This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101092889, Topic HORIZON-CL4-2022-HUMAN-01-14

SHARESPACE

Embodied Social Experiences in Hybrid Shared Spaces

Project Reference No	101092889
Deliverable	D4.1. Definition of SHARESPACE Avatar
Workpackage	WP4: Interaction-Aware Avatar Animation and Rendering
Nature	D (Deliverable)
Dissemination Level	PU - Public
Date	18-Dec-23
Status	V1.0
Editor(s)	Stéphane Donikian (Golaem) Panayiotis Charalambous (CYENS) Amir Azizi (CYENS) Andreas Panayiotou (CYENS)
Involved Institutions	Golaem, DFKI, UM, CYENS, DMU, INRIA, AE
Document Description	This deliverable presents the topology, skeleton and appearance of the SHARESPACE virtual human to allow for robust animation including locomotion, face and fingers.

CONTENTS

List of Tables

List of Figures

Table 1: List of Abbreviations

Term / Abbreviation	Definition
AI	Artificial Intelligence
Mocap	Motion Capture
FACS	Facial Action Coding System
APFS	Anatomically Plausible Facial System
VR	Virtual Reality
XR	eXtended Reality
AAA	Classification term used for games with the highest development budgets and levels of promotion
LED	Light Emitting Diode
DCC	Digital Content Creation
HMD	Head Mounted Display
GDPR	General Data Protection Regulation
PoP	Proof of Principle

1 INTRODUCTION

1.1 PURPOSE OF THE DOCUMENT

This deliverable presents the topology, skeleton and appearance of the SHARESPACE virtual human to allow for robust animation including locomotion, facial expressions and finger movement. Avatars will comply with participant requirements with minimum modeling effort and data input. Common practices in defining animated characters (e.g., Epic's Metahumans, Reallusion’s Character Creator) in combination with the methods developed in WP3 will be considered. Special care will be taken so that the defined avatar considers the skills and autonomy of each of the different levels of avatars (L1-L3), and the specificity of each scenario such as riding a bike.

1.2 STRUCTURE OF THE DOCUMENT

This document is structured as follows:

- The first part of this document includes the description of the state of the art in the technologies used to model and animate an avatar.
- The second part is dedicated to ethical and legal issues related to the representation of a real person through an avatar in a collaborative space involving social interactions.
- Finally, the third part is devoted to presenting the chosen approach, before concluding this document.

2 STATE OF THE ART

2.1 INTRODUCTION

A Participant is a real person immersed in the Shared Space and who collaborates with other agents, such as Virtual Humans, within this Shared Space. A Shared Space is an eXtended Reality environment shared by Participants located in at least two different physical locations. An Avatar is a 3-dimensional character which represents a real Participant in the Shared Space. While it may not look the same as a Participant physically, the Avatar is usually based on the Participant's appearance and acts similarly in motion with him or her. Virtual Humans are computer-based simulations of human beings. A Virtual Human has three different uses:

1) be an Avatar, as defined above;

2) be a figurative virtual character in the Shared Space;

3) be an Autonomous Virtual Character who will interact with the Participant(s) in the Shared Space.

In the SHARESPACE project, only the first and third cases will be addressed. In the SHARESPACE proposal, the term avatar is used for all levels (L1-L3) however, to be consistent with the literature, in the rest of this document, we will not only address Avatars but more generally Virtual Humans as for L3, no Participant is controlling the Virtual Human.

2.2 REVIEW OF THE APPROACHES

2.2.1 Introduction

The range of representation of an avatar can be very large as it is going from low-def cartoon avatars such as the ones sold on the Avatoon website[1] to hyper-realistic avatars such as the ones done by Digital Domain[2] (cf Figure 2-1).

Figure 2-1: Human reference and avatar 3D model (Avatoon on the left and Digital Domain on the right).

Both approaches are based on real data coming from a real person, however for Avatoon, the only input is a picture used by an expert illustrator to manually create a 3d avatar, while for the DigiDoug project performed by Digital Domain, the process is far more complex and time expensive: countless

[1] https://avatoon.net/services/custom-cartoon-avatar
[2] https://digitaldomain.com/technology/project-digi-doug

hours of voice recordings and facial captures, body motion, skin motion, subdermal blood flow, pore structure changes, skin reflectance capture. There is a huge gap between those two solutions, however both require manual work. We will now present different levels of complexity and of realism available to model an avatar. For each level of complexity, we will see how complex it is and if it is reachable by using a combination of off the shelf technologies.

2.2.2 Digital Doubles

A Digital Double is a computer-generated 3D replica of a real-life person able to represent him or her in a 3D virtual world, such as DigiDoug. The idea for a digital double is not to build a virtual human from scratch by using 3D modeling software but to use different techniques to capture data from a real person and try to obtain the better level of accuracy of his/her replica. This is used for many years by film, television and video game industries for various purposes:

- to create virtual versions of real-life athletes in sport games like *FIFA*[3] and *NBA 2K*[4];
- to create memorable and iconic characters in action games such as *Red Dead Redemption*[5], *Call of Duty*[6] or *Assassin's Creed*[7];
- to create memorable battle scenes between Neo and multiple instances of agent Smith in *The Matrix* trilogy[8];
- to create the entirely computer-generated world of Pandora and its inhabitants in the *Avatar* movie[9];
- to age or de-age a real actor in a movie such as Brad Pitt playing Benjamin Button at different ages in *The Curious Case of Benjamin Button* movie[10], or actors Robert De Niro, Al Pacino and Joe Pesci de-aged to portray younger persons in *The Irishman* movie[11];

[3] https://www.ea.com/games/fifa
[4] https://nba.2k.com
[5] https://www.rockstargames.com/reddeadredemption
[6] https://www.callofduty.com
[7] https://www.ubisoft.com/en-us/game/assassins-creed
[8] https://www.imdb.com/list/ls076404175/
[9] https://www.avatar.com/movies/avatar
[10] https://www.imdb.com/title/tt0421715/
[11] https://www.imdb.com/title/tt1302006/

To do so, it requires the use of high-end solutions based on complementary technologies:

- **3D Human Scanning System:** several techniques exist, but the most used solution is based on pipeline using multiple cameras for 3D reconstruction (cf Figure 2-2). The technique to reconstruct a 3D model form multiple images is called photogrammetry and several software packages exist to manage this task, either open source such as Meshroom[12] or commercial such as Agisoft Metashape[13], PhotoModeler[14], or RealityCapture[15].

Figure 2-2: Camera setup for a 3D scan.

- **PBR (Physically Based Rendering) textures:** to enhance the realism, another technique called photometric stereo adds light sources to the set of cameras that flash in order. With photometry you can get highly-detailed albedo, normal map, and height map from multiple images that are needed for the realistic rendering of characters in a CG scene (cf Figure 2-3). Even if some software exists such as Details Capture[16], there is no specific solution to manage this huge amount of data out of the box, and each studio has developed its own solution.

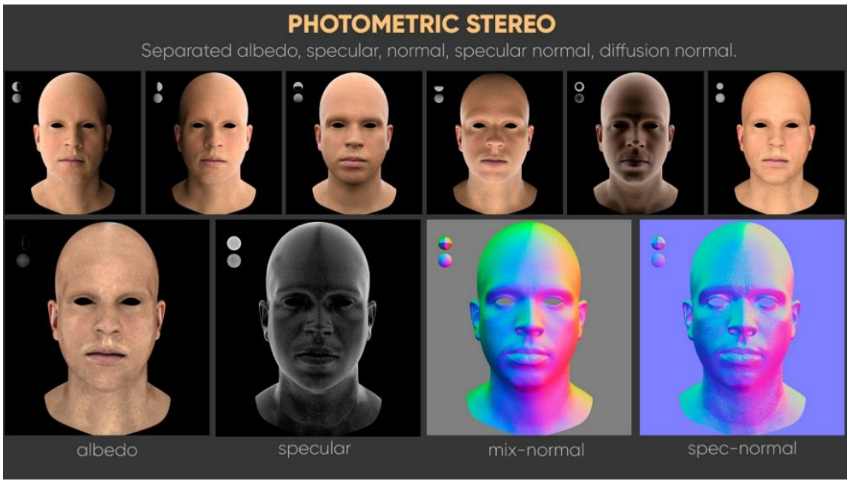

Figure 2-3: Photometric outputs.

[12] https://alicevision.org
[13] https://www.agisoft.com
[14] https://www.photomodeler.com
[15] https://www.capturingreality.com/realitycapture
[16] https://www.vfxgrace.com/details-capture

- **Full body motion capture**: Motion capture is the process of recording the movement of objects in general and humans in particular. In the case of human motion capture, the objective is to record actions of human persons and to use that information to animate corresponding digital characters in a 3D world. Usually, what is captured is related to the motion of the skeleton of the person, and s/he should wear markers near each joint to identify the motion by the positions and angles between markers. Depending on the capturing techniques used, markers can be acoustic, inertial, magnetic, emissive (LED) or reflective (examples shown in figure Figure 2-4) and they are tracked at a frequency optimally at least two times faster than the frequency rate of the motion to be recorded[17].

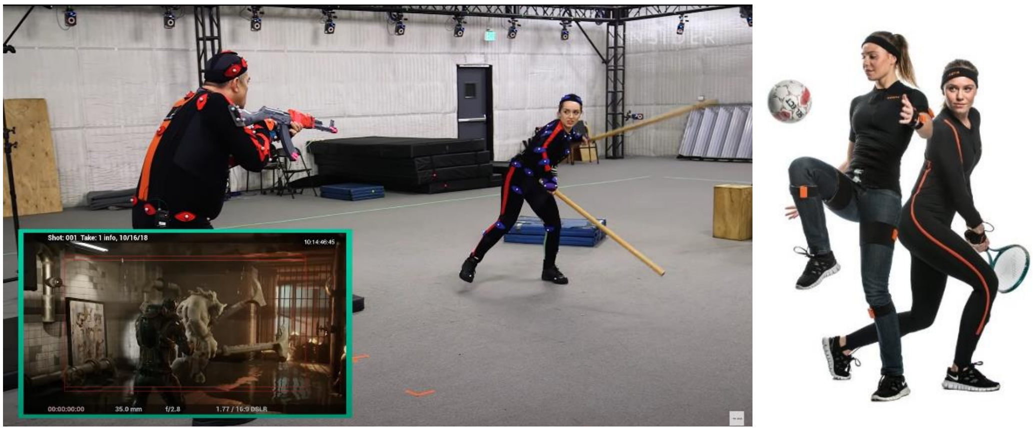

Figure 2-4: Optical motion capture (left), inertial and magnetic motion capture (right).

