# Requirements and Functionalities - Version B
## Period Covered: 2014/04/01-2015/12/31

| | |
|---|---|
| **Project ref. no.** | FP7-ICT-2013-7 - 610691 |
| **Project acronym** | BRIDGET |
| **Start date of project (duration)** | 1 November, 2013 (36 months) |
| **Document due Date:** | 2015/12/31 |
| **Actual date of delivery** | 2015/12/23 |
| **Leader of this document** | Leonardo Chiariglione |
| **Reply to** | leonardo@chiariglione.org |
| **Document status** | Final |

## Deliverable Identification Sheet

| | |
|---|---|
| **Project ref. no.** | FP7-ICT-2013- 610691 |
| **Project acronym** | BRIDGET |
| **Project full title** | BRIDging the Gap for Enhanced broadcasT |
| **Document name** | D2.5 Requirements and Functionalities.docx |
| **Security (distribution level)** | PU |
| **Contractual date of delivery** | 2015/12/31 |
| **Actual date of delivery** | 2015/12/23 |
| **Document number** | D2.5 |
| **Type** | Report |
| **Status & version** | Final v1.0 |
| **Number of pages** | 30 |
| **WP / Task responsible** | WP2/T2.1 |
| **Other contributors** | Alberto Messina |
| **Author(s)** | Leonardo Chiariglione |
| **Project Officer** | Alberto Rabbachin |
| **Abstract** | This deliverable contains the BRIDGET user and functional requirements and the functionalities provided by BRIDGET applications for the 2nd phase of the project (Version B) |
| **Keywords** | User, Functionalities, Requirements |
| **Sent to peer reviewer** | 2015/12/22 |
| **Peer review completed** | 2015/12/23 |
| **Circulated to partners** | |
| **Read by partners** | |
| **Mgt. Board approval** | |

| Version | Date | Reason of change |
|---------|------|------------------|
| Rev0.1 | 2015-11-08 | Leonardo Chiariglione – ToC |
| Rev0.2 | 2015-11-15 | Leonardo Chiariglione – revision with questions |
| Rev0.3 | 2015-11-16 | BRIDGET meeting |
| Rev0.4 | 2015-11-28 | LC – further revision |
| Rev1.0 | 2015-12-21 | LC - Proposed final version |
| Rev1.1 | 2015-12-23 | Stavros Paschalakis – Peer Review |

## Table of Contents

## Figures

## Tables

## 1    Executive Summary

This document is a revision of Deliverable D2.2 "Requirements and Functionalities - Version A" [1]. It substantially updates the User and Functional Requirements specified in the early months of the project by taking into account the developments that occurred until month 26 of the project, in particular the development of the BRIDGET Authoring Tool, the BRIDGET player, the ISO/IEC Media Linking Application Format (ISO/IEC 23000-18), the user trials and the 2nd BRIDGET workshop.

## 2    Introduction

This document collects the revised requirements and functionalities to be used as a basis for the 2nd phase of the BRIDGET project.

The updates maily concern

1) Alignment of the requirements and functionalities identified in D2.2.1rev (January 2015) with
    a.  The implementation of technologies integrated in the BRIDGET Authoring Tool and Player
    b.  The Media Linking Application Format being standardised by MPEG
2) Addition of requirements and functionalities to support the new version of  the BRIDGET Authoring Tool and Player for the new "Live bridgets" and "World in 3D" use scenarios (see D2.4)
3) Accommodation of feedback received from the BRIDGET User trials and the 2nd BRIDGET Workshop
4) Impact of standardisation activities (mainly the MPEG Media Linking Application Format).

The structure of this document is similar to the one adopted for D2.2. It contains the list of requirements and functionalities.  Those in normal font will be supported by the Authoring Tools and Player of Version B. Those in italic are required for commercial BRIDGET applications and services, but cannot be implemented because of the limited project resources.

## 3    Glossary

The glossary in this section builds on the glossary provided in D2.1 Version A [1].

*Table 1 - Terms and Definitions*

| Term | Definition |
|------|------------|
| Bridget | A data structure containing 1) data related to the source content 2) data related to the destination content 3) link from 1) to 2) and 4) optionally information on how the bridget should be presented to the user. |
| BRIDGET | The EU-funded project that defines format, technologies, systems and develops applications (authoring, playing etc.) based on them |
| Bridget application | Any software component that includes technologies to handle bridgets, e.g. software that creates bridgets, or software to play or interact with bridgets |
| Bridget editor | A person in charge of creating bridgets or bridget templates, they can be either a professional or end-user. |
| Bridget repository | A component (e.g. a database) able to store bridgets during their life cycle |
| Bridget template | A data structure containing characteristics (e.g. in form of a combination of descriptors) denoting the source and/or the destination content of a bridget and parameters describing the presentation layout. |

| | |
|---|---|
| Bridget temporal scope | The duration in time of a time aligned bridget, during which the media time of the main programme matches with its temporal validity as defined by the bridget editor |
| Content repository | A repository of destination content |
| Destination content | The content that is linked to by a bridget |
| Elementary bridget | A bridget created from ingested programmes that only contains destination content |
| Embedded bridget | A bridget multiplexed with some other source content for transport. *Transporting bridgets in the main programme is out of the scope of the project.* |
| Global bridget | A bridget which links from the entire source content |
| Media-time aligned bridget | A bridget which links from a specific time interval of the source content |
| Media time/space segments | A specific time instant/interval, portion of frame or combination thereof in the source content |
| Programme | The audio-visual broadcast content used as source for bridgets. |
| Source content | The content used to link to a piece of destination content |
| Temporal detection accuracy | The maximum absolute error required from a synchronisation technology in identifying a media time point |
| User | A human operating a BRIDGET-based system (initially authoring tool and BRIDGET player) |
| Visual object | A distinct object in a video scene. It may be represented by a bounding polygon or by a ROI segmentation map. |

## 4   Simple BRIDGET Reference Model

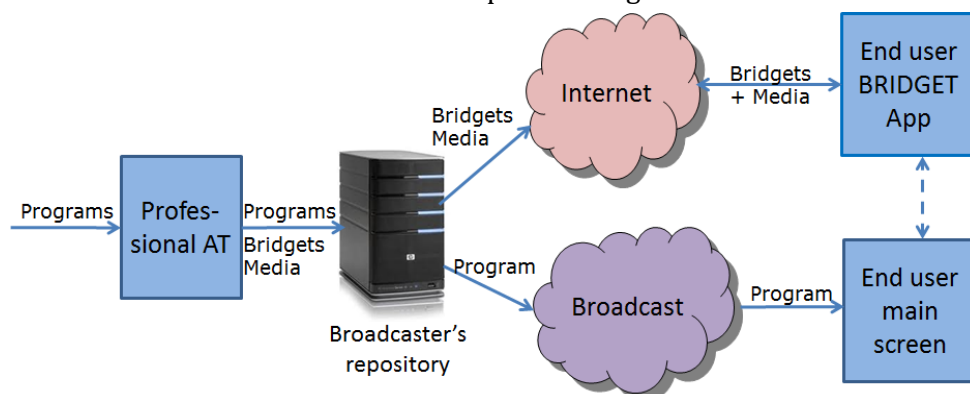A basic end-to-end BRIDGET reference model is depicted in Figure 1.



