, Barcelona, Spain

<sup>3</sup> University of Valencia, Valencia, Spain

# 1 Introduction

The recent explosion of immersive media technologies, like Virtual Reality (VR) and 360° video, opens the door to new fascinating opportunities and revenue models, not only in the entertainment sector, but also in other key sectors of society, like education and culture [24]. In this context, 360° videos have become a simple and cheap, yet effective and hyper-realistic, medium to provide VR experiences. In 360° videos, a view in every direction is recorded using an omnidirectional camera or a camera rig that captures overlapping angles simultaneously from a fixed point. The multiple views are then stitched together into a single, high resolution and seamless panoramic video. The users can then freely explore the omnidirectional space surrounding the capturing point, but at any moment they can only watch a portion of the sphere, mainly determined by the Field-of-View (FoV) of the display being used. This free 360° exploration cannot be augmented with a free navigation around 3D spaces (commonly known as 6 Degrees of Freedom, 6DoF), unlike in 3D VR environments, because the 360° videos are in fact spherical projections of flat images. However, 360° videos have the potential of providing immersive experiences, as well as a high degree of realism with less effort than using virtual modeling and reconstruction techniques in 3D VR, as real scenarios and characters can be directly captured with a 360° camera.

Due to their potential, the scientific community and industry have devoted significant efforts in the last years to providing improved solutions in terms of many relevant aspects, like capturing and consumption hardware, compression and streaming strategies, and display methods [9, 39]. Likewise, the demand for production and consumption of 360° videos has significantly increased in the last few years. The study in [24] reports on the impact in terms of content production, subscribers and consumption rates being achieved by many relevant media companies, institutions and online platforms that currently include 360° videos in their service offerings. Examples include: New York Time, BBC, Facebook and YouTube. As proof of relevance, more than 400.000360° videos were uploaded to YouTube in 2019 [12]. In general terms, the global market size of 360° films production is expected to be around 4 billion dollars in 2020, according to [11]. The interest in 360° videos is also being reflected in the broadcast sector, as confirmed by a report issued by the European Broadcasting Union (EBU) [7] that indicates that 49% of its members are exploring and devoting efforts toward the integration of immersive video services. Examples of broadcasters offering 360° videos include, among others: BBC (UK), BT (UK), RTVE (Spain), CCMA (Spain) and RBB (Germany).

The high potential of, and interest for, 360° videos has also led to the development of a wide range of 360° players for a variety of platforms and consumption devices [31, 33], like desktop computers, smartphones and Head Mounted Displays (HMDs). As for every (media) service, 360° consumption experiences need to be accessible. Typically, accessibility has been considered in the media sector as an afterthought, and mainly for mainstreamed services. The community is fully aware of the relevance of accessibility and of the existing regulation frameworks to guarantee accessible services. However, it seems that accessibility for immersive media services is still in its infancy. Likewise, the lack of standardized solutions and guidelines has catered for the development of own non-unified solutions, meeting just specific requirements.

This situation has served as a motivation to conduct the research study presented in this work, which is a *pioneering in-deep exploration and categorization of how and to what extent accessibility services are integrated in the key existing 360° players in the media landscape*. This main contribution of the paper is an initial, but necessary step, towards identifying what

are the advances and limitations in this field, complementing surveys focusing on other technological aspects for 360° video, like [9, 39]. Based on the insights from the conducted survey and analysis, a second contribution of the paper is the *description of the iterative process towards a user-centric design and development of a new 360° player* [23, 25] that aims at filling the existing gaps, providing tested accessibility features based on users' needs and preferences. That new accessibility-enabled 360° player is called the *ImAc player*, as it has been developed under the umbrella of the EU H2020 Immersive Accessibility (ImAc) project,<sup>1</sup> by combining the expertise of a cross-disciplinary team.

The structure of the paper is as follows. The rest of Section 1 indicates who the target audience of the paper is (Section 1.1), presents the accessibility services, features and guidelines taken into account in this study, and their target users (Section 1.2), and details the goal and followed methodology in this work (Section 1.3). Next, the key existing web based and executable 360° players are reviewed in Section 2, based on the criteria and guidelines introduced in Section 1.2. Section 3 provides a review of the existing approaches to interact with the access services, with the goal of reflecting the need for more uniform and appropriate solutions in this context. After having reviewed the solutions offered by existing 360° players, the ImAc player is presented in Section 4, by describing its accessibility and personalization features, providing solutions to the identified limitations and gaps. All surveyed players, including the ImAc player, are then categorized in Section 5 to provide a whole picture of to what extent accessibility features are provided by existing players, discussing the existing limitations and comparing them to the improvements and solutions provided by the ImAc player, as well as outlining some conclusions. Finally, some ideas for future work are provided in Section 6.

## 1.1 Focus and target audience

Everyone has the right to access and comprehend the information provided by media services in general, and by immersive 360° experiences in particular. Therefore, accessibility becomes a priority, not only to adhere to current worldwide regulations, but also to reach a wider audience and contribute to equal opportunities and global e-inclusion. Immersive experiences need to be fully inclusive across different languages, addressing not only the needs of consumers with hearing and vision impairments, but also of those with cognitive and/or learning difficulties, low literacy, newcomers, and the aged.

Among the requirements to enable universal access to immersive 360° content, this paper focuses on solutions to achieve an efficient integration of access services (like subtitling, audio description and sign language) with 360° content, and on appropriate User Interfaces (UIs) to mostly enable an effective interaction and usage of these services.

The studies in [24, 26] provide statistics about the percentages of the worldwide population with some form of audio-visual impairments, and about the increasing ageing process [8], which is also strongly related to accessibility needs. In addition, these studies reflect on the societal impact of media accessibility and review the existing regulatory framework to ensure equal participation in the current society where technologies and media play a key role.

Therefore, the current situation with regard to accessibility needs, in combination with the associated regulatory framework, are raising awareness and putting pressure on content producers and providers for fulfilling the missing requirements. On the one

---

<sup>1</sup> <https://www.imac-project.eu/> Last Access in August 2020.

hand, the presented study and categorization serve to motivate the need for further advancing on the accessibility field for immersive media. On the other hand, the contributions of the paper are meant to become a valuable resource for the interested audience in this field, including: 1) users with accessibility needs (e.g., deaf and hard-of-hearing users, vision impaired and blind users, users with cognitive impairments, etc.) in order to select the player that best fits their needs; 2) the development community and service/content providers, in order to improve their solutions; and 3) the standardization bodies and research community, in order to have an overall view of what is solved and what is missing, and/or determine to what extent the existing guidelines are met.

## 1.2 Accessibility guidelines

The World Wide Web Consortium (W3C) provides detailed guidelines for producing content for Web environments, such as the *Web Content Accessibility Guidelines* (WCAG) [36], and the presented study assesses the level of conformity of existing 360° video players with these guidelines. In particular, the WCAG defines four key services that are required in order for a 360° video to be regarded as accessible:

- A transcript, which is a written version of the spoken audio, and is required in order to provide the most basic level of accessibility.
- Subtitles (ST), which are essentially the transcript, broken into small sections (usually ~30 characters, 2 lines), synchronized with the video and/or audio. Although not an essential requirement, ST are beneficial for people who are deaf, hard of hearing or non-native speakers. By using ST, the text can be read whilst also viewing facial expressions, body language and actions which are used to indicate the context and intent of the specific characters who are speaking. Usually, different colours are used to identify each speaker, and special characters are often used to identify sounds or actions. ST are particularly important when consuming immersive content, as reading a separate transcript document would break the immersion on a desktop computer, and even be impossible when wearing a HMD.
- Audio Description (AD), which is an additional audio track played over the top of the main audio, and typically produced by a skilled audio describer to fit a description into the gaps in the main dialogue/action. The AD must provide a description for any relevant visuals which have not yet been discussed in the dialogue. The AD must describe what is being visually represented in order to address the needs of those who are partially sighted or blind. Technically, AD is commonly mixed with the main audio track/stream, which is not ideal for personalisation (a key aspect for accessibility).
- Sign Language (SL), which uses the movement of the hands, facial expression and body language to convey meaning. Sign languages are full-fledged natural languages with their own grammar and lexicon and, as such, they are not universal and mutually intelligible with each other. Often, a signer is overlaid onto the video (Picture-in-Picture) in order to translate the dialog into a sign language. SL is not directly required for accessibility, although translation and interpretation is recommended for people who are deaf, and especially for whom sign language is their primary language. Technically, SL is commonly burned with the main video track/stream, which is not ideal for personalisation (a key aspect for accessibility).

These WCAG guidelines have become a widespread and useful measure of accessibility in a web context. In the broadcast sector, other more relaxed guidelines have been provided, e.g. by International Telecommunication Union (ITU) [16]. In addition, an overview of contributions and initiatives towards standardizing accessibility guidelines can be found in [19].

The comparison aspects considered in this study are then based on the accessibility guidelines previously identified, which mostly rely on the support for the traditional access services: ST, AD and SL. In addition, two further key aspects for accessibility are examined:

- support for Accessible UIs, and voice control interfaces, which become key for an effective usage of these services.
- support for multi-screen scenarios, which allow for richer personalisation (e.g. presentation of the required access service on a companion device, with the proper settings) and overcoming distance barriers to the main screen [27].

These above aspects can provide relevant benefits, especially for partially sighted and blind users.

