Deliverable

Project Acronym:	**VRTogether**
Grant Agreement number:	762111
Project Title:	*An end-to-end system for the production and delivery of photorealistic social immersive virtual reality experiences*

D5.1- Market analysis report.

Revision: 1.0

Authors: Pascal Perrot (Viaccess-Orca), Laurence Leopold (Viaccess-Orca), Romain Alcaras (Viaccess-Orca)

Delivery date: M12 (09-18)

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement 762111

Dissemination Level		
P	Public	x
C	Confidential, only for members of the consortium and the Commission Services	

Abstract: This deliverable describes the market for audio-visual immersive products and provides a set of recommendations as input for the use cases and the exploitation and innovation transfer activities in T.5.1.

REVISION HISTORY

Revision	Date	Author	Organisation	Description
0.1	11 April 18	Pascal Perrot	Viaccess-Orca	First Draft
0.2	29 may 18	Laurence Leopold	Viaccess-Orca	Second Draft
0.3	18 June 18	Laurence Leopold	Viaccess-Orca	Third Draft
0.4	3 July 18	Romain Alcaras	Viaccess-Orca	Fourth Draft
1.0	16 August 18	Romain Alcaras	Viaccess-Orca	First Release
1.1	20 Sept 18	Pablo Cesar	CWI	Internal Review
1.2	24 Sept 18	Tom De Koninck	TNO	Review
1.2.1	26 Sept 18	Pascal Perrot	Viaccess-Orca	Review and update
1.2.2	02 Oct 18	Laurence Leopold	Viaccess-Orca	Re-organise and re-write section 2
1.2.3	10 Oct 18	Pascal Perrot	Viaccess-Orca	Review and update section 5
1.2.4	19 Oct 18	Tom De Koninck	TNO	Rewriting of sections 4.1 and 4.3
1.2.5	24 Oct 18	Marc Brelot	MSE	Internal second review

CONTRIBUTORS

First Name	Last Name	Company	e-Mail
Laurence	Leopold	Viaccess-Orca	laurence.leopold@viaccess-orca.com
Pascal	Perrot	Viaccess-Orca	pascal.perrot@viaccess-orca.com
Romain	Alcaras	Viaccess-Orca	romain.alcaras@viaccess-orca.com
Marc	Brelot	Motion Spell	marc.brelot@gpac-licensing.com
Romain	Bouqueau	Motion Spell	romain.bouqueau@gpac-licensing.com
Argyris	Chatzitofis	CERTH	tofis@iti.gr
Spyridon	Thermos	CERTH	spthermo@iti.gr
Francesca	De Simone	CWI	Francesca.De.Simone@cwi.nl
Tom	De Koninck	TNO	tom.dekoninck@tno.nl
Sergi	Fernandez	I2CAT	Sergi.fernandez@i2cat.net

TABLE OF CONTENT

TABLE OF FIGURES

LIST OF ACRONYMS

Acronym	Description
AR/VR/MR	Augmented Reality / Virtual Reality / Mixed Reality
B2B	Business to Business
B2C	Business to Consumer
B2B2C	Business to Business to Consumer
CAD	Computer-Aided Drawing
CAGR	compound annual growth rate
DoF	Degrees of Freedom
FoV	Field of View
GPU	Graphic Processing Unit
HMD	Head Mounted Displays
MMO	Massively Multiplayer Online
Mo-Cap	Motion Capture
MPEG	Moving Picture Experts Group
ODMs	Original Design Manufacturing
OEMs	Original Equipment Manufacturing
RGB-D	Red Green Blue - Depth
RPG	Role Playing Games
WebRTC	Web Real-Time Communication This is a free open-source project

1 INTRODUCTION

1.1 Purpose of this document

The objective of this document is to analyze the current and expected future market opportunities for immersive audio-visual products by:

- Shedding light on changing behaviors and associated expectations in audio-visual consumption.
- Assessing the market potential of the solutions developed within VRTogether both in terms of enhancing existing formats and in creating new types of content.
- Identifying technological trends, potential competing platforms and solutions as well as complementary stakeholders.
- Identifying the targets (B2C, B2B, etc.) of the solution and how they could be successfully exploited.

1.2 Status and final scope of this document

This document will be alive during the whole project period, that is, during the 3 iterations of the project. Three different versions will be formally submitted to the European Community and uploaded in the project website.

This first version will focus on providing an in-depth overview of the immersive audio-visual market.

The second and third versions will focus on the presentation of the VRTogether's solution, the enabled use cases and on a set of recommendations to ensure a successful launch on the market. The document will particularly highlight:

- How the solution could or should be transformed into a viable product (value proposition, functional set-up, value network and financial model).
- The conditions for a successful exploitation of the proposed solution and the associated optimal go-to-market strategy

1.3 Relation with other VRTogether activities

The objective of the Work Package 5 is to maximize the impact of the VRTogether platform on the audio-visual market and research domain. The results of the project are expected to have a significant impact on the audio-visual market. This WP will:

- Determine overall conditions for successful exploitation of the proposed solutions, such as standardization, additional stakeholder involvement, etc.
- Identify optimal go-to-market strategies of products using VRTogether technologies. The project should take into account specific requirements of each of the stakeholders, and be translated in feasible and viable exploitation plans.
- Organize maximal visibility for the proposed solutions by attending major relevant events and setting up direct contacts with potential customers.
- Implement a communication strategy aligned with the exploitation strategies of the consortium partners.

2 VR MARKET OVERVIEW

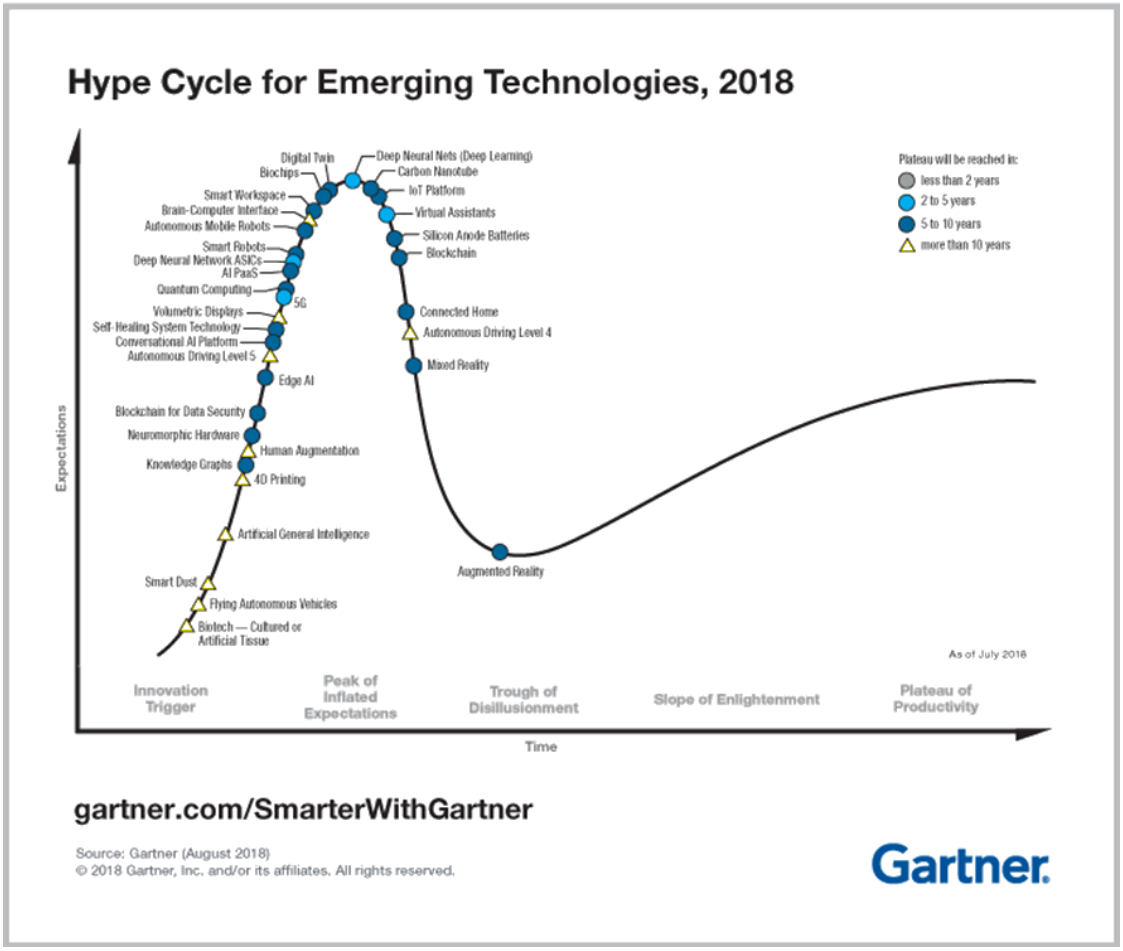
2.1 VR/AR hype cycle

Virtual reality was highlighted last year in Gartner's hype cycle for emerging technologies.

This year, Virtual reality does not appear in Gartner's cycle anymore as it has already proven many scenarios (in addition to gaming) where it gets employed with success in the Enterprise world like training, education, video collaboration, etc…

Instead, Gartner highlights Augmented Reality content and apps which show an even broader promise since mobile AR has many millions people that can use it.

Figure 1: Hype Cycle for Emerging Technologies, *Gartner*[1] *(2018)*



[1] GARTNER. Hype-Cycle 2018 **[online].** (published August 16, 2018) Available on : https://www.gartner.com/smarterwithgartner/5-trends-emerge-in-gartner-hype-cycle-for-emerging-technologies-2018/

2.2 VR/AR revenue forecasts

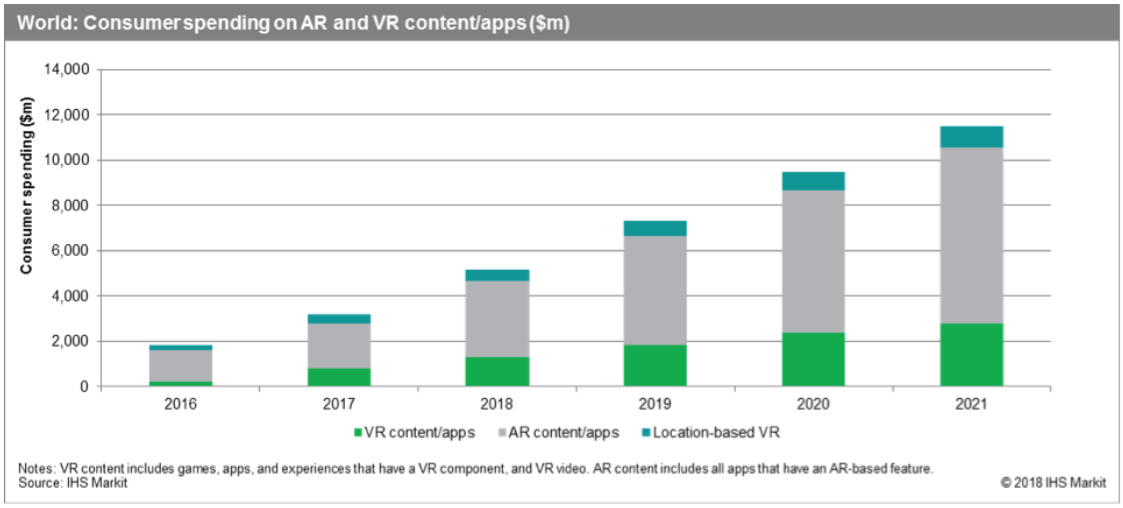
The global market for **consumer** AR and VR content and apps grew by 72%, reaching $3.2 billion in 2017, according to IHS Markit

World consumer spending on **VR games**, interactive experiences and video reached $803 million in 2017. IHS Markit forecasts that spending will reach $2.8 billion by 2021.

Consumer spending on **location-based VR** reached $385 million in 2017. This is forecast to grow to $906 million by 2021.

Consumer spending on all **AR-related mobile** apps is forecast to grow from $1.6bn in 2017 to $7.8bn in 2021.

Figure 2 : Consumer spending on AR and VR content/apps ($m) - *IHS Markit [2](April 2018)*

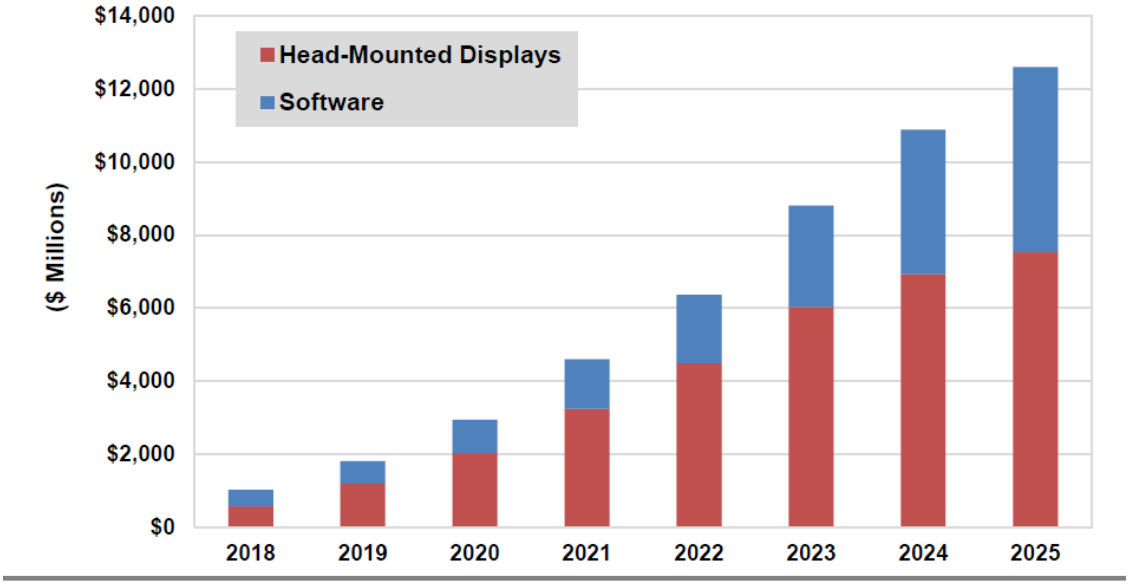


According to the analyst Tractica, the VR hardware and software market for **Enterprise** is expected to reach $ 4.6 billion in 2021. This market targets (see figure 3):

- education
- training
- medical therapy
- virtual prototyping / 3D modeling
- location-based entertainment

[2] IHS MARKIT. Augmented and Virtual Reality Consumer Content and Apps Market Hits $3.2 Billion in 2017, IHS Markit Says **[online].** (published April 10, 2018) Available on : https://technology.ihs.com/601858/augmented-and-virtual-reality-consumer-content-and-apps-market-hits-32-billion-in-2017-ihs-markit-says

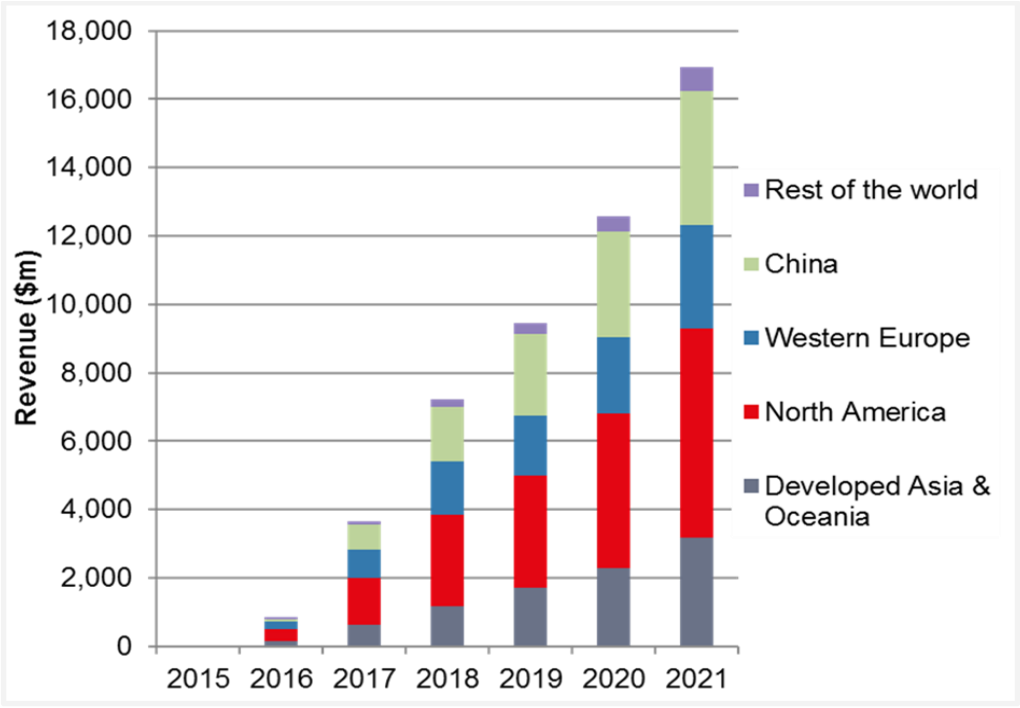
Figure 3 : Enterprise VR hardware and software revenue by segment, World markets: 2018-2025 *Tractica (2Q 2018)*



(Source: Tractica)

According to Ovum (August 2017), North America will generate the highest VR revenues in the next 3 years – driven by high-spending gamers with dedicated VR headsets (see Figure 4). China, while seeing massive VR interest and headset adoption, will lag behind in content revenues due to the low spending habits of consumers and a lack of marketplaces.