*Figure 1 - Basic BRIDGET reference model*

Programmes (e.g. archive programmes) ingested in the Professional Authoring Tool are used to create elementary bridgets that only contain references and metadata to the ingested programme and will eventually become bridgets when they will be complemented with information related to the source content. When a programme to be enriched is ingested in the Professional Authoring Tool, the bridget editor will use the Professional Authoring Tool to select portions of the programme, search for suitable elementary bridgets in the repository and create and store bridgets in the repository.

When the enriched programme is eventually broadcasted, the end user will see the programme on the main screen and activate the end user BRIDGET app on the companion screen. The app will synchronise with the programme on the main screen using appropriate technologies developed by the project and download all the bridgets related to the programme. When the app sees that a bridget has a time stamp equal to the actual time, it will alert the user that a bridget related to the specific time of programme is available. The user may decide to play to bridget or skip it.

Figure 2 shows an extended BRIDGET reference model where the end user can enjoy a subset of the Professional AT functionality on the end user BRIDGET app. The user could for instance create new bridgets and even use user generated content available on the personal repository as source and/or destination content. In BRIDGET, however, the end user will only be able to edit an existing bridget and personalise its content.
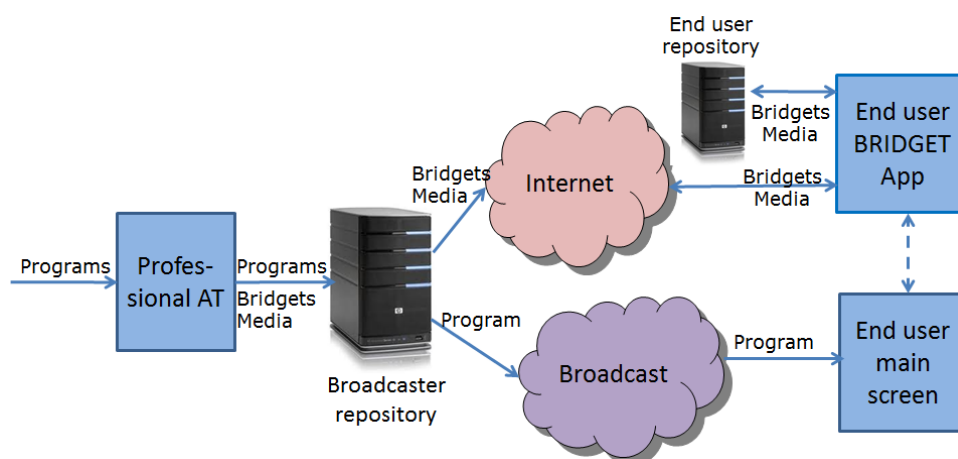


*Figure 2 - Extended BRIDGET reference model*

The BRIDGET technology is very flexible and its use is by no means limited to a broadcast scenario, even if enhanced with a companion screen. As internet becomes increasingly a viable delivery solution for broadcast content, it is important to make sure that the use of BRIDGET is not limited to broadcast. Figure 3 shows that this is not the case and indeed BRIDGET provides an example of a possible merge of a broadcast channel with the internet.
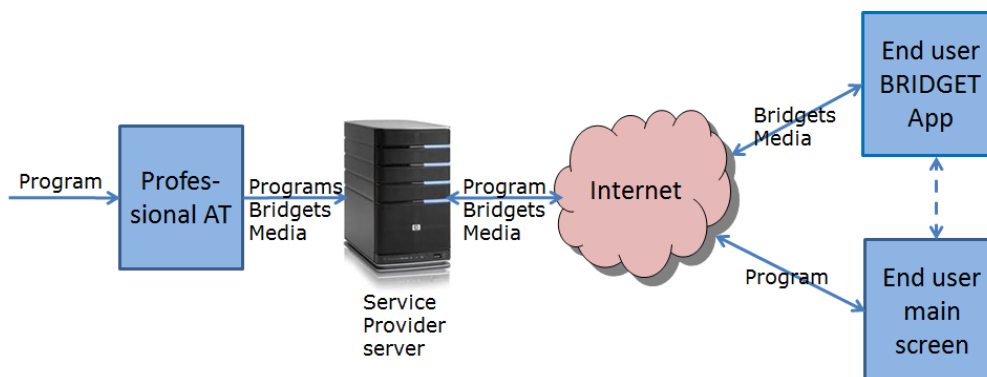


*Figure 3 - BRIDGET reference model in full-internet scenario*

Here we assume that the end user employs a different BRIDGET app running on a tablet to access some content enriched with bridgets and broadcasted as live. The tablet will send a push notification

to the end user's smart TV to get the program from the service provider server. Relevant bridgets will continue to be notified on the tablet. Of course this is just one of the many possible implementations of a bridget scenario on the internet.

The WimBridge service (www.wimbridge.tv) developed by one partner is an implementation of Figure 3. The Professional AT is accessible via the web by any user who can create their own bridgeted videos and stream them as a scheduled live service. Of course several possible configuration can be envisaged where the first and second screen merge. One possibility is that bridgets are displayed on the "first" screen or that the main program is merged with bridgets on the "second" screen.

Figure 4 represents yet another possible scenario where bridgets originate from Service Providers' programmes, are edited (e.g. with a consumer AT) and redistributed via social networks. Now bridgets can become a viral tool to promote the Service Provider's content. The project will not specifically handle distribution of bridgets via social networks but the figure highlights the potential of the BRIDGET technology beyond established business roles in the internet-based media value chain. The "Social live" scenario planned for the 2nd phase [22] will include elements of bridgets over social networks when bridgets are streamed in a live program and bridgets can promote user interaction.
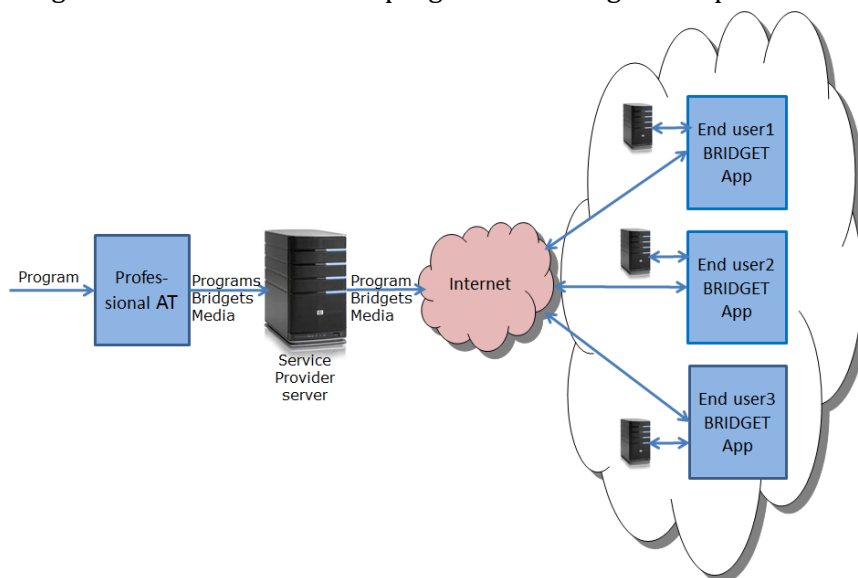


*Figure 4 - Bridgets in a social media environment*

The models described above is conceived to naturally fit into the currently deployed broadcast/broadband architectures for connected TVs and Set-Top Boxes. Relevant examples are HbbTV [4] and the Italian platform for digital television defined in HD-Book specifications [5, 6].