### 1.3 Goal and methodology

This research study aims at shedding some light on the existence, lack and/or limitations with regard to accessibility solutions for 360° media content consumption. After defining the target audience of the paper (Section 1.1) and selecting the accessibility guidelines and aspects to take into account (Section 1.2), a wide sample of existing 360° players have been surveyed, analyzed and categorized. The goal of the study is not to list all existing 360° players, but at least to compile and analyze a wide sample of them, including the key ones. The selection has been made by thoroughly searching for the existing 360° players, by analyzing the state of the art and existing online 360° video services, and by asking experts in the field, like content providers, broadcasters and researchers.

Then, a set of user-centric activities have been conducted to gather accurate requirements to be met by the ImAc player, and to evaluate/refine/validate them via user-testing. These activities will be briefly introduced in Section 4.

## 2 Overview of 360° video players

In this section, the key web-based and executable 360° video players are reviewed, and their approach to accessibility is discussed. The selection of the players has been based on the aspects identified in Section 1.2, but also on their support in both desktop mode (i.e. players run on desktops, laptops, smartphones, tablets, and other similar screen-based devices) and VR mode (players run on a VR display, like an HMD). The consumption of 360° videos in both types of devices is quite frequent, and thus the review of both modes becomes convenient. The section starts by reviewing web-based players in Section 2.1, which are supposed to provide cross-platform, cross-device, and even cross-browser support, and then continues by reviewing executable players in Section 2.2, by detailing the supported platforms. A qualitative study and discussion about the pros and cons of executable vs web-based applications can be found in [22]. The review is then categorized in Table 1, as part of Section 5, where a general taxonomy and discussion is provided.

**Table 1** Summary of directly supported accessibility features in key 360° video players

		Subtitle	Audio description	Sign language	Accessible UI	Voice control	Multi-screen	VR mode
Web based	JWPlayer	✓	✓					✓
	Omnivirt					✓	✓	✓
	YouTube	✓	✓					✓
	Facebook	✓						✓
	TheoPlayer	✓						✓
	Radiant Media Player	✓			✓			✓
	Google VR	✓				✓		✓
Executable	Vimeo	✓						✓
	ImAc Player	✓	✓	✓	✓	✓	✓	✓
	5KPlayer	✓				✓		
	VLC 360	✓					✓	
	GOM Player	✓				✓		
	Skybox VR	✓						✓
	Simple VR	✓						✓
	Deo VR	✓						✓
	Bitmovin	✓						
	Exoplayer	✓						
	RiftMax VR	✓						✓
	VR Player	✓				✓		✓
	LiveViewRift VR							✓
	GoPro VR Player							
	Marzipano							

<sup>1</sup> Apart from an adaptive and responsive interface, the ImAc player UI provides multi-language support and assistive methods, like visual feedback and magnification features while interacting with the menu

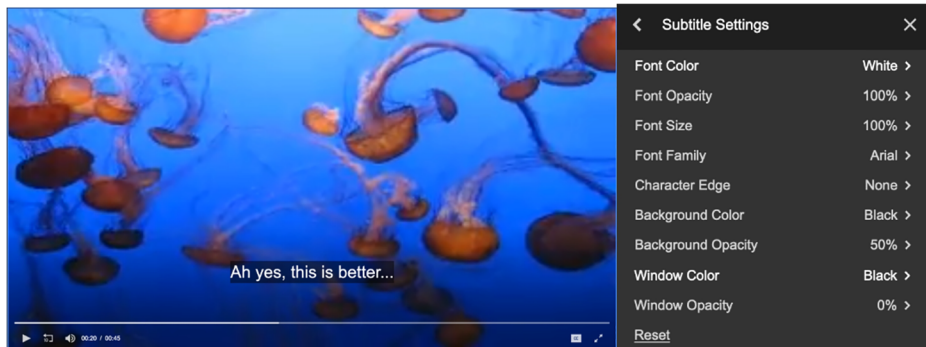
## 2.1 Web based 360° players

**JWPlayer**<sup>2</sup> provides an Application Programming Interface (API) specifically targeted at developers and designers who are building their own apps and websites for delivering video. The API is very extensive and provides either a commercial version integrated with a Content Delivery Network (CDN) or a free to use version for self-hosting. In particular, the API enables the setup of a 360° player, and its embedding in a web context. The UI of **JWPlayer** (see Fig. 1, left) is quite basic and standard, and a series of keyboard shortcuts are additionally provided.

The API is very limited in what it provides out of the box for accessibility. However, it does provide an extensible framework, and it is also possible to adapt the UI using JavaScript.

In terms of accessibility, **JWPlayer** supports three subtitle formats: Web Video Text Tracks (WebVTT) [38], SubRip Subtitle (SRT) [32] and Timed Text Markup Language (TTML) [37]. It also supports captions which are embedded in HTTP streams, like Dynamic Adaptive Streaming over HTTP (DASH) and HTTP Live Streaming (HLS). It provides the user with basic customisation controls for the rendering of the subtitles, but no control over position (see Fig. 1, right).

<sup>2</sup> [www.jwplayer.com](http://www.jwplayer.com) - All links provided in the article accessed for the last time in August 2020



**Fig. 1** JWPlayer: default interface (left), Subtitle Options (right)

Likewise, **JWPlayer** can support multiple audio tracks. Although there is technically no specific AD support, the support of multiple audio tracks can be used as a solution for delivering AD, by having an additional audio track which combines the main audio with AD. There is currently no support for SL, as it is generally accepted that the signer would be burned into the video, thus providing no option for personalisation.

In addition, it is worth mentioning that there are open source projects for adding voice control support to **JWPlayer**, such as the integration of Siri on iOS devices [17].

**Omnivirt**<sup>3</sup> was primarily developed for commercially displaying advertising banners, although the company developing it branched into 360° video, by developing a standalone 360° video player. The **Omnivirt** player is free to use with some limitations (10 k monthly views, 2GB files), after which users are required to upgrade to a premium version.

The UI of the **Omnivirt** player is very limited (see Fig. 2). However, it provides an interesting feature for 360° video consumption, which is a radar display indicating the current viewing position relative to the centre of the omnidirectional scene. This feature allows the user to re-centre the view by clicking on the radar icon. The **Omnivirt** player does not provide support for any of the access services considered in the study.

**YouTube**<sup>4</sup> is one of the largest players in the video hosting and delivery sector, which also moved into the 360° video space. YouTube allows users to upload their own content which is then shared through its network. Although it is free to use, revenue is made by placing adverts on the videos.

The **YouTube** player provides good support for ST, and although there is no option to customize the position of the ST, it is possible to control the font colour and style as well as the background (see Fig. 3, left). **YouTube** provides a service which attempts to automatically use speech recognition to generate ST for content. Although such a service produces nice results, they may not be good enough for the WCAG, which insists on professionally authored accessibility content. In the 360° video mode, the ST are fixed into the user's view (see Fig. 4, left).

Natively, the **YouTube** player has no support for AD or SL. It is simply assumed that a user would provide a specific version of the video with a burned in signer video or a specific AD soundtrack. There are however additional projects, such as YouDescribe [35], which allow for YouTube videos to be played through a third-party website, whose users collectively provide

<sup>3</sup> [www.omnivirt.com](http://www.omnivirt.com)

<sup>4</sup> [www.youtube.com](http://www.youtube.com)



Fig. 2 UI of the Omnivirt player (left), with a cropped view of the radar display showing the users position within the scene (right)

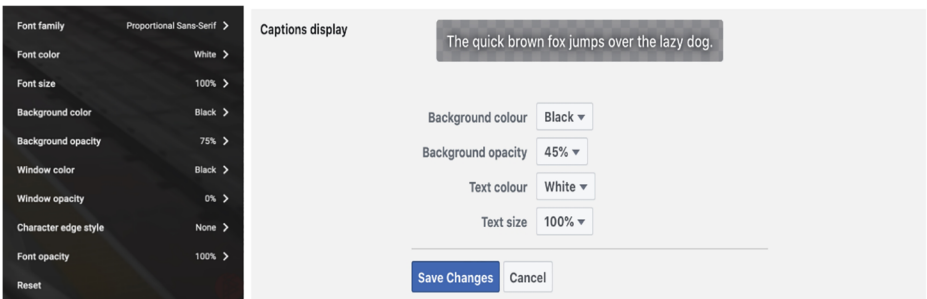


Fig. 3 Customizing the subtitle display in YouTube (left), Facebook (right)



Fig. 4 UIs of the YouTube (Left) and Facebook (Right) 360° players



Fig. 5 THEOplayer (left), Radiant Player (right)



their own AD for the videos. In this case, the YouTube video is paused at any point that a contributor has specified, and then a text-to-speech engine is used to read the description before resuming the video.

**Facebook**<sup>5</sup> is a large social media platform with over 2 billion active users and is primarily focused on sharing personal videos and photos. Facebook additionally integrates a 360° player (see Fig. 4, right), which does provide limited support for ST, by either uploading an SRT file or using their online tool (which includes an auto generate function based on voice recognition). There is however no support for AD and SL. The player also allows setting some basic style preferences for ST, such as font colour and size (see Fig. 3, right). Likewise, Facebook takes advantage of users having to be signed in and tracked, which enables the platform to remember the user's preferences and also allows the user to specify that ST should be displayed by default.

**THEOplayer**<sup>6</sup> is another 360° player, which consists of a growing portfolio of feature-rich Software Development Kits (SDKs) with wide video ecosystem pre-integration. It provides basic support for ST, including WebVTT and TTML file formats, and the presentation mode consists of fixing the ST in the user's view (see Fig. 5, left). The UI shares elements with other players, such as the controls at the bottom of the screen, with the play / pause button on the left and the display ST button on the right. However, the latter button is only displayed if ST are available, which can leave users searching for this control when it is not displayed. The user has a similar control of the subtitle rendering than in the **JWPlayer** and **YouTube** players, and has the ability to change the rendering style, but not the position.