More recently, markerless approaches have been developed removing the constraint of wearing specific equipment for tracking.

- **Full body rig to control body motions**: rigging is a technique that links together a mesh representing the body of a character and a set of bones forming the skeleton of the character (cf Figure 2-5), usually organized through a hierarchical structure of interconnected bones.

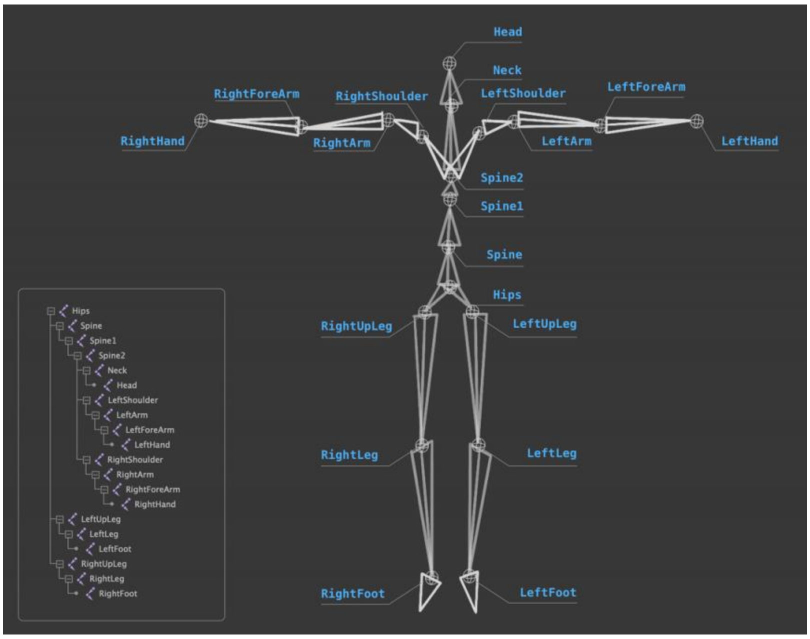

Figure 2-5: Example of a virtual human skeleton.

[17] i.e., the Nyquist Frequency.

Each bone has a three-dimensional transformation from the default bind pose (which includes its position, scale and orientation), and an optional parent bone. Each bone in the skeleton is also associated with some portion of the character's mesh in a process called skinning, and several bones can influence the same vertex of the mesh, especially near an articulation of the skeleton (cf Figure 2-6).

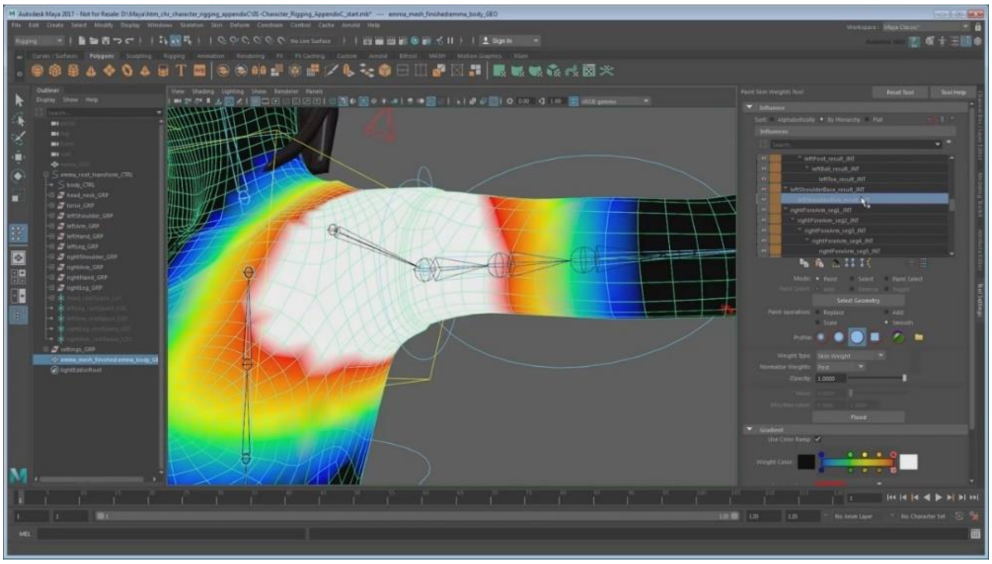

Figure 2-6: binding a modeled surface to a skeleton (skinning).

Animation controllers are used to animate the character through the modification of bones' transformation over time. Forward and inverse kinematics animation techniques (cf Figure 2-7) are usually used to compute the animation of a character. In Forward Kinematics, each joint is treated as a separate entity, and the animator specifies how they should move. However, as they are connected together, acting on a bone will have an influence on the others. For example, moving or rotating an arm will make the hand follow. With Inverse Kinematics, the control is on the end effectors (extremity of a hierarchical chain) and the algorithm computes the motion of all other joints of the chain based on the one of the end effectors.

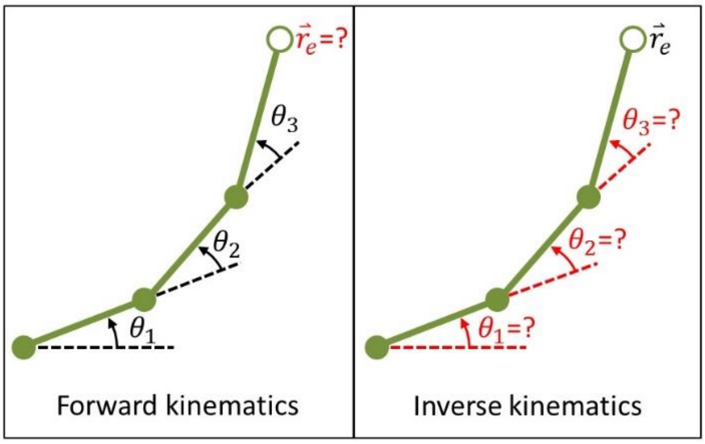

Figure 2-7: Forward vs Inverse Kinematics.

- **Facial motion capture**: Facial motion capture is the process of electronically converting a person's facial movements into a digital database using sensing peripherals (cf Figure 2-8). Traditional marker-based systems apply up to 350 markers to an actor's face and track their movements with a set of infrared high-definition cameras. Markerless solutions have also been developed allowing less preparation time but also less precision. For example, Rokoko markerless mocap solution captures only 52 different facial blendshapes.

Figure 2-8: Vicon Facial Marker System (left), Faceware Markerless Solution (right).

- **Facial rig to control facial expressions**: Facial rigging is the process of adding controls to a face for animating facial expressions. These controls are commonly bound to either bones or blend shapes, or a combination of them. Bones are invisible objects that form a hierarchy of joints and segments, that are used to deform the face mesh. A deformation bone is a bone that directly controls the deformation of a mesh, while a control bone does not have any vertices directly assigned to it but instead it controls other bones or blend shapes. Blend Shapes allow to deform directly the mesh of the face, the only restriction being not to add or remove vertices. After the blend shape has been created it is possible to blend between the original mesh and the blend shape mesh using linear interpolation. To generate complex facial expression, it is possible to use multiple blend shapes at the same time, and their effect on the vertices of the original mesh are calculated additively. It is possible to combine blend shapes and bones and it offers the best of both worlds: the flexibility of bones and the fine control of blend shapes. Many facial control rigs are based on the Facial Action Coding System (FACS). FACS refers to a set of facial muscle movements that correspond to a displayed emotion. Originally created by Carl-Herman Hjortsjö with 23 facial motion units in 1970, it was subsequently developed further by Paul Ekman, and Wallace Friesen in 1978[18], and was updated by them in 2002. As stated recently by Joe Letteri (Weta FX Senior VFX Supervisor), the FACS system is an emotion-based system to code facial expressions, but not a system to encode dialogue which is very important for talking characters. They propose[19] a new approach called Anatomically Plausible Facial System (APFS) which is based on the use of 178 muscle fiber curves that can contract or relax to provide fine-grained high-fidelity human facial expressions (cf Figure 2-9).

[18] Paul Ekman and Wallace V. Friesen. Facial Action Coding System. *Environmental Psychology & Nonverbal Behavior* (1978).
[19] B. Choi, H. Eom, B. Mouscadet, S. Cullingford, K. Ma, S. Gassel, S. Kim, A. Moffat, M. Maier, M. Revelant, J. Letteri, and K. Singh (2022). *Animatomy: an Animator-centric, Anatomically Inspired System for 3D Facial Modeling, Animation and Transfer*. In SIGGRAPH Asia 2022 Conference Papers (SA '22). ACM, Article 16, 1–9.

Figure 2-9: The muscle curves (L) and the tracking markers (R) in APFS.

Several companies have invested in this area and have developed a dedicated digital double production pipeline:

- The Digital Human Group at Digital Domain[20] is pioneering the art and technology of digital and virtual humans for all media delivery platforms. To create DigiDoug, the digital double of Dr. Doug Roble, senior director of Software R&D of the Company, Digital Domain integrated machine learning into its creative process, but needed enormous amounts of data for the technique to work. So, they started by taking thousands of images of Roble's face using different angles and lighting stages to capture as much data as possible. The team built a deep neural network that would take those images and learn all the details about Roble's face — from the mix of expressions to physical details — so it could compute all the information that would allow DigiDoug to act like a real person (cf Figure 2-10). DigiDoug is then rendered in real time using NVIDIA RTX technology and Unreal Engine, the game engine developed by Epic Games. Roble's movements and expressions are captured using a special motion-capture suit and camera. That data is transferred to the machine learning software, and then translated to the rendered digital human in real time.

Figure 2-10: Doug Roble wears the motion-capture suit and camera (image on the left) that brings DigiDoug to life (image on the right). ©Digital Domain.

[20] https://digitaldomain.com/digital-humans-lab/

- DI4D[21] is a 20 years old studio devoted to the development and use of stereo photogrammetry for 3D and 4D facial capture. They have been used over the years for video games such as *EA's FIFA, Far Cry 5*, or *Call of Duty: Modern Warfare II*, TV series such as *Love Death Robot* or movies such as *Blade Runner: 2049 (cf* Figure 2-11*)* or *The Matrix Resurrections*.

Figure 2-11: Rachael's return in Blade Runner 2049 movie.

- Based in Paris and Monaco, EISKO[22] is a company specialized in the technical production, legal management and artistic implementation of 3D Digital Doubles and Digital Humans of first-class celebrities (cf Figure 2-12) and top-tiers companies.

Figure 2-12: Digital Double of Lionel Messi, football player.