# 5    User Requirements

In this section the word "user" indicates any type of user of BRIDGET: bridget editor, 3D object creator, and user playing bridgets or destination content.

Nota Bene: requirements expressed in normal font have been or are planned to be implemented in a basic form. Some of them will be shown at the Y2 review in M28. Requirements in *italic font* are considered mandatory for a commercial service but will not be implemented in BRIDGET. Some requirements in *italic font* will be implemented in the commercial service run by a partner and will be demonstrated at the final project review.

## 5.1    Bridget Creation

This section collects requirements on how users create bridgets.

*Table 2 - Bridget creation user requirements*

| Req. ID | Name | Description |
|---------|------|-------------|
| UBC1 | Content link | The user shall be able to create a bridget linking a destination content to a programme. |
| UBC2 | Stored content | The user shall be able to create bridgets for stored programmes. |
| *UBC3* | *Live content* | *The user shall be able to create bridgets for live programmes.* |
| UBC4 | Bridget association – time-aligned | The user shall be able to associate one or more bridgets to a defined interval in the timeline of the programme (time-aligned bridget). |
| UBC5 | Bridget association – global | The user shall be able to associate one or more bridgets to the entire duration of the programme (global bridget). |
| UBC6 | Bridget made global | The user shall be able to transform, media time-aligned bridgets into global bridgets. |
| UBC7 | Bridget reuse | The user shall be able to store, retrieve and reuse the bridgets previously associated to a different programme. |
| UBC8 | Time–aligned bridget points candidates | The user shall be able to retrieve a list of manually and automatically generated candidate media time/space points or segments for bridget insertion. |
| UBC9 | Destination content candidates | The user shall be able to retrieve a list of manually and automatically generated candidate destination content for a programme segment. |
| UBC10 | Destination content | The user shall be able to create a bridget whose destination content is a 2D audiovisual content or a 3D model of a scene including 3D visual objects, 3D Audio objects or both. |
| *UBC11* | *Bridget information creation* | The user shall be able to edit a bridget, following a defined bridget representation structure. |

## 5.2 Bridget Access

This section collects requirements on how users access bridgets.

*Table 3 - Bridget access user requirements*

| Req. ID | Name | Description |
|---------|------|-------------|
| UBA1 | Devices for bridget access | The user shall be able to access bridgets and corresponding content using a device (e.g. a companion screen device). |
| UBA2 | Bridget retrieval | The user shall be able to retrieve bridgets based on the identification of the programme and of the media time point being watched on the main screen. |
| UBA3 | Retrieval of remote bridgets | The user shall be able to retrieve bridgets from the Internet. |

| Req. ID | Name | Description |
|---------|------|-------------|
| UBA4 | Synchronised access & presentation | The user shall be able to access and present bridgets synchronised with the programme displayed on the main screen. |
| UBA5 | Synchronisation independence | The user shall be able to access synchronised bridgets independently from the broadcast network (e.g. DVB-T/S/C or the internet) used to receive the programme |
| UBA6 | Recording support | The user may be able to access synchronised bridgets from a programme previously recorded using a PVR or Network PVR. |

## 5.3  Bridget Search

This section collects requirements on how users search bridgets and related content.

*Table 4 - Bridget search user requirements*

| Req. ID | Name | Description |
|---------|------|-------------|
| *UBS1* | *Bridget Search* | The user shall be able to search and retrieve bridgets using metadata |
| UBS2 | Content Search – production | The user shall be able to search and retrieve media content similar or related to a segment of the programme. |
| *UBS3* | *Content Search – consumption* | *The user shall be able to search and retrieve media content similar or related to the content linked through a bridget.* |
| UBS4 | Visual object selection | The user shall be able to select visual objects from the programme at any time to perform a search of related content in an image database resource. |
| UBS5 | Content search - metadata | The user shall be able to search for similar or related content using metadata. |
| *UBS6* | *Content search – visual queries* | *The user shall be able to search for similar content through visual queries.* |

## 5.4  Bridget Presentation

This section includes requirements on how bridgets are presented to users for navigation and selection.

*Table 5 - Bridget presentation user requirements*

| Req. ID | Name | Description |
|---------|------|-------------|
| UBP1 | Presentation of in scope time-aligned bridgets | The user shall be able to view a presentation of the bridgets according to their temporal scope on the media time interval of the programme. |

| Req. ID | Name | Description |
|---------|------|-------------|
| UBP2 | Presentation of out of scope time-aligned bridgets | The user may be able to view a presentation of associated time–aligned bridgets after their temporal scope of the media time interval of the programme. |
| UBP3 | Presentation of bridgets or groups of bridgets | The user shall be able to decide the visual presentation of bridgets or groups of bridgets. |
| UBP4 | Presentation of global bridgets | The user shall be able to view a presentation of global bridgets associated with a programme. |
| UBP5 | Bridget presentation resources | The user shall be able to define how to notify the availability of bridgets using different media resources such as icons, images, text, video, and audio for appropriate presentation. |
| UBP6 | Bridget notification | The user shall be able to enable and disable the notification of available active bridgets on the second screen at any time (for both media-time aligned and global bridgets). |
| UBP7 | Multiple bridget notifications | The user shall be able to receive notifications of more than one available bridget at the same time. |
| UBP8 | Filter & Selection – manual | The user shall be able to filter and select presented bridgets manually. |
| *UBP9* | *Presentation preferences* | *The user shall be able to set and update preferences concerning bridgets presentation (e.g. preferred kind of content: 2D video, 3D model, etc...).* |
| *UBP10* | *Filter & Selection – automatic* | *The user may be able to filter and select presented bridgets automatically based on preferences.* |

## 5.5   Bridget Consumption

This section includes requirements on how users consume bridgets.

