All Too Real: A Typology of User Vulnerabilities in Extended Reality*

User Vulnerabilities in Extended Reality

James J. Cummings†
Division of Emerging Media Studies
Boston University
Boston, MA, USA
cummingsj@bu.edu

Alexis Shore
Division of Emerging Media Studies
Boston University
Boston, MA, USA
ashore@bu.edu

ABSTRACT

The metaverse promises to blur digital and physical boundaries of communication, presenting novel contexts of previously contended risks. We present a typology of individual and relational vulnerabilities in networked XR, proposing examples of threats to users' agency, safety, and privacy.

CCS CONCEPTS

• Human-centered computing→ Human computer interaction (HCI)

KEYWORDS

Extended reality, Privacy, Agency, Safety, Vulnerability

ACM Reference format:

James J. Cummings and Alexis Shore. 2022. All Too Real: A Typology of User Vulnerabilities in Extended Reality. In *Proceedings of the 1st Workshop on Novel Challenges of Safety, Security and Privacy in Extended Reality*, April 29-May 5, 2022, New Orleans, LA, USA, 4 pages.

1 Introduction

Recent accounts of an impending metaverse anticipate an adoption and usage of extended reality (XR) that is ubiquitous, networked, and regularly engaged with by the general public, blurring the boundaries between offline and online experience [1, 2]. This vision foresees the application of and reliance on XR technologies across different social domains, whether formal workplace exchanges, professional and commercial services, or casual hangouts. Even non-social communications or solitary experiences — such as news and entertainment — are expected to take on a more immersive character, likely with direct consequences for how users consume, process, evaluate, and share content. As such, this vision of the metaverse, broadly defined, is one in which XR technologies will merge the best affordances and worst harms of face-to-face (F2F) communication, computer-mediated communication (CMC), and human-computer interaction (HCI) into a singular digital user experience.

Notably, enthusiasts often emphasize the key special affordances of XR technologies that will permit such a vision to come to fruition. In particular, they highlight the uniquely high levels of spatial and social presence conferred by such technologies [3, 4], which can lead users to feel physically located within digital environments or co-located with wholly virtual objects and the digital representations of other users from anywhere in the world. However, such immersive experiences, whether social or solitary, may also present certain threats to users' wellbeing. What such bullish accounts often fail to note is that — to the extent that it serves to mediate various forms of F2F communication or substitute direct experience with immersive simulation — the initial success and long-term prosperity of the metaverse will not hinge solely upon the degree of presence inherently afforded. Equally requisite, if not more so, will be the ability to elicit users' trust — namely, trust in their ability to freely and safely share information and trust in the fidelity and credibility of the immersive content they experience.

Whether socializing with other users, accessing and consuming certain content, or even simply operating the associated XR hardware itself, engaging with the metaverse will require users to share information about themselves of variable breadth and depth of sensitivity. In both offline and online environments, users implicitly engage in a privacy calculus to compare the perceived benefits (e.g., self-expression, social rewards) and risks (e.g., privacy violations) of sharing information [5]. As a venue for mediated-yet-incredibly-vivid interpersonal exchanges, the metaverse will require that users consider this trade-off under entirely new circumstances: the digitalization of rich, multimodal behaviors akin to those found offline will include system tracking of a variety of user inputs — verbal, gestural, semantic, biometric — and rendering them on screen to other users. Users will

need to trust that the information they expressly disclose and implicitly provide will be safely shared — whether computationally or conversationally — in a manner that does not place privacy at risk.

Beyond confidence in the security of the information they disclose, users will also need to trust the information they consume. In XR, social exchanges often consist of virtual representations that may or may not genuinely reflect users' real-world selves, simultaneously affording creative self-expression as well as deception. Even in non-social contexts, the immersive character of the metaverse will pose risks for users. Immersion within a mediated message is thought to augment media effects through a heightened sense of presence, such that the messages encountered are all the more impactful. While there has been a good deal of popular press and empirical work into the prosocial outcomes of XR [6-8] there has been virtually no attention given to the capacity of XR technology to amplify less desirable effects. In an age of fake news and misinformation, the plausibility [9] and sense of "being there" [10-11] conferred to messages conveyed through XR may enhance implicit trust in depicted content, regardless of authenticity, credibility, or accuracy.