Although not directly supporting AD, **THEOplayer** supports multiple audio and language tracks for one single video – both for live and on demand streaming. The support for multiple audio streams can be then used to provide multiple language tracks for a specific video clip, but also to provide specific AD tracks. However, switching between audio tracks is needed, which is completely opposed to having a specific AD track in addition to the main audio.

**Radiant Media player**<sup>7</sup> is a commercial video player which is extensive in its implementation. It provides a quite elaborated UI (see Fig. 5, right), including controls specifically designed for 360° video, allowing reorientation and zooming through the interface. It also provides a mature support for ST (including WebVTT, WebVTT and TTML formats), but the standard implementation contains no method for the user to customise the display of the ST.

**GoogleVR**<sup>8</sup> is in fact not a 360° player, but an API for developers to create 360° video experiences, by Google. One of the main features of this API is the ability to use it in order to directly embed 360° videos onto a web page, which is done via a JavaScript application that creates and controls the contents of an HTML *iframe* element, or by explicitly declaring the *iframe* itself. Although the API supports an extensive selection of file formats, there is no consideration for ST, AD or SL (see Fig. 6, left). In addition, it by default only provides a minimal UI for playing video. However, the general design behind it can be extended to create more advanced UIs, using JavaScript and HTML.

**Vimeo**<sup>9</sup> provides similar features than YouTube, although it is more targeted for professional content creators who want to demonstrate their content in high quality. Relevant differences between YouTube and Vimeo are that the Vimeo allows playback of content that

<sup>5</sup> [www.facebook.com](http://www.facebook.com)

<sup>6</sup> [www.theoplayer.com](http://www.theoplayer.com)

<sup>7</sup> [www.radiantmediaplayer.com](http://www.radiantmediaplayer.com)

<sup>8</sup> [developers.google.com/vr/discover/360-degree-media](http://developers.google.com/vr/discover/360-degree-media)

<sup>9</sup> [www.vimeo.com](http://www.vimeo.com)



**Fig. 6** GoogleVR (left), Vimeo (right)

is not compressed, and it also provides professional delivery such as faster load times. This option is great for videographers and content producers who want to show off their work in the highest possible quality. Like the **YouTube** player, the **Vimeo** player requires videos to be equirectangular and it supports resolutions up to 8 K. The **Vimeo** player provides a radar display to help the user orientate themselves with the video. Vimeo provides basic support for traditional ST, although there is no option to customize their display. The ST also contains no spatial information and so the radar does not provide an indication of the direction of the action (see Fig. 6, right). It offers no support for AD or SL.

## 2.2 Executable 360° video players

In this subsection, the key 360° video players developed to run as native applications for specific platforms are reviewed with regard to their accessibility features, being categorized into three main groups based on their primary target platforms: desktop computers, HMDs, and players provided as part of an API.

### 2.2.1 Desktop computer

360° video players like 5kPlayer,<sup>10</sup> VLC<sup>11</sup> and GOM Player<sup>12</sup> (see Figs. 7 and 8) are designed for running on a desktop computer. By using them, the 360° video can be viewed in either a window or a full screen, and all interactions with the environment are done with the mouse and keyboard, such as clicking and dragging the video to move the users' FoV. These players generally come with support for a wide range of file formats (such as .MPG, .MP4, .WEBM, .AVI, .MOV, .OGG) and share a relatively common UI. Likewise, these players typically come from extensions to existing 2D video players for adding the ability to render 360° videos, and therefore simply take advantage of the features already provided.

Many of these players are free to use, although **GOM Player** generates revenue by the placement of adverts into the UI, which can be removed by the purchase of a commercial version. **VLC** is the only player that is fully open source, with the software being developed as a community project.<sup>13</sup> These players typically can be run on a wide range of platforms. For instance, **5kPlayer** and **VLC** player run on Windows, Mac and Linux. **GOM Player**,

<sup>10</sup> [www.5kplayer.com](http://www.5kplayer.com)

<sup>11</sup> [www.videolan.org](http://www.videolan.org)

<sup>12</sup> [www.gomlab.com/gomplayer-media-player](http://www.gomlab.com/gomplayer-media-player)

<sup>13</sup> <https://github.com/videolan/vlc>

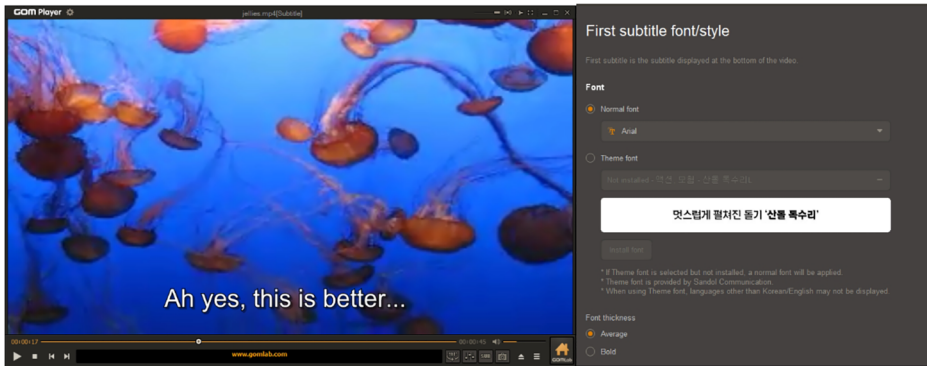


Fig. 7 GOM Player: default interface (left), Subtitle options (right)

however, is only available for Windows. In addition, **VLC** provides an additional product, ‘*VLC\_VR*’, which provides similar functionalities for Android, IOS and Xbox.

All of these players provide some support for ST. Although only **VLC** player allows for the ST position to be moved, they do all allow for the font and style to be changed. **GOM Player** is unique by offering the ability to load two separate subtitle files, with one displayed at the top and the other at the bottom of the screen. However, this support is directly provided by their “predecessors” 2D video players. All of these players allow loading a ST file and rendering it as a 2D overlay onto the video window. As there is no standard ST file specifically designed for 360° videos, there is no special information about where the subtitle relates within the omnidirectional scene. **GOM Player** and **VLC** player offer some basic customisation options for the presentation of subtitles, such as style and position.

These players additionally provide support for selecting alternative audio tracks. These tracks can be used for alternative languages but could also be used for the AD service. However, there is no mechanism for seamlessly mixing the AD over the existing audio track in the player.

None of these players provide a mechanism for adding SL or overlaying an additional video stream as e.g. in Picture-in-Picture mode, which could be used for the SL service. Therefore, the only way to add a signer would be to burn the additional SL video into the main video and provide both videos as a single stream. However, it then causes the position of the signer to be fixed and gives no opportunity for the signer position to be fixed in the users view, or to customize its presentation, based on particular needs and/or preferences.

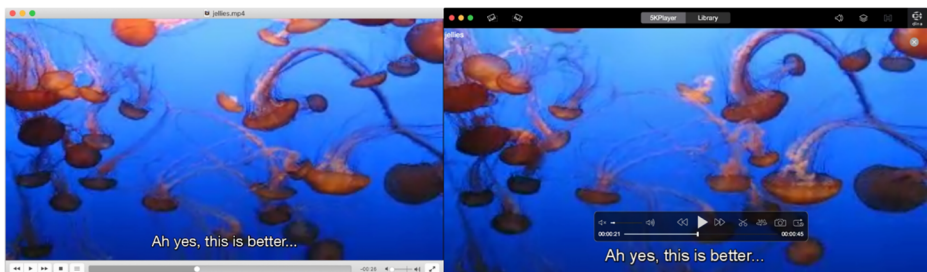


Fig. 8 VLC (left), 5 k player (right)



**Fig. 9** VLC Rendering modes: Default (top left), Zoomed (top right), Little Planet (bottom left) and Reverse Little planet (bottom right)

VLC player is by far the most advanced player in this category. It provides a number of additional rendering modes for 360° videos, such as ‘zoomed’, ‘little planet’ and ‘reverse little planet’ (see Fig. 9). It is possible to smoothly scale between these modes, giving partially sighted users further control over how much of the scene they have in view. Although it is primarily a desktop video player, it does also support rendering on connected HMD displays.

Finally, there are other existing players specifically designed for stitching and reviewing footage taken from video cameras, as it is the case of **GoPro VR Player**,<sup>14</sup> a Windows application for GoPro cameras. However, **GoPro VR Player** does not provide any accessibility functionality.

### 2.2.2 HMD

There are two types of HMD’s available for consuming immersive VR content. The first type of headsets is ‘tethered’ to a powerful desktop computer host, where the graphics are rendered and simply displayed on the HMD. This includes HMD’s like Oculus Rift, HTC Vive/Pro and PlayStation VR. The second type of headsets is standalone and has its own processing capabilities, like Oculus Go, Oculus Quest and Samsung Gear VR (embedded with smartphones). It is worth noting that each of the standalone HMD’s have a web browser provided by their Operating System (OS). This enables the usage of any of the web based 360° players previously reviewed. This is a strong advantage for the web-based players [22], as previously highlighted.