*Table 6 - Bridget consumption user requirements*

| Req. ID | Name | Description |
|---------|------|-------------|
| UBU1 | Bridget consumption | The user shall be able to consume available bridgets (i.e. bridget presented to the user) and corresponding destination content on a device (e.g. a companion screen device). |
| UBU2 | Global bridget consumption | The user shall be able to consume a global bridget and corresponding destination content at any time during the associated programme. |
| UBU3 | Independent consumption | The user shall be able to consume bridgets and corresponding destination content without interrupting the programme on the main screen. |
| *UBU4* | *Bridget bookmarks creation* | *The user shall be able to bookmark time aligned and global bridgets for future use.* |

| Req. ID | Name | Description |
|---------|------|-------------|
| *UBU5* | *Bridget bookmarks consumption* | *The user shall be able to consume bookmarked bridgets outside the context of the specific programme.* |
| *UBU6* | *Bridget context reset* | *The user shall be able to reset the context; deactivating bookmarked bridgets that are not related to the specific programme.* |

## 5.6    3D Object Creation

This section collects requirements on how users create 3D objects.

*Table 7 - 3D object creation user requirements*

| Req. ID | Name | Description |
|---------|------|-------------|
| U3C1 | 3D reconstruction | The user shall be able to create a 3D model representation reconstructed from a set of videos or images. |
| *U3C2* | *Candidate selection for reconstruction.* | *The user shall be able to enable manual and automatic selection of content items for 3D reconstruction.* |
| U3C3 | Initial 3D reconstruction | The user shall be able to request the 3D reconstruction engine to generate a coarse 3D scene geometry. |
| U3C4 | 3D reconstruction refinement | The user shall be able to request the 3D reconstruction engine to refine the 3D scene geometry with new images and videos of the scene. |
| *U3C5* | *3D input content* | *The user shall be able to provide 3D content to the 3D reconstruction engine in order to either generate a coarse, initial 3D model of a real scene or refine an already existing 3D model.* |

## 5.7    3D Object presentation

This section includes requirements on how 2D views from 3D objects are presented to users

*Table 8 - 3D object presentation user requirements*

| Req. ID | Name | Description |
|---------|------|-------------|
| U3P1 | Viewpoint rendering | The user shall be able to request 2D views from the reconstructed 3D scene model at original viewing positions. |
| U3P2 | Viewpoint navigation | The user shall be able to request 2D views from the reconstructed 3D scene model at intermediate viewing positions, to allow navigation inside/around the model |
| U3P3 | Device independence | The user shall be able to view [almost] identical 2D images on devices with different capabilities, e.g. HW-based acceleration for 3D graphics, although the result may be obtained with very different response times or frame rates. |

| Req. ID | Name | Description |
|---------|------|-------------|
| U3P4 | 3D audio rendering | The user shall be able to listen to 3D spatial audio through headphones. |
| *U3P5* | *3D audio navigation* | *The user shall be able to listen to 3D spatial audio adapted to the virtual position selected through the user interface.* |
| U3P6 | 3D audio/video rendering | The user shall be able to view 2D views at intermediate viewpoints synchronised with the spatial audio rendered according to the corresponding position. |

## 6    Functional Requirements

Nota Bene: requirements expressed in normal font have been or will be implemented for Version B delivery at M32. Requirements in *italic font* are considered important for a commercial deployment.

### 6.1    Bridget Representation

This section includes requirements on how bridgets are digitally represented. The definition of this structure implies that functionalities to set, modify, and read the related element shall be supported.

*Table 9 - Bridget representation functional requirements*

| Req. ID | Name | Description |
|---------|------|-------------|
| FBR1 | Bridget identification | The bridget format shall support unambiguous and persistent domain-based identification of a bridget. |
| FBR2 | Bridget authoring information | The bridget format shall support inclusion of bridget authoring information (e.g. title, date, copyright, etc...). |
| FBR3 | Bridget source programme | The bridget format shall allow specification of the associated source programme. |
| FBR4 | Bridget source programme description | The bridget format shall support inclusion of textual descriptions related to the associated source programme. |
| FBR5 | Bridget types | The bridget format shall allow specification of whether the bridget is associated to a specific time interval of the programme (time–aligned) or to the entire programme (global). |
| FBR6 | Bridget destination content | The bridget format shall allow specification of one destination content item. |
| FBR7 | Bridget destination content types | The bridget format shall allow specification of audiovisual, audio, 3D video, 3D audio, or 3D video/audio scenes as destination content. |
| FBR8 | Bridget extended destination content types | The bridget format may allow specification of any kind of media or combination thereof as destination content (e.g. images, text, web pages, etc...). |
| FBR9 | Bridget destination content description | The bridget format shall support inclusion of descriptions related to the destination content (e.g. content type, media duration, MIME |

| Req. ID | Name | Description |
|---|---|---|
| | | type, file size, etc...). |
| FBR10 | Bridget multiple sources for destination content | The bridget format shall support inclusion of a list of alternative sources for the destination content. |
| FBR11 | Bridget default icon | The bridget format shall allow the definition of a default icon or image that can be used to visually present the bridget to the user. |
| FBR12 | Bridget presentation description | The bridget format may allow specification of the visual and the acoustic presentation of the destination content. |
| FBR13 | Bridget consumption-related information | The bridget format shall support inclusion of consumption-related information, (e.g. if destination content should be watched during or after the programme), parental rating, accessibility and device requirements (e.g. stereo headphones). |
| FBR14 | Bridget rights | The bridget format may support inclusion of rights and payment information for destination content consumption. |
| FBR15 | Bridget modifications | The bridget format shall allow modification of all its representation data values excluding those related to identification. |

## 6.2   Content and Bridget Management

This section defines requirements on how bridgets and content are handled by the system.

*Table 10 - content and Bridget management functional requirements*

| Req. ID | Name | Description |
|---|---|---|
| FCM1 | Content ingestion | The content management system shall allow media content ingestion and storage in content repositories. |
| FCM2 | Content types | The content management system shall support the following media types: images, audiovisual content, audio content, and 3D models. |
| FCM3 | Content metadata | The content management system shall allow storage of metadata and annotations related to content. |
| FCM4 | Content removal | The content management system shall allow deletion of content and related metadata and annotations from a content repository. |
| FCM5 | Content publishing | The content management system shall allow publication of content available on a content repository to the Internet. |
| *FCM6* | *Protected Content publishing* | *The content management system may allow publication of content protected with a content protection system (e.g. DRM) and available on a bridget repository to the Internet.* |
| FCM7 | Indexed Content | The content repository shall maintain a list of indexed content. |

| Req. ID | Name | Description |
|---------|------|-------------|
| *FCM8* | *Bridget ingestion* | *The content management system shall allow bridget ingestion and storage in bridget repositories.* |
| *FCM9* | *Bridget publishing* | *The content management system shall allow publication of a bridget available on a bridget repository to the Internet.* |

## 6.3    Synchronisation

This section defines requirements on how bridgets are synchronised with the programme.