[21] https://di4d.com
[22] https://eisko.com

- Based in Seoul, EVR Studio[23] is a digital content developing company focused on creating realistic environments and lifelike digital humans (cf Figure 2-13) for AAA console games, Metaverse platforms and XR.

Figure 2-13: Project TH in Unreal by the EVR Studio.

- Weta FX has developed an entirely new face pipeline for James Cameron movie Avatar 2[24] called APFS (as presented above). APFS is an animator-centric, anatomically inspired system for facial modeling, animation, and re-targeting transfer (cf Figure 2-14).

Figure 2-14: Actress Sigourney Weaver retargeted to younger character Kiri in Avatar: The Way of Water movie.

This level of quality requires both a huge amount of time and money for each single digital double and is not adapted for an XR experiment in which we want to model the avatar for many participants in a short period of time and at a reduced cost.

[23] https://www.evrstudio.com/digital-human/?lang=en
[24] https://www.imdb.com/title/tt1630029/

2.2.3 Modeling photorealistic Virtual Humans from preset characters

Digital Content Creation software offers solutions to model and animate photorealistic virtual humans by using as a basis a preset character instead of the specific data captured on a real person.

2.2.3.1 MetaHuman

MetaHuman[25] from Epic Games is a complete framework that gives anyone the power to create, animate, and use highly realistic digital human characters, however it is restricted to be used in an Unreal Engine project. MetaHuman presets are based on pre-existing scans of real people and only physically plausible adjustments can be made through a lot of control parameters (cf Figure 2-15).

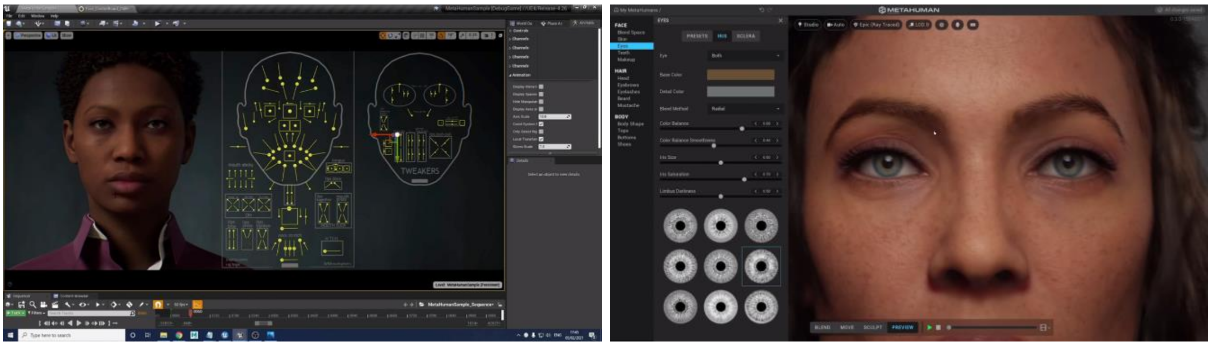

Figure 2-15: Example of control parameters within MetaHuman.

With a huge range of facial features and skin complexions, plus many different choices for hair, eyes, clothes, and more, it is possible to create a truly varied array of characters (cf Figure 2-16). Many variations upon genre, ethnicity and age are possible:

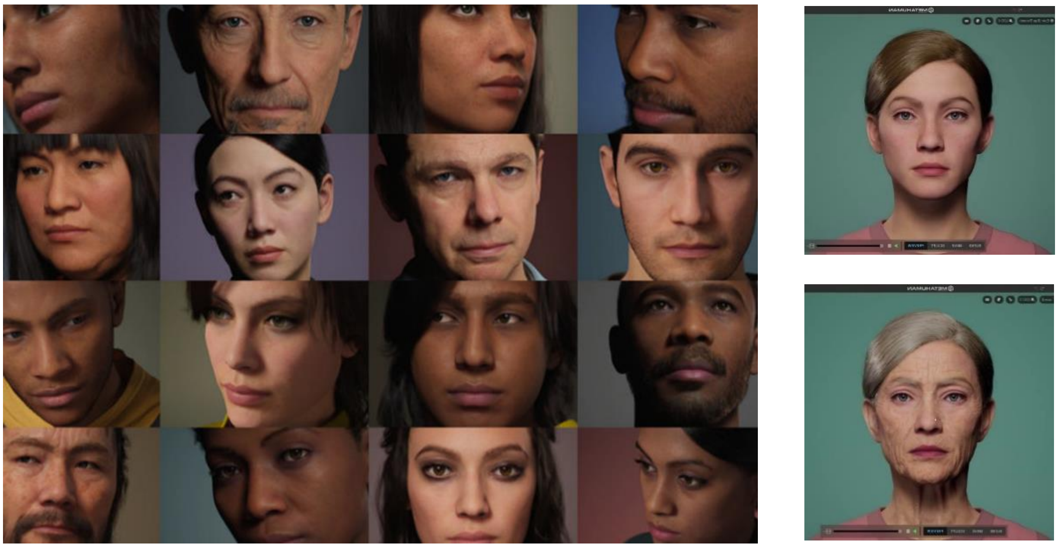

Figure 2-16: Possible variations upon genre, ethnicity and age within MetaHuman.

To customize a MetaHuman to be morphologically and visually similar to a real participant calls for a long manual process. From a reference image of the face of a person, the process is fully manual by playing with all control parameters. A feature exists to map a head mesh obtained by photogrammetry onto a MetaHuman[26]. For body adaptation, there are eighteen unique body shapes available (9 for male and 9 for female), made up of different height (Short, Average, Tall) and body proportion

[25] https://www.unrealengine.com/en-US/metahuman

[26] https://dev.epicgames.com/documentation/en-us/metahuman/mesh-to-metahuman-for-unreal-engine

(Underweight, Medium, Overweight) combinations, so it is not possible to reproduce precisely the morphology of a real person.

Once created, a MetaHuman can be animated through many different ways by using all character technologies available within Unreal Engine, including physics-based animation and ML Deformer. By using Live Link, it is possible to do performance capture (cf Figure 2-17) by connecting the character to a full body or facial mocap system:

Figure 2-17: Performance Capture.

2.2.3.2 Character Creator

Character Creator is a full character creation solution, developed by the Reallusion company, for designers to easily generate, import and customize stylized or realistic character assets for use with iClone, Maya, Blender, Unreal Engine, Unity, or any other 3D tools.

Character Creator provides a modular character design with many complementary dedicated tools (shaders, skin & makeup, morphs, dynamic wrinkles, hair &beard, outfits & accessories, expressions) and for each of them a huge amount of control parameters (cf Figure 2-18). While MetaHuman is mainly focused on the realism of the head of a character, Character Creator is more versatile and provide a lot of complementary tools for the different parts of the body (hands, eyes, mouth, hair, body, clothes, …). It is even possible to export the full body to a MetaHuman character.

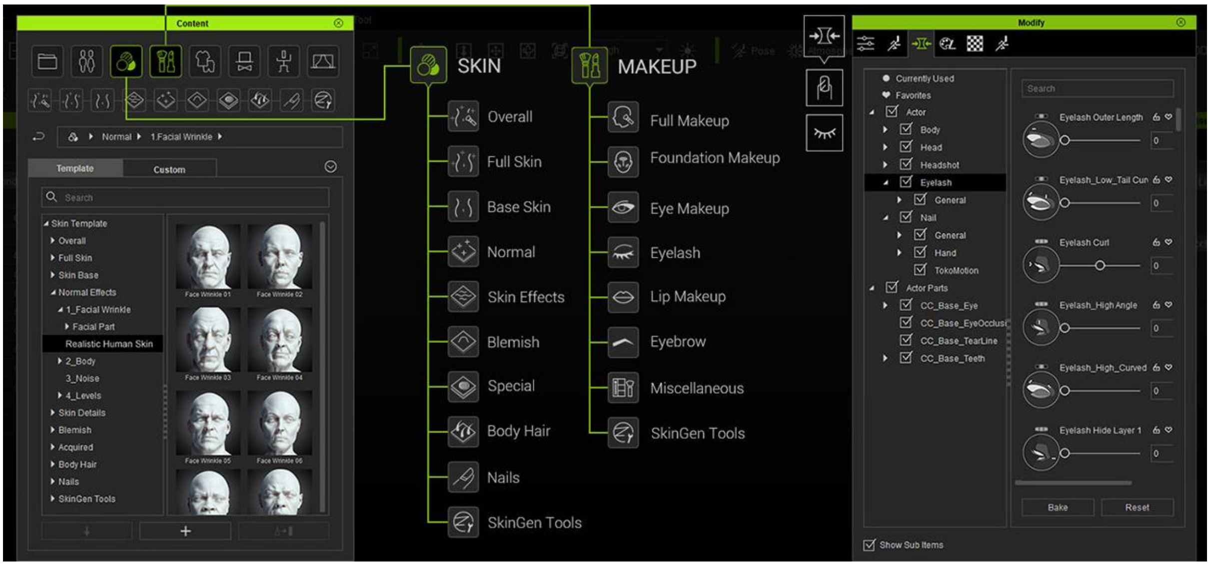

Figure 2-18: Example of control parameters within Character Creator.

The Headshot plugin contains several workflows to adapt the head of a character to data coming from a real person including:

- Image workflow (cf Figure 2-19): the Auto Mode generates a complete 3D head and a 3D hair mesh based on the image input. The Pro Mode provides a fully control of the head generation process with loads of advanced features and super sculpting tools, and it produces 3D heads with up to 4096 texture resolution while the Auto Mode is restricted to 1k texture resolution.

Figure 2-19: Auto Mode on the left, Pro Mode on the right.

- Mesh workflow (cf Figure 2-20): This AI based process allows designers to handle diverse mesh conditions (complete head, covered face, incomplete mesh, full body mesh and stylized mesh) to generate a fully-rigger head of the character.

Figure 2-20: Mapping of a 3D Scan mesh of a head within Character Creator.

Character Creator provides free addons to speed up the character design workflow with major DCCs: GoZ for ZBrush, CC Tools for Blender, and CC AutoSetup (shader and bone structure) for Unreal UE4 & UE5 and Unity Mecanim. UDIM compatible with Substance Painter, and rigged character compatibility with Marvelous Designer. For example, the CC AutoSetup tool automates the task of Digital Human shader assignment and characterization for Unreal Engine 4 & 5 (cf Figure 2-21).