Figure 4: Revenue of past, current and future VR market by regions, Ovum (August 2017)



2.3 VR market adoption

Making money with VR content by targeting directly consumers will continue to be a challenge.

Indeed, with a limited adoption of high-end PC and VR headsets so far, many independent VR companies targeting directly consumers have found difficult to obtain a return on their investment. For instance AltspaceVR[3] – one of the leading pioneers of social VR – was finally sold to Microsoft.

Except for large social networks like Facebook and YouTube that integrates VR content as part of a global advertising-based business model, all other companies, it's unlikely that any significant advertising revenue can be generated. Advertising is still a long way from being viable as the installed base of VR headsets does not provide a sufficient audience.

Nevertheless, a continued but moderate headset adoption is expected over the upcoming years, as hardware prices drop, technology improves, content becomes more appealing and business usages develop.

The following table present a first inventory of the SWOT matrix (Strengths, Weaknesses, Opportunities and Threats) regarding the VR .

Figure 5: VR SWOT Analysis

Opportunities ● HMD : wireless mobility, resolution increasing ● PC, TVM encoding and distribution technologies more mature ● Capture technologies more mature ● Video 360 capture device more mature and affordable ● Use of mobile phone as HMD ● Larger IP bandwidth available	**Threats** ● Consumers sustainably VR adoption? ● Headset acceptation? (still low resolution, heavy, expensive, not good looking) ● Risks of psychological deviations and sickness
Strengths ● VR rebuilds the reality ● Efficiency of VR for education and training has been proven ● Helps people with disabilities & older people to develop their social interactions ● Enables cost savings for companies (fewer travels & logistics, less storage, time saving,..)	**Weaknesses** ● Not enough VR content today ● lake of synchronization and realism ● Poor self-representation / avatar ● Lack of interoperability in the ecosystem

Regarding this SWOT table, the VRTogether project has a very interesting positioning since it aims at offering **photorealistic immersive virtual reality** content which can be experienced together with friends or colleagues, and demonstrate its use for domestic and business VR

[3] https://altvr.com/

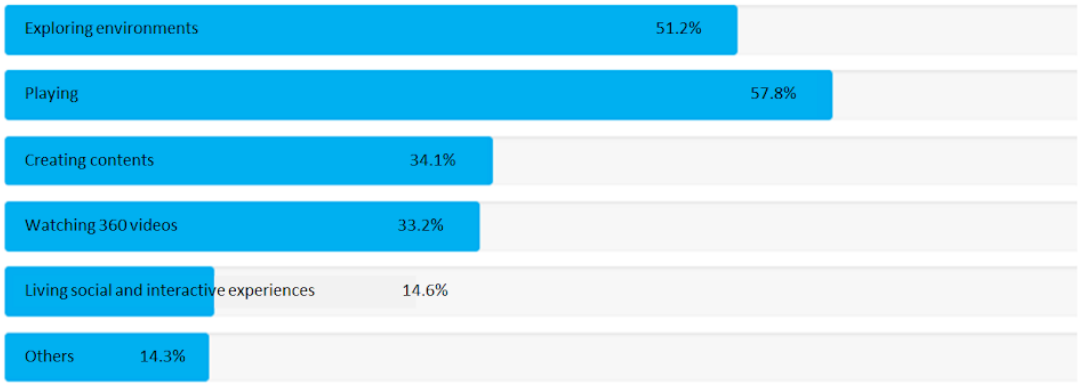
consumption. Then, VRTogether project will try to take advantage of the new opportunities to get ride (or lower) the weaknesses of the VR.

2.4 VR usage

Sketchfab[4], a digital community for creating 3D and virtual reality content, has released a report in Q2 2017 on the state and the future of VR. This report is based on data from its platform of over 700,000 contributors and one million scenes in VR.

The figure below measuring the VR usage confirms - like many other sources - that gaming is the first usage. In addition, social VR is clearly visible (14%) among other VR usages.

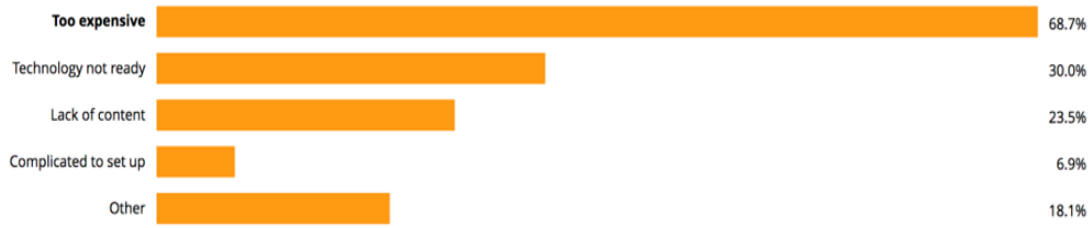
Figure 6: Main VR usages, Sketchfab (Q2 2017)



While most VR experiences are currently single-user, some recent development has promoted social experience such as Ubisoft's Star Trek: Bridge Crew[5] or the Ciné VR[6] Social VR Movie Theatre application. The ability to communicate with others through VR has already proved to be very popular through co-operative games and social networks. But social VR users are always represented as artificial avatars. And this might not be appropriate for many communication settings such as business meetings, or sharing experiences with family or friends.

According to Sketchfab, the majority (54%) of the interviewed population believes that VR has potential while some barriers will have to be addressed.

Figure 7: Reasons for non-adoption of the VR technology, *Sketchfab*



To understand users' expectations, VRTogether's project teams led a survey at the VR DAYS 2017 exhibition in Amsterdam to understand user's expectations in social VR.

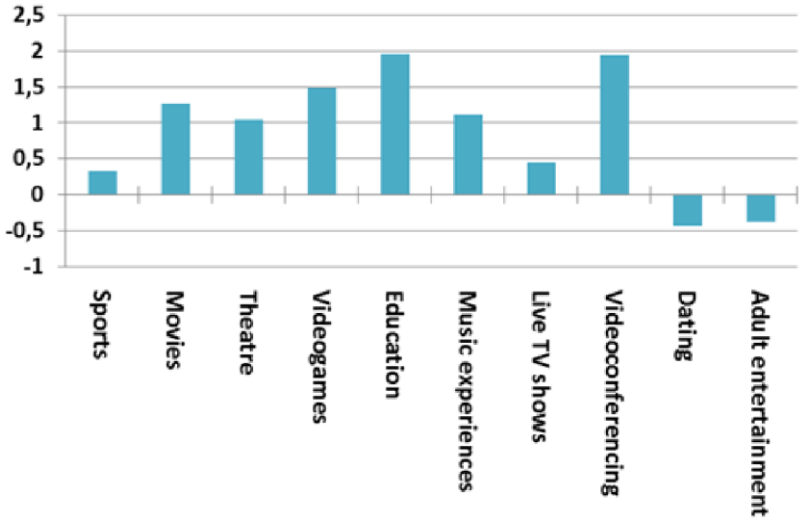
[4] https://sketchfab.com/
[5] https://www.ubisoft.com/en-gb/
[6] https://cinevr.io/

The participants to our survey had an interest in VR or/and had experience using VR applications. Before filling out the questionnaire, they were asked to try out our social VR demo so that they had an understanding of our social VR concept.

47% of the participants expressed that they were extremely interested in Social VR experiences.

When they were asked what, for them, the important factors in social VR experiences would be, "interaction within the experience" and "enjoyment of overall experience" seemed to be considered extremely important by more than half of them.

Figure 8: Users interest in Social VR for different application contexts



In Figure 7 above, we see that the highest interest is shown for video conferencing and education applications, followed by video games, movies and music experiences.

We are still at the initial stages of creating social VR experiences for general uses. Further experiments will be performed in order to evaluate the most relevant use cases and, consequently, define the appropriate enabling technologies.

2.5 VR applications and platforms

Since the outset of virtual reality, the VR user experience has been a very solitary one. As a consequence some developers have begun to include a social dimension to their content and to diversify the usages: games, content platform, social Networks, sports, music & arts, education, tourism and of course business applications.

This section is divided first present the VR for the consumer market, secondly VR for business and then a focus on Social VR.

2.5.1 VR for the consumer market

2.5.1.1 Games

According to many market researches, games are the main application of VR. Even if they are mostly single player games without any social interaction, some game developers have begun to include this dimension, or even encompass all their games around that.

The VR Gaming market in 2016 was estimated at $3.6 billion and is expected to reach $22 billion by 2020, according to market research from Statista[7].

Multiplayer VR is a growing feature which is due to consumer desire to pay for sociable experiences with friends. We can segment the game market into three categories:

- The first category of games is the **MMO game** adapted to social VR. It is an online game with large numbers of players, typically from hundreds to thousands, on the same server. In order to illustrate this category, there are two well-known games :
  - **REC ROOM[8]**, from the studio *Against Gravity*, launched in June 2016. The game uses full 3D motion via the motion capture system of a virtual reality headset and two hand-held motion controllers. Players can explore the space around them within the confines of their physical floor-space, while roaming further by using controller buttons to teleport a short distance, with minimal or no virtual reality sickness. Motion controllers are required to pick up and handle objects in the game world, including balls, weapons, construction tools, and other objects. Rec Room consists of separate multiplayer games, including cooperative action role-playing games, first-person shooters, a charades-like game, and four sports games (Soccer, Paddleball, Disc Golf, Dodgeball, 3D Charades, Paintball, Laser Tag, Quests, Rec Royale). [CTA: liste a puces + nom des jeux en face de chaque categorie].
  - **VRChat**:[9] it offers an endless collection of social VR experiences by giving the power of creation to its community. It is the most played game in social VR with more than 3 million players in 1 year. However the experience is not realistic.
- The second one is **RPG game** adapted to social VR: it is a game in which players assume the roles of characters in a fictional setting. Players take responsibility for acting out these roles within a narrative, either through literal acting or through a process of structured decision-making of character development. The best example of RPG game in VR is **The Elder Scrolls V: Skyrim VR[10]**, but there is no social interaction. Conversely, a good example of RPG Social VR game is **Star Trek: Bridge Crew[11]**, it allows the user to fulfil a storyline in a starship with real-time discussion and interactions inside the cockpit.
- Finally, there are the social games which are similar to a social network; we will simply call them **social game**. The only goal of this type of game is to meet new people and interact with them in a different way (detailed in 4.5.5). It also enables users to attend free live events, comedy show, dance party, and play interactive games with friends in a virtual scene with avatars. In this category, three games are very popular: AltspaceVR, High Fidelity, and Sansar.
  - **AltspaceVR[12]**: Both mobile and PC game. Enables to attend free live events, a comedy show, a dance party, and play interactive games with friends in virtual scene with avatars. Audio conversation is possible between the avatars in the same environment. However in this case also the experience is not realistic.
  - **High Fidelity[13]**: This game allows the user to go in virtual space hopping, attend events, shop the Marketplace, build worlds of any size, use any kind of avatar, create complex physical objects, import from standard formats, and share with others through the High Fidelity Marketplace, and check what others are up to in the "Metaverse". High Fidelity raises $35 million with the goal of bringing VR to 'a billion people'. The platform leverages blockchain technology to enable VR developers to create experiences that can be easily

[7] https://www.statista.com/statistics/499714/global-virtual-reality-gaming-sales-revenue/
[8] https://www.againstgrav.com/rec-room/
[9] https://vrchat.net/
[10] https://bethesda.net/en/article/4YSRrlXtpuKeWIK6EMiG84/skyrim-vr-comes-to-steamvr
[11] https://www.ubisoft.com/en-gb/game/star-trek-bridge-crew/
[12] https://altvr.com/
[13] https://highfidelity.com/

deployed across platforms that also adhere to the open, transparent, and secure ledger technology.

- **Sansar**[14]: it is a social virtual reality platform developed and owned by Linden Lab and launched in "creator beta" to the general public on July 31, 2017. The platform enables user-created 3D spaces where people can create and share interactive social experiences, such as playing games, watching videos, and having conversations in VR. Each participant is represented by a detailed avatar that is the graphical representation of the user including speech-driven facial animations and motion-driven body animations.

Figure 9: Sansar in-game capture, *Linden Labs*

2.5.1.2 TV & movie distribution

In the premium TV & movie arena, several Pay-TV operators and content owners made VR trials:

- **Orange**: In October 2016, Orange launched its own-brand virtual reality VR headset. The VR headset is only one of the VR value chain activities in which Orange is involved. The operator has either developed solutions internally or has invested in companies that provide capabilities across the entire VR content creation, distribution, and consumption value chain. The operator has identified the potential of VR to strengthen the value proposition posed by its combination of network, cloud, TV, DRM, and live events solutions.
- **SK Telecom**: in South Korea, enabled their customers to invite other peers in the virtual world to share and jointly watch full HD content together.
- **FOX** N.Y.-based tech startup LiveLike has partnered with Fox Sports to produce the 2018 World Cup for fans in social VR with all 64 games during the tournament available through the Fox Sports VR app for free. The production also includes sponsorship integration with Adidas.
- **CNN Broadcaster**: CNN has created a dedicated section on their website to provide high quality 360 Video content to VR users
- **NBC** provided more than 50 hours of live VR coverage powered by Intel True VR during the 2018 Olympic Winter Games .
- **SKY** pay-TV operator. SKY has launched its own SKY VR app allowing its customers to enjoy immersive content produced by Sky and its partners. The SKY VR app is compatible with 360° video player and can be used with HMD or directly on smartphones

[14] https://www.sansar.com/

Most of the above mentioned actors eg Orange or NBC use a white label media platform to stream their VR content.

For example Livelike[15]: this is a white label platform that allows TV broadcasters to stream regular or social VR content under their own brand with or without a VR headset.
Livelike has many famous customers: La Liga, Fox Sports, Roland Garros, BT, Audi, Buffalo Wild Wings, Samsung, UEFA, francetv sport, Jeep, Sky, and Manchester City.

