Human-Centered Metaverse and Digital Environments, Vol. 99, 2023, 112–121
https://doi.org/10.54941/ahfe1003937

Anonymous Collaboration in Metaverse

Taneli Nyyssönen[1], Olli I. Heimo[1], Seppo Helle[1], Teijo Lehtonen[1], Tuomas Mäkilä[1], and Jussi S. Jauhiainen[2,3]

[1]Department of Computing, Faculty of Technology, University of Turku, 20014 University of Turku, Finland
[2]Department of Geography and Geology, Faculty of Science, University of Turku, 20014 University of Turku, Finland
[3]Institute of Ecology and the Earth Sciences, Faculty of Science and Technology, University of Tartu, 51003 Tartu, Estonia

ABSTRACT

This paper introduces a method enabling anonymous collaboration in the metaverse. As the Internet is turning to its next revolution in which time and space are incorporated more thoroughly to the everyday communication, visual and aural communication are given more prominent roles. This might lead to more biased interaction since people may possess pre-judgmental attitudes regarding other people's age, gender, and ethnical attributes, which could impact collaboration. These attributes are rather easy to visually change within the metaverse by simply changing the avatars' appearance. This research focused on the real-time aural communication in the metaverse, and the risks of harmful profiling based on people's language use and voice. The research setting was a custom-made metaverse, where two anonymized participants interacted with each other to achieve the goal through collaboration. In the study, their interaction was monitored. Afterwards, they answered to questions in which told about their experiences and tried to de-anonymize the other participant. This preliminary research indicated that via modulating the sound, it was possible to anonymize users to a reasonable range within the custom-made metaverse.

Keywords: Metaverse, Ethics, Anonymization, Virtual reality, Linguistic profiling, Collaboration

INTRODUCTION

Metaverses are a quickly developing field of communication and interaction within the Internet. A metaverse is a rather complex and not unilaterally agreed concept where the future of the Internet is seen embedded within virtual or mixed reality. Metaverses can be understood as post-reality universes, places within the digital realm where the users can interact with the environment, digital objects, and each other. Some challenges in the development of a metaverse can be its environmental, social, and economic sustainability, including innovative and equal possibilities for collaboration. (Lee, Braut, Zhou, et al. 2021; Mystakidis, 2022; Jauhiainen et al. 2023.)

In online worlds, using an avatar is an easy way to hide one's identity. However, as metaverses rely not only on visual but also aural methods for communication, it is rather easy to make presumptions, linguistic profiling, on the other actors' identity through their voice. These presumptions can

alter one's behavior in the group and make the group dynamics less equal through pre-assumed race, gender, or age roles. In many cases this can lead to unintentional or even purposeful discrimination (Baugh, 2016).

To counter these presumptions, this paper introduces the concept of *anonymized collaboration in metaverse*. In this setting, the purposeful design of the metaverse environment hampers the participants' ability to make prejudgements about the collaboration partner(s) one has. This is achieved by anonymizing the users behind identical pre-made avatars (as people may aim to design avatars to represent their visual appearance) and altering their voice (audio signal), thus making the actors' age, gender, ethnicity, and social status less identifiable. By taking these actions, it can be assumed that the collaboration between different actors in metaverse becomes more equal and less focused on accustomed or pre-determined societal roles.

The test was conducted with a tailor-made metaverse built upon the Unity game engine utilizing Meta Integration SDK (Unity Technologies, 2023a). The task of the test-subjects was to collaborate in a virtual game of solving a labyrinth. The hardware in use was Meta Quest 2 VR glasses (Meta, 2023) for all users. The user interface utilized hand tracking with inside out facing cameras which approximated the users' hand positions using visible light. The voice communication between the users could be modulated by increasing the pitch. The virtual environment was designed to be static and visually neutral so that the participants would focus on the task instead of the metaverse itself. In the metaverse, the participants accomplished tasks with the support of one facilitator.