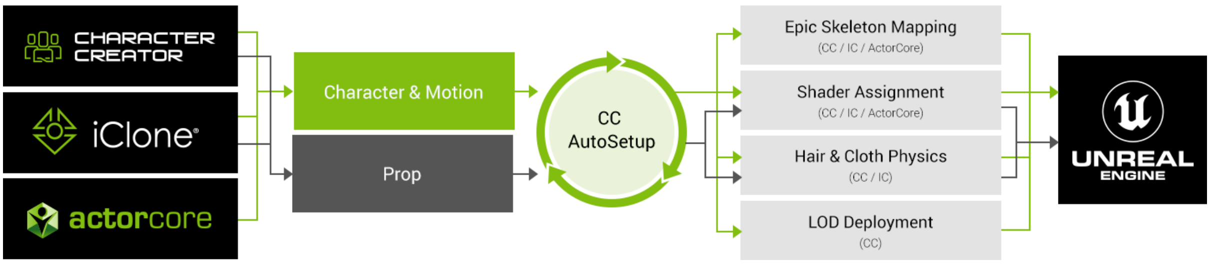

Figure 2-21: Pipeline between Character Creator and Unreal Engine.

2.2.4 Possibility to use an anonymized picture

For SHS participants who don't want to use their own data, it is possible to use instead an anonymized picture. For example, the anonymizer website[27] provides a set of AI generated pictures that are morphologically similar to the provided picture of the real person. Figure 2-22 and Figure 2-23 present the result for two members of the SHS project:

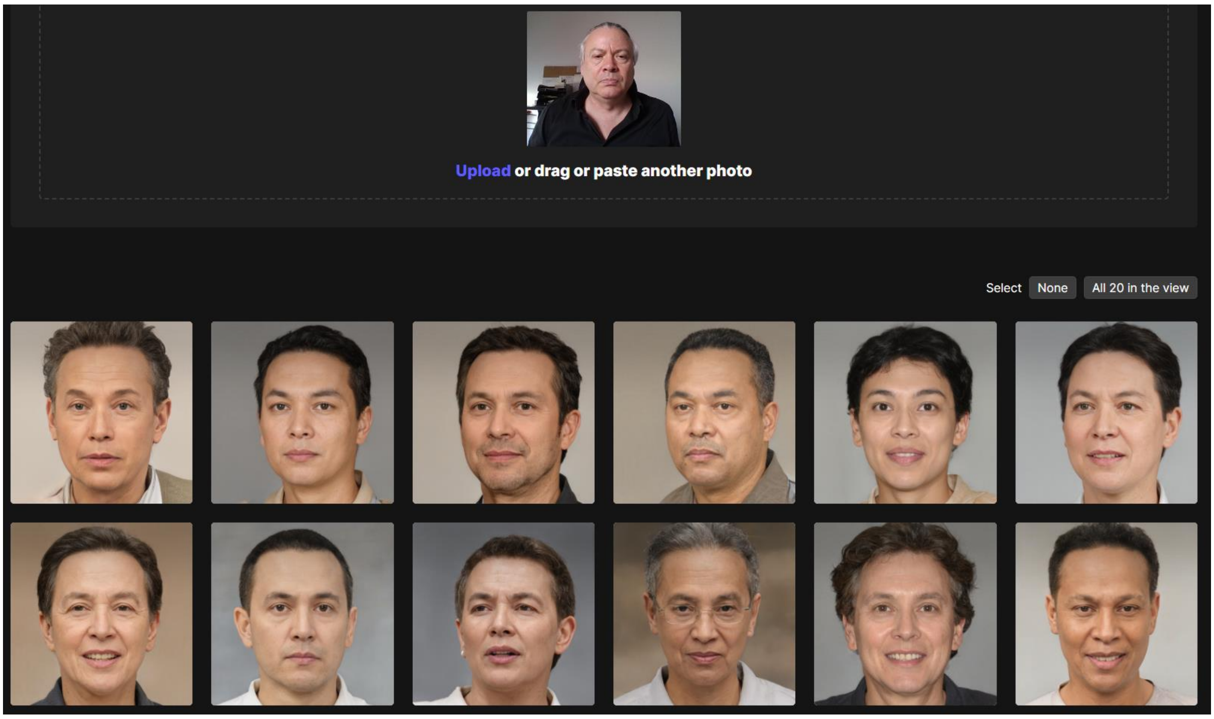

Figure 2-22: Proposed anonymized pictures for Stéphane Donikian (Golaem).

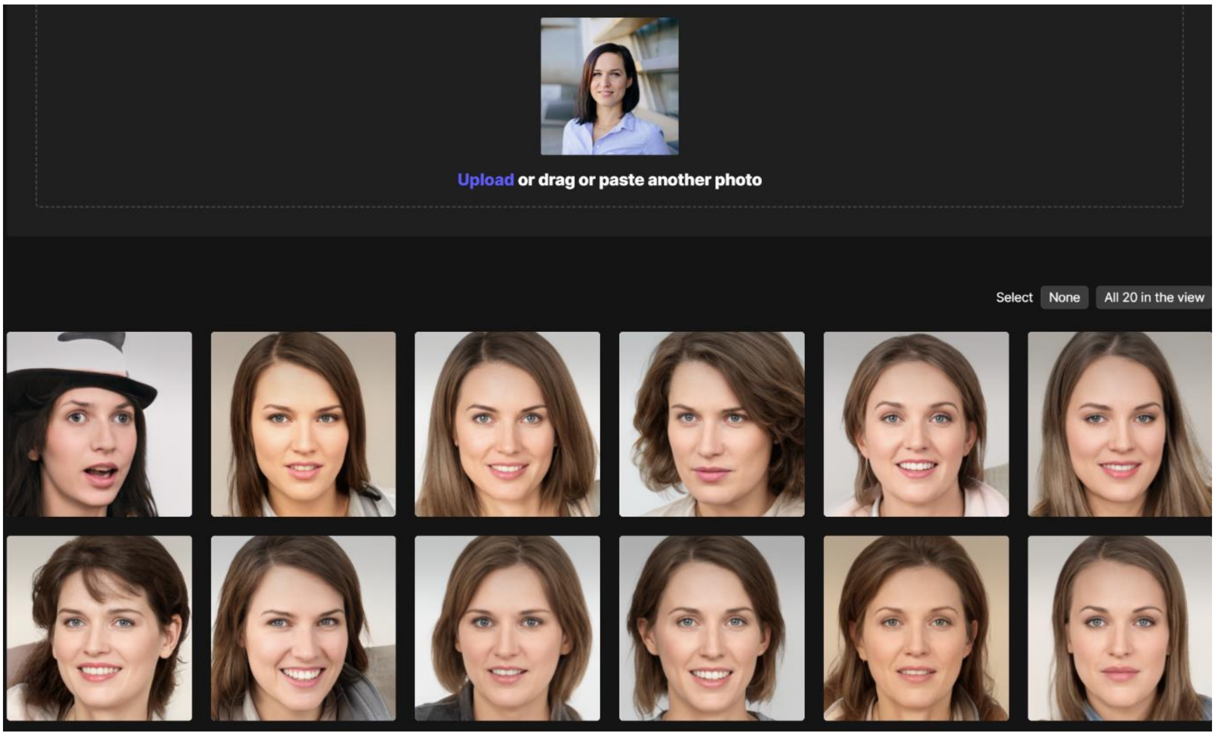

Figure 2-23: Proposed pictures for Marta Bienkiewicz (UM).

[27] https://generated.photos/anonymizer

2.3 IMPACT OF THE REALISM OF AVATARS ON USER INVOLVEMENT

In the first study related to this topic in 2017, M. Latoschik et al.[28] provide evidence that avatar realism impacts the quality of social interactions. The VR application involved self-perception in a mirror and perception of other avatars. When users interact with photorealistic avatars, generated with a state-of-the-art photogrammetry 3D scan pipeline, they develop a higher sense of ownership of their own body than when they interact with a wooden mannequin, and this affects engagement with others. The three different avatar representations are shown in Figure 2-24. Measurement concerns humanness, acceptance and change factors. The study has also shown that the appearance of other avatars influences our self-perception in Virtual Reality.

Figure 2-24: avatar representation used by M. Latoschick et al.

Weidner et al. [29] propose a systematic review of the 72 research papers published between 2015 and 2022 related to the visualization of avatars and agents in Augmented Reality and Virtual Reality displayed using Head-Mounted Displays. Among contributions of the paper, they provide an analysis of the effects of rendering styles and body parts. Among findings of this review, those of interest for SHARESPACE are listed below:

- Articles comparing various body-part visibility configurations report significant advantages of a full-body avatar compared to other configurations. Advantages range from task performance, communication behavior, user experience, presence and social presence;
- Realistic rendering outperforms other rendering styles and improves embodiment, body ownership, presence and social presence;
- User experience and task performance benefit from realistic representations;
- While more realistic avatars come with a higher technical cost, their benefits are tangible;
- Some task-focused activities (e.g., gaming or education) do not benefit much from highly realistic avatars or are in some cases even hampered by them.

---

[28] M. E. Latoschik, D. Roth, D. Gall, J. Achenbach, T. Waltemate, and M. Botsch (2017). *The effect of avatar realism in immersive social virtual realities*. In Proceedings of the 23rd ACM Symposium on Virtual Reality Software and Technology, Article 39, 1–10.

[29] F. Weidner et al. A Systematic Review on the Visualization of Avatars and Agents in AR & VR displayed using Head-Mounted Displays. in *IEEE Transactions on Visualization and Computer Graphics*, vol. 29, no. 5, pp. 2596-2606, May 2023

When it comes to the perception of the biological motion of characters, other studies have shown very different results. Chaminade et al. [30] discover that the biological motion of characters, rendered in different levels of detail, was easier to recognize when their appearance was less detailed. More recently, Zibrek et al.[31] have shown that while the same animation was used for all render styles (cf Figure 2-25), the movement was perceived more realistic on a sketch character.

Figure 2-25: Render styles of the character used by Zibrek et al.

2.4 CONCLUSION

In SHARESPACE we aim to find a good compromise to obtain realistic avatars without using a long and tedious generation process. For this reason, a complex process based on 3D full body scanning should be discarded for the benefit of an approach using a configurable virtual human modeling tool, such as MetaHuman or Character Creator, taking as input data on the face and body dimensions.

However, as it is important, for a motion fidelity purpose, to approach as much as possible the morphology of the real person, MetaHuman is not adapted for this task as it allows only 9 different body shapes per gender. Character Creator is far more adaptable, and it is possible to adjust the morphology of the avatar based on body measurement and full body pictures as reference of the real person. Concerning the face of the avatar both techniques are based on processing a single picture or a full head scan, the second one (head scan) being more complex and time consuming than the first one (single picture).