**Deo VR**<sup>15</sup> and **VR Player**<sup>16</sup> are two 360° video players designed to run on all of the commonly available HMDs (including implementations for HTC, Oculus, Android, IOS and Windows). This is achieved by having native builds for the stand-alone HMD’s and for desktop computers acting as hosts. They are both free to use and support basic support for

<sup>14</sup> [gopro.com/pt/br/news/gopro-vr-player-2-now-available](https://gopro.com/pt/br/news/gopro-vr-player-2-now-available)

<sup>15</sup> [deovr.com](https://deovr.com)

<sup>16</sup> [vrplayer.com](https://vrplayer.com)

ST and multiple audio streams. The developers for **Deo VR** player are aware that its subtitle support is very limited, and it is made clear throughout their support forums that early on, due to limited resources, subtitles were not a priority to them. **VR Player** sets itself apart from other players by integrating successfully with the controllers provided with many of the HMD's. These controllers can be moved freely in space, allowing the user to use gesture control to operate the player. The interface is also extended through both simple voice control and integration with Bluetooth controllers.

It is important to note that the existing players in this category, such as **VR Player**, were in general initially designed for projecting 2D video into the immersive space, by providing a virtual cinema with the video playing on a virtual screen. This is mainly due to HMD's being available sooner than 360° cameras and there being more 2D content available. Other players that operate in this manner include **RiftMax VR**<sup>17</sup> and **Simple VR**,<sup>18</sup> which both run under both Windows and Mac hosts with implementations designed for the HTC Vive/Vive Pro. **RiftMax VR** player is free and simulates a giant multiplayer Imax-style theater that is able to play 2D and 360° videos. Using this player, users can interact with each other and have a shared experience while watching the video. **Simple VR** player is a commercial product but focuses heavily on having an easy to use interface and also allows the use of hand gestures. All of these players are however limited in what they provide for accessibility, and none go beyond the traditional ST rendering.

**Skybox VR**<sup>19</sup> provides a player which runs on standalone headsets. One of the issues faced with the standalone HMD's is that they are fairly limited for storage. **Skybox VR** gets around this by providing a Windows or Mac file server, which streams 360° video files to the headset. It supports all of the common stand-alone HMD's, such as Oculus Go, Oculus Quest and Samsung Gear VR. **Skybox VR** supports traditional ST fixed in the user's view. Although there is no opportunity for customisation, **Skybox VR** player provides basic controls for resynchronising the subtitle timing, by delaying or advancing the timings, improving the temporal alignment if needed.

Finally, it should be noted that many of the reviewed web based 360° players, like **YouTube** and **Vimeo**, are also provided as native applications for specific platforms, like Android.

### 2.2.3 Development API's

Several API's designed specifically for building 360° players also exist. As a first example, **Bitmovin**<sup>20</sup> is a multimedia technology company that provides full-fledged solutions for content producers to make their 360° content publicly available. These solutions include cloud storage and encoding servers, analytics features, and playout solutions. **Bitmovin** provides a feature rich API for Web, Android and IOS development (see, Fig. 10), but there is currently no direct support for accessibility services.

There are other API's, such as **Exoplayer**,<sup>21</sup> which provides a direct alternative to Android's MediaPlayer API (see, Fig. 11). However, there is no direct consideration for

---

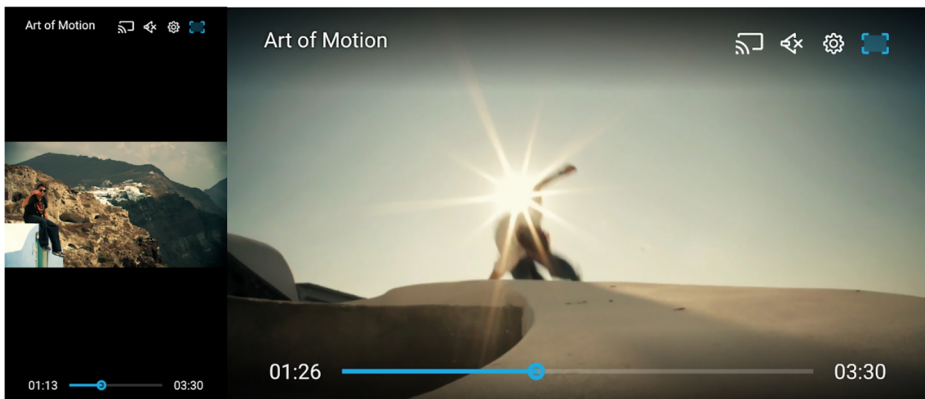
<sup>17</sup> [riftmax.com](http://riftmax.com)

<sup>18</sup> [simplevr.pro](http://simplevr.pro)

<sup>19</sup> [skybox.xyz](http://skybox.xyz)

<sup>20</sup> [bitmovin.com](http://bitmovin.com)

<sup>21</sup> [exoplayer.dev](http://exoplayer.dev)



**Fig. 10** Bitmovin Player on Android: portrait orientation (left), landscape orientation (right)

accessibility. Other notable development tools include **Marzipano**<sup>22</sup> which although not being a 360° video player, allows users to create immersive tours from 360° photos, and thus it is worth to be mentioned in this study. However, once again there is no consideration for accessibility, and certainly no opportunity to integrate AD.

### 2.3 Ad-hoc players by content providers

In the race to provide innovative and engaging 360° experiences to their audiences, content providers started developing ad-hoc players to offer their produced 360° content. Relevant examples of bespoke web-based players built by content providers are the ones by the BBC<sup>23</sup> (see Fig. 12, left), NYT<sup>24</sup> (see Fig. 12, right), and RTVE<sup>25</sup> (see Fig. 13), developed to meet their specific requirements whilst still trying themselves to understand how 360° video production could be accomplished and massively offered.

Although these solutions do not provide support for loading subtitles, the BBC, NYT and RTVE partially addressed the requirement by adding creative captions burned into the video (see Figs. 12, 13 and 14). These captions are not verbatim and give the user no control on how they are displayed. This does however also mean that the broadcaster can maintain control on the creative look of the ST and guarantee that they will be rendered as designed. It also allows for the designer to provide captions that are fully integrated with the scene in terms of position and style. However, by using this approach, if the caption is burned into the video, it could appear behind the viewer and be missed. Therefore, different broadcasters have addressed the need for navigation in different ways. For example, the BBC took an approach which replicates the ST at 120° intervals around the user to ensure it is always visible [5]. Similarly, NYT adopted a radar in order to give the user spacial awareness as to where they are looking relative to the video. RTVE additionally developed a VR app for the consumption of their produced 360° videos (see Fig. 13, right).

More recently, with the widespread adoption of 360° video services, content providers, including the mentioned ones, have commonly moved towards using major full-fledged

<sup>22</sup> [marzipano.net](http://marzipano.net)

<sup>23</sup> <https://www.bbc.com/reel/playlist/reel-world-360-videos-from-the-bbc>

<sup>24</sup> <https://www.nytimes.com/2016/11/01/nytnow/the-daily-360-videos.html>

<sup>25</sup> <https://www.rtve.es/lab/realidad-virtual/>



Fig. 11 Exoplayer on Android: portrait orientation(left), landscape orientation (right)

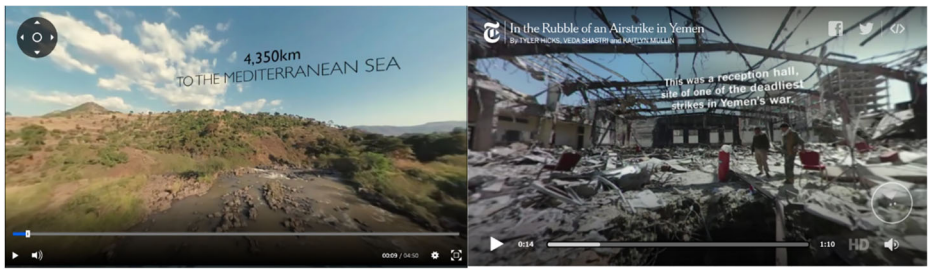


Fig. 12 BBC player (left), NYT player (right)

platforms, like YouTube, for hosting and distributing their 360° videos to their audiences. An example is provided in Fig. 14 for a branded version of the YouTube player for RTVE.

### 3 Access to access services

In order for an access service to be effective and useful, it needs to be easy to access, and interact with. This is relevant for all user profiles, but especially for users with visual impairments, as graphical UIs are the most common interaction modality. From this study, it

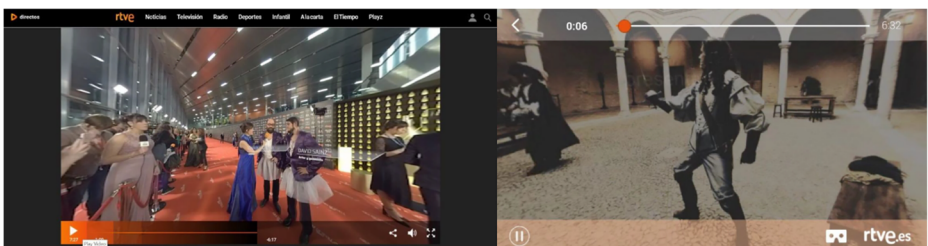


Fig. 13 RTVE player: web-based player with creative captions burned in (left), VR app (right)

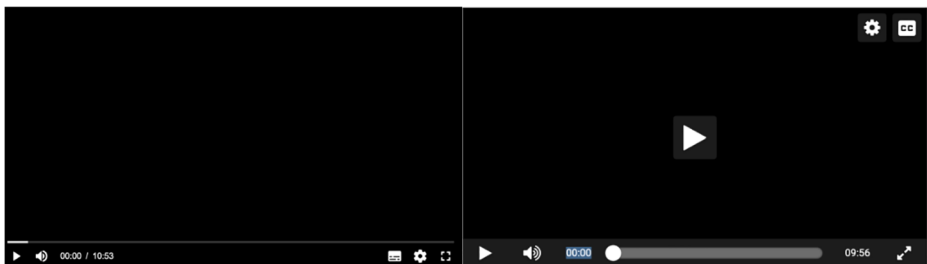


**Fig. 14** RTVE using a branded YouTube player, with textual burned in (attached to the 360° scene)

becomes clear that there is no standard and unified approach to designing the 360° player UI, including the controls for accessing the access services and for setting the available features. In addition, it is often not clear if the access services are available at all. For example, in all reviewed players, if subtitles are not provided/supported, the activation button simply does not display, causing the user to search around for the options. These findings are inline with the ones in [21], focused on reviewing traditional 2D players.