Figure 10: Capturing a software using Livelike[16]

2.5.1.3 Sports

VR is revolutionizing sports in nearly every possible field, ranging from consuming sports content to training and recruiting athletes

In 2017, the chipmaker Intel has broadcast Major League Baseball, NCAA tournament and PGA Tour events in VR; Boston Red Sox and Detroit Tigers have installed VR batting cage experiences among other fan entertainment uses; and Facebook announced it was developing a VR app called Venues that lets users watch live sports and concerts in a shared environment.

Italian Soccer Juventus unveiled in October 2017 its new Virtual Reality App. Juventus VR is also a crucial way of appealing to a younger demographic of fans, who enjoy interacting with their team increasingly in a digital context. “These days, watching matches on TV is not enough for clubs to grow their fan base.” said Federico Palomba - Juventus’ co-chief revenue officer and marketing head[17].

Fox Sports has provided, with their partner Livelike, VR content and 360-degree “virtual suite” overlooking a sporting event. The Social VR Sports Startup LiveLike raised $9.6 Million. The funding round brings LiveLike, founded in early 2015, to about $15 million raised to date. The New York-based company has developed a live-streaming platform that uses VR, as well as augmented reality and mixed reality, to produce an interactive, 360-degree “virtual suite” overlooking a sporting event for users to connect with friends or fellow fans via Facebook to

[15] http://www.livelikevr.com/
[16] LIVELIKE. Website **[online]**. Available on : http://www.livelikevr.com/
[17] https://www.sporttechie.com/juventus-virtual-reality-app-facebook-oculus-google-cardboard/

watch the game. Users can select different camera angles and access stats, highlights and replays. Fans can view broadcaster content with or without a VR headset.

In 2017, Orange experienced Holotennis enabling its customers to virtually play tennis with their friends on the main Roland Garros tennis court. Orange built the solution with the help of start-ups Mimesys and Emissive; it also used Unity's player app.

Sports performance training has also been gaining a foothold for years through companies such as STRIVR[18] and EON Reality[19].

2.5.1.4 Music & arts

Oculus Venues[20] provides a large cultural VR programming on both mobile and PC platform, such as concerts, sporting events, comedy shows and more. Thousands of viewers meet each other, or watch in solo mode from a box seat high above the crowd.

The WaveVR has been able in April 2018 to raise $6 million in venture funding for its social VR platform and community for VR music and visual arts experiences. TheWaveVR is a platform for people who love music, enabling them to view, host, and socialize in shows worldwide, anytime, anywhere.

A similar platform is VRTIFY. Currently more and more content is presented in 360°, like music videos and museum and art gallery walk-throughs. However this is lacking interaction with the presented content.

2.5.1.5 Education

Figure 11: Example of educational usage of VR, *The Synapse*[21]

VR is more and more used in the educational domain.

[18] STRIVR. Website **[online]**. https://www.strivr.com/

[19] https://www.eonreality.com/

[20] https://www.oculus.com/experiences/go/1555304044520126/

[21] THE SYNAPSE. Adopting Virtual Reality for Education **[online]**. (Published on June 20, 2015) Available on : https://medium.com/synapse/adopting-virtual-reality-for-education-401e7456c7a5

- Virtual fields trips: The students can discover many environments in the world and even further in space with the solar system.
- Medical studies: brain surgery and virtual anatomy. The Royal London Hospital, has conducted the first brain surgery on an aneurysm recorded in virtual reality. Working with FundamentalVR, the team used both VR 360-degree cameras in the operating theatre and GoPros strapped to the heads of the surgeons to capture the surgery procedure. FundamentalVR provides medical personnel with simulation solutions enabling them to rehearse, practice and test themselves
- Content creation: It allows students to create environments and structures and then explore them with a VR Headset.

2.5.1.6 Tourism

The sense of sight is very important for the travel industry as the consumers appetite for buying travel services depends on visual stimulation. Several businesses in tourism have already picked up the VR technology to design virtual tours. However, these, to the most part, consist of stitched panorama photos of the premises in 360 video captures. This is essentially not true VR because there is only limited, if at all, ability for the customer to navigate, and no possibility to interact with the content shown. These applications are considered VR-type applications, demonstrating that there is an interest in the industry to use VR-based instruments to promote products and services. However, they do not fulfill the necessary characteristics to be considered true VR yet.

One of the main benefits of using VR systems is the fact that clients can 'sample' beforehand a destination. Destinations can provide some previews for clients to enjoy at fairs, in travel agencies or even in their respective homes. Experiencing these snapshots of a destination aids the customers in making an informed decision. Even though the client might not have decided yet, images of the viewed destination stay in the clients' mind and possibly induce a desire to visit the destination in real life.

2.5.2 VR for businesses

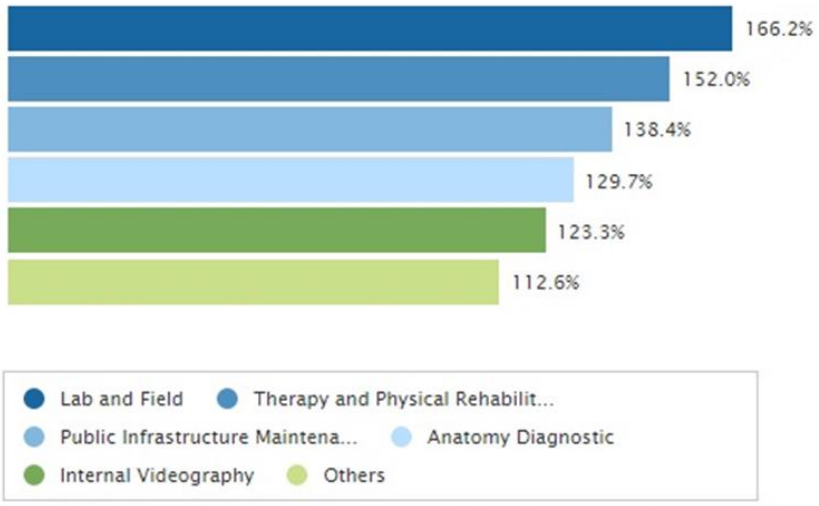
According to IDC's "Worldwide semiannual AR and VR spending guide 2016 H2", the fastest growing business use cases are lab & field experiments; then engineering & design, VR Business Conferencing, education, remote assistance, medical/health care sector, Biomechanical Research (see figure 11).

With a few improvements related to co-spacing abilities and headset deficiencies, VR technology could provide companies with a full range of services including the cases shown on the above graphic and hereafter:

Participants in a business meeting could manipulate 3D models, files and data in the VR space. They could meet virtually in a videoconference to discover a prototype (medical equipment, luxury watch, ..); they could even transfer each other the prototype files to print the object on a 3D printer.

Enterprises could also be served with other applications such as telepresence, architecture, healthcare, engineering, education, tourism).

Figure 12: Top use case based on 5 year CAGR (2016 - 2021) (Value (Constant Annual)), *IDC [22](2016H2)*



2.5.2.1 The relevant industries

- **Car manufacturers**: VR is expected to quickly revolutionize the car production by turning 3D drawings into VR prototypes or providing a real time and personalized support to workers (source Volkswagen). In addition, VR will improve the efficiency of the distribution network avoiding the distributor to stock cars locally for demonstrations. Car repair, and particularly car accident repair industry will benefit from VR training. Indeed the latter enables a quicker acquisition of the employees' repair skills, as the repair methods constantly evolve and the solutions to train mechanics need to quickly adapt.

Figure 13: Example of VR usage in the manufacturing industry, *Harman a SAMSUNG company[23] (March 2018)*

[22] IDC. Worldwide Spending on Augmented and Virtual Reality Expected to Double or More Every Year Through 2021, According to IDC **[online]**. (Published August 3, 2017) Available on : https://www.idc.com/getdoc.jsp?containerId=prUS42959717

[23] HARMAN. ARand VR – e Rise of the Two Spectrums of Immersive Technology. (Published March 23, 2018) Available on : https://services.harman.com/blogs/ar-and-vr-rise-two-spectrums-immersive-technology

- **Engineers and Designers**: The 3D modeling can be done inside the application and presented to other members of the group. Visualization of the models is more accurate because VR allows for real-life scale compared to computers, and everyone participating can interact with the model. The company middleVR[24], has developed such a tool called: improov3[25], a virtual collaborative VR teleconferencing system for CAD teams, which is used by NASA and CNES France for rocket development.

- **Health care** : The Global Augmented and Virtual Reality in Healthcare Market was valued at USD 510.36 million in 2016 and is projected to reach USD 8,313.29 million by 2025, growing at a CAGR of 36.35% from 2017 to 2025 (market insights report).

  Healthcare industry is the highest adopter of VR. The advantage of VR in healthcare is that it allows healthcare experts to learn new skills and update existing skill sets in a safe environment. VR is used in various areas of healthcare that ranges from diagnosis, treatment medical education and training, surgery using robots, telemetering, counseling and more.

  For example behavioral health care includes disorders such as depression, treatment of phobia and anxiety is interesting to explore. The global market of behavioral healthcare is valued at $75.40 billion in 2016, and will rise to $88.10 billion in 2022, at a Cumulative Annual Growth Rate (CAGR) of 2.6%. Behavioral healthcare is struggling due to the lack of skilled personnel and sufficient infrastructure (1% of the global healthcare workforce can deal with behavioral health issues). VR and AR is actually offering a perfect solution to face this lack of skilled personnel by providing digital health technology (HMD, application and contents).

  There are many other fields in health care where VR is more and more used:

  VR Medical Training: VR enables medical students to acquire in-depth understanding about human anatomy by means of interaction within a virtual environment. Students can perform hands-on procedures in a safe and controlled way where they are able to learn from mistakes.

  VR Dentistry: Dentistry is an area where VR will thoroughly impact its subjective and practical training. VR can be used to perform a range of procedures, e. g. filling using virtual

[24] miidleVR: https://www.middlevr.com/home-2/
[25] improove3 : https://www.improovr.com/home-v2/

drills which replicates the pressure and precise movement of a real drill by means of force feedback.

VR Preventive Medicine: VR can be used to educate patients with positive lifestyle choices such as quitting smoking, "moderating" alcohol intake, healthy eating, and consecutive exercise.

VR Counselling: VR has been widely used in areas of psychological counseling, treating fear psychosis and phobias such as fear of public speaking, amongst others. This is done by setting up virtual environments to build confidence amongst patients and address concerns related to fear.

VR surgery: VR can be used to help surgeons in surgical operations (tele-operation through robots etc.). In some applications, 3D images can be converted into virtual holograms enabling a more detailed control and in-depth surgery (source Plain Concepts tech vendor).

- **Architecture & construction**: Virtual reality is increasingly used in the field of architecture and design because it allows photo-realistic 3D modeling while having a real-scale point of view by moving inside the structure using the VR headset. Beyond that, real estate agencies can use VR for visits to apartments and homes without the customer having to move from home and without wasting time to visit, thus increasing productivity.

  The company REinVR[26] creates stunning visuals and interactive animations for architects. Using cutting-edge technology developed for the video game industry and rendering engine, REinVR develops photo-realistic experiences and immersive visual environments that help architects for designing, marketing, and sales. The REinVR's team consists of many talented designers who have spent their careers making some of the most innovative video games ever created. This gives REinVR a technological advantage over its competitors due to the speed at which changes can be made. With a simple click, a floor plan layout can be changed and different color palette combinations can be created for interior design.

2.5.2.2 Solution providers for businesses

Hereafter are the main providers of VR solutions for businesses; they all provide social VR features enabling companies to increase their productivity by simplifying trainings and meetings of the staff all around the world.

- **Innoactive Hub[27]** : This software is very complete and offers three different services :
  - VR Training: An employee can train in virtual environment prototyped by the company itself with a toolkit. It allows the employees to train wherever they are in the world.
  - VR Collaboration: Collaborative VR allows teams to work together in a completely new way. If everyone can join a VR session from their workplace, from home or a meeting room somewhere in the world, the team can meet and visualize relevant 3D data from anywhere. It's a VR business meeting.
  - VR Showrooms: Virtual shopping experience.
- **Mimesys[28]**: The Mimesys platform is a one-stop solution to puts real people and their interactions into any application. Mimesys has developed a complete pipeline, capturing real people in real-time using an off-the-shelf depth sensor, transmitting it over their cloud platform to anywhere in the world. They combine this true 3D 'holographic' representation with natural interactions on shared virtual objects to give to the user an unparalleled sense

[26] https://www.reinvr.com/
[27] https://innoactive.de/hub/
[28] http://www.mimesysvr.com/

of co-presence. This software is the only one to offer photorealistic virtual business meetings.

- **Bigscreen**[29]: The main focus of this app is to help people collaborate and allow friends to hang out together. It can allow colleagues to meet and converse in VR. This can help them share ideas and work on projects together without leaving the comforts of their home.
- **Doghead Simulations' rumii**[30]: Rumii is similar to Bigscreen but is more focused on business and productivity. Doghead's environment is made for official office work rather than for friends.
- **Improov**[31]**:** This app focuses on the designers. Improov focuses more on CAD and BIM data rather than charts and graphs. The app can be used individually or in groups. The 3D modeling can be done inside of the app and presented to other members of your group. The models tend to be more accurate because VR allows for real-life scale compared to computers.
- **Cisco Spark**[32]: Cisco Spark's main focus is to help team members work together even if they are miles apart with group chats and sharing files.
- **meetingRoom**[33]**:** meetingRoom proposes a solution allowing remote teams to use digital rooms to collaborate and get face to face from anywhere.
- **Objective Theory**[34]: it is an MR software company focused on developing software for HoloLens and Windows Mixed Reality headsets. They created collaboration software that allows several people to work together from different locations.
- **VR Medusa**[35]: Business communication platform, developed by TeleSoftas and offering connection between multiple individuals, VR Medusa provides immersive and effective communication from any part of the world inside virtual reality.

2.6 Focus on social VR

Social VR enables people getting together in a virtual world, be it a lifelike or a fantasy world.

Social VR software is also used for professional purposes such as training new employees, selling items in a virtual marketplace or business meetings with participants all over the world reproduced in the same virtual office.

2.6.1 Type of interactions

[29] https://bigscreenvr.com/
[30] https://www.dogheadsimulations.com/
[31] https://www.improovr.com/home-v2/
[32] https://www.cisco.com/c/fr_be/solutions/spark/index.html
[33] https://meetingroom.io/
[34] http://objecttheory.com/
[35] https://www.telesoftas.com/portfolio/vr-medusa/

There are different forms of interactions: participants may just interact in a discussion, share or contribute to a content, use a 3D object,...

- **Real time discussion**: Users in the same virtual place with a microphone can communicate as if they were in a phone conversation.
- **3D Drawings**: With virtual markers, users are able to draw or erase everything they want in 3 dimensional spaces. The other users next to the drawer are able or not to see the drawing.
- **Content sharing**: Online if the software allows to access the internet or personal files, the users can share contents on live with others.
- **Creation sandbox**: this software allows users to create their own virtual environment, starting from an empty space where they can add 3D objects to create their own world.
- **Instant messaging**: Chat box integrated to the software.
- **Interactive games**: collective mini-games where users interact with others, for example paintball or dodgeball.
- **Object interactions**: in this case, users may use any kind of 3D object: a proof-of-concept car for a business meeting, a sofa for social networking,...
- **Skype**: Video chat but integrating skype.
- **Video chat**: It enables users without VR headset to participate to the meeting with a video conference tool.
- **Virtual photos**: With a virtual camera, users are able to take a picture of them (or the avatar) in the virtual environment.