*Table 11 - Synchronisation functional requirements*

| Req. ID | Name | Description |
|---------|------|-------------|
| FSY1 | Synchronisation presentation delay | The synchronisation tool shall allow bridget presentation, synchronised to a point in the media timeline of a programme, with a maximum delay of 5s. |
| FSY2 | Detection accuracy | The synchronisation tool shall be able to identify the programme watched and the relative point in the media timeline with a temporal detection accuracy comparable to the average length of the visual shots. |
| FSY3 | Bridget identification | The synchronisation tool shall be able to identify the bridgets to retrieve and present at specific points in the media timeline of a programme. |
| FSY4 | Programme media time identification | The synchronisation tool shall be able to retrieve the media time from the main programme at specific points in time. |
| FSY5 | Independence from main screen interactions | The synchronisation tool shall be able to synchronise the presentation of a bridget to a programme without any interaction with the main screen (e.g. using audio and/or video fingerprinting mechanisms). |
| FSY6 | Automatic synchronisation | The synchronisation tool shall be able to identify and retrieve synchronisation information automatically from the programme stream (i.e. without the intervention of the user). |
| FSY7 | Synchronisation information production | The synchronisation tool shall be able to create synchronisation information from a segment of a programme. |

## 6.4    Content Access

This section defines requirements of how bridgets and content are delivered to and through the system.

*Table 12 - Content access functional requirements*

| Req. ID | Name | Description |
|---------|------|-------------|
| FCA1 | Content delivery | The content access system shall support access to content delivered through HTTP *and HTTP adaptive streaming.* |
| *FCA2* | *Network QoS* | *The content access system shall be able to access content with or without a defined Quality of Service (through managed or unmanaged networks).* |
| FCA3 | CDN | The content access system shall be able to access content delivered through a Content Delivery Network (CDN). |
| FCA4 | Bridget delivery | The content access system shall support access to bridgets delivered through HTTP. |
| *FCA5* | *Audience measurement* | *The content access system may collect audience measurement data related to bridgets and content accessed by the user.* |

## 6.5   Media Analysis Tools

This section defines the requirements of the functionalities exposed by the Media Analysis tools. These tools perform the temporal segmentation and structure analysis of audiovisual content [7], as well as content description [8] and content quality assessment [9]. More specifically, the media analysis tools will enable the correct temporal association of the destination content within the source content, for example allowing the availability of destination content only at specific automatically determined intervals of the original content, aligned with the shot or chapter structure of the programme. Meanwhile, the media analysis tools will also greatly improve the efficiency and results of the visual search and 3D reconstruction tools, for example, by filtering content with specific characteristics (e.g. daytime shot versus night-time shot, interior versus exterior) or quality (e.g. professional quality versus user-generated content).

*Table 13 - Media analysis tools functional requirements*

| Req. ID | Name | Description |
|---------|------|-------------|
| FMA1 | Multiple content types | The media analysis tools shall be able to process audiovisual content comprising images and/or video and/or *audio*. |
| FMA2 | Coding independence | The media analysis tools shall be able to process audiovisual content independently of the content coding format. |
| FMA3 | Coding awareness | The media analysis tools may be able to process audiovisual content according to specific content coding formats, but shall produce comparable results to the media analysis tools which are able to process audiovisual content independently of the content coding format. |
| FMA4 | Content duration independence | The media analysis tools shall be able to process audiovisual content of arbitrary duration. |
| FMA5 | Content modification | The media analysis tools shall be robust to audiovisual content modifications comprising temporal resampling, spatial resampling, |

| Req. ID | Name | Description |
|---------|------|-------------|
| | robustness | spatial aspect ratio changes, A/D conversion, and I/P conversion. |
| FMA6 | Content description self-containment | The media analysis tools shall provide audiovisual content descriptions which are self-contained, in that their use does not require the audiovisual content from which they were generated. |
| FMA7 | Content non-alteration | The media analysis tools shall leave the audiovisual content on which they operate unaltered. |
| FMA8 | Real-time content description extraction | The media analysis tools shall be able to generate audiovisual content descriptions in real time or near-real-time. |
| FMA9 | Faster than real-time content description extraction | The media analysis tools may be able to generate audiovisual content descriptions at very high speeds, far exceeding real time, but shall produce comparable results to the media analysis tools which are able to generate audiovisual content descriptions in real time. |
| FMA10 | On-the-fly content description extraction | The media analysis tools may be able to generate audiovisual content descriptions on the fly, processing video and audio only in a small temporal window around the current media time. |
| FMA11 | Low-level structural segmentation | The media structure analysis tools shall be able to provide a shot-level temporal segmentation and keyframe representation of audiovisual content. |
| FMA12 | *Linear hierarchical structural segmentation* | *The media structure analysis tools shall be able to provide a hierarchical linear temporal segmentation and keyframe representation of audiovisual content.* |
| FMA13 | *Non-linear hierarchical structural segmentation* | *The media structure analysis tools shall be able to provide a hierarchical non-liner temporal segmentation and keyframe representation of audiovisual content.* |
| FMA14 | *Genre-specific structural segmentation* | *The media structure analysis tools shall be able to provide a temporal segmentation and keyframe representation of audiovisual content based on multi-modal computational scene models (e.g. face, speaker, audio, etc.) according to genre-based heuristics for a limited number of genres.* |
| FMA15 | *Content-based annotation* | *The media annotation tools shall be able to provide low-level descriptions (e.g. dominant colour, colour layout, colour temperature, motion activity, and motion type) and high-level descriptions (e.g. indoor, outdoor, town, landscapebased on overall scene classification or depicted object classification) of audiovisual content.* |
| FMA16 | Content-based quality assessment | The media quality assessment tools shall be able to provide quality descriptions (e.g. dark, bright, noisy, shaky, many coding artefacts) of audiovisual content. |
| FMA17 | Fast filtering/ search | It shall be possible to perform faster than real-time filtering / search of audiovisual content using the descriptions provided by the media annotation tools. |

| Req. ID | Name | Description |
|---------|------|-------------|
| FMA18 | Unambiguous content description syntax | The media analysis tools shall present the content description they generate to the rest of the system in an unambiguous format, namely AVDP. |

## 6.6  Search Tools

This section defines the requirements of the functionalities exposed by the Search tools [10]. The search tools allow automated Visual Search in image [11] and video databases [12], supporting fast creation of content-based links for the Authoring Tools (ATs) and user-originated search capabilities for the player. They include tools for off-line descriptor extraction, database indexing and on-line search suitable for ultra-large databases of images and videos. The user will be able to select a specific object of interest in the query via Region of Interest (ROI) selection. The tools will also support feature selection and matching for use in 3D model extraction in the WP6.