BACKGROUND

Metaverses

The term Metaverse is a complex and not unilaterally agreed upon. It consists of the idea around the on-going revolution of the Internet where the virtual and mixed reality become increasingly superimposed to services of the Internet. It can be seen as a post-reality space within the digital realm where users interact with environments, digital objects and each other, and in which the time and space are incorporated more thoroughly to the concept of the Internet (Lee et al., 2021; Zuckerberg, 2021; Fernandez & Hui, 2022; Mystakidis, 2022; Wang et al., 2022). According to Lee et al. (2021):

> *"The term 'metaverse' has been coined to further facilitate the digital transformation in every aspect of our physical lives. At the core of the metaverse stands the vision of an immersive Internet as a gigantic, unified, persistent, and shared realm."*

The Internet, also when consisting of metaverses, is an information system. Every time we alter a part of the information system, including the ethical basis behind the system, we change how the whole system works, including the ethical ramifications (Leavitt, 1964; Nurminen, 1986, Nurminen & Forsman, 1994; Heimo, Kimppa & Nurminen, 2014; Heimo, 2018, pp. 46 – 47). Hence, as the metaverse revolution alters the methods on how people conduct their daily digital lives, the ethical issues arisen should be taken into consideration before the revolution has already occurred.

Grandiose predictions have been announced about the metaverses' future: current Internet activities can move to various metaverses en masse, public and private services included. Parts of both business and leisure can be managed in metaverses as they become marketplaces, workplace meeting rooms, communication arenas, and platforms for gaming. People will be able to meet each other in 3D spaces, have discourse upon matters of their fields of interest, educate themselves, participate in decision-making, create new art and science, and enjoy themselves, all under the same service umbrella. Yet the details on how the concept of metaverse revolution will turn out is unclear at best. (Lee et al., 2021; Zuckerberg, 2021; Mystakidis, 2022; Time, 2022; Wang, 2022; Wang et al., 2022.)

Bias in Interaction

Bias and discrimination towards other people are rather common phenomena amongst humans. Moreover, people are rather accustomed towards their social structures that promote those voices that are in power at the expense of those who have the role to listen or submit. This can occur in various ways regarding on the social structure of the situation, for example in work places the older, more experienced, and more educated workers tend to have more social standing and thus can dominate the interaction (Schieman, Schafer & McIvor, 2013). Additionally, Eagly and Karau state that the overlap with a person's perceived characteristics and a job role correlates positively with the perceived competence in that role (Eagly and Karau, 2002; Brown Stulhmacher, and Keegin, 2014). Moreover, within our tribal mindset people tend to treat differently – or even mistreat – those who act and communicate outside of their perceived norms or ideals (Takashi, Yamagishi, Liu, et al. 2008; Baugh, 2016; Kumar, Tsoi, Lee, et al. 2021).

The linguistic profiling – inadvertent or otherwise – can lead to unwanted and unethical situations as the Internet turns more to real-time aural communication. Whereas the textual communication can reveal a lot from the writer by the choice of words and phrases, avoiding spelling errors, and general good use of language, the real-time speech and accents are even more difficult to mask (Baugh, 2016). Hence, the situations where the background of the internet-user will affect the social situations they encounter will multiply within the metaverse settings.

Research Question

As it is clear that the visual ques of the persona can be mostly anonymized by changing the avatar, the anonymization discourse focuses on aural themes. Hence, to counter the aforementioned biases and possibilities to outright discrimination, the research question of this paper is as follows:

> *"Can the level of anonymization in a metaverse be increased significantly by modulating the voice of the user without detracting from the experience of the interaction?"*

To solve this question a metaverse with suitable setting was to be constructed.