2.6.2 Avatars representations

Avatar definition: an avatar is a character representing a user on the Internet and in video games. Originally, it is the digital incarnation of an individual in the virtual world of an online game. By extension this term has arrived in the discussion forums and in the current language; it refers to the representation of a character on websites and social networks. There are 5 categories of avatars:

- **Unrealistic**: The character can be anything, for example a slice of pizza.
- **Realistic**: The character has a human shape.
- **Photo-realistic**: The character is a retranscription of the user.
- **Head/Bust/Body**: the avatar may either be represented by a head or a bust or a body
- **User similarity**: An avatar closely looking like the user.

2.6.3 Social VR players

This Figure 14 compares the different approach proposed by major VR players that provides solution for social VR functionalities. We focused our comparison on the type of market they are targeting (B2C/B2B), the type of avatar proposed and its moving capabilities, the type of environment and the type interaction between users.

Figure 14: Social VR players

SOFTWARE	B2C/B2B	AVATAR (cf 2.6.2.2)	RELOCATION/MOVING ? IN THE VIRTUAL WORLD	VIRTUAL ENVIRONMENT	INTERACTIONS BETWEEN PEOPLE (cf 2.6.2.1)
ALTSPACEVR	B2C	Unrealistic	YES	Indoor environments	**Real-time discussion** 3D Drawings Content sharing (Online) Interactive games

BIGSCREEN	B2C & B2B	Realistic (Human chest) User similarity	NO	Indoor environments	**Real-time discussion** Content sharing (Online) Virtual photos
CISCO SPARK	B2B	Realistic (Human head)	YES	Indoor environments	**Real-time discussion** 3D Drawings Content sharing Instant messaging Object interactions Video chat
DOGHEAD SIMULATIONS' RUMII	B2B	Realistic (Human chest) User similarity	NO	Indoor environments	**Real-time discussion** Content sharing (Personal & Online) Object interactions
FACEBOOK SPACES	B2C	Realistic (Human body) User similarity	NO	Outdoor fixed environments	**Real-time discussion** 3D Drawings Content sharing (Facebook & Personal) Virtual photos
HIGH FIDELITY	B2C	Unrealistic	YES	Different ambiances "Free world"	**Real-time discussion** Virtual photos
IMPROOV	B2B	Realistic (Human head) User similarity	NO	Created by users	**Real-time discussion** 3D Drawings Content sharing (Personal & Online) Object interactions Skype Virtual photos
INNOACTIVE HUB	B2B	Realistic (Human body)	YES	Created by users	Real-time discussion 3D Drawings
LIVELIKE	B2C	Realistic (Human head)	NO	Living-room	**Real-time discussion** 3D Drawings Object interactions
MEETINGROOM	B2B	Realistic (Human head) User similarity	YES	Indoor environments	**Real-time discussion** Content sharing (Personal & Online) Personnal content sharing 3D Drawings
MIMESYS	B2B	Photo-Realistic	YES	Office	**Real-time discussion** Object interactions
OBJECT THEORY	B2B	Realistic (Human body) User similarity	YES	Indoor environments	MIXED REALITY **Real-time discussion** Content sharing (Personal & Online) Object interactions
OCULUS VENUES	B2C	Realistic (Human body)	NO	Living-room and theater	**Real-time discussion**
REC ROOM	B2C	Unrealistic	YES	Game shaped environments	**Real-time discussion** Interactive games
SANSAR	B2C & B2B	Realistic (Human body) User similarity	YES	Different ambiances "Free world" created by users	**Real-time discussion** Creation Sandbox
VR MEDUSA	B2B	Realistic (Human body) User similarity	YES	Indoor environments	**Real-time discussion** Content sharing (Personal) Interactive games Skype
VRCHAT	B2C	Unrealistic	YES	Different ambiances "Free world"	**Real-time discussion** 3D Drawings
VTIME	B2C & B2B	Realistic (Human body)	NO	Outdoor fixed environments	**Real-time discussion** Virtual photos

3 RELEVANT TECHNOLOGIES IN VRTOGETHER'S SCOPE

In this section we will describe first the social VR value chain and then we will look at several technologies, both hardware and software, that are relevant for the VRTogether project, based on the architecture and on the objectives. We thus find several categories of hardware to enjoy

VR content oriented towards the objectives of the project as well as the comparison of several applications having developed social character rendering capabilities.

3.1 Introduction to the social VR value chain

The value chain is made up of Hardware & Software, and Content.

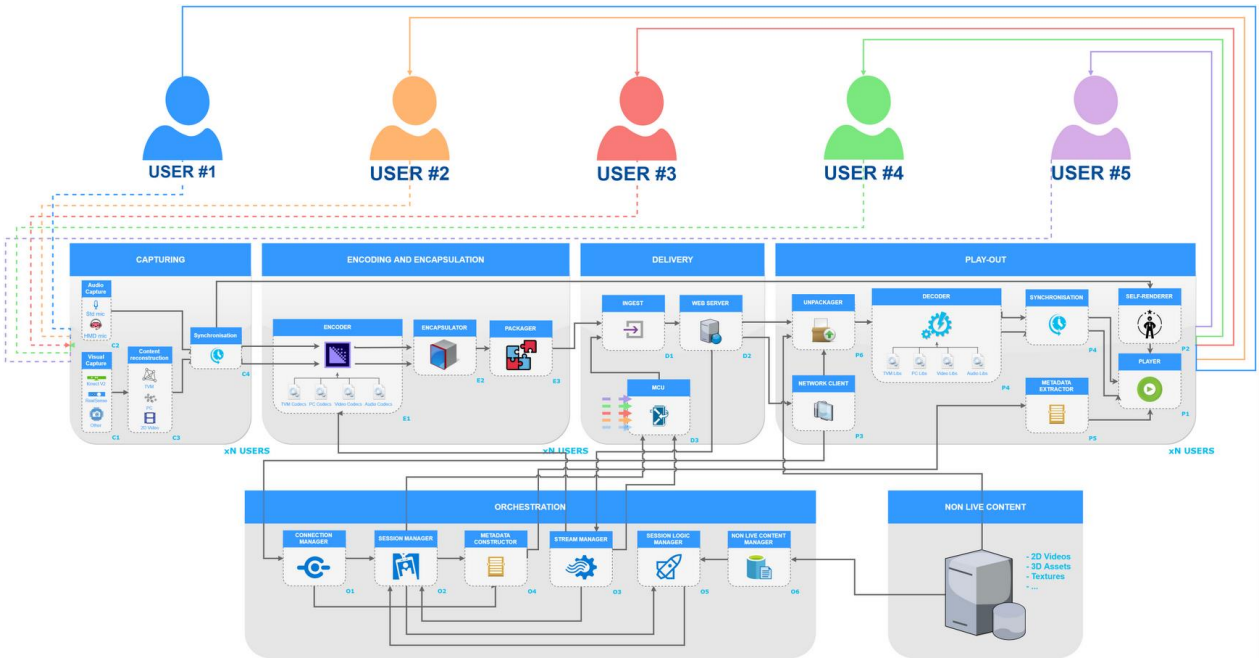
The experience consists of an offline reconstructed content merged with user representation. The user is captured live. Both are then encoded and distributed before being displayed. Every step of this content chain relies on hardware and software:

- Cameras for live capture
- Software to ensure the processing of the captured content
- Software allowing to edit the processed content
- Content delivery networks to ensure distribution
- Various devices to display and control the experience (playout)

Each of these hardware or software elements rely on components, some of which being specifically designed for producing and displaying immersive content.

Hereafter an initial architecture description of the different software components in the VR system and how these components are orchestrated. More information can be found in the document D2.4 Integrated Software Platform

Figure 15: Architecture of different software components in the VR system



A selection of best-in –class technology vendors will be detailed in this report. This selection will include mono-technology developers and end-to-end platform vendors. The latter usually developed one or several technologies and integrate it with third parties to provide an end-to-

end white label service to other companies. This enables them to easily provide a VR service to their partners & employees or to consumers.

Figure 16: Input technologies for VR, *Virtual Reality in tourism*[36]

3.2 Capture

In VR together we use mo-cap to record the movement of the actors, which is post-processed and added to the 3D/360° environment, as well as live capture of the participants in the virtual environment. For this purpose we use or investigated 360° cameras, 3D motion capture and depth cameras.

3.2.1 360 Capture

A 360° camera system consists of a single camera, usually with multiple lenses, or a rig of multiple cameras pointing in different directions. The system can capture a whole 360° scene in 1 single shot.

Since the full scene is captured through several lenses or with multiple cameras (multiple viewpoints), additional post-processing (video stitching) is required to create a seamless omnidirectional video.

A 360° camera system can capture a 360° environment (photo) or video.

[36] ALTHINGSVR. Infographic for input technologies for VR **[online]**. (Published on August 21, 2015) Available on : http://www.allthingsvr.co/post/127234136817/infographic-for-input-technologies-for-vr

Figure 17: 360 capture system of Gopro, *Gopro*

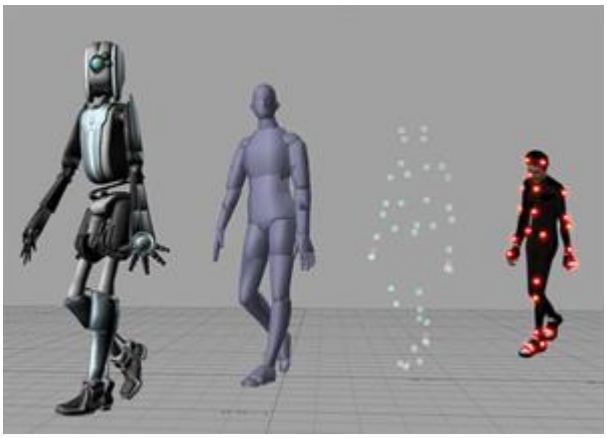
3.2.2 3D motion capture

A 3D motion capture system records the movement of humans or objects and maps this into a 3D digital model. This 3D model can recreate the recorded movements.

The system captures the movements using markers on the body or it may be markerless.

A marker-based system can use acoustic, mechanic, magnetic or optic markers (Nogueira, 2011)[37] on each limb or body part which are captured by a sensor system. For instance in case of optical markers the sensor system is a high-speed camera which captures the light reflected or emitted by the markers. A markerless system only uses volumetric capture to proceed (see 3.2.3).

Figure 18: 3D motion capture principle, *Wikipédia*[38]



[37] Nogueira, P. (2011). *Motion Capture Fundamentals, A Critical and Comparative Analysis on Real-World Applications.* Porto: Faculdade de Engenharia da Universidade do Porto. Retrieved from https://paginas.fe.up.pt/~prodei/dsie12/papers/paper_7.pdf

[38] WIKIPEDIA. Motion Capture **[online]**. (Published on July 8, 2010) Available on : https://en.wikipedia.org/wiki/Motion_capture

3.2.3 Live volumetric capture system

A live volumetric capture system uses special video cameras that record the depth of the recorded image. Several technologies exist to capture the depth picture (Medium.com[39]).

3.2.3.1 Passive stereo camera

Using 2 or more lenses, a stereo camera mimics the binocular vision of humans: the camera uses the difference in location of an observed object (disparity) between the left and right view to measure the depth of the object.

3.2.3.2 LiDAR or pulsed TOF

Light Detection and Ranging (LiDAR) emitters and sensors send a pulsed laser to a specific point of the object and receive the reflected pulse. Then, they derive the distance using the Time of Flight (TOF) and wavelength of the laser. By rapidly scanning the target area point by point a depth map of the scene can be determined.

3.2.3.3 Continuous wave TOF

While LiDAR has a lot of benefits, its technology is expensive and bulky (since it needs to scan the object). Continuous Wane TOF overcomes this by illuminating the full scene with continuous wave modulated light. The light reflected on the object is captured using standard camera sensors. By measuring the phase shift of the received light wave, the distance between the camera and the reflecting object can be derived.

3.2.3.4 Structured light

Pre-defined light patterns are sequentially projected onto the object. The patterns are deformed by the geometric shape of the object. A camera then observes the deformed patterns. By analyzing the distortion of the observed patterns, i.e. the disparity from the original projected pattern, depth information can be extracted.

3.2.3.5 Monocular methods

Conventional monocular cameras can be used in combination with multi-view photogrammetry methods to achieve 3D capture and reconstruction of objects and scenes. One approach is Structure from Motion (SfM). By moving around a scene/object and capturing numerous 2D images with multiple camera views, the SfM algorithm calculates the parallax between the different views and reconstructs a detailed 3D representation of the object.

3.2.3.6 Comparison

Here we introduce a shallow comparison of the different volumetric capture techniques.

[39] Medium.com **[online]** Choose the Right 3D Vision Camera For Your IoT Device: https://medium.com/iotforall/choose-the-right-3d-vision-camera-for-your-iot-device-962d95c581cb

	Strengths	Weaknesses	Example applications
LiDAR (Pulsed ToF)	• No need for ambient light • Wide field of view	• High Cost • Relative Bulky • Adversely affected by reflective properties of materials (e.g. translucent, water) • Lower refresh rate	• Mapping • Autonomous vehicles • Vendors (Velodyne, LeddarTech, Quanergy, Continental, etc)
Passive Stereo	• Use traditional simple low cost camera • Rich visual data for computer vision analytics • Works in both indoor and outdoor settings	• Poor low light performance • Does not work well with textureless surfaces • Require high processing power to derive depth map	• Robotic vision • Autonomous vehicles • AR/VR • Vendors(StereoLabs, Point Grey, etc)
Continuous Wave ToF	• Simple and compact hardware • Require low processing power • High refresh rate • No need for ambient light	• Poor outdoor performance (under sunlight) • Adversely affected by reflective properties of materials (e.g. translucent, water) • Interference by the presence of other ToF cameras	• Gesture control • Body tracking • Industrial • Vendors (Microsoft Kinect V2, SwissRanger/Heptagon)
Structured Light	• No need for ambient light • High resolution and accuracy then ToF • Relatively short range than ToF	• Poor outdoor performance (under sunlight) • Adversely affected by reflective properties of materials (e.g. translucent, water) • Laser speckle pattern on the target surface may not be desirable	• Face recognition • 3D scanner • AR/VR, body tracking, Industrial • Vendors (Structure,Orbbec,Intel RealSense F200, etc)

3.2.4 VRTogether's live capture scope

According to the specifications of the project VRTogether, three capture systems deserve a deeper study among dozens of them. The choices are oriented for the RGB-D cameras because they are the best candidates for live capturing of volumetric user reconstruction with minimum drawback (no material to wear, only a camera to connect).

3.2.4.1 Camera RGB-D Intel RealSense

Intel RealSense Technology is a suite of depth and tracking technologies designed to give devices depth perception capabilities that will enable them to "see" and understand the world. There are many uses for these computer vision capabilities including autonomous drones, robots, AR/VR, smart home devices amongst many other broad market products. RealSense technology is made up of Vision Processors, Depth and Tracking Modules, and Depth Cameras, supported by an open source, cross-platform SDK called "librealsense" that simplifies supporting cameras for third party software developers, system integrators, ODMs and OEMs.

In January 2018, new Intel RealSense D400 Product Family[40] was launched with the Intel RealSense Vision Processor D4.

[40] https://realsense.intel.com/introducing-intel-realsense-d400-product-family/

3.2.4.2 Camera RGB-D StereoLabs Zed

The ZED Stereo Camera[41] is a 3D sensor for depth perception and motion tracking. The ZED enables developers to build advanced systems that understand their environment with artificial intelligence algorithmic.

Stereo cameras allow to capture two separate images taken by optical sensors positioned several inches apart. Then computer software looks at the distance in pixels between similar features in each image and uses that to estimate the depth or distance from the camera to objects in the scene. These algorithms require precise calibration of the cameras in order to work.

Based-on smart phone camera technology, the ZED camera was designed to be small, lightweight, low cost, and still have high quality output.