*Table 14 - Search tools functional requirements*

| Req. ID | Name | Description |
|---------|------|-------------|
| FST1 | Visual search in image library | The visual search engine shall support visual search in a pre-indexed image library based on a query image. |
| *FST2* | *Visual search in image library where query specifies ROI in the query image* | *The visual search engine shall support visual search in a pre-indexed image library based on a query image with a ROI specified as bounding polygon or ROI segmentation map.* |
| FST3 | Visual search in video library utilising video key frames | The visual search engine shall support visual search in a pre-indexed video library based on a query image. Search and access points are restricted to video key-frames. |
| *FST4* | *Visual search in video library utilising video key frames where query specifies ROI in the query image* | *The visual search engine shall support visual search in a pre-indexed video library based on a query image with a ROI specified as bounding polygon or ROI segmentation map. Search and access points are restricted to video key-frames.* |
| FST5 | Off-line indexing of image library | The visual indexing engine shall support off-line creation of a database/index for a library of image resources. |
| FST6 | Off-line indexing of video library based on video key-frames | The visual indexing engine shall support off-line creation of a database/index for a library of video resources where search and access points are restricted to video key-frames. |
| FST7 | Ranking of the search results based on matching confidence | The visual search engine shall support ranking of the search results based on match confidence. |
| *FST8* | *Spatial Localisation of the matched region* | *The visual indexing and search engines shall support coarse spatial localisation of the matching regions.* |
| FST9 | Multiple content types | The visual search and indexing engines shall be able to process audiovisual content comprising images *and/or video.* |

| Req. ID | Name | Description |
|---------|------|-------------|
| FST10 | Coding independence | The visual search and indexing engines shall be able to process audiovisual content independently of the content coding format. |
| FST11 | Content modification robustness | The visual search and indexing engines shall support visual search to be robust to audiovisual content modifications such as temporal resampling, spatial resampling, spatial aspect ratio changes, A/D conversion, and I/P conversion. |
| FST12 | Robustness to visual deformations | The visual search and indexing engine shall support visual search that is robust to viewing aspect, illumination changes, partial occlusion and similar artefacts. |
| FST13 | Content description self-containment | The visual search and indexing engines shall derive and use content descriptions which are self-contained, in that their use does not require the access to the content from which they were generated. |
| FST14 | Content non-alteration | The visual search and indexing engines shall leave the audiovisual content from which they were derived unaltered. |
| FST15 | Integration with content-based search | The visual search and indexing engines shall execute queries based on content input by the user. |

## 6.7    3D Tools

This section defines the requirements of the functionalities exposed by the 3D tools. These tools shall be able to perform computer vision [13] tasks, including 3D scene reconstruction, coding, decoding and rendering. In particular, the tools shall support the following workflow for 3D scene reconstruction:

1. Taking as input several pictures of a given 3D scene shot from different viewpoints (and possibly with different cameras or zoom factors), both the intrinsic and extrinsic camera calibration parameters can be estimated. This requires feature point matching [14], but not necessarily the use of calibration patterns [15], and yields as well, thanks to the disparities between feature points across different pictures, depth data [16] leading to a cloud of isolated, coloured 3D points.
2. From the calibrated cameras and the sparse 3D point cloud, a denser 3D model can be extracted, e.g., a polygon mesh to be multi-textured with the input pictures of step 1 [17, 18].
3. The initial 3D scene model obtained in the previous steps can then be progressively refined with similar material from other sources, like the web, where individuals or user groups will also upload new content, such as pictures taken from yet other viewpoints, or even synthetic 3D models.

Furthermore, the 3D tools shall be able to generate a description of the acoustic properties of the 3D scene which allows performing an adaptive auralisation of the environment for arbitrary positions of the listener [19]. First, the 3D tools provide a geometric description of the environment by exploiting the 3D model and define initial positions of all sound sources within the environment. Then, room acoustic parameters such as the reverberation time and early reflections are derived from the description of the 3D scene [20].

*Table 15 - 3D tools functional requirements*

| Req. ID | Name | Description |
|---------|------|-------------|
| F3T1 | Feature point extraction | The feature point extraction tool shall extract a set of robust feature points in several images corresponding to the same 3D scene. |
| F3T2 | Generation of matching pairs | The generation of matching pairs tool shall identify pairs of 2D points in different images known to correspond to the same 3D point in the related 3D scene. |
| F3T3 | Camera calibration | The camera calibration tool shall provide both, the intrinsic and extrinsic parameters of the user's camera which took particular pictures. |
| F3T4 | Reconstruction of isolated 3D point locations | The reconstruction of isolated 3D point locations tool shall provide sets of reconstructed isolated 3D points. |
| F3T5 | Dense mesh reconstruction | The dense mesh reconstruction tool shall create a dense 3D reconstruction starting from a sparse representation. |
| F3T6 | 3D model update | The 3D model update tool shall be able to update a precomputed 3D model by adding additional images/*videos*. |
| F3T7 | 3D model rendering | The 3D engine shall be able to render a 3D model through a sparse representation (e.g. point clouds) or dense representation (e.g. meshes). |
| F3T8 | 3D viewpoint synthesis | The 3D engine shall be able, starting from a 3D model of a scene, to create a synthetic 2D view of that scene from any possible viewpoint. |
| F3T9 | 3D model coding | The 3D model coding tool shall be able to *efficiently* compress/decompress sparse/dense 3D model representations. |
| F3T10 | Audio scene modelling | The audio scene modelling tool shall be able to generate a description of the position of the acoustic sources *and acoustic properties* of a 3D scene. |
| *F3T11* | *Audio scene projection* | *The audio scene projection tool shall be able to project the Audio scene model onto the scene geometry described by a 3D model* |

## 6.8   Application-related Presentation

This section defines the requirements of the functionalities exposed by the presentation layer of the BRIDGET applications.

*Table 16 - Application-related presentation functional requirements*

| Req. ID | Name | Description |
|---------|------|-------------|
| FAP1 | Bridget presentation | The presentation system shall present time-aligned and global |

| Req. ID | Name | Description |
|---|---|---|
| | | bridgets associated with a programme using the information provided by the synchronisation system. |
| FAP2 | Multiple bridgets presentation | The presentation system shall be able to present more than one bridget at the same time. |
| *FAP3* | *Bridget filtering* | *The presentation system shall provide a way to filter the retrieved bridgets based on type (time-aligned or global) or destination content type.* |
| *FAP4* | *Bridget filtering mode* | *The presentation system shall provide a way to select automatic or manual bridget filtering.* |
| *FAP5* | *Groups of bridgets* | *The presentation system shall provide a way to interact with a group of bridgets allowing collapsing or expanding the related items.* |
| *FAP6* | *Bridget bookmark* | *The presentation system shall provide a way to save a bridget for future use, consume and remove it* |
| FAP7 | Content presentation | The presentation system shall be able to decode and render destination content linked by a bridget. |
| FAP8 | Programme presentation and navigation | The presentation system shall provide at least one way to navigate the source programme timeline using information provided by the Media Analysis tools. |
| FAP9 | Search results presentation | The presentation system shall provide at least one way to present the content or the bridgets provided by the Search tools. |
| FAP10 | Multi-view scene on-line reconstruction | The presentation system shall provide scene rendering and reconstruction methods. |
| FAP11 | 3D model rendering – sparse | The presentation system shall be able to render 3D models through a sparse representation (e.g. point clouds) using the 3D engine. |
| FAP12 | 3D model rendering – dense | The presentation system shall be able to render 3D models through dense representation (e.g. meshes) using the 3D engine. |
| FAP13 | Binaural audio rendering | The presentation system shall be able to render audio tracks binaurally. |
| *FAP14* | *Enable/disable bridget presentation* | *The presentation system shall provide a way to enable or disable at any time the presentation of new bridgets.* |

## 7   Requirements of Applications

This section provides a map of the functionalities required by the BRIDGET applications that will be developed for the Phase 2 of the Project: the Professional Authoring Tool (PAT), the Consumer Authoring Tool (CAT) and the Player. For each of the 3 applications the list of relevant functionalities are provided. The list elements do not represent the status (implemented or otherwise) of the functionalities.

