

It's NOT "Only a Game"! Strategies Against Sexual Harassment in Immersive VR Environments

Sibylle Kunz, Claudia Hess and Adrienne Steffen

IU International University, Erfurt, Germany

sibylle.kunz@iu.org

claudia.hess@iu.org

adrienne.steffen@iu.org

Abstract: Sexual harassment and cyber grooming have been happening in all digital media since their beginning. Recently, more users are experiencing 3D worlds in Virtual reality (VR), and immersive multiplayer environments allow talking and interacting with other users' avatars or even "touching" each other's virtual body. This can lead to sexual harassment situations. Especially female and queer VR users are frequently "contacted" in inappropriate ways. Since the identity of a user is often not revealed, the barriers for offence up to virtual rapes are relatively low. There is no clear jurisdiction against the offenders if no physical contact has been made, but the victims can severely suffer from this experience. This research describes different forms of virtual abuses and discusses what counter measures software companies have already taken and whether these are effective. It then discusses the necessity of digital house rules and different technical measures against offenders.

Keywords: Virtual Reality, Sexual Harassment, Avatar, Digital House Rules, Virtual Rape, Metaverse

1. Introduction

Virtual Reality (VR) experiences and the Metaverse are designed as three-dimensional graphical spaces striving for a maximum degree of immersion and presence that users can perceive. These spaces lead to all kinds of interaction formats between users of virtual experiences – including negative side effects like cyber-mobbing and sexual harassment ranging from written or oral communication to virtual touching, grabbing, or even performing sexual movements on a user's avatar. What some users might perceive as harmless gestures might lead to traumatizing experiences in others. According to Outlaw & Duckles (2018) almost every second female-identifying user in VR has already experienced virtual sexual harassment. Companies providing virtual rooms for social VR like Meta's Horizon World, VRChat or RecRoom have started reacting to this phenomenon with different technical features.

With regard to the Metaverse as a space that "transcends borders and countries" (Dwivedi et al., 2022), we need to discuss whether guardrails, rules and laws from the real world can be used as a foundation for international standards, so that technical solutions can be based on a common ground. Software companies or platform enterprises could even be obliged to provide security mechanisms to protect their customers.

This paper discusses the phenomenon of digital sexual harassment especially in 3D-environments and identifies classification frameworks. It analyzes technical approaches and how they fit into this classification. Finally, the idea of digital house-rules and technical or legal sanctions in case of their violation are discussed.

2. Sexual Harassment in VR: Examples and Countermeasures

Sexual harassment (SH) has originally been defined as "unwelcomed sexual [advances], or . . . unwelcome [requests] for sexual favours, or . . . other unwelcome conduct of a sexual nature in relation to the other person" (Mason & Chapman, 2003, p. 206), mostly found in workplace environments maintaining social inequality. Barak (2002) looks at sexual harassment as "gender-based harassment" (like telling chauvinistic jokes or posting pornography), "unwanted sexual attention" (like sexual insults), and "sexual coercion" (using physical or psychological force or threat to gain sexual cooperation). Meanwhile the definition has broadened to include the use of digital media. Slaughter and Newmann (2022, p.7) define online harassment as "interpersonal aggression or offensive behavior(s) that is communicated over the internet or through other electronic media" (Slaughter and Newmann, 2022, p.7).

In online environments, SH happens mostly in the form of active (sending explicit content to a user) or passive (storing content where an unsuspecting user encounters it) graphic or verbal harassment. Technology-assisted abuse on the internet has happened for a long time, because anonymity protects the perpetrators from being discovered, so the barriers are low. What's new in VR environments is the graphical and haptical representation

(e.g. if controllers can give force-feedback using vibration or if the user wears data-gloves). 3D experiences are far more immersive and emotional than flat 2D representations on a screen (Blackwell et al., 2019).

Jordan Belamire (20.10.2016) was one of the first to write about her experiences, when an anonymous player tried to grope her avatar after only three minutes in a multiplayer VR game. The programmers of the game reacted instantly (Stanton, 25.10.2016). Although there was a “personal bubble” function, this function only worked if other players tried to block your view but did not if they touched the avatar’s body. They improved the feature, but the problem persisted. Therefore, the programmers turned it into a power gesture that fades out all players. As a power gesture, it can immediately be carried out instead of pausing the game and selecting it from a menu. It fades out all players coming too close and mutes their voices. They made the code open source so it could be used by other developers to become a standard “911 gesture”. A similar experience was reported by Nina Jane Patel, a female beta tester in Meta’s Horizon worlds (Clayton, 29.01.2022). In both cases, discussion arose whether virtual sexual harassment in VR counts as such. One recommendation to avoid harassment situations was removing the VR headset, thus ending the immersive experience. But this does not solve the problem, since the situation would continue as soon as the user re-enters the VR world. Distorting users’ voices would make females more difficult to recognize but would also limit desired user interactions in normal conversation situations in social VR.