3.2.4.3 Camera RGB-D Kinect V2

Kinect is a line of motion sensing input devices that was produced by Microsoft for Xbox 360 and Xbox One video game consoles and Microsoft Windows PCs. Based on a webcam-style add-on peripheral, it enables users to control and interact with their console/computer without the need for a game controller, through a natural user interface using gestures and spoken commands.

On February 1, 2012, a new hardware version of the device called “Kinect for Windows” was released. It was similar to the existing Xbox 360 device but tested and supported under warranty for commercial Windows applications. At the same time, version 1.0 of the SDK was released, allowing commercial use, with license terms requiring the use of the Kinect for Windows hardware even for non-commercial use.

3.2.4.4 Comparaison of the live capture systems

Figure 19: Technical table of the live capture systems

CAPTURE DEVICE	Camera Intel RealSense D415	Camera Stereolabs Zed	Camera Microsoft Kinect V2
TYPE OF CAMERA	RGB-D Camera	RGB-D Camera	RGB-D Camera
TYPE OF CAPTURE	Live IR Laser	Live Double RGB Camera	Live IR Laser
RESOLUTION	1920 x 1080 (RGB) 1280 x 720 (Depth)	4416 x 1242 (RGB & Depth) 3840 x 1080 (RGB & Depth) 2560 x 720 (RGB & Depth) 1344 x 376 (RGB & Depth)	1920 x 1080 (RGB) 512 x 424 (Depth)
RATE	90 FPS (Depth) 30 FPS (RGB)	15 FPS (2.2K RGB & Depth) 30 FPS (1080p RGB & Depth) 60 FPS (720p RGB & Depth) 100 FPS (WVGA RGB & Depth)	30 FPS (RGB) 30 FPS (Depth)
MINIMUM DEPTH DISTANCE (Meters)	0.16 m	0.5 m	0.5 m
MAXIMUM DEPTH DISTANCE (Meters)	10 m	20 m	8 m
FOV	Depth: 63.4° x 40.4° (+/-3°) RGB: 69.4° x 42.5° (+/- 3°)	90 x 60 x 110°	70°
LATENCY	Unknown	Unknown	60 ms
CONNECTOR	USB 3.0 Type C	USB 3.0	USB 3.0

[41] https://www.stereolabs.com/

DIMENSIONS	99 x 20 x 23 mm	175 x 30 x 33 mm	127 x 355 x 152 mm

Figure 20: Market table of the live capture systems

RGB-D Camera	Camera Intel RealSense D415	Camera Stereolabs Zed	Camera Microsoft Kinect V2
Sales (units)			45M (all kinect versions mingled)
Launch date	January, 2018	January, 2015	February 1, 2012
Profile	Leader	Follower	Leader
Pros	Great resolution Size	Long range Great resolution Size	Great resolution High depth quality
Cons	No official software support Low FOV Low depth quality	Low depth quality	Short range Low FOV Size

3.3 3D rendering engines

The market of display technologies is led by 3 major players: Unity, Unreal and Cry Engine.

3.3.1 Majors players of display technologies

- **Unreal Engine**

This 18 year-old company has developed an aggregation engine integrating 3D graphics with VR interactivity. The technology is used by several industries: gaming, architects, manufacturers, movie producers, etc. Although primarily developed for first-person shooters, it has been successfully used in a variety of other genres, including stealth, fighting games, MMORPGs, and other RPGs. With its code written in C++, the Unreal Engine features a high degree of portability and is a tool used by many game developers today. It has won several awards, including the Guinness World Records award for "most successful video game engine.

There are four versions of the Unreal Engine; the first one was released in 1998 and the fourth one in 2012.

The currently supported platforms in virtual reality are: SteamVR/HTC Vive, Oculus Rift, PlayStation VR, Google Daydream, OSVR and Samsung Gear VR.

- **Unity 3D**

Unity is a cross-platform game engine developed by Unity Technologies, which is primarily used to develop both three-dimensional and two-dimensional video games and simulations for computers, consoles, and mobile devices. First announced only for OS X at Apple's Worldwide Developers Conference in 2005, it has since been extended to target 27 platforms. Six major versions of Unity have been released.

It offers a wide range of features and is easy to use. Its strong point is cross-platform integration; games can be quickly and easily ported to android, iOS, Windows Phone 8 and BlackBerry, making it an ideal game engine for mobile game development.

- **Cry Engine** [42]

[42] https://en.wikipedia.org/wiki/CryEngine

CryEngine is a game engine designed by the German game developer Crytek. It has been used in all of their titles with the initial version being used in Far Cry, and continues to be updated to support new consoles and hardware for their games. It has also been used for many third-party games under Crytek's licensing scheme. Ubisoft maintains an in-house, heavily modified version of CryEngine from the original Far Cry called the Dunia Engine, which is used in their later iterations of the Far Cry series.

According to various anonymous reports in April 2015, CryEngine was licensed to Amazon for $50–70 million. Consequently, in February 2016, Amazon released its own reworked and extended version of CryEngine under the name of Amazon Lumberyard.

3.3.2 Business model of the rendering engines

- **Unreal Engine**:
  - Free: Royalty equal to 5% of all worldwide gross revenue actually attributable to each Product, regardless of whether that revenue is received by you or any other person or legal entity. No royalty is owed on specific forms of revenue including: The first $5,000,000.00 in gross revenue for each Product from the Oculus Store.
- **Unity**:
  - Free: For beginners, students and hobbyists who want to explore and get started with Unity.

    Limitation : the user must not make more than $100k in annual gross revenues, regardless of whether Unity Personal is being used for commercial purposes, or for an internal project or prototyping.
  - $35/month: For creators who are serious about bringing their vision to life. Features: Unity Success Advisor, Unity Game Dev Courses ($144 value), 20% off Asset Store Subscriber Benefit, Customizable Splash Screen, Performance Crash Reporting, Extended Unity Analytics, Flexible Seat Management, Pro Skin Editor UI.

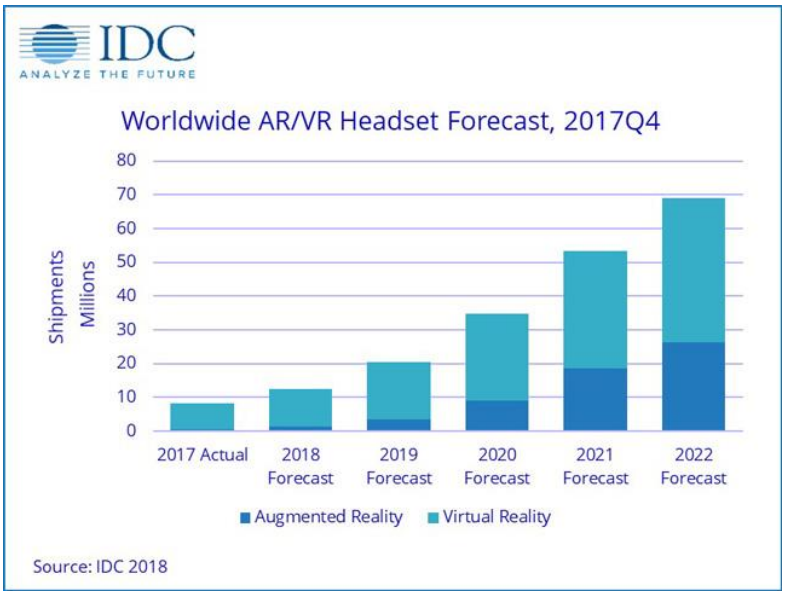
    Limitation: annual revenue or funds rose by $200k or less.
  - $125/month: For professionals who need complete flexibility and crave advanced customization. Features: all the precedent with Premium support and source code access plans available.

    No Limitation on annual revenue or funding.
- **CryEngine** :
  - Free: Royalty equal to 5% of all worldwide gross revenue actually attributable to each Product, regardless of whether that revenue is received by you or any other person or legal entity. No royalty is owed on the first $5,000 of annual revenue per project.

3.4 Head Mounted Displays (HMD)

HMD improvements are closely scrutinized as they are instrumental in the market adoption. The scale market has not started yet, and could take Apple entering for an inflection point. According to IDC, the number of shipments will increase drastically by 2022. Most of the devices can display any type of image, be it CGI or 360 live.

Figure 21: Worldwide AR/VR HMD Forecast 2017Q4, *IDC [43](2018)*



New HMDs have been presented at the CES 2018, solving real issues (increased field-of-view, better displays, less motion-sickness, autonomous).

The mobile HMDs currently supports only 3DoF but has the largest installed base whereas the computer/console market supports full 6DoF, has a higher revenue potential per user but a lower installed base.

While it is exciting to start seeing 6DoF standalone VR HMD like the Vive Focus and Oculus GO, both unfortunately include a 3DoF controller which in many cases functions as an immersion barrier compared to a fully 6DoF VR system. Recently, standalone HMDs with 6DoF capabilities are emerging like the Pico Neo or the HMD from GameFaceLabs[44], it is considered as the next step of the VR experience.

Unlike the standalone HMDs, some are connected to either a PC or mobile; their role is to display the VR immersive image; the intelligent features are located in the mobiles & PCs.

3.4.1 HMD for mobiles phones

Passive stereoscopes that work in conjunction with a smartphone work well as entry-level HMD: the smartphone is placed in the HMD and serves both as screen and motion capture device. Some of these HMDs are smartphone or brand specific like the Samsung Gear VR. Others can accept any (mid- or high-end) smartphone: Google Cardboard, Google Daydream and many clone brands like Shinecon[45] and Homido[46].

VRTogether will explore if current mobile phones are powerful and capable enough to process the VRTogether immersive content.

[43] IDC. Augmentede Reality and Virtual Reality Are on the VRge of Growwth, Says IDC **[online]**. (Published on June 19, 2018) Available on : https://www.idc.com/getdoc.jsp?containerId=prUS44001618

[44] https://gamefacelabs.com/

[45] http://www.shinecon.com/

[46] https://homido.com/fr/

3.4.2 Headsets connected to PCs

The HMD is connected to the PC through a cable.

With devices that work in conjunction with computers, the VR experience is far more immersive than a mobile HMD. For this usage, some empty space is needed. Motion sensors are then placed at the edge of this space. The HMD is connected to a powerful computer capable of rendering the virtual environment in real-time: images are not pre-rendered. With the HMD on, users can walk through the room or teleport themselves to different virtual places using a controller in their hand.

There are two leaders of PC-based HMDs: the HTC Vive (and premium device: Vive Pro) and the Oculus Rift. Both of them are implemented in the market since the early times of VR games in 2016.

3.4.3 Wireless standalone headsets

The new stand-alone VR HMD does not need a computer or video game console to run; it's battery powered. It doesn't need cameras or sensors in the room; instead, the sensors and cameras built into the stand-alone HMD track the movement in six degrees of freedom. It supports Google WorldSense[47]. The intelligent features are embedded in the HMD and they are connected to the internet through Wi-Fi.

This new type of HMD is considered as the future of VR HMD because it resolves one of the big obstacles for many consumers, the connectivity. As it is an all in one device, the consumers only have to buy the device and turn it on. But it amplifies the other problems: the technology is not ready and is still very expensive for the consumers. Indeed, the tracking precision without external device is less efficient and the controllers are systematically only 3DoF (except the Pico Neo)

3.4.4 VRTogether's HMD scope

According to the specifications of the project VRTogether, four HMDs deserve a deeper study amongst dozens of them.

3.4.4.1 HTC Vive & Vive Pro

"The HTC Vive is a virtual reality HMD developed by HTC and Valve Corporation. The HMD uses "room scale" tracking technology, allowing users to move in 3D space and use motion-tracked handheld controllers to interact with the environment.

The HTC Vive was unveiled during HTC's Mobile World Congress keynote in March 2015. Development kits were sent out in August and September 2015, and the first Consumer version of the device was released on April 5th, 2016."[48]

The HTC Vive is made up of four hardware items/pieces which work together to optimize the VR experience:

- **Vive HMD**: The Vive HMD uses two OLED panels. Safety features include a front-facing camera avoiding collision issue in the real world. Inside the HMD there are dozens of

[47] https://developers.google.com/vr/discover/worldsense
[48] https://en.wikipedia.org/wiki/HTC_Vive

infrared sensors that detect the base stations' IR pulses to determine the HMD's current location in a space.

- **Vive Controllers**: The controllers have multiple input methods and a use per charge of about 6 hours. Across the ring of the controller are 24 infrared sensors.
- **Vive Base Stations**: Also known as the Lighthouse tracking systems are two black boxes that create a 360° virtual space up to 15x15 foot radius. The base stations emit timed infrared pulses that are then picked up by the HMD and controllers. Wireless syncing lowers the amount of wires.
- **Vive Tracker**: A motion tracking accessory; it is designed to be attached to physical accessories and controllers, so that they can be tracked via the Lighthouse system.

Vive for business[49]: "In June 2016, HTC announced the release of their 'Business Edition' of the Vive for $1,200 USD which would include a Professional Use License, a 12-month Commercial Warranty, access to an exclusive support line, a 16-foot cable extension kit, and it included the Deluxe Audio Strap" (it added integrated over-ear headphones as well as improved the HMD's comfort through better weight distribution).

Figure 22: HTC Vive headset, *HTC*

3.4.4.2 Oculus Rift[50]

"The Oculus Rift is an HMD developed and manufactured by Oculus VR, a division of Facebook Inc.

Oculus initiated a Kickstarter campaign in 2012 to fund the Rift's development, after being founded as an independent company two months before. The project proved successful, raising US$2.5 million. In March 2014, Facebook purchased Oculus for $2 billion.

The Rift has gone through various pre-production models since the Kickstarter campaign, around five of which were demonstrated to the public. Two of these models were shipped to backers, labelled as 'development kits'; the DK1 in mid-2013 and DK2 in mid-2014, to give developers a chance to develop content on time for the Rift's release. However, both were also purchased by a large number of enthusiasts who wished to get an early preview of the technology."

The Oculus Rift embeds three hardware items which work together to optimize the VR experience:

- **The Oculus HMD**: The HMD has a Pentile OLED display, one per eye.
- **The Oculus Touch**: It consists of a pair of handheld units, one for each hand, each containing an analog stick, three buttons, and two triggers. The controllers are fully tracked in 3D space by the Constellation system, so they may be represented in the virtual environment, and

[49] https://enterprise.vive.com/us/BE/
[50] https://en.wikipedia.org/wiki/Oculus_Rift

each controller features a system for detecting gestures users may perform while holding them.

- **Constellation (tracking system)**: The positional tracking system is performed by a USB stationary infrared sensor that picks up light emitted by IR LEDs. They are integrated into the head-mounted display. The sensor normally sits on the user's desk. This creates 3D space, allowing the user to use the Rift while sitting, standing, walking, or even jumping around the same room.

Rift for Business: During Oculus Connect in June 2017, Oculus VR announced and released their Oculus Rift for Business bundle for $900 USD which included the Rift HMD, Oculus Touch controllers, and expanded warranty, preferential customer service, commercial use license, three constellation sensors, an Oculus remote, and three Rift Fits.

Figure 23: **Oculus Rift HMD, *Oculus***

3.4.4.3 Samsung Odyssey

The Samsung Odyssey was launched early 2018. It is positioned as a Mixed Reality HMD supporting Windows Mixed Reality. The Odyssey has 2 cameras pointing to the front to capture the real environment, so mixing the real and virtual world is possible.

With the 2 cameras the HMD is capable of supporting 6DoF and inside-out tracking. No external sensors like with the Vive or Oculus are needed for calibration or detection of your position in a room.

The Odyssey has dual high-quality AMOLED displays with a resolution of 1,440 x 1,600 pixels per side and a refresh rate of max 90Hz.

The Odyssey comes with 2 controllers, each with a thumbstick, a touchpad and 4 buttons: Trigger, Menu, Grab and Windows.