Generally, as shown in Fig. 15, the controls are positioned at the bottom left of the video window. However, there are no standards to define this, so developers typically position the controls anywhere they choose. For example, the menu and service controls are positioned at the top right of the screen in the Radiant Media Player (Fig. 15, left).

The positioning of the controls at the bottom or top is generally a good practice in 2D players to provide a cleaner interface for video viewing, minimizing blocking issues caused by the UI. These control positions are typically mapped into the VR mode of these players, replicating the same locations. However, this is not an appropriate approach, especially when using HMDs. This is due to the fact that HMDs typically have a reduced FoV, which is determined by their specific hardware and can range from 90°-110° depending on the specific brand and model and placing visual elements around the edges of the FoV leads to uncomfortable viewing experiences [23]. Although personalisation of the UI layout could help to overcome this issue, this is not provided in the reviewed players.



**Fig. 15** Typically, the controls to open the access services are located on the bottom right, such as in THEOplayer (left), with exceptions such as Radiant Player which positions the controls top right (right)



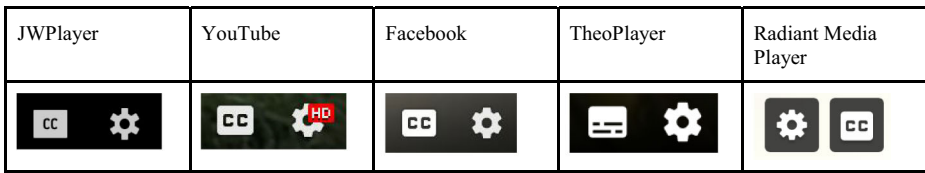


Fig. 16 Examples of the button designed for opening subtitles

With regard to UI elements, there is no standardized solution, although web-based players follow similar approaches. As shown in Fig. 16, different symbols have been chosen for the ST service, which is made even more confusing by the common choice of a ‘CC’, taken from the US ‘Closed Caption’ description. The use of the service acronym is sometimes also used for this, which besides varies depending on the active UI language. In addition, while this is already an issue for ST, it becomes more evident and diverse for the less developed access services, like AD and SL, which although being very scarcely supported in 360° players, their support is a bit more frequent in traditional 2D players [21].

The executable 360° players make it even more difficult to access the ST service. For example, in a desktop environment, both **5 K Player** and **VLC** player require the user to select subtitles through the menu, as shown in Fig. 17. Although **VLC** player does offer a keyboard shortcut, this key varies between OS, and can be hard to identify if the correct subtitle track has been selected.

Most of the executable players also replicate the desktop UI in the HMD mode. As shown in Fig. 18, the interface for selecting ST on the **Skybox VR** player is very similar to those found in the other surveyed desktop applications. However, it is significantly more difficult to control a menu within an immersive HMD environment, making it even more challenging to find the access services controls. Other players, like **Deo VR**, offer no control for ST. Using **Deo VR** player, the ST are loaded and displayed by default, if available. Although this offers a direct solution to users requiring ST, it could be annoying to users who prefer to not use them, and no presentation control is provided.

Another relevant aspect in this context is to minimize the number of clicks/interactions required to (de-)activate the access services and to set the available personalisation options. Whilst typically 2D players provide quite efficient solutions for these purposes [21], these issues do not yet apply to 360° players, due to their scarce support for access services presentation and personalisation. However, they need to be taken into account when addressing these gaps/requirements.

Due to the diversity in terms of UI elements and icons for the access services [21], DR Design has proposed four universal icons to identify access services (Fig. 19), to be adopted as

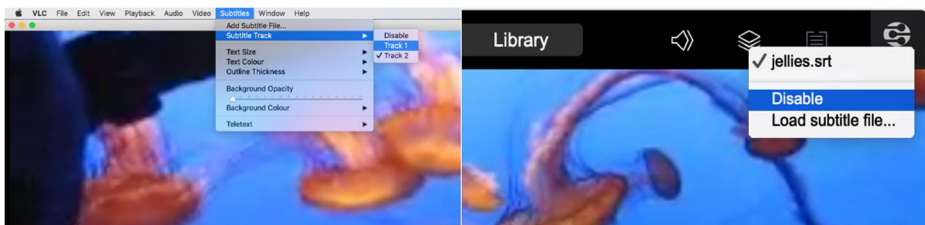


Fig. 17 Selecting subtitles on a desktop 5 K Player (left) and VLC player (right)

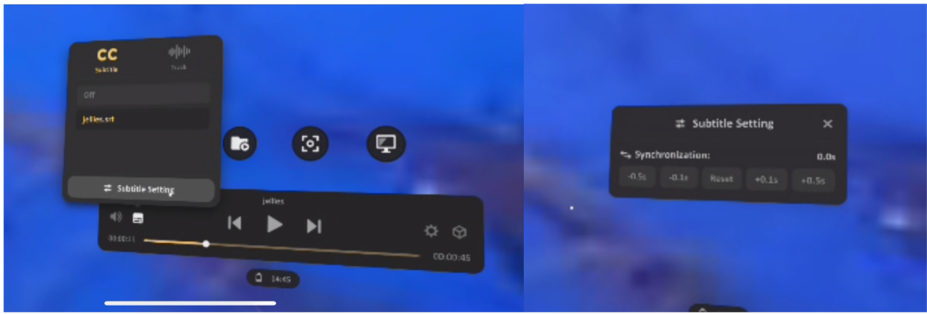


Fig. 18 Selecting and controlling subtitles on SkyboxVR in HMD mode

a standardized solution for player UIs [6]. By standardizing, the icons would not only become easily recognisable, but also independent of any specific language, preventing non-native language speakers from using the services. The Danish Broadcast Corporation (Fig. 20) has already started adopting the icons for their content, and the European Broadcasting Union (EBU) through its Access Services Group of Experts is supporting this initiative for standardized universal icons. Since the icons are text-based, they can be typed on a keyboard, included in descriptive metadata and read out aloud (e.g. using a screen-reader). The ultimate goal is to increase simplicity, quality and usability through standardization.

Finally, having a traditional graphical menu may not be appropriate at all for visually impaired users. The use of assistive methods in this context can contribute to a better accessibility. Examples are visual feedback when navigating over the menu options, when setting the desired features/options, as well as magnification features (e.g. an enlarged version of the menu, or magnifying visual elements once having the focus). Beyond some basic visual feedback features, the available 360° players do not fully address these important issues.

Due to the possible limitations with regard to visual interactions, the players need to adequately inter-operate with existing screen-readers. This is commonly not an issue for web-based and executable 360° players when running on non-VR mode, thanks to the use of standardized metadata that are typically recognised successfully by screen-readers. However, this is an issue in the VR context, requiring an ad-hoc development of bridges between the screen-reader and the VR engine, or of specific screen-reader-like features. Similarly, voice control becomes a very useful interaction modality. Despite initial efforts towards the integration of this feature in existing players (not just at the platform level), like for the **JWPlayer** [17] and **VR Player**, a full integration is unfortunately not yet commonplace in existing players.

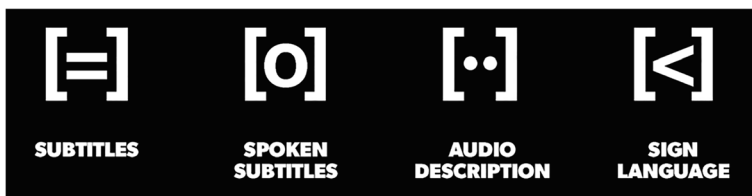


Fig. 19 A standard approach to identifying accessibility services, proposed by DR Design [6]



Fig. 20 Standardized symbols used by the Danish Broadcasting Corporation (2D player)

## 4 ImAc player

Based on the gathered insights and identified limitations in the conducted survey (Sections 2 and 3), the ImAc project members developed a modular end-to-end toolset to allow the integration of immersive and accessibility content in current broadcast-related services [31], including the necessary components from media authoring to media consumption. A key component of the ImAc platform is the open-source web-based 360° player, which enables an interactive and hyper-personalized presentation of access services for 360° content, combined with a set of assistive technologies. The specific details about the ImAc player are provided in [23, 24], but they are also briefly reviewed in this section for the sake of comparison with the other surveyed players.

In particular, this Section provides details about the adopted user-centric methodology for the development of the ImAc player (Section 4.1), identifies its related technological aspects (Section 4.2), describes the designed accessible and responsive UI (Section 4.3), and finally details all presentation options for access services and assistive methods developed in the ImAc player (Section 4.4). Therefore, this section covers all accessibility services and guidelines identified in Section 1.2, as reviewed for all other players in Section 3, but paying more attention to those novel features to address identified limitations and gaps.

### 4.1 Research methodology

The integration of accessibility solutions into new technologies from the start contributes to a more effective deployment and exploitation. Based on this premise, the process towards the design and development of the ImAc player has been built on three key pillars: 1) requirements gathering, 2) development and integration; and 3) validation and dissemination. In such a process, a *user-centric methodology* has been followed to accurately gather the accessibility, interaction and personalization requirements. This in turn requires the involvement of end-users, professionals and stakeholders at every stage of the project, through the organization of workshops, focus groups, tests, and the attendance to events, thus closely adopting the “*Design for users with users*” motto.