## 7.1   Professional Authoring Tool

| | Version B – Professional Authoring Tool functionalities |
|---|---|
| | **Bridget Representation** |
| FBR3 | Bridget source programme |
| FBR5 | Bridget types |
| FBR6 | Bridget destination content |
| FBR7 | Bridget destination content types |
| FBR8 | Bridget extended destination content types |
| FBR9 | Bridget destination content description |
| FBR10 | Bridget multiple sources for destination content |
| FBR11 | Bridget default icon |
| FBR15 | Bridget modifications |
| | **Content and Bridget Management** |
| FCM1 | Content ingestion |
| FCM2 | Content types |
| FCM3 | Content metadata |
| FCM4 | Content removal |
| FCM5 | Content publishing |
| FCM7 | Indexed Content |
| | **Synchronisation** |
| FSY1 | Synchronisation presentation delay |
| FSY2 | Detection accuracy |
| FSY3 | Bridget identification |
| FSY4 | Programme media time identification |
| FSY5 | Independence from main screen interactions |
| FSY6 | Automatic synchronisation |
| FSY7 | Synchronisation information production |
| | **Content Access** |
| FCA1 | Content delivery |
| FCA3 | CDN |

## 7.1   Professional Authoring Tool

| FCA4 | Bridget delivery |
|------|------------------|
| | **Media Analysis Tools** |
| FMA1 | Multiple content types |
| FMA2 | Coding independence |
| FMA3 | Coding awareness |
| FMA4 | Content duration independence |
| FMA5 | Content modification robustness |
| FMA6 | Content description self-containment |
| FMA7 | Content non-alteration |
| FMA8 | Real-time content description extraction |
| FMA9 | Faster than real-time content description extraction |
| FMA10 | On-the-fly content description extraction |
| FMA11 | Low-level structural segmentation |
| FMA17 | Fast filtering/ search |
| FMA18 | Unambiguous content description syntax |
| | **Search Tools** |
| FST1 | Visual search in image library |
| FST3 | Visual search in video library utilising video key frames |
| FST5 | Off-line indexing of image library |
| FST6 | Off-line indexing of video library based on video key-frames |
| FST7 | Ranking of the search results based on matching confidence |
| FST9 | Multiple content types |
| FST10 | Coding independence |
| FST11 | Content modification robustness |
| FST12 | Robustness to visual deformations |
| FST13 | Content description self-containment |
| FST14 | Content non-alteration |
| FST15 | Integration with content-based search |
| | **3D Tools** |
| F3T1 | Feature point extraction |
| F3T2 | Generation of matching pairs |

| F3T3 | Camera calibration |
|------|--------------------|
| F3T4 | Reconstruction of isolated 3D point locations |
| F3T5 | Dense mesh reconstruction |
| F3T6 | 3D model update |
| F3T7 | 3D model rendering |
| F3T8 | 3D viewpoint synthesis |
| F3T9 | 3D model coding |
| F3T10 | Audio scene modelling |
| | **Application-related Presentation** |
| FAP1 | Bridget presentation |
| FAP2 | Multiple bridgets presentation |
| FAP7 | Content presentation |
| FAP9 | Search results presentation |
| FAP10 | Multi-view scene on-line reconstruction |
| FAP11 | 3D model rendering – sparse |
| FAP12 | 3D model rendering – dense |
| FAP13 | Binaural audio rendering |

## 7.2   Consumer Authoring Tool

| | **Version B – Consumer Authoring Tool functionalities** |
|------|--------------------|
| | **Bridget Representation** |
| FBR3 | Bridget source programme |
| FBR5 | Bridget types |
| FBR6 | Bridget destination content |
| FBR7 | Bridget destination content types |
| FBR8 | Bridget extended destination content types |
| FBR9 | Bridget destination content description |
| FBR10 | Bridget multiple sources for destination content |
| FBR11 | Bridget default icon |
| FBR15 | Bridget modifications |

| | | |
|---|---|---|
| | **Content and Bridget Management** | |
| FCM1 | Content ingestion | |
| FCM2 | Content types | |
| FCM3 | Content metadata | |
| FCM4 | Content removal | |
| FCM5 | Content publishing | |
| FCM7 | Indexed Content | |
| | **Synchronisation** | |
| FSY1 | Synchronisation presentation delay | |
| FSY2 | Detection accuracy | |
| FSY3 | Bridget identification | |
| FSY4 | Programme media time identification | |
| FSY5 | Independence from main screen interactions | |
| FSY6 | Automatic synchronisation | |
| FSY7 | Synchronisation information production | |
| | **Content Access** | |
| FCA1 | Content delivery | |
| FCA3 | CDN | |
| FCA4 | Bridget delivery | |
| | **Media Analysis Tools** | |
| FMA1 | Multiple content types | |
| FMA2 | Coding independence | |
| FMA3 | Coding awareness | |
| FMA4 | Content duration independence | |
| FMA5 | Content modification robustness | |
| FMA6 | Content description self-containment | |
| FMA7 | Content non-alteration | |
| FMA8 | Real-time content description extraction | |
| FMA9 | Faster than real-time content description extraction | |
| FMA10 | On-the-fly content description extraction | |
| FMA11 | Low-level structural segmentation | |

| | |
|---|---|
| FMA17 | Fast filtering/ search |
| FMA18 | Unambiguous content description syntax |
| | **Search Tools** |
| FST1 | Visual search in image library |
| FST3 | Visual search in video library utilising video key frames |
| FST5 | Off-line indexing of image library |
| FST6 | Off-line indexing of video library based on video key-frames |
| FST7 | Ranking of the search results based on matching confidence |
| FST9 | Multiple content types |
| FST10 | Coding independence |
| FST11 | Content modification robustness |
| FST12 | Robustness to visual deformations |
| FST13 | Content description self-containment |
| FST14 | Content non-alteration |
| FST15 | Integration with content-based search |
| | **3D Tools** |
| F3T1 | Feature point extraction |
| F3T2 | Generation of matching pairs |
| F3T3 | Camera calibration |
| F3T4 | Reconstruction of isolated 3D point locations |
| F3T5 | Dense mesh reconstruction |
| F3T6 | 3D model update |
| F3T7 | 3D model rendering |
| F3T8 | 3D viewpoint synthesis |
| F3T9 | 3D model coding |
| F3T10 | Audio scene modelling |
| | **Application-related Presentation** |
| FAP1 | Bridget presentation |
| FAP2 | Multiple bridgets presentation |
| FAP7 | Content presentation |
| FAP9 | Search results presentation |

| FAP10 | Multi-view scene on-line reconstruction |
| FAP11 | 3D model rendering – sparse |
| FAP12 | 3D model rendering – dense |
| FAP13 | Binaural audio rendering |