Figure 24 : Samsung Odyssey - Mixed Reality HMD

3.4.4.4 Mobile VR HMD

For VRTogether we will investigate the possibility to use a mobile VR HMD. The Samsung Gear VR and the Google Daydream both provide a passive HMD with 2 lenses and a support for smartphone.

Some additional functionality is provided by the 2 mobile VR HMDs:

- The Samsung Gear VR connects with the phone using a USB port, providing additional features like gyro and positional sensor.
- The Google Daydream detects a Daydream compatible smartphone when inserted through NFC and will automatically start the Daydream environment.

3.4.4.5 VRTogether HMD comparaison

Figure 25: Technical table of the VR HMDs

HMD	HTC Vive	HTC Vive Pro	Oculus Rift	Samsung Odyssey
TYPE OF CONNECTION	PC Based	PC Based	PC Based	PC Based
HMD DOF	6 DoF	6 DoF	6 DoF	6 DoF
CONTROLLER	Vive Controller	Vive Controller	Oculus Touch Controllers	Samsung Controllers
CONTROLLER DOF	6 DoF	6 DoF	6 DoF	6 DoF
RESOLUTION	2160x1200	2880x1600	2160x1200	2880x1600
REFRESH RATE	90 Hz	90 Hz	90 Hz	90/60 Hz
DISPLAY	OLED	OLED	OLED	AMOLED
FOV	110°	110°	110°	110°
SENSORS	Accelerometer Gyro Sensor Laser position sensor Front facing Camera	Accelerometer Gyro Sensor Laser position sensor Front facing Camera	Accelerometer Gyro Sensor Magnetometer Proximity Sensor	Accelerometer (6 Axis), Gyro Sensor (6 Axis), Compass (3 Axis), Proximity Sensor, IPD Sensor
AUDIO	Microphone	Microphone Headphones	Microphone Headphones	Microphone Headphones
DIMENSIONS/WEIGHT	19.2 x 11.7 x 12.1 (cm) 555 g		470 g	20,2 x 13,2 x 11,1 (cm) 645g

Figure 26: Market table of the VR HMDs

Headsets	HTC Vive	HTC Vive Pro	Oculus Rift	Samsung Odyssey

Sales (units)	2016 : 420K 2017 : 297K		2016 : 243K 2017 : 326K	
Sales Region	North America, China, Western Europe, Developed Asia & Oceania			
Turnover /y (Average 2016 and 2017)	$178,9 M (estimation)		$113,5 M (estimation)	
Launch date	April 5, 2016	April 5, 2018	March 2016	October 2017
Market Share (source : Statista[51])	12%		20%	
Sales forecast in 2021 (using Market Share) (source : IDC[52])	3,36M units		5,6M units	
Profile	Leader in PC Based		Leader in PC Based	
Pros	Room-scale tracking Comfort Intuitive interface	Room-scale tracking Comfort Intuitive interface Display	Price Comfort Speakers	Price Inside-out tracking Windows MR compatible
Cons	Price Bulky cable Hard on neck and back		Small room-scale tracking No pass-through camera No accessory platform	

3.5 Encoding & Delivery

VRTogether platform will offer much more than a typical VR platform since it deals with the capturing, encoding, delivery and rendering of complex scenes mixing VR, 360 live video, 3D audio, and 3D characters captured in real time. In this sense, the encoding and the delivery are two key steps of the entire chain.

In this part, we will consider first encoding and delivery separately and then we will deal with the value chain of the combination of both encoding and delivery.

3.5.1 Encoding (Live Content)

In the VRTogether project, the encoding step will have to actually deal with very different contents in terms of nature: 360 live video, 3D audio, 3D objects described by point cloud and meshes and scene description to compose all contents. The 360 video and 3D audio encoding is quite similar to classical video and audio encoding. On the contrary, encoding 3D objects in real time requires newer technologies of 3D objects compression that will open new market opportunities. Scenes description, allowing the composition of these objects inside immersive environments, will also drive some market opportunities.

[51] https://www.statista.com/statistics/755645/global-vr-device-market-share-by-vendor/

[52] https://thejournal.com/articles/2018/03/20/report-ar-and-vr-headset-sales-to-return-to-strong-growth-following-lackluster-2017.aspx#

3.5.1.1 3D characters

3.5.1.1.1 Time Varying Meshes (TVM)

Time-Varying Mesh (TVM) TVM is a sequence of polygonal mesh models that represent dynamic 3D entities. TVM is often used to capture and reproduce human behavior and natural movements as a dynamic 3D full-body representation. 3D Mesh (like TVM) encoding is used from VR/AR companies that offer holographic solutions or 3D content streaming. But TVM that requires huge storage space and computational cost needs to be encoded and compressed. There are currently many well-known, publicly available and open-source libraries that can be utilized for TVM or in general 3D mesh compression. We can mention for example:

- Draco[53]: Draco is a library for compressing and decompressing 3D geometric meshes and point clouds. It is intended to improve the storage and transmission of 3D graphics.
- Corto[54]: Corto is a Mesh compression library, designed for rendering and speed.

Many companies propose also an SDK or a full software solution including capture and 3D mesh compression and delivery, like:

- 8i[55]
- Owlii[56]
- Mimesys[57]
- Holobeam[58]

3.5.1.1.2 Point-Cloud Compression (PCC)

A point cloud is defined as a set of points in 3D space with specified geometry coordinates and associated attributes such as a color. Point clouds are an important part of emerging VR applications because of their simplicity and versatility. There are no restrictions on the attributes associated with every point in the cloud. Point clouds are usually captured using multiple cameras or depth sensors and can contain millions of points in order to create a photorealistic reconstruction of an object. Compression of point cloud geometry and attributes is essential in order to efficiently store and transmit point cloud data for applications such as tele-immersive video and free viewpoint sports replays.

Thus, live Point Cloud Compression (PCC) has been recently attracting attention by many major companies. Even a specific PCC activity has begun in MPEG that launches an ad hoc group in order to start the standardization activity on point cloud compression. In April 2017 MPEG published a call for proposals on point cloud compression (Reference: 3D Graphics Subgroup, "Call for proposals for point cloud compression v2," ISO/IEC JTC1/SC29/WG11 MPEG, output document N16763, Apr. 2017). It divides the activity into three categories for static frames (category 1), dynamic sequences (category 2) and dynamically acquired/used point clouds (category 3).

Leading technology companies responded to the call for proposal and the proposals were evaluated in October 2017. In addition to objective metrics, each proposal was also evaluated through subjective testing at GBTech and CWI. The winning proposals were selected as test models for the next step of the standardization activity.

[53] https://google.github.io/draco/
[54] https://github.com/cnr-isti-vclab/corto
[55] https://8i.com/technology/
[56] https://www.owlii.com/
[57] http://www.mimesysvr.com/
[58] https://valorem.com/post/holobeam-ces2018/

For the compression of dynamic sequences, it was found that compression performance can be significantly improved by leveraging existing video codecs after performing 3D to 2D conversion using a suitable mapping scheme. This also allows the use of hardware acceleration for video codecs such as HEVC that is supported by many current generation GPUs. In this manner, synergies with existing hardware and software infrastructure can allow rapid deployment of new immersive experiences.

The next step was to identify and investigate methods to optimize the test models by performing core experiments. The core experiments for dynamic sequences explored the use of different mapping schemes from 3D to 2D, hybrid codecs that use 3D geometry compression (using octrees) as well as video codecs and the use of motion field coding.

In the recent MPEG meetings (MPEG 121 and 122), the adhoc group (AhG) was particularly crowded, attracting attention from many industrial partners. During the meetings, the core work was to cross check test models and to review the results of the core experiments. In addition, a number of new datasets are being created by commercial companies and there was a proposal to converge between TMC1 and TMC3 (static point clouds). The most interesting development is the creation of a group to explore the delivery and transmission of point clouds based on the current encoding approach (3D to 2D conversation) by using 3 ISOBMFF tracks: 2D video, depth, metadata.

3.5.1.2 Audio

The encoding of audio for the VR market has pushed forward some old technologies, like Ambisonics but has pushed also the birth of new technologies like object-based Audio.

Ambisonics was developed in the UK in the 1970s within the British National Research Development Corporation. Despite its solid technical foundation and many advantages, Ambisonics had not until recently been a commercial success, and survived only in niche applications and among recording enthusiasts. The successful market introduction of home theatre surround sound systems since the 1990s and later the explosion of the VR market put Ambisonics technologies in front of the audio scene. Ambisonics appears quite powerful to encode spatialized audio contents but suffers from the fact that it is a Channel-based audio.

Recently, Ambisonics has been adopted by the VR industry to deliver 360 audio for 360 videos, gaming and virtual reality experiences. The Ambisonics Channel-based audio containing a mix of several channels but each audio channel in the final product has to be reproduced by a loudspeaker at a well-defined position. This fixed audio mix is transmitted to end-users with basically no means to adapt it to their needs, which may be a specific playback device or their personal preferences. Some extension of Ambisonics technologies allows producing binaural encoding which is perfect for the end user experiencing VR via headphones and a head-mounted display.

Beside the classical Channel-based approach, the Object-based audio concept takes its origin from MPEG 4 system specifications around 1999. It allows to represent a set of individual assets together with metadata describing their relationships and associations. Thus, media objects can be assembled in many ways to create new user experiences. Nowadays, Object-based audio approach seems to be the new revolution for creating and deploying interactive, personalized, scalable and immersive content.

With the introduction of HTML5 and the Web Audio API, an important prerequisite was made for native rendering of object-based audio in modern browsers since object-based audio needs much more processing power to render the final audio scene. The Object-based audio VR market appears to drive many industrial companies like Dolby Laboratories[59] (Atmos

[59] https://www.dolby.com/en/index.html

technology), Fraunhofer IIS[60], DTS Inc[61]. (DTS-X technology), Qualcomm[62], Merging Technologies[63]. New Audio Technology[64] or Gaudio[65] Labs actors like IRT, IRCAM are also involved.

3.5.1.3 360 video (On-demand)

The 360 videos are more and more used into VR to enrich the end- user experience. In the scope of the VRTogether project, 360 video is a fundamental type of content for obtaining a final rich 3D scene. The main challenge about 360 videos is to deal with the huge resolution and so the huge size of the files. This implies to find the optimal way to encode those contents going from classical 4K video to 16k stereo 360 video at 60 Fps. Finding the right encoder and configuring it in the right way open the door to a brand new market where many actors offer optimized small form factor encoders that usually rely on ISO/IEC/ITU-T MPEG standard-based H.265/HEVC compression.

- Advantech 3311 based on SocioNext chip M30
- NGCODEC[66]

3.5.1.4 360 video (live)

To deal with live 360 video is an even higher challenge since the goal is still to deliver high quality video, under a variety of network conditions. Proposing always more powerful encoders can be one solution to face the delivery of very high-resolution 360 video at good visual quality. Another solution has also emerged to face the challenge of delivering high quality video: 360 Video tiling technologies which are capable of transmitting a complete UHD field of view. Then, tiling technology that only transmits the pixels that are watched allows to actually save bandwidth. VR tiling technology works with all relevant devices, including the Oculus Rift, HTC's Vive, and Samsung's Gear VR.

Live 360 video encoding challenges push the apparition of new technologies that will drive a new market around the VR market.

3.5.1.5 Scene description

To have a scene format that describes how to compose 2D or 3D objects or contents is not new. There are many scene description formats that exist. Already in 1994, VRML (Virtual Reality Markup Language) has been proposed as a standard file format for representing 3-dimensional (3D) interactive vector graphics, designed particularly with the World Wide Web in mind. Nowadays, It has been superseded by X3D.

In the 1990s, we also saw the apparition of the BIFS format (an abbreviation for "Binary Format for Scenes"). BIFS provides a complete framework for the presentation engine of MPEG-4 terminals. BIFS enables to mix various MPEG-4 media together with 2D and 3D graphics, handle interactivity, and deal with the local or remote changes of the scene over time. BIFS has been designed as an extension of the VRML 2.0 specification in a binary form. In the scope of the VRT project, the need to have a scene description-based system may appear. BIFS seems to be a good candidate since there are already some connections with Object-based audio concepts.

[60] https://www.iis.fraunhofer.de/
[61] https://dts.com/
[62] https://www.qualcomm.com/
[63] https://www.merging.com/
[64] http://newaudiotechnology.com/
[65] https://www.gaudiolab.com/
[66] https://ngcodec.com/markets-vr-ar-mr/

3.5.1.6 Whole market-chain

At a technical level, to be able to propose encoding for all kind of media is an actual challenge. As shown by the early works at MPEG, there are many research opportunities around mastering the whole encoding chain. We are convinced that mastering the whole encoding chain, as we would like to do in the VRTogether project, will open new business opportunities.

3.5.2 Delivery

After the encoding of the different media into elementary stream, the delivery deals with the packaging and distribution of the different content stream through the appropriate network. The packaging phase will ensure that the content is segmented appropriately and that will include all necessary metadata regarding signaling and synchronization. The distribution will leverage the state-of-the art technology to deliver the best possible quality to all users depending on their capabilities like MPEG DASH.

From the delivery point of view, any media is seen as a stream whatever the underlying media. However, some particular encoding of media, like scalable video or audio, needs to be considered specifically at delivery level.

3.5.2.1 3D characters

As we explained in the introduction, the delivery of Time Varying Meshes (TVM) and Point Cloud (PC) is not really different from the transmission of video.

Regarding TVM delivery, it appears that there is no specific market since any company announced a media serving platform (similar to traditional media servers) dedicated to meshes.

On the contrary, Point Cloud is a field with active players at MPEG. Point cloud delivery may drive some market opportunities in the future.

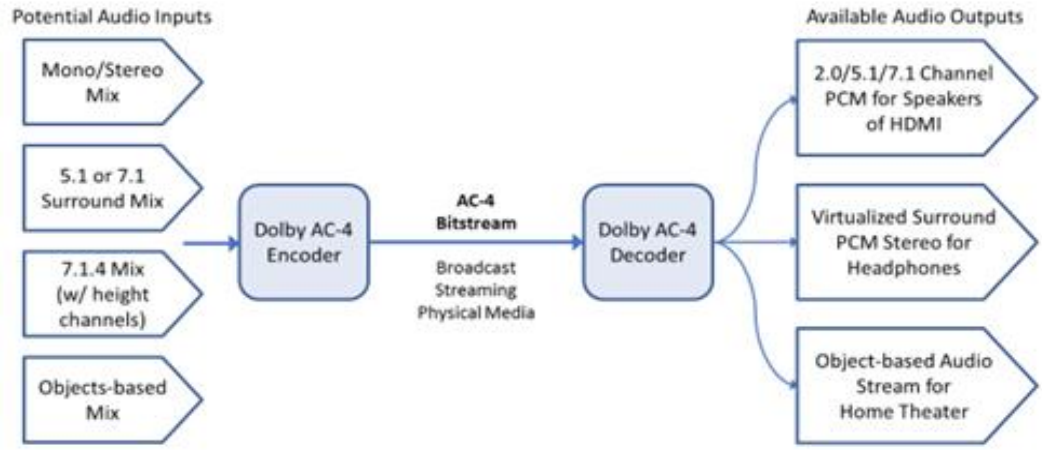
3.5.2.2 Audio

The delivery of the newly encoded audio content is a challenge similar to scalable video. As scalable video, scalable audio offers adaptation capabilities to the end-user needs. Audio will be only encoded once with several layers and the end-user player will be able to retrieve only the streams it needs according to the user constraints. Nowadays, the market of the "3D-enabled codecs" is shared by three main actors: Fraunhofer leading MPEG-H 3D Audio[67], Dolby pushing AC-4[68], and DTS promoting DTS-X[69].