RESEARCH SETTING

Research Setup

We utilized the Unity game engine version 2021.3.15f1 and Oculus Integration SDK v46.0, (Unity Technologies, 2023a) extended with the Meta Avatars 2 SDK (Meta, 2023) to develop our testing scenario. The scenario consists of two Unity-scenes, a matchmaking lobby and the test scene. The lobby is used to connect the users to a shared virtual space, and also to determine which users are "clients" and which are "admins". The networking solution is built upon Photon Unity Networking (Exit Games, 2023) and the voice connection is handled through Photon Voice 2 (Unity Technologies, 2023b). The roles of the clients and the admin included the following abilities:

- The admin was able to move around using gesture commands (locomotion and teleporting), and also to pull client users to their current location. Additionally, the admin was able to alter the pitch of their own voice and their avatar appearance (detailed, basic, or invisible) using a menu. The admin was also able to spawn objects into the scene and move them around the maze. The admin started the test with their normal voice and with a basic avatar consisting of a sphere as a head and a cube and a sphere as hand models to be clearly distinguishable from the clients.
- The clients (the test participants) had a detailed humanoid avatar (generic and not representative of the client, represented in Figure 1.) with hand tracking feature for the simulation of the users' real hand and arm movements with the avatar. The clients were only able to move physically in the room space and had no gesture command functionality. The clients' voices had their pitch increased using Unity's Audio Pitch Shifter Effect component so that their voices became indistinguishable. The pitch increase (from default 1 x multiplier to the maximum of 2 x) made the clients' voices sound higher without speeding them up.

Figure 1: Test subjects discussing.

The test contained one admin, who instructed the test, and two participating clients. The clients were in separate rooms during the test and had no knowledge of each other in the test to preserve the anonymity. The clients were connected to the test server through the lobby scene with the help of external test supervisors situated in their respective testing rooms. The admin connected first and waited until the clients had connected to instruct both at the same time. The main task for the clients was to direct a cube controlled by the admin through a small maze, which was occluded from them using fog particle systems. The clients were situated on a high up platform overseeing the entire maze to make navigation choices easier to visualize and to make room space movement unnecessary. The platform contained a railing to make the clients feel safer. Before this the clients had to decide on a color out of four options: green, red, blue, and purple. This decision determined the color of their cube which they needed to direct to the goal marked with the corresponding color, visible to them.

The navigation was accomplished by the clients via discussing and deciding what direction should the cube go to next. The possible directions were visible on the cube's current location with arrows pointing to nearby tiles in the maze and additionally on the far end wall as text. The decision was communicated verbally to the admin, who made the movement command afterwards. Each time a cube moved to an undiscovered tile in the maze, the fog was lifted from that tile permanently, revealing the movement options from that tile to the surrounding ones.

The maze was 4x6 tiles large, not including the starting tile which offered options to take one of the four starting paths illustrated in Figures 2 and 3. Some tiles afforded back and forth movement, while others were monodirectional only. The possible movement options were LEFT, RIGHT, UP, and DOWN, and these directions matched the clients' viewing directions, apart

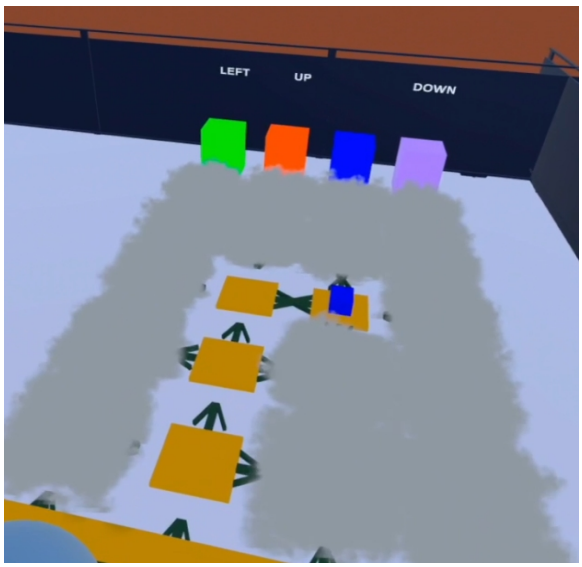

Figure 2: "Fog of war".