## 7.3    Player

| | Version B – Player functionalities |
|---|---|
| | **Bridget Representation** |
| FBR3 | Bridget source programme |
| FBR5 | Bridget types |
| FBR6 | Bridget destination content |
| FBR7 | Bridget destination content types |
| FBR8 | Bridget extended destination content types |
| FBR9 | Bridget destination content description |
| FBR10 | Bridget multiple sources for destination content |
| FBR11 | Bridget default icon |
| | **Synchronisation** |
| FSY1 | Synchronisation presentation delay |
| FSY2 | Detection accuracy |
| FSY3 | Bridget identification |
| FSY4 | Programme media time identification |
| FSY5 | Independence from main screen interactions |
| FSY6 | Automatic synchronisation |
| FSY7 | Synchronisation information production |
| | **Content Access** |
| FCA1 | Content delivery |
| FCA3 | CDN |
| FCA4 | Bridget delivery |
| | **3D Tools** |
| F3T7 | 3D model rendering |

| F3T8 | 3D viewpoint synthesis |
|------|------------------------|
| F3T9 | 3D model coding |
| F3T10 | Audio scene modelling |
| | **Application-related Presentation** |
| FAP1 | Bridget presentation |
| FAP2 | Multiple bridgets presentation |
| FAP7 | Content presentation |
| FAP11 | 3D model rendering – sparse |
| FAP12 | 3D model rendering – dense |
| FAP13 | Binaural audio rendering |

## 8   Conclusions

The list of functional requirements organised and presented in this document has guided the development of the extended versions of the Professional Authoring Tool and the Player and will continue to play this role in the 2nd phase when also an instance of the Consumer Authoring Tool will be implemented.

## 9   References

[1]      BRIDGET, Requirements and Functionalities - Version A, Deliverable D2.1v1, 2014/03/31

[2]      Abran et al. - 2004 - Guide to the software engineering body of knowledge

[3]      IETF, RFC 2119 (1997-03), IETF, "Key words for use in RFCs to Indicate Requirement Levels", https://www.ietf.org/rfc/rfc2119.txt

[4]      ETSI, TS 102 796 1.2.1 Hybrid Broadcast Broadband TV, November 2012

[5]      HDFI/DGTVi, HD-Book DTT: "Compatible High Definition receivers for the Italian market", ver 2.1.1, October 2012

[6]      HDFI/Tivù, HD-Book SAT: "Compatible High Definition receivers for the Italian market", ver 2.0, October 2012

[7]      Tapu, R.; Zaharia, T., "A complete framework for temporal video segmentation,", 2011 IEEE International Conference on Consumer Electronics, vol., no., pp.156,160, 6-8 Sept. 2011

[8]      Jun Yang, Yu G. Jiang, Alexander G. Hauptmann, Chong W. Ngo, "Evaluating bag-of-visual-words representations in scene classification", Proceedings of the International Workshop on Multimedia Information Retrieval, pp. 197-206, 2007

[9]      Chikkerur, S.; Sundaram, V.; Reisslein, M.; Karam, L.J., "Objective Video Quality Assessment Methods: A Classification, Review, and Performance Comparison", IEEE Transactions on

Broadcasting, vol.57, no.2, pp.165-182, June 2011

[10]    Bernd Girod, Vijay Chandrasekhar, Radek Grzeszczuk, Yuriy A. Reznik, "Mobile Visual Search: Architectures, Technologies, and the Emerging MPEG Standard," IEEE Multimedia, vol. 18, no. 3, pp. 86-94, July-September 2011.

[11]    International Standard ISO/IEC 15938-13 Compact Descriptors for Visual Search

[12]    ISO/IEC JTC 1/SC 29/WG 11, w50140 - Compact Descriptors for Video Analysis: Requirements for Search Applications

[13]    R. Hartley and A. Zisserman, "Multiple View Geometry in Computer Vision", Cambrigde Universitity Press, 2003.

[14]    N. Atzpadin, P. Kauff, and O. Schreer, "Stereo Analysis by Hybrid Recursive Matching for Real-Time Immersive Video Conferencing", IEEE Trans. on Circuits and Systems for Video Technology, Special Issue on Immersive Telecommunications, vol. 14, no. 3, pp. 321-334, Mar. 2004.

[15]    P. Carballeira, J.I. Ronda and A. Valdés, "3D reconstruction with uncalibrated cameras using the six-line conic variety", Proc. IEEE Intl. Conf. on Image Processing (ICIP'08), pp. 205-208, Oct. 2008.

[16]    W. Waizenegger, I. Feldmann, and O. Schreer, "Real-time Patch Sweeping for High-Quality Depth Estimation in 3D Videoconferencing Applications", Proc. SPIE Conf. on Real-Time Image and Video Processing, vol. 7871, pp. 78710E-10, Jan. 2011.

[17]    A. Smolic, K. Müller, P. Merkle, T. Rein, M. Kautzner, P. Eisert, T. Wiegand, "Free Viewpoint Video Extraction, Representation, Coding, and Rendering", Proc. IEEE Intl. Conf. on Image Processing (ICIP'04), vol. 5, pp. 3287-3290, Oct. 2004.

[18]    F. Morán (ed.), ISO/IEC 14496 16:2011/Amd.1:2011, "Animation Framework eXtension (AFX)", Amendment 1, "Efficient Representation of 3D Meshes with Multiple Attributes", ISO, 2011.

[19]    Michael Vorländer, "Auralization - Fundamentals of Acoustics, Modelling, Simulation, Algorithms and Acoustic Virtual Reality", Springer Verlag, Berlin, 2008.

[20]    J. H. Rindel, "Modelling in auditorium acoustics: From ripple tank and scale models to computer simulations", Forum Acusticum, Seville, Revista de Acústica XXXIII, no. 3-4, pp. 31-35, 2002.

[21]    BRIDGET, Requirements and Functionalities - Version A, Revised version for Y1 review, D2.1 rev, January 2015

[22]    BRIDGET, Second BRIDGET Workshop and Use Scenarios, D2.4, November 2015

[23]    ISO/IEC 23000-18 DIS Media Linking Application Format (October 2015)