The schema below shows a Dolby AC-4 delivery chain with adaptation capabilities that improve bandwidth consumption.

[67] https://www.iis.fraunhofer.de/en/ff/amm/prod/digirundfunk/digirundf/tvaudio.html#tabpanel-5
[68] https://hub.dolby.com/hearing-is-believing-comparing-dolby-ac-4-to-dolby-ac-3/
[69] https://dts.com/dtsx

Figure 27: Dolby AC-4 delivery chain with adaptation capabilities thanks to scalability



3.5.2.3 Video (live and on-demand)

After the encoding step, to deliver live video or on-demand video is similar and is based on existing delivery systems for 360 video. Like scalable audio, scalability brings also here a new flexibility for the end-user. A delivery format like MPEG-DASH offers all functionalities needed to deliver content in an efficient way like viewport-aware adaptive 360° video.

3.5.2.4 Scene description

Scene description can also be considered as a media stream to be transmitted. However, the scene description is in this case static since it has been generated at production level. A new market seems to appear around dynamic scene created at user level. HbbTV can provide some end user dynamic scene adaptation.

3.5.2.5 Whole market-chain

There is a business opportunity in proposing a synchronous transport solution for any type of media. Standardization bodies such as MPEG are currently extending their specifications “systems” format for MPEG-DASH. This effort in standardization for OMAF Application Format or MPEG-MORE shows that industry is pushing in that direction and is ready to propose new services based on those formats.

3.5.3 Market opportunity to combine encoding and transport of any media

To have a complete system that combines the whole technical “Broadcast” value chain seems to be a real market trend that many companies propose. This chain includes encoding, delivery and play-back as well with a combination of a client and server. We mention some companies that provide this chain:

- Startups: Tiledmedia[70], NextVR[71], LiveLike, Texel[72], Pixvana[73], Imeve[74]
- Companies: Viaccess-Orca[75], Harmonic[76]

[70] https://www.tiledmedia.com/
[71] https://www.nextvr.com/
[72] http://texelvr.com/
[73] https://pixvana.com/
[74] https://imeve.com/imeve-live/
[75] https://www.viaccess-orca.com/
[76] https://www.harmonicinc.com/

4 STANDARDIZATION BODIES & TECHNOLOGIES

4.1 Relevant standardization groups

Standardization of VR production, distribution and rendering tool, protocols and interfaces can help drive the large-scale deployment and adoption of new VR services. The VRTogether consortium partners have indicated relevant standardization bodies and industry forum to monitor and/or actively contribute to.

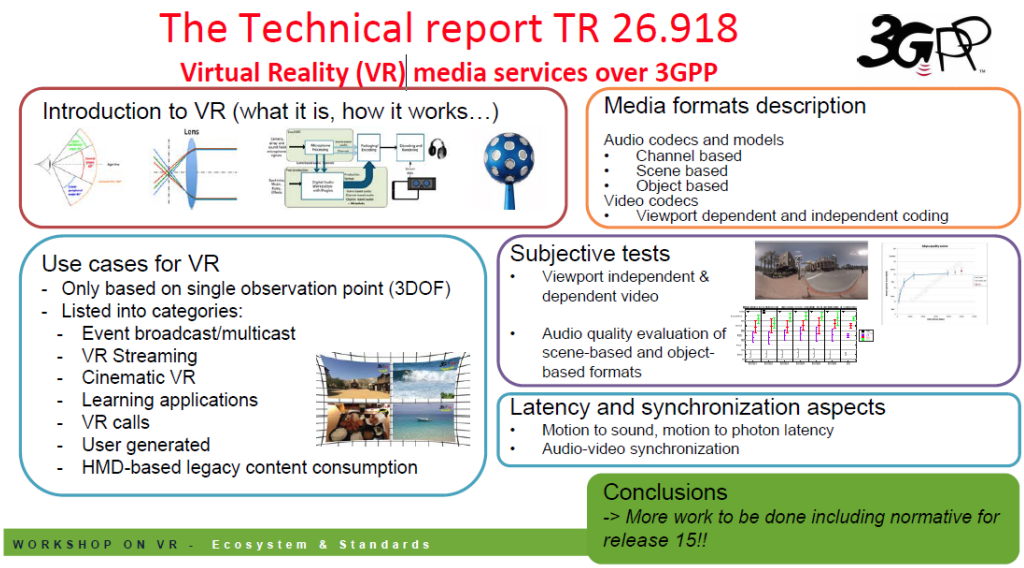
4.1.1 3GPP:

The 3rd Generation Partnership Project (3GPP) covers cellular telecommunications network technologies, including radio access, the core transport network, and service capabilities - including work on codecs, security, and quality of service - and thus provides complete system specifications. The specifications also provide hooks for non-radio access to the core network, and for interworking with Wi-Fi networks.

In 3GPP SA1 (service requirements), VR aspects are addressed under efficient content delivery and low latency and high reliability and cover requirements on the data rates, motion-to sound and motion-to-photon latencies, and audio/video synchronization.

3GPP SA4 (codecs) conducted a study from April 2016 to June 2017 on (360-3DOF) VR. The image below provides a summary of the study results. This work was adopted by ETSI and resulted in the technical report TR126 918.

Figure 28 : 3GPP Technical report TR26.918



4.1.2 DVB:

Digital Video Broadcasting (DVB) is an industry-led consortium of the world's leading digital TV and technology companies, such as manufacturers, software developers, network operators, broadcasters and regulators, committed to designing open technical standards for the delivery of digital TV and other broadcast services. Standardization work is only undertaken after rigorous requirement setting in the Commercial Modules (CM).

After a study mission on VR in 2016 [REF DVBVR], DVB started work on requirements in DVB CM-VR, with a focus on VR/3DoF. The CM-VR group aims at ensuring the effort is driven toward those who provide content services, and has started with developing end-user stories / use cases, and running a survey among its members starting end of December 2017.

4.1.3 Khronos:

The Khronos Group members are dedicated to the creation of royalty-free open standards for graphics, parallel computing, vision processing, and dynamic media on a wide variety of platforms from the desktop to embedded and safety critical devices. Khronos APIs are key technologies in their respective markets, such as Vulkan and OpenGL in graphics and gaming, WebGL in 3D web graphics, and OpenVX and OpenCL in embedded vision and compute. The OpenXR working group – previously known as the Khronos VR Initiative - is creating an open and royalty-free standard for VR and AR applications and devices.

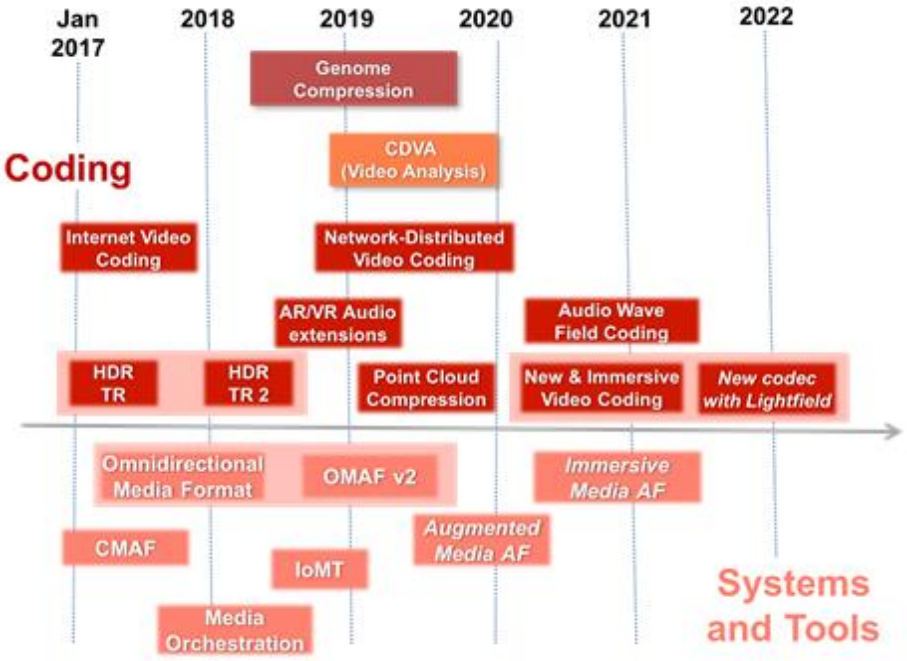
Khronos VR aim is to create a cross-platform VR standard that eliminates industry fragmentation by enabling applications to be written once to run on any VR system, and to access VR devices integrated into those VR systems to be used by applications.

OpenXR defines two levels of API interfaces that a VR platform's runtime can use to access the OpenXR ecosystem. Apps and engines use standardized interfaces to interrogate and drive devices. Devices can self-integrate to a standardized driver interface. Standardized hardware/software interfaces reduce fragmentation while leaving implementation details open to encourage industry innovation.

4.1.4 MPEG:

The Moving Picture Experts Group (MPEG) is a working group of ISO/IEC with the mission to develop standards for coded representation of digital audio and video and related data.
Since mid-2017, MPEG has started working on MPEG Immersive (MPEG-I). MPEG-I targets future immersive applications. The goal of this new standard is to enable various forms of audio-visual immersion, including panoramic video with 2D and 3D audio, with various degrees of true 3D visual perception (leaning toward 6 degrees of freedom). This full standard has already reached a relevant state that forces us to take it into account in VRTogether. The MPEG-I roadmap below points to many topics that are relevant to our project.

Figure 29 : Source: https://mpeg.chiariglione.org/docs/mpeg118-version-mp20-roadmap



4.1.5 VR-IF:

The VR Industry Forum (VR-IF) was established as a non-profit organisation during CES 2017, after a year of informal meetings. Its aim is to further the widespread availability of high quality audio-visual VR experiences, for the benefit of consumers. VR-IF focuses on content that is transmitted as audio and video, and it will monitor complementary technologies for inclusion in its scope, including those that enable augmented reality and mixed reality.

VR-IF targets immersive experiences that typically require head-mounted devices, understanding that immersive VR content may also be consumed on “2D flat screens” (like tablets, mobile phones, PC screens, TVs) with navigation capabilities.

VR-IF builds on available specifications and standards and recommends selections where needed, provides guidelines that go beyond standardization, supports interoperability tests and demos, and advocates interoperable solutions and standards. It is currently in the process of developing test vectors that can be used to validate implementations.

While an initial set of guidelines was published at CES 2018, the following steps will consider additional use cases, with a close relationship to OTT streaming and download, including some live distribution aspects, considering broadcast and 5G networks, considering UHD use and looking beyond 3DoF.

4.1.6 W3C:

The World Wide Web Consortium (W3C) is an international community working to develop Web standards.

WebVR is an open specification that makes it possible to experience VR in your browser. The goal is to make it easier for everyone to get into VR experiences, no matter what device you have.

The WebVR API (or rather; the WebXR API that will soon replace it) provides interfaces to VR hardware to allow developers to build compelling, comfortable VR experiences on the web. It

provides access to input and output capabilities commonly associated with Virtual Reality (VR) and Augmented Reality (AR) hardware like Google's Daydream, the Oculus Rift, the Samsung Gear VR, the HTC Vive, and Windows Mixed Reality headsets. More simply put, it lets you create Virtual Reality and Augmented Reality websites that you can view with the appropriate hardware like a VR headset or AR-enabled phone. The API is used by web-based virtual reality frameworks such as A-Frame.

4.1.7 VQEG

The Video Quality Experts Group (VQEG) provides a forum, via email lists and face-to-face meetings for video quality assessment experts to exchange information and work together on common goals. The general motivation of VQEG is to advance the field of video quality assessment by investigating new and advanced subjective and objective methods for assessing quality. VQEG activities, such as validation tests, are documented in reports and submitted to relevant ITU Study Groups (e.g., ITU-T SG9, ITU-T SG12, ITU-R WP6C), and other SDOs as appropriate. Several VQEG studies have resulted in ITU Recommendations.
IMG is a group of VQEG that is currently looking at the quality assessment of immersive media, involved in virtual and augmented reality applications.

4.2 Participation in standardization groups (SDO)

VR-Together is monitoring and/or actively contributing to the SDO's in the table below. Several partners have initiated or are continuing their presence in the most relevant SDO's

	3GPP	DVB	Khronos	MPEG	VRIF	W3C	VQEG
CWI				X			X
Motion Spell				X		X	
TNO	X			X	X		
Viaccess-Orca		X			X		

4.3 Standardization strategy and goals

The technology partners are mainly focusing on following standardisation topics relevant to VRTogeher:

- Point Cloud Compression
- QoE / QoS and related test and assessment methods
- Network Based Media Processing
- Social VR

4.3.1 Point Cloud Compression

For the project CWI is providing a solution for lossy compression of dynamic point clouds, based on the open source software for Point Cloud Compression developed at CWI[77]. This solution will not compete in the standardization race, but it serves as an open source tool to benchmark different solutions and experiment research ideas. Currently, it is being integrated into the VRTogether DASH-based point cloud communication pipeline that will allow multiple users to see each other point cloud representation, captured in real time, and rendered in the

[77] https://github.com/cwi-dis/cwi-pcl-codec

same virtual environment. The software package is distributed as open source (http://www.dis.cwi.nl/pointcloud/) and as part of the initial set of components of the VRTogether platform.

4.3.2 QoE, Performance Assessment and Testing

CWI is involved in the current activity of IMG Group in VQEG, which focuses on the definition of a joint test plan for the design of a subjective test campaign concerning subjective quality assessment of 360-degree content. The group has also established a link with the ITU-T Question 13, on Quality of experience (QoE), quality of service (QoS) and performance requirements and assessment methods for multimedia. The next face-to-face meeting of the VQEG IMG is scheduled for November 12 to 16. CWI is planning to participate in the meeting and present the current activities concerning quality assessment of point cloud signals and user's QoE in social VR applications.

The activities of IMG VQEG match with Objective 4 _ Develop appropriate QoE metrics and evaluation methods _ as defined in the Grant Agreement.

CWI has also been selected as the Test Lab for the MPEG subjective evaluation of the coming proposals on point cloud compression, where we analyzed the technologies submitted by nine industry leaders.

4.3.3 Network Based Media Processing

Network-Based Media Processing (NBMP) is a framework that allows service providers and end-users to describe media processing operations that are to be performed by the network. NBMP describes the composition of network-based media processing services out of a set of network-based media processing functions and makes these NBMP services accessible through Application Programming Interfaces (APIs).

Motion Spell has decided to attend sessions on NBMP since this new activity that allows building media workflows has generally been ignored so far. One of the main use-cases covered by the output of the MPEG meeting in Gwangju included ingesting media for distribution. Unified Streaming, co-chaired, wants to standardize ingest. Indeed, this is very exciting for VRTogether; Motion Spell will be interested in implementing an ingest component.

TNO, also a partner of the VRTogether project, proposed a contribution that puts the focus on extending scene description for 3D environment in the scope of NBMP. This contribution also exposed a tentative list of NBMP functions that could be useful for Social VR:

- Background removal.
- User detection and scaling.
- Room composition without users.
- Room composition with users.
- Low-latency 3DOF encoding.
- Network-based media synchronization.
- 3D audio mixing functionality.
- Lip-sync compensation.

4.3.4 Social VR

Social VR has become important in MPEG. For instance some of the requirements can be fulfilled by the MORE specification. This notably applies to the simpler forms of Social VR, where images of users are composited into a VR360 experience. This requires both temporal synchronization (multiple users should experience the same scene simultaneously) and spatial coordination (the composited images for all users need to have consistent location and size for the experience to be perceived as realistic and compelling). MORE defines the necessary metadata and protocols for this.

TNO started working on VR and in particular Social VR in 2016 as a use case for the activities on 360° video delivery based on tiled streaming. In April 2017 this work was presented in the 118th MPEG meeting, showing Social VR as an important use case.