In particular, the design and development process of the ImAc player have undergone two iterative cycles, including the following key activities:

- **Focus groups and interviews** with end-users to derive users' needs and preferences in terms of access service presentation modes, personalisation features, and interaction modalities. Examples are focus groups conducted for ST [1] and AD [10]. Further details can be found in [20].
- Novel **technological contributions** to implement these required features. Published results in this context include e.g. [13, 23, 24].
- **Subjective testing with different profiles of end-users** (e.g. deaf and hard of hearing, partially sighted and blind, elderly, etc.) to validate and refine the technological contributions and accessibility features with professionally produced content, by measuring key aspects like usability, immersion, preferences, Quality of Experience (QoE). Published results in this context include e.g. [2, 3, 29]. Further details can be found in [18].

## 4.2 Technology

In addition to the user-centric methodology, another important premise has been adopted in the development of the player: to guarantee backward-compliance with current formats, technologies, infrastructures and practices in the broadcast/media ecosystem. This is key to maximise re-usability, interoperability and the changes of successful deployment and exploitation.

In particular, the technology developed to allow the signaling of access services and their appropriate and personalised presentation in 360° environments is based on extending the MPEG DASH<sup>26</sup> and W3C Internet Media Subtitles and Captions (IMSC) subtitles<sup>27</sup> standards. On the one hand, MPEG DASH has become the dominant technology for adaptive media delivery. On the other hand, IMSC is a TTML profile drawing the most attention of standardization bodies and industry agents in the last few years, being standardized by W3C, and being compliant with relevant services, like Digital Video Broadcasting (DVB), Hybrid Broadcast Broadband TV (HbbTV) and Netflix [14]. The specific extensions to IMSC to accommodate the subtitling features for 360° video are detailed in [34]. Despite the selection of these standards, it should be noted that the same or similar extensions could be also adopted for similar solutions, like HLS for media delivery, and WebVTT or other TTML variants for subtitles formats. The work in [34] provides guidelines to achieve this.

Likewise, the player has been built by making use of HTML5 and JavaScript, and by adopting widespread web components and APIs, such as: *dash.js*<sup>28</sup> (the reference player for DASH), *three.js*,<sup>29</sup> WebXR,<sup>30</sup> and IMSC rendering libraries.<sup>31</sup> In addition, the player includes the necessary technological solutions to enable multi-screen scenarios in a synchronized and interactive manner, in both fully web-based and Hybrid Broadcast Broadband TV (HbbTV) [15] scenarios, as described in [24].

<sup>26</sup> <https://mpeg.chiariglione.org/standards/mpeg-dash/>

<sup>27</sup> <https://www.w3.org/TR/ttml-imsc1.0.1/>

<sup>28</sup> <https://github.com/Dash-Industry-Forum/dash.js/wiki>

<sup>29</sup> <https://threejs.org/>

<sup>30</sup> <https://www.w3.org/TR/webxr/>

<sup>31</sup> <https://github.com/IRT-Open-Source>

The use of web technologies and components guarantees cross-device, cross-platform, and even cross-browser support [22], which means that the player can be effectively on traditional consumer devices (e.g. Connected TVs, PCs, laptops, tablets and smartphones) and on VR devices (e.g. HMDs).

Demo video showcasing the ImAc player features can be watched at: <https://bit.ly/2Wqd336> and [tiny.cc/imac3](http://tiny.cc/imac3) Its current version and a wide sample of 360° videos with access services can be accessed via this URL: <http://imac.i2cat.net/player/> Finally, the source code can be downloaded from: <https://github.com/Fundacio-i2CAT/ImAc>

### 4.3 ImAc player UI

The ImAc player includes a landing page, i.e. portal, for the selection of the available 360° content, and for initially personalizing the media experience (see Fig. 21).

Once a video is selected for playout, the player menu is run. The player menu (see Fig. 22, left) can be opened by a single click, looking down for a period, or via voice control. The menu has been designed to be adapted for all range of potential consumption devices (TVs, desktop computers, smartphones, HMD's...), taking special care on the limited FoV in smartphones and HMDs in order to provide a comfortable viewing experience. This issue is one of the key lessons learned during the first round of tests [18], in which a menu expanding all over the FoV was designed (see Fig. 22, right). With regard to the access to the access services, the menu has adopted the previously introduced universal icons (see Fig. 22, left), unlike its first version that added acronyms that depended on the active UI language (see Fig. 22, right) [23]. The menu also provides visual feedback when interacting with it (see Fig. 23) and to indicate the current settings (see e.g. the activated ST service, the menu section with the focus, the magnification features and the menu item with the focus in the screenshots in Fig. 23). In addition, that magnifier control, which is the element at the top left of the menu, allows opening an enlarged version of the menu (Fig. 24, left), which is more suited for users with sight loss, and also when using small screens, like in smartphones. This enhanced-accessibility variant of the menu has also significantly evolved since its first version presented in [23] (Fig. 24, right). For blind users, the player includes a voice control feature, by having developed a gateway that communicates with Amazon Echo (i.e. Alexa). Conducted user tests have proven the satisfactory usability of the UI when using smartphones and HMDs, when interacting with the menu options, especially for the ST and AD services, and when using the voice control feature [18].

Finally, as the Omnivirt and Facebook players which include a radar to indicate where the centre of the scene is, the ImAc player also includes guiding methods to assist the users in not

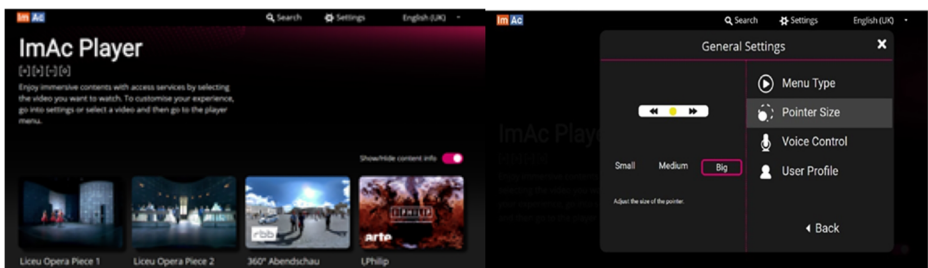


Fig. 21 The ImAc portal UI

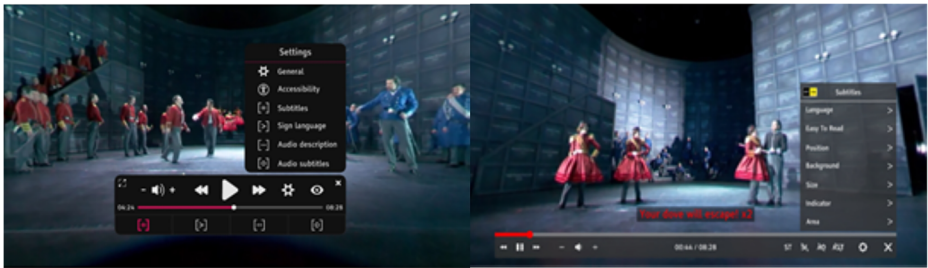


Fig. 22 The ImAc player UI (latest version on the left; first version on the right)

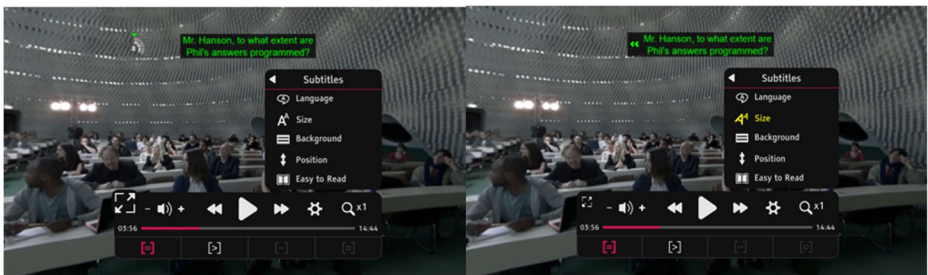


Fig. 23 Visual Feedback when interacting with ImAc player UI

losing track of the main action, additionally indicating where the active speaker is in the 360° space. This is provided by means of a radar or arrows, as preferred by the users (see Fig. 23).

#### 4.4 Presentation of access services in the ImAc player

**Subtitles (ST)** Unlike the existing 360° players that mainly present ST fixed in the user view, the ImAc player allows different presentation modes for ST:

- Fixed in the user view (aka always-visible ST), as most of the existing players;
- Fixed to scene, by replicating the ST at 120° intervals around the user, as implemented by the BBC [5].
- Fixed to speaker [13], which is similar to the above mode, but by rendering the ST close to the associated speaker, and additionally using always-visible guiding methods (e.g. arrows) to indicate where the speaker is, if he/she is outside of the user’s FoV. As long as the

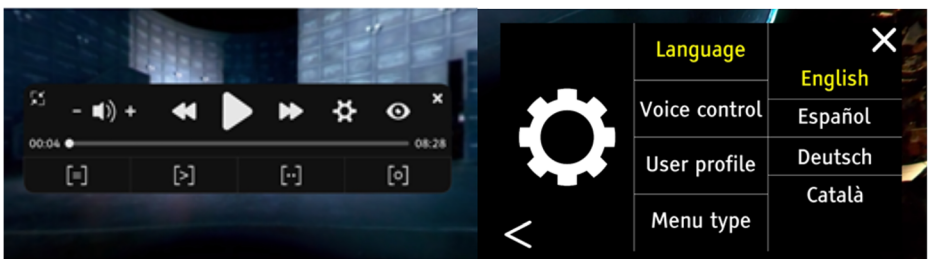
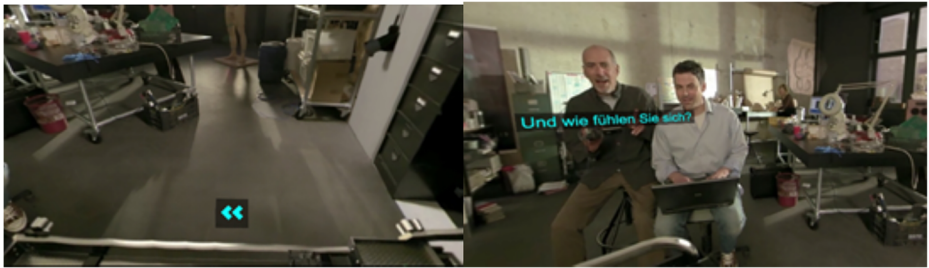


Fig. 24 Enlarged version of the ImAc player UI (latest version on the left; first version on the right)



**Fig. 25** Subtitles attached to the speaker with always-visible visual indicators

speaker is within the user's FoV, the visual indicator is automatically hidden. This presentation mode is outlined in Fig. 25.