[30] Thierry Chaminade, Jessica Hodgins, and Mitsuo Kawato (2007). Anthropomorphism influences perception of computer-animated characters' actions. *Social Cognitive and Affective Neuroscience* 2, 3, 206–216.

[31] Katja Zibrek, Sean Martin and Rachel McDonnell (2019). Is Photorealism Important for Perception of Expressive Virtual Humans in Virtual Reality? *ACM Transactions on Applied Perception*, Volume 16, Issue 3, 1-19.

3 ETHICAL AND LEGAL ISSUES

3.1 INTRODUCTION

An avatar being the virtual body representation of a real person proposed in a shared space to other participants, it is important to discuss ethical and legal issues of the use of such virtual self. A review of the existing literature is proposed in the next section, followed by proposed requirement for SHS in the conclusion section. A more general and extensive review of literature concerning ethical aspects is proposed in Deliverable D1.4[32].

3.2 REVIEW OF LITERATURE

K Szita in[33] is studying the links between the digital manifestation of body characteristics and social behavior in extended reality environments. This work focuses on how the graphical representation of users epitomizing various body types and demographic characteristics—such as gender or ethnicity—may impact behaviors and identity. The more a virtual body visually corresponds to a physical body in its motion or appearance, the stronger the sensation of ownership of the represented body. The use of avatars may impact the connection (or even divide) between digital and physical world identities and transport identity-related issues to physical realms: for example, avatars can create false identities and body images and can negatively impact trust and interpersonal connections. While some cultural, social, and behavioral frameworks are translated between one's physical-world and virtual identities, it is not to be forgotten that an online identity is tied to a constructed persona. This constructed persona involves a certain appearance, characteristics, and behavior which might be constrained by a particular XR platform and one's own personality, social roles, and/or personal choices.

Some of the important conclusions of this state-of-the-art executive summary that informed SHARESPACE design are the following:

- XR platforms can be considered "inclusive spaces," where one can choose their body representation and demographic characteristics as a response to the audience or a social group's dynamics;
- XR spaces may endow negative social behaviors, such as discrimination or harassment, based on avatar representation and the demographic characteristics an avatar or virtual persona presents;
- changing one's real-life gender or other demographic characteristics in XR spheres may help to avoid stereotypes or reinforcement of gender and social roles.

J. Rueda et al. [34] study the ethical aspects of the use of VR in enhancing empathy. The body transfer to an avatar consists of replacing our physical body with a virtual one. An accurate tracking of the congruent visuo-motor correlations between the real body and the "fake" virtual body is important to induce embodiment. Thus, in a first-person perspective through the use of HMDs, the use of mirrors is important to familiarize participants with their novel virtual appearance, and to confirm that the virtual body moves synchronously with the physical one. In their study they reviewed the scientific body of literature related to what they call "Virtual Reality Embodied Perspective-Taking" (VREPT), which

[32] Sharespace Deliverable D1.4: Ethical Framework. December 2023.

[33] K. Szita (2022). A Virtual Safe Space? An Approach of Intersectionality and Social Identity to Behavior in Virtual Environments. *Journal of Digital Social Research*, 4(3), 34-55.

[34] J. Rueda and F. Lara (2020). Virtual Reality and Empathy Enhancement: Ethical Aspects. *Frontiers in Robotics and AI* 7:506984.

means the use of VR technology to change participants' perspective through virtual embodiment. The range of VREPT studies includes experiences such as embodying other skin tones, genders, ages, members of disabled groups, people in situations of extreme social exclusion, and even other species. For example, racial embodiment has been widely studied with both positive[35,36,37] and negative[38,39] outcomes. The authors propose that empathetic perspective-taking is the psychological process of imagining the world from another person's point of view (sometimes known as Theory of Mind[40]), and it is dependent on individual capabilities that differ from person to person. They discuss ethical issues of the use of VREPT with three criteria:

- *Requirement of target specificity:* a subject should be assisted in the effort to empathize with a specific target with whom he or she finds it more difficult to empathize when there are good moral reasons to do so.
- *Requirement of context dependency:* empathy should be especially strengthened in situations in which empathy is morally important to make socially informed decisions.
- *Requirement of complementarity:* even in cases where promoting empathy is necessary to be a competent moral agent, it is not enough on its own to make an informed choice with social consequences. VREPTs should not serve as isolated interventions or as substitutes for other initiatives.

Maloney et al.[41] present potential risks posed by virtual avatars, which mean that avatars may produce unwanted effects on users:

- Perceptual manipulation: if there is a perceptible distance between the motion of physical body and the way the virtual body behaves in the virtual world, it may generate dissonance in users resulting in cybersickness, physical discomfort or frustration. After such a VR experience, if users are not re-calibrated back within the physical world, mental and physical discomfort may occur.
- Negative self-embodied avatar experience: Embodiment of a virtual avatar can provide either positive or negative changes to the user's behavior and can affect attitudes and self-objectification. For example, Fox et al.[42] have shown that embodying a highly sexualized avatar produced changes in women's behavior online and offline resulting in increased self-objectification. One of their findings is that women embodied in sexualized avatars that looked

---

[35] T. C. Peck, S. Seinfeld, S.M. Aglioti and M. Slater (2013). Putting yourself in the skin of a black avatar reduces implicit racial bias. *Conscious. Cogn.* 22, 779–787.

[36] D. Banakou, P. D. Hanumanthu and M. Slater (2016). Virtual embodiment of white people in a black virtual body leads to a sustained reduction in their implicit racial bias. *Front. Hum. Neurosci.* 10:601.

[37] B. S.Hasler, B. Spanlang and M. Slater (2017). Virtual race transformation reverses racial in-group bias. *PLoS ONE* 12:e0174965.

[38] V. Groom, J. Bailenson and C. Nass (2009). The influence of racial embodiment on racial bias in immersive virtual environments. Soc. Infl. 4, 231–248.

[39] R. L. Bedder, D.Bush, D. Banakou, T. Peck, M. Slater and N. Burgess (2019). A mechanistic account on bodily resonance and implicit bias. *Cognition* 184.

[40] Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind?. *Behavioral and brain sciences*, *1*(4), 515-526.

[41] D. Maloney, S. Rajasabeson, A. Moore, J. Caldwell, J. Archer and A. Robb (2019). *Ethical Concerns of the Use of Virtual Avatars in Consumer Entertainment*, IEEE Conference on Virtual Reality and 3D User Interfaces (VR), Osaka, Japan, pp. 1489-1492,

[42] J. Fox, J. N. Bailenson, and L. Tricase (2013). The embodiment of sexualized virtual selves: The proteus effect and experiences of self-objectification via avatars. *Computers in Human Behavior*, 29(3):930–938.

like themselves demonstrated greater acceptance of the rape myth than women embodied in other avatars.

- Risks Presented by Sustained Immersion: No study is related to the effect of long-term immersion with an embodied avatar, and authors point out the need of a longitudinal study to understand such risks.

Aymerich-Franch et al. [43] propose a self-guiding tool for scholars conducting responsible research with embodiment technologies. This tool is devoted to help them engage in ethical and responsible research and innovation (RRI) in the area of embodiment technologies in a way that guarantees all the rights of the embodied users and their interactors, including safety, privacy, autonomy, and dignity. They have established a list of 36 questions related to the five RRI principles: inclusion, anticipation, reflection, responsiveness and transparency. They indicate that the answer to these questions may trigger the researcher to take action. The full list of questions is presented in Appendix 1.

Lin and Latoschik[44] present a review of research works related to digital body, identity and privacy in social virtual reality. They indicate that unlike for physical bodies, digital bodies can be altered or changed at the users' will, allowing them to disguise their identity and that identity verification is still lacking in professional social VR applications; therefore, the authenticity of the identity cannot be guaranteed. They have selected and analyzed 49 research papers related to their topic, which have been published between the years 2008 and 2022. Some papers report the existence of violence and sexual harassment, such as an offender maliciously touching and rubbing body parts of the victim's avatar, or using techniques to offensively modify the victim's avatar, such as making the avatar appear naked or exaggerating the avatar's appearance features. Other papers report that identity theft is one of the most common concerns stemming from the use of avatars, especially personalized avatars, and that unfortunately, this impersonation may be legally permitted in some countries, such as United States, as it may be considered as a parody, commentary or entertainment. Lake[45] mentioned the inconsistencies of dated internet laws related to identity misappropriation in social VR. Concerning accountability and referring to the work by Gorini et al. [46], they indicate that "anonymity might negatively impact patients during e-therapy, as anyone can interact with the patients and there is no guarantee of who the patients are interacting with". Referring to further research work [47,48,49,50,51], they indicate that "the relationships between users and avatars are central to the understanding of identity

---

[43] L. Aymerich-Franch and E. Fosch-Villaronga (2020). A Self-Guiding Tool to Conduct Research With Embodiment Technologies Responsibly. *Frontiers in Robotics and AI,* Volume 7, N°22.

[44] J. Lin and M. E. Latoschik (2022). Digital body, identity and privacy in social virtual reality: A systematic review. *Frontiers in Virtual Reality* 3:974652.

[45] Lake, J. (2020). Hey, you stole my avatar!: Virtual reality and its risks to identity protection. *EMORY LAW J*. 69, 48.

[46] A. Gorini, A. Gaggioli, C. Vigna and G. Riva (2008). A second life for eHealth: Prospects for the use of 3-D virtual worlds in clinical psychology. *J. Med. Internet Res*. 10, e21.

[47] Graber, M. A., and Graber, A. D. (2010). Get your paws off of my pixels: Personal identity and avatars as self. *J. Med. Internet Res*. 12, e28.

[48] Freeman, G., Zamanifard, S., Maloney, D., and Adkins, A. (2020). *My body, my avatar: How people perceive their avatars in social virtual reality*, in Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems, CHI EA '20.

[49] Freeman, G., andMaloney, D. (2021). *Body, avatar, andme: The presentation and perception of self in social virtual reality*. In Proc. ACM Hum. Comput. Interact. 4, 239–23927

[50] Kanamgotov, A., Koshy, L., Conrad, M., and Prakoonwit, S. (2014). *User avatar association in virtual worlds*, in 2014 International Conference on Cyberworlds, 93–100.

[51] Carruth, A. D., and Hill, D. W. (2015). Identity and distinctness in online interaction: Encountering a problem for narrative accounts of self. *Ethics Inf. Technol.* 17, 103–112.

and privacy issues and are a prerequisite for ethical and legal discussions. Only when the nature of user-avatar relationship is clarified, can the violations to avatars be characterized and corresponding protection mechanisms be developed.”. The traceability of an avatar is crucial to its acceptability in the social VR[52]. User authentications are the key to combating identity theft and achieving secured identity management. For this purpose, from their review of literature, they mention four types of authentication methodologies: knowledge based, biometric, multi-model and gaze based. For example, Bader et al.[53] proposed a watermarking solution: a coded biometric information is inserted into the avatar’s face as an invisible watermark.