Through papers, demonstrations and standardization contributions in MPEG-I, 3GPP and VR-IF, both TNO and MotionSpell further work on disseminating this use case through feeding in the use case requirements in several standardization efforts: Media Orchestration, Metadata Definitions, Scene Description, etc.

5 VALUE PROPOSITION FOR VRTOGETHER

5.1 Market perspectives for social VR

VRTogether will offer **photorealistic immersive virtual reality** content which can be experienced together with friends or colleagues, and demonstrate its use for domestic and business VR consumption. We will deliver VR social experiences by the orchestration of innovative media formats (video, a blend of videos, point cloud representations and 3D mesh interpolations). VRTogether addresses 5 specific objectives: (i) Develop and integrate new media formats, (ii) Adapt the existing production pipeline, (iii) Re-Design the distribution chain, (iv) Develop appropriate Quality of Experience (QoE) metrics and evaluation methods, (v) Maximize the impact of VRTogether.

5.2 Recommended solution for VRTogether

5.2.1 Value proposition

Positioning

VRTogether aims at being **the new reference** in terms of realistic VR conferencing software in domestic and business market by providing photorealistic rendering of end- users. The first step is to develop and integrate cost-effective production and delivery techniques to provide photorealistic experiences.

User's expectation is high. It is impossible for now to virtually meet friends or colleagues with a real-time photorealistic representation of themselves and, by extension, impossible to perceive the emotions of the person in front of them. Prior to VRTogether platform release, end-users seem to prefer typical video conferencing, since the existing immersive VR applications do not offer realistic content.

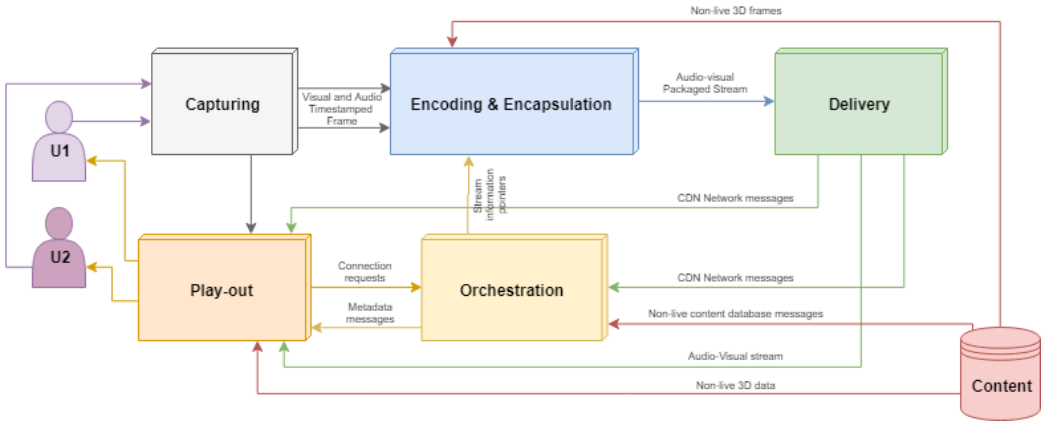
5.2.2 Ecosystem

The document D2.1 describes the high level Software components that will compose the platform are:

- Capturing (CERTH, CWI, TNO)
- Encoding & Encapsulation (MSE, CWI, TNO)
- Delivery (MSE)
- Play-out (TNO, i2CAT)
- Orchestration (TNO, i2CAT)

In VRTogether, cutting-edge technologies are integrated to an efficient pipeline, synthesizing a multifunctional, end-to-end VR platform, starting from raw data capturing to end-user standalone and web players.

Figure 30: Ecosystem for the VRTogether project



- 3D Capture and Reconstruction

This module is responsible for RGB-D and audio data acquisition from multiple sensors. The acquired visual stimuli are reconstructed and then synchronized with the corresponding audio signal to facilitate the play-out.

- Encoding & Encapsulation

The visual or audio content of one user is encoded in order to reduce the bitrate needed to represent the visual and audio signal. Once encoded, the signals are multiplexed and encapsulated to a media format. In order to be transmitted to the delivery module, the signals of multiple users are packaged.

- Delivery

VRTogether is a multiplayer VR platform consisting of multimedia (audio, 2D video, point cloud, time-varying mesh) delivery technologies, offering low-latency transmission between remote users. The ingest brick adapts the multi-media data to the format needed for storage in the "Web-server" module that transmits the data on the network.

- Orchestration

Based on an efficient orchestration module which spatiotemporally aligns and manages the user-representation media and the content of the pilots, VRTogether constitutes a smooth and unified VR experience.

The module receives the user's connection information and transfers the relevant ones to the metadata constructor which constructs the metadata in order to project the video stream of the users that has been managed upstream, and also the non-live content.

- Play-out

VRTogether platform supports standalone and web-based players based on cutting-edge VR rendering technologies, where the unified content from the orchestrator module is consumed.

The module receives the unpackaged audio-visual content, the self-created stream from the user and the metadata pointers in order to project the desired media content after being decoded, split and synchronized according to the universal timestamp.

5.2.3 Addressed targets and use-cases

5.2.3.1 B2C (Business to Consumer)

The project is not intended to directly reach the general public. The B2C target is not a priority at this time. Indeed, in the case of a B2C development the project would lead to an application grouping people with a photo-realistic representation and exploiting the technology 6DoF, which would be distributed on a platform such as Steam VR store or Oculus Experiences. This would also require the consumer to bring an RGB-D-type camera in addition to the entire ecosystem for broadcasting VR. Finally, in order to align with all other software today it would be necessary that the software is offered free of charge on the stores, with an advertising service or a freemium-type content.

The B2C target is therefore not a priority today because the solution envisaged is not easily compatible with distribution to the general public. Nevertheless B2C remains a potential target.

5.2.3.2 B2B (Business to Business)

A big problem for businesses today is that not all actors in a meeting can be in the same place. The arrival of the VR allowed to create software of conference and virtual prototyping with the intervention of several actors at the same time. It also allows teaching new employees by trainers who are not in the same place. Some even have the opportunity to join this virtual conference directly with Skype if they are not equipped with VR technologies. But for now all these conferences are far from being realistic because the actors are represented by avatars or simple busts.

VRTogether would allow employees to have a feeling of presence of other people and the 6DoF is a more intuitive technology than a mini-teleportation system which is essential to improve the productivity and handling of prototypes.

5.2.3.3 B2B2C (Business to Business to Consumer)

- **Games**

Social VR games are rising since only few months. The goal is to overcome the problem of isolation of VR players and allow interaction between multiple players to make the experience even more realistic. Escape games have been a precursor to this kind of game, allowing players to help each other solve complex problems. MMOs and MMORPGs have also emerged with the possibility for players to meet in large numbers and large spaces to discuss and interact with each other and embody role plays.

Having a realistic photo representation would add immersion to whatever game you have.

- **TV/movie**

In the TV / film industry, VR has already made its appearance with social features that make it possible to watch sports together in a virtual salon such as LiveLike for example.

The preferred targets would be the video content platforms and TV broadcasters with the photorealistic representation of their friends, or even a person interviewed by a journalist, and the 6DoF to immerse end- users in their virtual living room.

- **Education**

Education must contain an essential social component and VR tends to isolate people in spite of the social aspect. Still, VR with a social dimension where all students are in the same virtual place would allow them to visit different parts of the world as a class trip, attend a class from another part of the world, or interact with others on virtual objects to learn.

The photorealistic representation of the students would allow the children to remain connected to the real world and not be destabilized; moreover the addition of the technology 6DoF would allow them a total freedom in their interactions in this virtual environment.

- **Social networks**

A bit like games but with a very marked tendency for the social, it is simply an extension of our communication system on social networks like Facebook with the appearance of a small conference between several friends. Today it is limited to a representation of friends through avatar and they are often static in space.

The addition of a photorealistic representation and 6DoF technology would increase the user's immersion to reach the goal of having the impression that their friend is really in the same room as us. Nevertheless social networks are not a privileged target because the actors often have their own tool (Facebook Spaces for example).

5.2.4 Conditions for successful exploitation

The conditions for a successful exploitation is linked to two factors: the current and future market (global and detailed), and the solution itself.

As seen in the SWOT analysis of the VR market, the conditions for a successful exploitation of our solution are: A better adoption of the headsets which is the base stone of the VR, the development of WiFi and 5G allowing to enrich VR content, and the improvement of the VR technology to make it more realistic. But every target of the VR-Together project has its own conditions for successful exploitation which are recounted below. Also, hardware and network infrastructure costs must decrease to provide an ecosystem solution for a correct price.

Business meetings: Companies look for productivity over the sensation of reality. A photo-realistic representation of a colleague will thus have no big interest because today the majority of the conferences are made by telephone without even the video. The real interest is to facilitate the interactions on a product between colleagues separated by distance, in particular for design phases. Also an interest could be to push the training of new employees at the level of reality as high as possible so that the training in VR is the most faithful to the real situation with whom the employee will be confronted.

Games: The only influential factor is the speed with which technology evolves. A photo-realistic rendering is directly related to the ability of game hardware.

TV/movie: In the live video broadcast field (Sport events, Music events, VOD, etc.), the VR (360 and point-cloud) streams needs to be synchronized with the classic TV stream to maximize the user experience and provide a seamless content visualization experience between various clients and platforms.

Education: For education, the educational content must follow. Schools cannot offer VR to students if it does not teach them anything, but if there is enough content to educate students with VR, the photo-realistic representation of students among them would certainly be appreciated to connect students to reality, so that they are not isolated in their VR content.

Social Networks: As previously explained, social networks are not a privileged target. Nevertheless for the exploitation to be successful it is necessary to sell a technology of photorealistic representation and 6DoF for a total immersion to big actors of the social network such as Google or Facebook rather than to create our own social network.

6 CONCLUSION

This marketing analysis document has firstly shed light on the strengths and weaknesses of VR to assess the potential of the market; secondly to identify the ecosystem of existing technologies to associate most relevant ones with the operation of the project; and lastly, a value proposition by identifying our ecosystem, our potential targets and the strategy to address each of them.

We have seen that VR is a market at the very beginning of life, only a small portion of the population took a step towards VR by equipping themselves with current technologies. Unfortunately for the majority of the population, the technology is too expensive for the experience it delivers; some believe that the technology is just not mature enough. There is a disillusion for the general public; nevertheless there is a growing expectation in professional uses. In particular, this data enables us in Part 6 to define our potential targets for the VRTogether project.

Today's technologies remain dominated by a handful of actors regarding the headsets: Facebook's Oculus and HTC. At least one of them is used in the VRTogether project, as well as for deep capture systems: Intel and Microsoft with Kinect are the two main actors. Furthermore, a section on 3D rendering engines and encoding and delivery provide essential information for future choices that consortium will make on the ecosystem.

Finally, a value proposition hints at the position to adopt to best exploit the objectives of the VRTogether project with an ecosystem scheme that details the role of each brick. Depending on the sectors of activity targeted, a discussion on the market potential of each of them was effected by stating that the majority of the targets would be B2B. For all potential targets, the study ends with proposals for strategies to make the most of each of them.

7 ANNEXE 1: STANDARDIZATION ACTIVITIES 2017-2018

7.1 3GPP SA4 / VR-IF Santa Clara

The Industry Forum Workshop on VR Ecosystem and Standards took place in Santa Clara from 4 - 6 December. Close to 150 participants from companies such as Apple, AT&T, Dolby, Ericsson, Intel, Movielabs, Oculus, Samsung, Qualcomm and Verizon attended sessions on standardisation coordination, VR content production, service provider and VR Hardware/device manufacturer challenges.

The workshop was co-chaired by Rob Koenen, VRIF President and Principal at VRTogether partner TNO. TNO further participated in the event by promoting Social VR, with Emmanuel Thomas showcasing an early-stage demo that allowed visitors to get a first feel for how the project will enable shared and social VR experiences. The participants provided useful feedback on the relevance of social VR and the demonstrated technology, and were keen to test and use further evolutions of the technology. Interestingly, many presenters talked about the need for making VR a more social experience. More info at http://www.3gpp.org/news-events/3gpp-news/1903-virtual-reality-ws

7.2 DVB : Geneva

The DVB Steering Board meeting in Geneva on 5 July 2018 was attended by Viaccess-Orca. At that meeting it was decided, due to lack of contribution and redundancy work done by other fora like VR-IF, to pause work on VR for 6 months and to revisit the topic at SB91 (February ’19).

7.3 MPEG

7.3.1 MPEG 121 Gwangju

TNO participated and actively contributed in the MPEG 121st meeting that took place in Gwangju, China from 22 January to 26 January. At this meeting, MPEG promoted its “Media Orchestration” (MORE) standard to Final Draft International Standard (FDIS), the final stage of development. MPEG expects that the Media Orchestration standard to be especially useful in immersive media settings. This applies notably in social virtual reality (VR) applications, where people share a VR experience and are able to communicate about it. Media Orchestration is expected to allow synchronising the media experience for all users, and to give them a spatially consistent experience as it is important for a social VR user to be able to understand when other users are looking at them. TNO provided input showing past experiments on social VR.

7.3.2 MPEG 122 San Diego

The MPEG 122st meeting that took place in San Diego from 16 April to 20 April was highly attended, by CWI, Motion Spell and TNO. Work done by the VRTogether partners on point cloud compression, network based media processing (standardised ingest), Omnidirectional Media Format (OMAF), MPEG-More (Media Orchestration) and timed metadata. TNO showed a demo of the current state of a Social VR experience, attended by over 40 visitors.

7.3.3 MPEG 123 Ljubljana

The MPEG 123rd meeting that took place in Ljubljana, Slovenia from 16 July 2018 to 20 July 2018 was attended by both CWI and TNO. Looking at priorities and capabilities, CWI was now focusing on objective and subjective testing of point cloud compression. TNO expressed the need for a general Social VR framework, instead of scattered metadata contributions.

7.3.4 VR-IF

7.3.4.1 VR-IF Santa Clara

The VR-IF Santa Clara workshop that took place from 4 December 2017 to 7 December 2017 was a joint workshop with 3GPP (see 7.1). TNO demonstrated Social VR and the Social VR use case was recognized as a top three major next step. This resulted in the study of Social VR in the Requirements Workgroup.

7.3.4.2 VR-IF Reading

The VR-IF Reading (UK) workshop that took place from 7 and 8 March 2018 was attended by TNO. The Social VR use case and requirements form a major part of VRIF's liaison letter to MPEG.

7.4 WorldWideWeb Consortium (W3C) Brussels:

W3C organized a workshop on WebVR authoring from December 5-7, 2017 in Brussels. The primary goal of the workshop was to bring together WebVR stakeholders to identify unexploited opportunities as well as technical gaps in WebVR authoring. Topics that were discussed included the status of WebVR support in browsers, its performance, audio support, accessibility, control and interaction, and the future of WebVR developments. TNO participated in the workshop, in particular in the discussion on social VR. Focus was on today's avatar-based social VR approaches, with discussions on avatar consistency over various platforms and applications. TNO further presented 3 talks, on i) usage of WebRTC in combining communication with WebVR, mostly talking about the blend of WebRTC with WebVR and introducing our SocialVR system, ii) limits of high-resolution omnidirectional video, on the current limits of textures and videos in VR, focusing on some hands-on experience in their usage, and iii) a presentation on the performance problems (e.g. CPU / bandwidth) we face on our web-based platform. More info at https://www.w3.org/2017/09/webvr-authoring-workshop/

7.5 VQEG: Madrid

The VQEG workshop in Madrid that took place from 19 March 2018 to 23 March 2018 was attended by CWI. The Immersive Media Group focused on Quality of Experience of immersive and interactive VR. CWI is actively participating in this working group.

<END OF DOCUMENT>