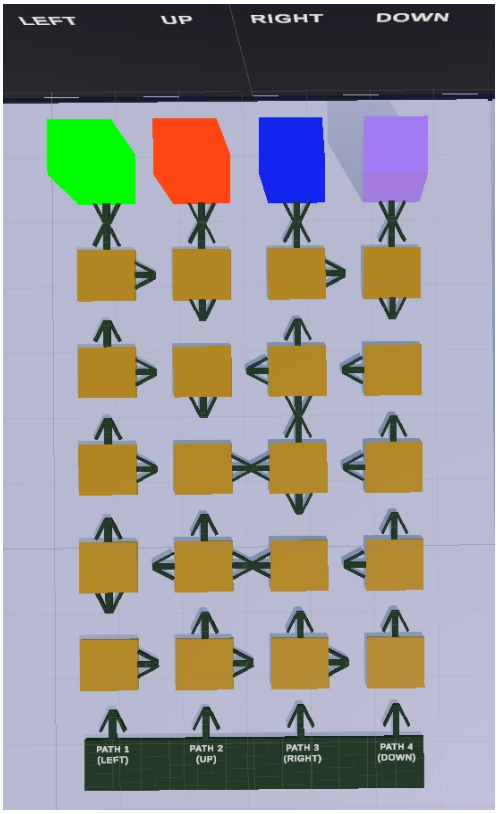

Figure 3: Game setup.

from the first decision which contained four paths going up. These were the main options the test subjects were to discuss upon.

Test Subject Selection and Guidance

In total, 8 participants completed the test. The test groups consisted of people from different ethnicities, language proficiencies, and genders. The user tests were conducted for four groups each with two test subjects, three for modulated voice and one for reference group without voice modulation.

Some participants were pre-selected to increase the variety among the test users, while others were recruited using convenience sampling. The test user selection was never simultaneous, as the goal was to keep the participants (test users) unaware of each other. In practice, one participant was selected first and brought into the research room, and then the other was discreetly looked for, and brought into a different room. This was all done without alerting any other people in the facility about the ongoing test, as this could inform participants of each other.

Once a participant entered a research room, they were instructed by the research supervisor designated for that room. The instructions included the basics of VR interaction: the play area and how it is utilized to avoid collisions with room objects during the experiment, the potential for VR sickness to occur and how to operate in case it does, and information about adjusting the

headset and how the hand-tracking functions. Additionally, the supervisors emphasized the test users to not share any personal information about themselves, as that would have compromised the anonymity aspect. The test users were also told to only speak in English once the test had started; this was required as some of the test participants came from different linguistic backgrounds, and accidentally using their native language would have revealed parts of their identity to the other participant.

One participant shared their physical testing room with the admin of the test, who instructed the participants within the software and took care of the functionality of the metaverse application. Both test subjects were accompanied by supervisors who did not enter the metaverse application. The purpose of the physically but not virtually present supervisors for both participants were to ensure their safety: they were observing that the participants did not injure themselves accidentally by going outside the play area bounds, and also, they were ready to help in case of VR sickness occurrences.

Interview Questions

Semi-structured interviews were used to answer the research questions and to find out a) whether the experience with the anonymized metaverse experience was pleasing and b) whether they could de-anonymize the other participant. The test was followed by c.a. 10–15 minutes interview with each participant. The questions are shown in Table 1.

Research Ethics

The security of all participants was of utmost priority, thus an assistant was present to guide and help at all times. The VR glasses may cause nausea or other unwanted side effects, so the test was to be halted if the participant would feel ill. All participants were adults and they volunteered with no pressure nor benefit to participate. The test subject set is biased towards academic participants due to the convenience sampling used.

Table 1. Interview questions.

Q1	Name, age, and gender	Background question
Q2	Nationality and language proficiency	Background question
Q3	Profession and level of education	Background question
Q4	How did you find the gameplay?	Open question + Easiness 1–5 (hard 1 – easy 5).
Q5	Did you find VR experience pleasing?	Open question + 1-5 (not at all 1 – very pleasing 5)
Q6	How communication went? (easiness)	Open question + 1-5 (very cumbersome 1 – very easy 5)
Q7	How did you find the communication with the other player? (pleasantness)	Open question + 1-5 (very unpleasant 1 – very pleasant 5)
Q8	Who was the other player? Gender? Ethnicity? Education level? On what basis?	Guess + Level of sureness for each: unsure – quite sure – very sure.
Q9	Open discussion.	Open question