Nowadays, “safety bubbles” evoked by a power gesture are built into most social VR spaces like RecRooms, Horizon World or Roblox. Apart from regulating a minimum distance between avatars or blocking other users in one’s own account, some systems also have tried installing invisible moderators present in the environment, watching the situation and coming for help if a user signals a problem (Ryan-Rosley, 2023). But this is time-consuming and cost-intensive, so human moderators might be exchanged for AI-based bots, trading in reduced costs for a higher danger of misunderstandings.

Zheng et al. (2022) discuss six categories of countermeasures. The first, “intimacy proxemic”, describes the forementioned safety bubble as a standard boundary feature between a user and other characters. An “intimacy rank” (used in VR chat) would allow for different levels of trust (trusted user, known user, new user etc.). “Social retreat spaces” where nobody else can enter are a feature in RecRoom. “Image control” allows to hide or show other avatars or change one’s own avatar concerning race, gender etc. VR Chat has built a whole “Trust System”, where the users can adjust microphone settings, graphical effects etc. Some systems also use some kind of “Age Check” to especially protect children.

Nowadays, users can normally choose between three ways of quick responses: immediate safety reactions, turning on a longer-lasting safe mode or safety reports to the platform owners for heavier sanctions. But what can be done to prevent critical situations in advance?

3. A Sanction-based Approach for Virtual “House-rules”

It has become clear that a common set of rules should be installed and accepted when users create their account for a virtual world. Blackwell et al. (2019, 100:23) suggest a responsive regulatory pyramid ranging from persuasion and counselling to license revocation for users who commit (sexual) harassment in digital environments. Kunz and Skerra (2023) outline a set of digital house-rules for virtual E-commerce environments, regulating the behavior of customer avatars visiting 3D online shops. Some of these rules are also applicable in Social VR, e.g. concerning spatial proximity of avatars, free, but polite speech and protection from harassment and fraud. Based on these ideas, we suggest the following technical sanctions, depending on the different types of harassment committed:

Table 1: Technical measures against different forms of (sexual) harassment in virtual environments

Content presentation mode	Content Media Type		
	Textual	Graphical	Acoustical
Passive (content placed for users to see)	- AI-based text recognition, then deleting critical content from display	- marking the aggressor’s avatar, e.g., with a warning sign above his head - labelling content with user’s name	- (impossible, since vocal expression can only be uttered in real-time)

Content presentation mode	Content Media Type		
	Textual	Graphical	Acoustical
Active (carried out on another user's avatar in real-time)	<ul style="list-style-type: none"> - AI-based text recognition (e.g., by identifying stop words) - Blocking text entries (e.g., by blocking input devices) 	<ul style="list-style-type: none"> - marking the aggressor's avatar, e.g., with a warning sign above his head - "paralyzing" the aggressor's avatar for a certain time - "mutilating" the aggressor's avatar, e.g. by removing an arm or leg or changing the texture, color etc. 	<ul style="list-style-type: none"> - AI-based speech recognition - muting the harassing user for the victim - increasing duration of muting with every encounter - muting to all other users

In all cases, a warning message should be sent to the aggressor to inform him about the violation of rules. There should be some kind of scoring system counting every moral failure and displaying that score to the aggressor or even other users. Depending on the score, accounts could be temporarily disabled or deleted. Re-opening an account could be prohibited by using identification procedures for new users. The owners of digital platforms could be forced to establish such guard rails by the Digital Services Act (DSA) that – among other goals- aims to create a secure digital environment that is free of illegal content.

But this opens up another dilemma: recognizing critical content requires a common and cultural-independent understanding of what is classified as offensive or illegal and also a continuous surveillance of all user actions. And there might also be environments (e.g., for adults) where a certain behavior will still be tolerated (or even desired) that might be inappropriate in other applications.