Shravishtha Ajaykumar[54], discusses rights and obligations of a virtual self and mentions an ethical issue with the user identity confirmation when creating an avatar:” Avatars are a manifestation of the self, beyond the realms of the physical. As user data is a requirement for avatar creation and data privacy is an essential pillar of any facet of the internet, confirming user identity for avatar creation becomes unethical. Therefore, avatar identity is self-defined and not pre-ordained. [...] identity-based actions of the avatar need platform-specific norms to avoid any incongruities concerning user data privacy. [..] To hold users directly responsible will require regulatory authorities to pre-develop rules in tandem with the ethics of confirming real-life communities and identities. Regulations in this area will invoke two main principles:

- To create an avatar’s identity in cyberspace as a separate individual and limit user creation of avatars to a certain number per user, and
- To treat avatars as separate entities (congruent to companies’ and proprietors’ existence). Both methods hinge on creating a distinct legal identity for avatars.

[…] Aside from these regulations on avatar autonomy, regulators will need to create globally relevant ethical guidelines around their usage and identity. Such ethical guidelines must include a code of conduct that are platform-specific and nurture user experience.”

Ben Chester Cheong[55] discusses some legal issues related to the avatar of a real-life person in a metaverse. The problem of the responsibility of actions of an avatar in the metaverses, would require attributing a legal persona to the avatar, but then a second question arises, which is how to chain the responsibility from the avatar to a real person as some avatars can be AI driven.” The law would need some mechanism to address actions taken against an avatar that may affect a human being behind the avatar, and to address actions taken by an avatar that may affect other avatars or people.” The author suggests that activities in the metaverse should be subject to regulation based on company law principles, and that all avatars in a metaverse should be subject to registration (like the incorporation process of companies). However, only for serious breaches in the law should the identity be disclosed as anonymity should be the standard mode. The paper is then addressing four different problems that

---

[52] This is developed in the Privacy section of SHARESPACE Deliverable 1.4.

[53] S. Bader and N. E. Ben Amara (2014). *A securing access approach to virtual worlds based on 3D mesh watermarking of avatar’s face*. In 4th International Conference on Image Processing Theory, Tools and Applications (IPTA 2014), 1–6.

[54] Ajaykumar, S. (2022). Avatars in Cyberspace: Rights and Obligations for a Virtual Self. *Observer Research Information: Expert Speak Digital Frontiers*. https://www.orfonline.org/expert-speak/avatars-in-cyberspace/ Accessed 14 December 2023.

[55] B.C. Cheong (2022), Avatars in the metaverse: potential legal issues and remedies. *Int. Cybersecur. Law Rev.* 3, 467–494.

can be created by avatars in the metaverse: fraud, identity theft, defamation and crime. The author then suggests that:

> "Where real-world harm has been caused by avatars in the metaverse, it would seem that the first step in legal proceedings would be a pre-action discovery against an incorporated avatar to force the disclosure of the real person's identity behind the avatar in order to commence real-world legal proceedings. In the event that the veil cannot be lifted to disclose the person behind the avatar, a litigation representative could be appointed to sue or defend an action on behalf of the avatar in the real World."

Concerning data protection and privacy, the author argues that the EU's General Data Protection Regulation (GDPR) could apply to the metaverse. As all peripherals used in an XR experiment will process a range of data about users and their private environment, such data are relevant to article 9 of the GDPR and should require special care in processing.

3.3 Conclusion

Character Creator seems to be the best option to model SHARESPACE avatars. Three possible options will be used for the avatar face:

- Front image of the face of the person;
- Choose an AI generated picture instead of the real picture of the person;
- Head scan based on several pictures taken from different view angles around the person.

For the body part of the avatar, to allow the best accurate motion replay, it is also needed to know the dimensions of all body segments. A non-invasive and fast solution should be used for this measurement.

To avoid negative self-embodied avatar experience, we should always validate the avatar appearance with each participant and avoid any highly sexualized clothes. For this purpose, a predefined list of clothes will be proposed for participants of each SHARESPACE Experiment. Based on the VREPT established ethical criteria, within SHARESPACE we should minimize changes of skin tone, gender, and age between the physical and the virtual body.

Concerning perceptual manipulations mentioned above, in SHARESPACE we will try to minimize them by using a measurement protocol to adapt the morphology of each avatar to that of the real user. Motion deformation that will be done intentionally by the cognitive architecture will be limited in amplitude and will locally target only a few joints. However, we will have to check if it is confusing for participants.

Concerning body motions, as they will come from the motion capture of the participant him/herself, it should not generate negative experience. We will have to however take this into account in task 4.3, which is devoted to style transfer. Facial expressions will be AI generated in real-time from user data input (audio and eye-tracking) but will only affect gaze and the area around the mouth. We will keep a neutral facial expression, except if we are able to recognize a specific expression from the audio and, in this case, we will have to validate that it is not negatively perceived by the participant.

In SHARESPACE, due to the accountability requirement, we should avoid anonymity of the therapist in the health scenario. He/she should be recognizable both by his/her voice and by the facial resemblance between the physical and the virtual head.

Concerning the needed level of realism of avatars, it should be defined per scenario and per PoP as it is needed to integrate the positive influence of high realism on the user involvement but at the same time the potential negative impact on his/her perception of the biological motion.

To facilitate the embodiment of participants of the SHARESPACE platform, we suggest integrating before each XR first experiment of participants a presentation phase of their virtual body by using mirrors to allow them to familiarize with their novel virtual appearance.

In the absence of regulations of XR collaborative platforms, we need to create ethical guidelines and a code of conduct specific to our SHARESPACE platform

As all peripherals used in an XR experiment will process a range of data about users and their private environment, such data are relevant to article 9 of the GDPR and should require special care in processing.

4 AVATAR AUTHORING PROPOSED SOLUTION FOR SHARESPACE

4.1 DATA COLLECTION

For each participant in a SHARESPACE Scenario or PoP it is required to collect data to create his/her avatar. Before any data collection, each participant will be asked to read and sign the protocol for avatar creation (cf Appendix 2).

This data collection consists in:

- One picture of the face or several pictures of the head of the Participant;
- full body pictures (front and side view) taken in front of a monochromatic and light background;
- Measure of the main body segment parts

The participant should be asked to select clothes of his/her avatar from among a few numbers of proposed clothes adapted to his/her gender, avoiding highly sexualized clothes.

The participant should decide to use his/her own picture or to select an AI generated similar picture by using the anonymizer website.

To avoid circulation of personal data between partners of the project, all data should be stored by using a unique identification number and only the SHARESPACE partner in charge of collecting data and running experiments will know the correspondence between the identity of a real person and the identification number.

For the Sport Scenario, due to the specificity and the fact that each participant will not see the face of other participants, it has been decided to use generic virtual bicyclists acquired on a specific 3D model store. The participant will only have to choose between one of the proposed 3D avatars.

4.2 AVATAR CREATION

Based on the provided data, an avatar will be created by using both Character Creator and Headshot, taking into account skin tone, gender and age of the real person. The avatar will be then exported in a FBX file format to be used afterward in Unreal Engine by the dedicated Scenario or PoP application. The name of the FBX file will integrate the unique identification number.

In order to create avatars, we need to create both body and face. Body measurements are very important to minimize artifacts such as hands passing through virtual bodies, end effectors not having precise contacts with floor or other objects, etc. The process of collecting the appropriate resources to enable the correct creation of body and face are presented in the following subsections. In order to generate full body avatars, Reallusion's Character Creator 4, in combination with the Headshot 2 plugin will be used. This software leverages cutting-edge AI technologies, such as deep learning algorithms, to produce remarkably detailed and lifelike 3D avatars. By analyzing vast datasets of facial features and structures, this system can generate highly realistic and customizable 3D heads from 2D photos or meshes. This process involves intricate pattern recognition and mapping, enabling users to create diverse characters with expressive details. The generated 3D model of the face is fully rigged and animatable, it can be attached to any body type and shape to form a fully rigged character, and then can be exported in a variety of formats including FBX, OBJ, and USD.

4.2.1 Body Representation

4.2.1.1 Body Measurements

Experiments are currently being made by DFKI to find the best compromise to obtain a good accuracy of the body measurement in a minimum of measurement time, without requiring specific hardware equipment.

4.2.1.2 Rigging

As explained in section Digital Doubles2.2.2, Rigging is crucial for creating realistic and controllable movements, serving as a bridge between the model and the animation, allowing for lifelike and complex animations. Character Creator 4 tool automatically creates a rig once the desired body is selected. Likewise, there is the functionality of adjusting the length of different body parts, something that will keep the skeleton's structure but will change its dynamics. When trying to animate multiple rigs with different skeletons with the same animation, it usually yields to strange motions. The solution to this problem is animation retargeting, where the nodes of a new skeleton are mapped to the original skeleton of the animation. This is one of the features of the Golaem's animation engine that will be used within SHARESPACE for character animation.

4.2.2 Avatar Faces through single Images

As has been mentioned above, the Headshot 2 plugin offers two options for generating avatar's head. The first, and more straightforward, way of generating the face of an avatar, is using a single image. Single images are easier to acquire, easy to edit but might not be accurate enough whereas a complete 3D model (mesh) is more time consuming but also of higher quality. However, to maximize the potential of this method, the process of obtaining the portrait photo has to follow some guidelines.

First, proper lighting conditions are quite important to acquire high quality results; to minimize shadows in the generated texture, the subject should be lit with as uniform lighting as possible (ideally with the use of an LED rig). Second, the image must capture the frontal face of the head and be as straight as possible. Third, the image must reflect as much as possible the real structure of the face, so stretched or distorted captures will yield unusual results.

Note that some things like facial hair and earrings will affect the end result, as will be baked into the generated texture. Thus, if the original texture is used, instead of a single skin tone, these constraints have to be considered.

Headshot 2 supports two face generation modes with a single image as input, Auto and Pro. The Auto mode uses a simple and smart process to create complete 3D heads while also has the functionality of separating the hair, creating separate 3D hair models. It is good at removing hair from images, adding natural textures to the sides, and dealing with shadows on faces. The Pro mode offers full control over making 3D heads and lots of advanced tools to sculpt them. It creates heads with super high-quality details, offering up to 4096 texture resolution. The user can manage how textures blend, adjust the mesh to match the original photo, and project textures accurately. Pro mode is perfect for creating virtual humans where we need close-up views and high-quality results that resemble real scans. In most cases, if the input image has been captured with ideal conditions, the functionalities offered by the Auto mode are enough.