RESULTS

The results indicate that the aural and visual anonymization of participants was possible in the metaverse environment that was designed and tested for the purpose of this paper. Of six tested participants, five participants could not guess other participants' identity and were rather unsure with their guesswork, whereas one was quite correct and certain about their guess

The gameplay was considered rather easy (4.2/5) as was the overall communication (4.3/5). The tested participants found the communication with the voice modulation pleasant (4.5/5), and such was also the VR experience (4.0/5), even though some compatibility issues were found among users with the need to use glasses with the VR headset. The use of standardized avatars was evaluated to mask the identity of the participants, thus anonymizing them even more.

Some tests took shorter time than others and thus limited the discussion between the participants, reducing the possibility to identify the other user. Therefore, in further studies, we aim to use more complex and more discussion-provoking tasks, such as jigsaw puzzles, or other commonly understood brainteasers. Furthermore, as the maze was easy to understand, it reduced the participants' need to discuss upon the possible options on how to proceed during the test.

DISCUSSION AND CONCLUSIONS

From the small test sample, we can conclude that anonymization in the metaverse by standardizing the avatars and altering the voice is possible – at least to some extent – and anonymization of collaboration can be implemented there. Further studies are required to confirm this observation. Such testing needs to be extended to more varied groups of people and more complex tasks. The accomplishing of the test was rather easy, which limited the participants' need for oral communication. For the future, a more complex and discussion-provoking test-setting, such as a common 9-piece jigsaw puzzle, could serve the research aim better.

The results indicate that it is possible to create anonymized metaverses or anonymized spaces within metaverses so that people's prejudices do not matter in their collaboration activities, including the judgements based on the voice of (an)other participant(s). This facilitates several situations where equality between participants would be an advantage. This would allow that the character and knowledge of a user in the metaverse could triumph over other users' prejudices and perceptions. As the Internet of tomorrow is forming and the metaverses are emerging, the possibility to anonymize the users to bring them to a more equal level of social interaction could indeed be an ethical choice – at times.

Moreover, this approach gives a number of possibilities for multidisciplinary applications in the fields of psychology, social sciences, and medicine, to further understand the human as an individual and as a member of a group. The social structures we build with each other can be studied and improved by understanding their formation and this tool could indeed be a valuable one.

The next step in the study regarding anonymization in the metaverse, in accordance to the multidisciplinary research, is to create test situations in which more communication and collaboration are needed, and to include more participants in these situations.

ACKNOWLEDGMENT

The authors would like to acknowledge the financial support from the Kone Foundation and the global network FAPI at the University of Turku. This study belongs to broader development of sustainable and responsible metaverse platform SURE (sites.utu.fi/sure).

REFERENCES

Baugh, J., (2016). Linguistic profiling and discrimination. *The Oxford Handbook of Language and Society*, pp. 349–368.

Brown, S. G. U., Stuhlmacher, A. F. and Keegin, M. M. (2014). Job Role Congruence and Gender, in Michalos, A. C. (ed.) *Encyclopedia of Quality of Life and Well-Being Research. Dordrecht: Springer Netherlands*, pp. 3448–3451. doi: 10.1007/978-94-007-0753-5_4089.

Eagly, A. and Karau, S. (2002). Role Congruity Theory of Prejudice Toward Female Leaders, *Psychological review*, 109, pp. 573–598. doi: 10.1037//0033-295X.109.3.573.

Exit games (2023). *Photon Unity 3D Networking Framework SDKs and Game Backend | Photon Engine* Available at: https://www.photonengine.com/pun (Accessed: 31.1.2023).

Fitzgerald, C. and Hurst, S. (2017). Implicit bias in healthcare professionals: a systematic review, *BMC Medical Ethics*. BMC Medical Ethics, 18(1). doi: 10.1186/s12910-017-0179-8.