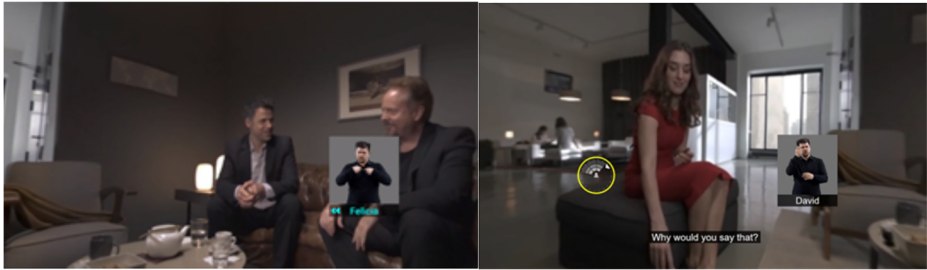
Conducted tests in [2, 3, 18] have shown that always-visible subtitles are clearly preferred, mainly because they were easier to find and to read, less distracting, and users perceived a higher freedom to explore the 360° environment without missing the subtitles.

With regard to the indicators, it is worth mentioning that: 1) the arrows are only shown if the speaker is outside the FoV and are automatically hidden when the speaker is again visible; 2) the radar indicates the current user's FoV and the relative position of the speaker, by using a mark of the same color as the subtitles for a better identification. Conducted user tests have proven that users tend to prefer the arrows than the radar, as they are simpler and more intuitive, although the radar was also welcome by especially young users and proven to provide a more comprehensive information about both the current FoV and speaker's position [2, 3, 18]. Therefore, the two indicators have been kept in the player, letting the users choose their preferred one.

Unlike the reviewed 360° players which provide few personalisation options, except YouTube, the ImAc player allows a personalised presentation of ST in terms of: size (three size levels); background (outlined text or a semi-transparent background box); position (top and bottom); and language. These need for these options were identified in the conducted user-centric activities [20], and the related UI items were proven to be intuitive and easy to use in the conducted user tests [18].

Finally, the presentation of Easy-to-Read subtitles [4] is supported. Results from conducted tests have preliminarily proven that Easy-to-Read subtitles are preferred over traditional subtitles by elderly participants when watching 360° clips from an opera performance [29]. To the best of our knowledge, no other research studies in this topic have been conducted so far. However, it is assumed that the presentation of Easy-to-Read subtitles on mobile devices can also provide benefits, due to the fact that they are typically shorter than traditional subtitles and the screen size is typically small, so this allows enlarging the text size.

**Sign language (SL)** The ImAc player supports the presentation of SL, but not just as a burned in video, as in the scarce solutions supporting SL (Section 2), but also as an independent DASH stream signalled as part of the main 360° video service [34]. On the one hand, this allows presenting the SL video fixed in the user view, and not fixed to the scene, thus being always visible regardless of where the user is looking at. On the other hand, this allows a dynamic personalisation of the SL service, in terms of activation/deactivation, and of language, size (three levels) and position settings (left, right). Two visual indicators (arrows and radar)



**Fig. 26** Presentation of SL with indicators, and simultaneously with ST (right)

can be also enabled, as for subtitles (see Fig. 26). In order to provide a better identification of the target speaker, his/her name (or even a descriptive info text) can also be added below the video window (see Fig. 26).

Two further innovation features of the ImAc player can be also highlighted:

- It enables the simultaneous presentation of ST and SL. In such a case, if the subtitles are moved at the top via the associated option of the player menu, then the sign language videos will be also moved at the top for a better visual alignment between both access services. Likewise, ST is considered as the master service for indicators.
- It allows dynamically showing/hiding the video window based on the signer's activity, based on metadata added at the production side.

Finally, all visual elements on screen, including ST, SL and indicators, can be dynamically placed at the preferred position, thanks to a developed drag & drop feature (see Fig. 26, left, where the radar is being moved, and this is indicated by a yellow outline).

The need for these features and personalization options for presentation of the SL service alone or in coordination with ST was derived from the conducted user-centric activities [20], and such features were proven to be easy to use, welcome, as well as to provide good immersion and QoE [18].

**Audio description (AD)** The ImAc player also provides does not just provide support for AD as independent streams, but leverages the availability of spatial audio technology (Ambisonics) to provide different forms of presentation modes and narratives in a personalized manner (if available), like:

- Classic Mode: no audio positioning.
- Static Mode: audio from a fixed point in the scene (e.g. like a friend whispering in your ear).
- Dynamic Mode: audio coming from the direction of the action.

Likewise, the fact that multiple independent streams allows for: 1) adding different scripting and narrative modes; 2) dynamic personalization in terms of language, presentation mode and volume levels.

The need for these features for presentation of the AD service was derived from the conducted user-centric activities [10, 20], and such features were proven to be easy to use, welcome, as well as to provide good immersion and QoE [18].



The same features have been implemented for the audio subtitles (AST) service, which is a much less developed access service, but that can provide very relevant benefits, as discussed in [30].

The ImAc player has been adopted by European broadcasters, like RBB (Germany), CCMA (Spain) and À Punt (Spain) to provide online access to their produced 360° videos augmented with accessibility features, even through HbbTV multi-screen services in the case of CCMA: <https://www.ccma.cat/experiencia-immersiva-accessible/>.

## 5 Taxonomy and conclusions

This study has reviewed the key existing 360° players with regard to relevant accessibility related aspects. This overall review is categorized in Table 1 to show an overall picture of to what extent the identified accessibility features and guidelines are met, and what the existing gaps are.

In general, the existing 360° players provide very little support for accessibility. The majority of players do support ST, but only as a traditional television rendering into the 360° projection. This means that they have some mechanism for displaying text which can be turned on and off, however there is no consideration for the space where it is to be rendered or its location within the scene. This is generally because they follow the principles used within traditional television broadcast, providing two lines of text, ~30 characters wide. There is also no mechanism within standard subtitle file formats to store position information, which would allow implementing a logic to indicate where the active speaker is in the 360° space at any moment, relative to the user's FoV, as proposed in [34].

There is absolutely no support for SL as a separate stream within any of the existing players. It has been generally accepted that if a signer is required, it can be burned into the video. However, this gives no scope for customisation, such as allowing the user to turn the feature on or off, customizing position and size, or keeping the signer within the user's viewpoint. In 360° video this can be a serious limitation, as the producer of the video has to choose where to put the signer and therefore forcing the user to follow a specific view. In case of actions happening around the 360° scene, it could also mean that the SL had to be re-located or that some relevant scenes could be blocked by the SL video.

Some players provide support for AD, but only as an alternative audio track, not an additional soundtrack which can be balanced against the main audio. This causes a major disadvantage for those with hearing impairments, as being able to control the AD level against the background can massively improve their experience. The only 360° players that natively support AD as an additional audio stream are JW Player and YouTube player. However, as far as authors know, they do not include specific personalisation options for AD, beyond those common for audio. In this context, no 360° players providing support for AST exist.

Although some of the players, such as JWPlayer, provide an API for developers to implement their own UI, there is very little consideration for the default UIs to be adaptive to meet the specific needs for: an omnidirectional environment projected onto a sphere; the limited FoV in HMD; and accessibility. This includes the ability to extend the media experience onto multiple screens. This is only provided by Omnivirt, thanks to community developed extensions.

Voice control is becoming increasingly popular, mainly due to the expanding market for voice activated digital assistants, such as Amazon Echo, Google Home and Apple Siri. The advantages of voice control to partially sighted and blind users are clear as it negated the need to directly see or operate the UI. Currently, there are available mechanisms to connect these digital assistants to most platforms, connecting at the system level. This can allow some control, such as starting the video player. However, the lack of standards for connecting the voice control to the UI prevents from actually providing an effective control over the actual player.

Finally, it was found that generally the 360° players are free to use. Although the development costs are high, revenue is normally generated through both advertising and the sales and licensing of the authoring tools.

In general terms, the surveyed 360° players hardly meet the necessary accessibility requirements, and the provided ones seem to be inherited from the traditional 2D world, instead of addressing the specificities of 360° environments. In terms of the WCAG guidelines, the existing web-based 360° players still present important limitations.

All these limitations have been overcome by the ImAc player, which has successfully addressed the recommended accessibility requirements and beyond.