The process of creating a face avatar through a single photo in Character Creator and Headshot 2 is as follows and is shown in Figure 4-1:

1. Photo Capturing: The camera of a smartphone can be used.
2. Photo Uploading: Users upload the photo (up to three photos can be selected).
3. Body/Skin Type Selection: Users select initial body and skin type (further modifications can be applied after the generation).
4. Key-points Selection: Certain key points on the face, such as the eyes, nose, and mouth will be automatically identified. Users can manually adjust these key points if necessary.
5. Avatar Generation: AI is used to generate a 3D model of the face. The processing time will vary depending on the complexity of the photo and hardware resources (typically takes a few minutes).
6. Avatar Refinement: Once the avatar is generated, it can be refined using a variety of tools. These tools allow users to adjust the facial features, hair, and clothing.

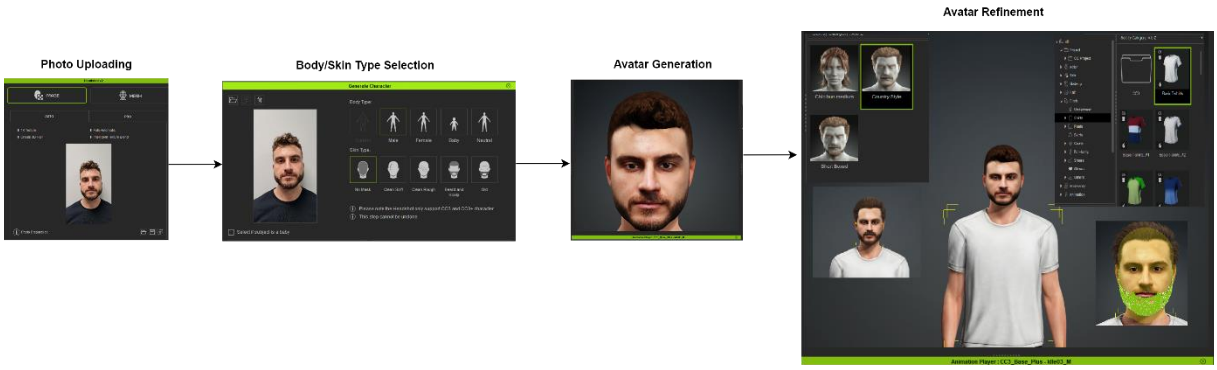

Figure 4-1: steps of the avatar creation from a single photo.

This approach proves to be highly efficient, yielding satisfactory outcomes, especially considering the small amount of time needed. However, the generated mesh is strongly affected by the quality and the content of the input image. Thus, if the result is not the desired one, the user can improve the generated model either by retaking a more appropriate photo or by spending some time post-processing the face model. Likewise, within the content section, diverse adjustments can be made, encompassing attire, footwear, hairstyle, beard, lighting, and more, enabling the generation of avatars with the desired quality. Furthermore, the Character Creator and Headshot tools afford the option to fine-tune facial features such as eye shape, nose structure, and lip size, thereby augmenting the avatar's fidelity to the original image.

4.2.2.1 Creating an avatar from AI-generated images

Using AI-generated portraits for 3D avatars is ethically considerate, especially for individuals who prefer not to use their own likeness due to privacy concerns or personal preferences. It allows users to maintain anonymity while engaging in digital spaces, offering a layer of protection against potential misuse or recognition. This approach respects individual choices and comfort levels regarding personal image exposure, making digital interactions more inclusive and secure for everyone.

The latest tools that are based on technologies such as Generative Adversarial Networks (GANs)[56] and Diffusion models[57], can create digital content using various types of inputs and conditions. Specifically,

[56] https://doi.ieeecomputersociety.org/10.1109/TPAMI.2020.2970919
[57] https://arxiv.org/abs/2112.10752

focusing on image generation, users can generate photos using as input: text prompts, images, conditions, or a combination of those. There are several websites, tools and platforms that take advantage of those models. Most of them are very intuitive to use, as the user just uploads a photo, selects a style, or enters a textual condition, and receives the generated image in real-time. Some of these models are open, meaning that users can download and use them on their own machine, something that eliminates ethics and personal data security issues. However, these models are usually computationally expensive, thus appropriate computer systems are needed. On the other hand, following an evaluation of several platforms, we opted for two tools that make the process easy and straightforward, and they are presented below.

LightX[58]: is a versatile photo editing and graphic design app that caters to both professional and casual users. It can generate images of AI avatars using advanced algorithms and machine learning techniques. Users can upload a photo and select one of the predefined styles that will condition the generated image. This website follows European regulations about storing and processing personal data. Results illustrated in Figure 4-2.

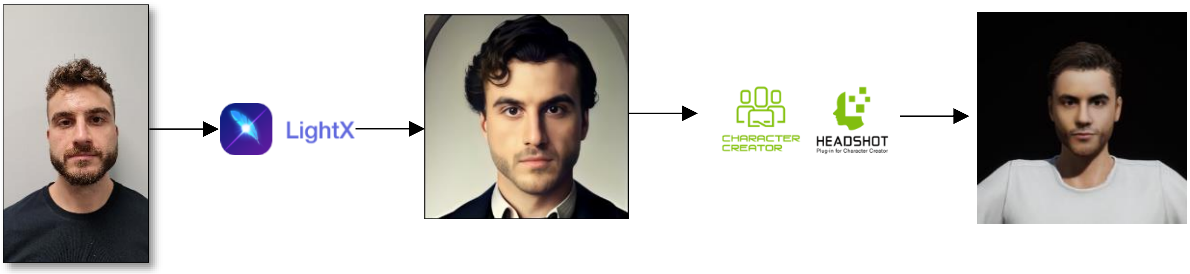

Figure 4-2: Generated 3D model character using image generated by LightX.

Anonymizer[59]**:** is a tool that can generate look-a-like images from real photos. Users upload a photo of their face, and the tool runs several iterations generating several similar images. This website follows European regulations about storing and processing personal data. Results shown in Figure 4-3

Figure 4-3: Generated 3D model character using image generated by Anonymizer.

Subsequently, we applied this approach to convert AI generated images into an avatar, utilizing the Character Creator and Headshot 2 plugin. As mentioned before, the efficacy of avatar creation is markedly influenced by the deliberate choice of lighting and poses during the procedural stages. Thus, while both platforms can be used for generating images, we argue that Anonymizer may be a better option as it generates images that are closer to real. Likewise, while LightX can produce interesting photos, the resulting images are more artistic, usually containing variable lighting and effects, that affect the performance of the Headshot 2 plugin.

[58] https://www.lightxeditor.com/photo-editing/ai-avatar-generator/
[59] https://generated.photos/anonymizer

4.2.3 Avatar Faces through Mesh Models

The second way of generating a face model for a 3D avatar, using Headshot 2 plugins, is to use a mesh as input. While there are various methods for creating 3D meshes of real objects, we chose to use Photogrammetry, because it is a relatively easy process and no specialized equipment is needed; it can be achieved using the camera of a smartphone and some open-source tools. Photogrammetry is a technique used to capture precise measurements from photographs, enabling the creation of 3D models of real-world objects or environments. The process begins by taking a series of overlapping photographs from different angles of the subject. Special software then analyzes these images, detecting common points or features across the photos and using the overlap to calculate distances and spatial relationships. This data is then used to reconstruct the subject, in a digital three-dimensional model, accurately capturing its size, shape, and texture. Specifically, for the process described above, the Polycam[60] app has been used.

As for previous section, where images must follow certain specifications, the creation of an appropriate model has to meet some conditions. First, regarding the environment, capturing should take place on a spot with uniform white lighting, using a neutral background, ideally with a color that makes high contrast with subject's skin tone. Likewise, the subject should stay still, maintaining a neutral pose throughout the capturing process. Second, regarding the capturing process, around 50 clear overlapping-images have to be captured in order to produce the desired results. These images must capture the whole head of the subject from every possible angle, both horizontally and vertically. Then, the generated model may need some cleaning, such as removing unnecessary captured elements to enhance the end results; this process can be done using an open-source software like Blender[61].

Subsequently, after the captured model is processed, is ready for import in the Headshot 2 plugin. The steps for generating a 3D avatar using a mesh as input are described below (cf Figure 4-4).

1. **Mesh Importing:** Users import the 3D model. Note that is critical to be aligned as straight as possible before starting generation process.
2. **Feature Detection**: The software can automatically detect several alignment points (24, 32, or 35). If these points are misaligned, users can manually enter the corresponding points between the reference face and the face of the imported 3D model. We noticed that 30 to 40 points are enough if the model is clean.
3. **Generate 3D Heads Based on Effective Areas**: The tool generates a 3D model of the face and gives the ability of selecting an effective area from it for the final generation. For example, the back of head may not be captured correctly, so users can select only the front part.
4. **Refine Topology with 3D Brushes**: Then, 3D brushes can be used to refine face parts including eyes, nose, mouth, ears and more.
5. **Assign Body Shape & Bake Textures:** Finally, the generated face can be attached to a desired body. Further modification is available through the software.

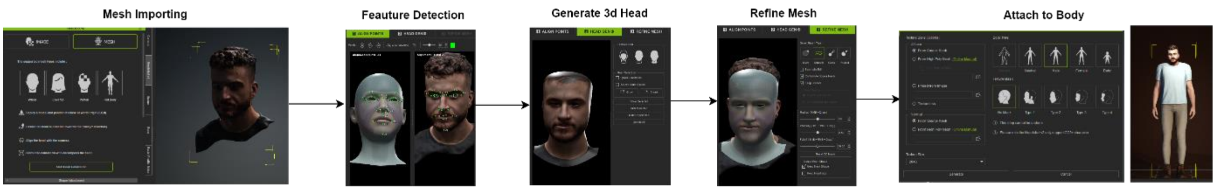

Figure 4-4: Mesh 2 Avatar steps.

[60] https://poly.cam/
[61] https://www.blender.org/

Results and process of this method are shown in Figure 4-5.