4. Conclusion

Sexual harassment in VR is a problem that cannot be technically avoided. Apart from countermeasures that require the victims to take action like power-gestures or incident reporting, there is a need for ethical and legal rail guards that are valid cross-border between countries and cultures. In some ways, 3D-E-Commerce environments are facing similar problems and therefore could be a driver to formulate digital house-rules that users have to accept when creating their accounts. Further research needs to analyze how legal constructs like the DSA can be used as a basis for sanction systems.

References

- Barak, A. (2005) "Sexual Harassment on the Internet", *Social Science Computer Review*, Vol. 23, Issue 1, pp.77–92.
- Belamire, J. (20.10.2016) "My First Virtual Reality Groping", [online], <https://medium.com/athena-talks/my-first-virtual-reality-sexual-assault-2330410b62ee>.
- Blackwell, L., Ellison, N., Elliott-Deflo, N. and Schwartz, R. (2019) "Harassment in Social Virtual Reality: Challenges for Platform Governance", *Proceedings of the ACM on Human-Computer Interaction*, Vol. 3, No. CSCW, Article 100 (November 2019), pp.1-26. <https://doi.org/10.1145/3359202>.
- Clayton, M. (29.01.2022) "Mother says she was virtually groped by three male characters within seconds of entering Facebook's online world Metaverse", Daily Mail, [online], <https://www.dailymail.co.uk/news/article-10455417/Mother-43-avatar-groped-three-male-characters-online-Metaverse.html>.
- Digital Services Act, REGULATION (EU) 2022/2065 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 19 October 2022 on a Single Market For Digital Services and amending Directive 2000/31/EC, [online], <https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32022R2065>.
- Dwivedi, Y.K., Hughes, L., Baabdullah, A.M., Ribeiro-Navarrete, S., Giannakis, M., Al- Debei, M.M., Wamba and S.F. (2022) "Metaverse beyond the hype: Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy", *International Journal of Information Management* 66, 102542.
- Kunz, S. and Skerra, C. (2023) "Corporate Digital Responsibility im Metaversum: Ein E-Commerce-Szenario", in: Lucas, C. and Schuster, G. (2023), *Innovatives und digitales Marketing in der Praxis*, DOI: [10.1007/978-3-658-38210-0_23](https://doi.org/10.1007/978-3-658-38210-0_23).
- Mason, G. and Chapman, A. (2003) "Defining sexual harassment: A history of the commonwealth legislation and its critiques", *Federal Law Review* 2003, 31, pp. 195–224.
- Noval, S. M. R. (2023) "Metaverse, Avatar and Illusion of Lawless World: Rethinking Boundaries in Virtual World", in: bin Abdullah, M. K. et al. (eds.), *Proceedings of the International Seminar on Border Region (INTSOB 2023)*. *Advances in Social Science, Education and Humanities Research* 823, https://doi.org/10.2991/978-2-38476-208-8_29.
- Outlaw, J. and Duckles, B. (2018) "Virtual Harassment: The Social Experience of 600+ Regular Virtual Reality (VR) Users", *The Extended Mind Blog* 4, [online], <https://www.extendedmind.io/the-extended-mind-blog/2018/04/04/2018-4-4-virtual-harassment-the-social-experience-of-600-regular-virtual-reality-vrusers>.

- Ramirez, E.J., Jennett, S., Tan, J., Campbell, S. and Gupta, R. (2023) "XR Embodiment and the Changing Nature of Sexual Harassment", *Societies* 2023, 13, 36, <https://doi.org/10.3390/soc13020036>.
- Ryan-Mosley, T. (01.05.2023), "Undercover in the metaverse. Human moderators in the metaverse are proving essential to digital safety", [online], <https://www.technologyreview.com/2023/05/01/1072463/undercover-in-the-metaverse/>.
- Slaughter, A. and Newman, E. (2022). New Frontiers: Moving Beyond Cyberbullying to Define Online Harassment. *Journal of Online Trust and Safety*, 1(2), <https://doi.org/10.54501/jots.v1i2.5>.
- Stanton, A. (25.10.2016), "Dealing With Harassment in VR", [online], <https://www.uploadvr.com/dealing-with-harassment-in-vr>.
- Zheng, Q., Do, T.N., Wang, L. and Huang, Y. (2022), "Facing the Illusion and Reality of Safety in Social VR", *CHI Conference on Human Factors in Computing Systems (CHI '22)*, April 29-May 5, 2022, New Orleans, LA, USA. ACM, New York, NY, USA.