Heimo, O. I., (2018). Icarus, or the idea toward efficient, economical, and ethical acquirement of critical governmental information systems. *Turun yliopisto*.

Heimo, O. I., Kimppa, K. K. and Nurminen, M. I., (2014). Ethics and the inseparability postulate: How to make better critical governmental information systems. *In Proceedings of Ethicomp*.

Jauhiainen, J. S., Krohn, C. & Junnila, J. (2023). Metaverse and sustainability: Systematic review of scientific publications until 2022 and beyond. *Sustainability* 15, 346. doi.org/10.3390/su15010346

Kumar, M. M. et al. (2021). Nationality dominates gender in decision-making in the Dictator and Prisoner's Dilemma Games, *PLOS ONE*. Public Library of Science, 16(1), pp. 1–21. doi: 10.1371/journal.pone.0244568.

Leavitt, H. J., (1964). Applied organization change in industry: structural, technical and human approaches.

Lee, L. H., Braud, T., Zhou, P., Wang, L., Xu, D., Lin, Z., Kumar, A., Bermejo, C. and Hui, P., (2021). All one needs to know about metaverse: A complete survey on technological singularity, virtual ecosystem, and research agenda. *arXiv preprint arXiv:2110.05352*.

Meta (2023). *Meta Quest 2: Immersive All-In-One VR Headset | Meta Store | Meta Store*. Available at https://www.meta.com/fi/en/quest/products/quest-2/ (Accessed: 31.1.2023).

Mystakidis, S., (2022). Metaverse. *Encyclopedia*, 2(1), pp. 486–497.

N. Ruggs, E., R. Hebl, M., Singletary Walker, S. and Fa-Kaji, N., (2014). Selection biases that emerge when age meets gender. *Journal of Managerial Psychology*, 29(8), pp.1028–1043.

Nurminen M. I. & Forsman, U. (1994). Reversed Quality Life Cycle Model. Human Factors in *Organizational Design and Management* - IV, pp. 393–398. Elsevier Science B. V., North-Holland, Amsterdam.

Nurminen, M. I. (1986). People or Computers: *Three Ways of Looking at Information Systems*. Studentlitteratur, Cartwell Brat Ltd. Lund, Sweden.

Schieman, S., Schafer, M. H. and McIvor, M. (2013). The Rewards of Authority in the Workplace: Do Gender and Age Matter?, *Sociological Perspectives*, 56(1), pp. 75–96. doi: 10.1525/sop.2012.56.1.75.

Takahashi, C. Yamagishi, T. Liu, J. H. Wang, F. Lin, Y. Yu S. (2008). The intercultural trust paradigm: Studying joint cultural interaction and social exchange in real time over the Internet, *International Journal of Intercultural Relations*, 32(3), pp. 215–228. doi: 10.1016/j.ijintrel.2007.11.003.

Time (2022). *Time: Into the metaverse | August 15, 2022*. July 25, 2022, Time USA LLC.

Unity Technologies (2023a). *Oculus Integration | Integration | Unity Asset Store*. Available at: https://assetstore.unity.com/packages/tools/integration/oculus-integration-82022 (Accessed: 31.1.2023).

Unity Technologies (2023b). *Photon Voice 2 | Audio | Unity Asset Store*. Available at: https://assetstore.unity.com/packages/tools/audio/photon-voice-2-130518 (Accessed: 31.1.2023).

Wang, C. (2022). VIVERSE: Your Open and Secure Metaverse Platform, HTC VIVE, 25 Feb 2022. https://www.youtube.com/watch?v=LNfGKRjNQKw (Accessed: 31.1.2023).

Wang, Y., Su, Z., Zhang, N., Xing, R., Liu, D., Luan, T. H. and Shen, X., (2022). A survey on metaverse: Fundamentals, security, and privacy. *IEEE Communications Surveys & Tutorials*.

Zuckerberg, M. (2021). Mark Zuckerberg Delivers Remarks at 'Facebook Connect' 2021 | LIVE, NotThis News, https://www.youtube.com/watch?v=bRhQ7bXX-z4. (Accessed: 31.1.2023).