## 6 Future work

The contributions and insights from this study can be used as a catalyst to improve the existing solutions, and to adopt the best suited ones by the interested agents (end-users, developers, content providers...). They can be also used to support standardization activities and future research efforts, thanks to the provided overall view of to what extent this research space is encompassed. Likewise, the development of the ImAc player is not meant to be completed. Its developers will continue monitoring the research topic and new opportunities to further refine and extend it. This can also be done by third-party agents, as the player is open source and released on Github.

Finally, a similar research study should be also conducted for 3D VR environments, with 6 Degrees of Freedom (6DoF), where additional challenges need to be addressed. The work in [28] is a starting point in that direction.

**Funding** This work has been partially funded by the European Union's Horizon 2020 program, under agreement n° 761974 (ImAc project). Work by Mario Montagud has been additionally funded by the Spanish Ministry of Science, Innovation and Universities with a Juan de la Cierva – Incorporación grant (with reference IJCI-2017-34611) and with the ALMA Excellence Network (with reference RED2018-102475-T).

**Code availability** The ImAc player is open source, and available from <https://github.com/Fundacio-i2CAT/ImAc> under a GPL-3.0 license.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

1. Agulló B, Matamala A (2019) Subtitling for the deaf and hard-of-hearing in immersive environments: results from a focus group. *Journal of Specialised Translation*, Núm, JoSTrans, p 32
2. Agulló B, Matamala A (2020) Subtitles in virtual reality: a reception study. *Íkala*, Forthcoming in 2020.
3. Agulló B, Montagud M, Fraile I (2019) Making interaction with virtual reality accessible: rendering and guiding methods for subtitles. *AI EDAM*, Volume 33, Issue 4, November 2019, pp. 416–428
4. Bernabé R, Orero P (2019) Easy to read as multimode accessibility service. *Hermeneus* 21:53–74
5. Brown A, Turner J, Patterson J, Schmitz A, Armstrong M, Glancy M (2017) Subtitles in 360-degree Video. In: In adjunct publication of the 2017 ACM international conference on interactive experiences for TV and online video (TVX '17 adjunct). Association for Computing Machinery, New York, NY, USA, pp 3–8. <https://doi.org/10.1145/3084289.3089915>
6. DR Design (DR) Smart icons design – common European standardization. <https://www.dr.dk/om-dr/about-dr/smart-icons-design-common-european-standardization> Last Access in August 2020.
7. European Broadcasting Union (EBU) (2017) Virtual reality: how are public broadcasters using it? [https://www.ebu.ch/publications/strategic/login\\_only/report/virtual-reality-how-are-public-broadcasters-using-it](https://www.ebu.ch/publications/strategic/login_only/report/virtual-reality-how-are-public-broadcasters-using-it) Last Access in August 2020.
8. European Commission (2018) The 2018 Ageing Report: Economic and Budgetary Projections for the EU Member States (2016–2070), Institutional Paper 079. May 2018. Brussels. ISBN 978–92–79-77460-7
9. Fan CL, Lo WC, Pai YT, Hsu CH (2019) A survey on 360° video streaming: acquisition, transmission, and display. *ACM Comput. Surv.* 52, 4, article 71
10. Fidyka A, Matamala A (2018) Audio description in 360° videos: results from focus groups in Barcelona and Kraków. *Translation Spaces* 7(2):285–303
11. <https://www.juniperresearch.com/researchstore/innovation-disruption/virtual-reality/subscription/innovation-disruption-future-prospects>. Last Access in August 2020
12. <https://www.statista.com/statistics/259477/hours-of-video-uploaded-to-youtube-every-minute/>. Last Access in August 2020
13. Hughes C, Montagud M, Tho Pesch P (2019) disruptive approaches for subtitling in immersive environments. *ACM TVX* 2019, June, Manchester (UK)
14. Hughes C, tho Pesch P, Montagud M, Brelot M, Bouqueau R, Torres E, (2019) D3.1. Architectural Design, Deliverable of EU H2020 ImAc project, <https://www.imac-project.eu/documentation/deliverables/> July 2019
15. Hybrid Broadcast Broadband TV (HbbTV) (2018) 2.0.2 Specification, HbbTV Association Resource Library, <https://www.hbbtv.org/resource-library> Last Access in August 2020
16. International Communication Union (ITU) (2011) Making television accessible. [http://stagingituint/en/ITU-D/Digital-Inclusion/Persons-with-Disabilities/Documents/Making\\_TV\\_Accessible-English.pdf](http://stagingituint/en/ITU-D/Digital-Inclusion/Persons-with-Disabilities/Documents/Making_TV_Accessible-English.pdf). Last Access in August 2020
17. JWPlayer Voicer, GitHub repository. [github.com/jwplayer/jwplayer-ios-bestPracticeApps/tree/master/JWBestPracticeApps/Voicer](https://github.com/jwplayer/jwplayer-ios-bestPracticeApps/tree/master/JWBestPracticeApps/Voicer) Last Access in August 2020
18. Matamala A (2020) D5.4 – pilot evaluation report, EU H2020 ImAc project deliverable, <https://www.imac-project.eu/documentation/deliverables/>, march 2020.
19. Matamala A, Orero P (2018) Standardising accessibility: transferring knowledge to society. *Journal of Audiovisual Translation* 1(1):139–154
20. Matamala A, Orero P (2019) D2.1. User centered design, EU H2020 ImAc project deliverable, <https://www.imac-project.eu/documentation/deliverables/>, July 2019
21. Mies R et al (2019) D3.4. Accessibility interface. Public deliverable of the EU H2020 ImAc project, DOI: <https://doi.org/10.5281/zenodo.3413288>, <https://zenodo.org/record/3413288#.XqvctagzY2w>
22. Montagud M, Boronaf F, Belda J, Marfil D (2016) Use of web components to develop interactive, Customizable and Multi-Device Video Consumption Platforms Communications in Computer and Information Science (CCIS), volume 605, Springer-Verlag, April 2016
23. Montagud M, Fraile I, Meyerson, Genis E, Fernández S (2019) ImAc player: enabling a personalized consumption of accessible immersive content. *ACM TVX* 2019, June, Manchester (UK)
24. Montagud M, Orero P, Matamala A (2020) Culture 4 all: accessibility-enabled cultural experiences through immersive VR360 content. *Pers Ubiquit Comput.* <https://doi.org/10.1007/s00779-019-01357-3>
25. Montagud M, Fraile I, Meyerson E, Fernández S (2020) Open-source accessibility-enabled V360 player. *ACM IMX* 2020, Barcelona (Spain), June 2020
26. Montagud M, Orero P, Fernández S (2020) Immersive media and accessibility: hand in hand to the future. *ITU Journal: ICT Discoveries*, Accepted for Publication in 2020

27. Montagud M, Boronat F, Marfil D, Pastor J (2020) Web-based platform for a customizable and synchronized presentation of subtitles in single and multi-screen scenarios. *Multimedia Tools and Applications*, Springer 79:21889–21923. <https://doi.org/10.1007/s11042-020-08955-x>
28. Montagud M, De Rus JA, Segura-García J, Fayos-Jordan R (2020) Towards an immersive and accessible virtual reconstruction of theaters from the early modern: bringing back cultural heritage from the past. *ACM IMX 2020*, Barcelona (Spain), June 2020
29. Oncins E, Bernabé R, Montagud M, Arnáiz V (2020) Accessible scenic arts and Virtual Reality: A pilot study with older adults about user preferences when reading subtitles in immersive environments, *MonTi journal*, Accepted for Publication in 2020
30. Orero P, Montagud M, Mata J, Torres E, Matamala A (2020) Audio subtitles or spoken subtitles: an ecological media accessibility service. In Dejica, Daniel & Carlo Eugeni, Anca Dejica-Cartis (eds.) *Translation studies and information technology - new pathways for researchers, teachers and professionals*. Timisoara: Editura Politehnica, translation studies series. ISBN: 978-606-35-0351-1
31. Papachristos NM, Vrellis I, Mikropoulos TA (2017) A comparison between oculus rift and a low-cost smartphone VR headset: immersive user experience and learning. 2017 IEEE 17th international conference on advanced learning technologies (ICALT). Timisoara 2017:477–481. <https://doi.org/10.1109/ICALT.2017.145>
32. Powers S (2011) *HTML5 media: integrating audio and video with the web*. O'Reilly Media, Inc, pp 82–83. isbn:978-1-4493-1531-3
33. Srivastava P, Rimzhim A, Vijay P, Singh S, Chandra S (2019) Desktop VR is better than non-ambulatory HMD VR for spatial learning. *Front Robot AI* 6:50. <https://doi.org/10.3389/frobt.2019.00050>
34. ttho Pesch P, Bouqueau R, Montagud M (2020) White Paper: Recommendations for immersive accessibility services. arXiv:2005.03756 [cs.MM], May 2020
35. Video Description Research and Development Center (2013) YouDescribe <http://youdescribe.org> Last Access in August 2020
36. World Wide Web Consortium (W3C), Web Content Accessibility Guidelines (WCAG), <https://www.w3.org/TR/WCAG20> Last Access in August 2020
37. World Wide Web Consortium (W3C) (2010). Timed Text Markup Language (TTML) 1.0 <http://www.w3.org/TR/tafl-dfxp/> Last Access in August 2020.
38. World Wide Web Consortium (W3C) (2013). WebVTT: The Web video text tracks format <https://www.w3.org/2013/07/webvtt.html> Last Access in August 2020
39. Xu M, Li C, Zhang S, Callet PL (2020) State-of-the-Art in 360° video/image processing: perception, assessment and compression. in *IEEE Journal of Selected Topics in Signal Processing* 14(1):5–26

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.