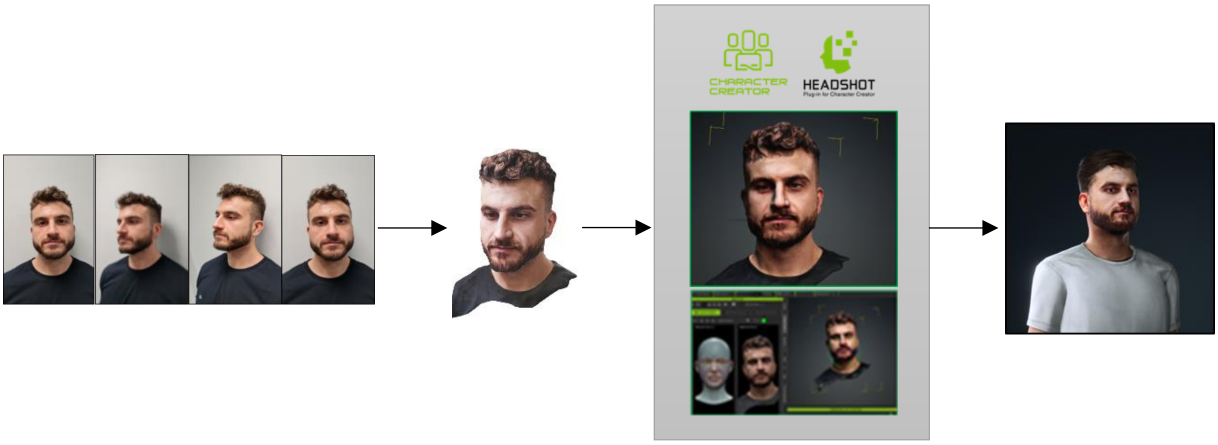

Figure 4-5 Avatar Creation using 3D Mesh.

This method, while yielding the highest quality outcomes among the techniques employed, demands a significant investment of time for design. Moreover, the utilization of multiple images significantly enhances the capacity to capture intricate details and dimensional accuracy in representing the head. Nevertheless, the time-intensive nature of this approach renders it unsuitable for projects operating under strict time constraints or constrained resources.

Figure 4-6 presents side by side the four different avatar representations of the same person animated within Unreal Engine:

Figure 4-6: Four different avatar representations (from left to right: lightX, 3D mesh, anonymizer, single picture).

5 CONCLUSIONS

This deliverable has presented the state of the art in the technologies used to model and animate virtual humans, and their consequences for the design of the SHARESPACE virtual humans. Ethical and legal issues of the use of avatars have then been presented through a review of the relevant literature. This allows us to list all issues that need to be taken into account within SHARESPACE. The next section was devoted to the description of the avatar authoring proposed solution with multiple options within the same software architecture.

For the purposes of the SHARESPACE project, the single image-based method will be used for avatar creation of both PoPs and for the patients of the Health scenario. However, for the therapists in the Health Scenario, a full scan of their head will be made to obtain the higher realism of their avatars for a better involvement of patients in the scenario.

For the Art scenario, it is too early to know which method will be used, while for the Sport scenario, a set of predefined 3D models of cyclists and bikes will be used.

APPENDIX 1: A SELF-GUIDING TOOL TO CONDUCT RESEARCH WITH EMBODIMENT TECHNOLOGIES RESPONSIBLY

We reproduce here the questionnaire built by L. Aymerich-Franch and E. Fosch-Villaronga[62] aiming at helping researchers in the field of embodiment technologies engage in ethical and responsible research and innovation.

ANTICIPATION

• What are the psychological, ethical, moral, and legal implications that the embodiment technology I am developing could bring?

• Could someone use the embodiment technology that I am developing for unintended practices in my or other fields? Who and how? If so, what measures have I developed to mitigate this?

• What are the risks and benefits of the embodiment technology that I am developing? How will they be distributed in society if the technology gets implemented?

• If I am designing avatars that will interact with other avatars or real people, do I provide enough protection mechanisms to my users so that they can protect their avatars in case of assault? Which ones?

• If I am working with physical avatars such as robots that could potentially harm someone in case of technical failure, have I considered emergency mechanisms so that the user can completely stop the action of the avatar or alternative control mechanisms to avoid causing harm to others or to the environment? Which ones?

• Who needs to take responsibility if something goes wrong with the embodiment technology that I am developing?

• Have I projected potential scenarios in which different aspects go wrong and have I determined who should take responsibility on each case? Which ones?

• What aspects of the avatars that I am developing could make them not socially desirable? How can I change that?

REFLECTION

• Who could be negatively affected by my research with embodiment technologies? How can I change that?

• How could the embodiment technology I am developing challenge the rights of future users? How can I avoid that?

• Am I equipped with enough knowledge to identify and address the implications of embodiment technologies by myself? If not, what experts could help me and how?

• Do I lead or participate in actions oriented at addressing concerns and fears regarding embodiment technologies for the general public? Which ones? Is there anything else I can do?

---

[62] L. Aymerich-Franch and E. Fosch-Villaronga (2020). A Self-Guiding Tool to Conduct Research With Embodiment Technologies Responsibly. *Frontiers in Robotics and AI,* Volume 7, N°22.

• Do my results contribute to providing useful insights for developers regarding how to commercialize embodiment technologies safely for the society from an ethical point of view? How exactly? Is there anything else I can do to ensure that?

• Do my results contribute to providing useful insights for regulators and legislators regarding how to regulate and legislate embodiment technologies? How exactly? Is there anything else I can do to ensure that?

• Do I discuss the ethical implications of my results when I report them?

• Have I conducted an actor analysis to understand which actors play an essential role for the embodiment technology that I am developing and identify people, industries, institutions, or organizations who can affect or are affected by the technology?

• Have I sufficiently reflected on the benefits and risks of the embodiment technology before starting its development: does it honestly and positively contribute to society? How exactly? Is it aimed to resolve a societal challenge? Which one/s?

• Have I organized discussion groups with the different stakeholders involved in the embodiment technology that I am working with to discuss potential ethical implications and create awareness of responsibility and accountability?

• Am I open to receive criticism and consider skepticism about the embodiment technology or experiments using it and integrate it into my research/design process? What could I do differently to encourage more feedback from the relevant stakeholders?

• Does my research with embodiment technologies have long-term consequences that could be potentially negative for society? How can I avoid them?

INCLUSION

• Am I familiar with the embodiment technologies that other researchers, companies, and start-ups are developing and the experimental work that other researchers are conducting with these technologies? What can I do to know them even better?

• Who are the relevant stakeholders in the development of my research in embodiment technologies? Have I included them in the process? How exactly?

• Could the virtual and robot avatars that I create cause gender, race, religious or age discrimination? How can I avoid this?

• Do I respect the principle of diversity when I design avatars so that I have enough avatar choices that can represent well all potential participants or users and do not create conflict in terms of gender, race, ethnicity, religion, and other demographics? (e.g., I have a participant that wears a hijab in real life and does not feel comfortable embodying an avatar that does not, do I have the right type of avatar for her?) How could I increase the range of options to ensure that?

• Have I talked directly to the stakeholders that my embodiment technologies target to enquire about how the technology could really improve their quality of life rather than making assumptions about it? (e.g., I am designing a robot avatar for persons with reduced mobility, but I actually never asked them directly if that is something that could be useful for them and how exactly would they want it) Who else should I talk to?

• Do I engage the target of the embodiment technologies I design in my R&I processes throughout all the stages? How exactly? What else can I do to ensure that?

RESPONSIVENESS

• What do I need to do to ensure social desirability for the research I conduct with embodiment technologies?

• What training am I receiving to conduct research with embodiment technologies responsibly?

• If I encounter ethical conflicts, are there any barriers (e.g., economic interests) that prevent me from changing the course of the development of the embodiment technology? Which ones? What can I do to overcome them?

• Is the technology of embodiment that I am developing prepared to evolve with the constantly evolving technological landscape? How exactly? What else can I do to prepare it better in this regard?

• Is the technology of embodiment that I am developing responding clearly to current societal needs? To which ones specifically?

• If I conduct experiments with embodiment technologies, are they meant to respond to societal needs and challenges? To which ones specifically?

TRANSPARENCY

• Are the motivations of the embodiment technology that I am developing transparent, honest, and geared toward the public interest? (e.g., I write a grant in which I argue that the embodiment technology will be useful for rescue operations just to get the grant, but I already know it will not be useful)

• Am I sharing the results not only with the scientific community but also with the target of the embodiment technology? How am I doing so?

• Am I openly sharing the uncertainties and limitations of the embodiment technology that I have developed with the full range of stakeholders as well as in my publications? All of them? (If this is something difficult for me, ask: What is the worst that could happen if I did so?)

• Do I share the lessons learned of the research with embodiment technologies with my community (including the negative aspects)? How exactly do I do that?

APPENDIX 2: PARTICIPANT AVATAR CREATION PROTOCOL:

Objective: This experiment which is part of the SHARESPACE project aims to create digital avatars representing participants. These avatars may be used for various purposes, including research, visualization, or artistic expression. The process involves capturing your appearance and characteristics for conversion into a digital representation.

Procedure:

1. **Consent and Privacy:**
   - Before proceeding, you will be required to provide informed consent for the creation and use of your digital avatar.
   - Your privacy will be respected, and any personal information will be handled in accordance with applicable privacy regulations and ethical standards.
2. **Data Acquisition:**
   - You will be scanned or photographed using camera, 3D scanning or photogrammetry technologies (we should leave the appropriate). This process may involve capturing your body shape, facial features, and clothing.
3. **Modeling and Rendering:**
   - The collected data will be processed to create a 3D digital model of your likeness.
   - Details such as textures, skin shading, and clothing will be added to make the avatar as visually accurate as possible – there might be slight inconsistencies due to technical limiations.
   - The avatar may also include basic animations for facial expressions and movements.
4. **Use and Dissemination:**
   - The resulting digital avatar may be used for research, artistic, or visualization purposes, but it will not be used for commercial or exploitative purposes without your explicit consent.
   - The avatar may be displayed in academic publications, presentations, or exhibitions, but any personal identifiers will be removed to protect your privacy.
5. **Feedback and Control:**
   - You have the right to provide feedback on the avatar's representation and appearance.
   - If you are uncomfortable with any aspect of the avatar's portrayal, you can request adjustments or withdrawal from the experiment.
6. **Storage and Security:**
   - Your scanned or photographed data will be securely stored and protected to prevent unauthorized access.
   - The data will be used solely for the purpose of creating the avatar and will not be shared with third parties without your consent.
7. **Duration:**
   - The avatar creation process may take several hours to complete, depending on the complexity of the avatar.
8. **Questions and Concerns:**
   - If you have any questions, concerns, or require further information about the experiment, you are encouraged to contact the experiment coordinator or researcher.
9. **Voluntary Participation:**
   - Your participation in this experiment is entirely voluntary, and you have the right to withdraw at any time without consequences.

Note: Please carefully read and understand this protocol before participating in the experiment. If you agree to proceed, your informed consent will be requested before data acquisition begins.