Thus, popular accounts of an impending metaverse instantiated through XR technologies entail a platform in which users may run the risk of several different threats. In an effort to inform and guide future technology design and policy research on the metaverse, this project seeks to provide a framework for typologizing the variety of vulnerabilities users may face. In doing so, it centers the notion of *vulnerability*, defining the metaverse not solely as a technology, but an institution capable of fostering particular types of vulnerable situations for users [12]. Similar to privacy [13], vulnerability has been studied at *individual*, *relational,* and *institutional* levels [14]. With respect to the metaverse, the latter is primarily a matter of system infrastructure (e.g., cybersecurity); in contrast, this piece focuses on individual and relational vulnerabilities — that is, the user experience affordances of a metaverse implemented through XR technologies and the associated threats to those users. In particular, individual and relational vulnerabilities — which can be exacerbated as a result of technological immersion — are likely to pose a range of threats to users' *agency*, *privacy* and *safety*. We briefly present examples of each of these threats, in an effort to spark conversations around the looming challenges to be faced in the metaverse as well as help identify and organize avenues for future research.

2 Individual Vulnerabilities

While many proposed use cases focus on social interaction and exploration, the metaverse will inherently subject users to individual vulnerability. This vulnerability may arise as a product of user status (e.g., children need special protections) or a lack of knowledge that renders users defenseless against institutional power [15].

2.1 Individual Agency

As part of an impending metaverse, XR technologies will increasingly permit users to experience content such as news, advertisements, and social media posts from within a message itself. To the extent that "seeing is believing," what might be the ramifications of spatially occupying a message? Similar to native advertisements, being perceptually immersed within a message may implicitly impair users' ability to discern authenticity, credibility, and authorial intent. This presents new levels of concern with respect to harmful messaging, such as disinformation or predatory content. Further, it has been suggested that by readily precluding juxtaposition of competing worldviews (literally), the immersive character of XR spaces may have "an unlegislated power to shape our politics" [16]. In these respects, users' capacity to critically evaluate the ideas they encounter in XR may be restricted.

In addition, individual agency is likely to be threatened as a result of the metaverse's inherently surveillant nature. Studies have concluded that surveillance chills behavior online and offline, stymieing individual agency in that users hide their authentic personality and behavior under the assumption that they are being watched, whether it be by institutional actors or other users [17]. Additionally, the technologies and techniques of surveillance capitalism, recently termed "surveillance technicity," minimize negative affective states so that users continue engaging, driving them further away from self-determination [18].

2.2 Individual Privacy

While privacy acts as a "shield" in the way of discovery or rendering vulnerability, it also can hide vulnerability [15]. For instance, privacy protections may inhibit a user from reporting an inappropriate encounter in the metaverse. A lack of privacy also creates vulnerability, as certain information required to enjoy services within the metaverse may also make users more vulnerable to exploitation and manipulation. In the metaverse, it is imagined that a user can engage with their body and voice in real-time. The metaverse extends the panoptic nature of traditional online mediums from measuring behavior through clicks to that closely resembling offline behavior, including non-verbal data which increases targeting and monetization potential [19]. Unlike other online mediums, XR technology may reduce anonymity and invisibility through the inclusion of additional information channels, such as voice or avatar, as default elements of the interface [20]. While many do not care about their online privacy, claiming that they have "nothing to hide," [21] this argument may be substantially weakened when a wider assortment of offline behaviors becomes digitized.

2.3 Individual Safety

The metaverse extends a looming threat to individual safety which already exists in online spaces. For example,

given that XR has already been leveraged for educational and pediatric purposes [19], designers can anticipate its use by children, thus requiring implementation of special protections that satisfy the requirements of relevant laws such as the Children's Online Privacy Protection Act (COPPA). While COPPA requires verifiable parental consent for use of online platforms, the FTC recognizes that it is nearly impossible to account for children lying about their age or forging consent mechanisms [22]. This becomes increasingly problematic when children are immersed in an environment that resembles the offline world, putting them directly at risk of safety harms, including being exposed to explicit content or coming into contact with dangerous actors.

An additional safety concern within the metaverse is that produced by dark patterns, or design intended to manipulate users toward a particular decision which maximizes shareholder value [23]. While dark patterns have been widely studied with respect to online environments at large, less is known about the operation of dark patterns in XR in particular.

3 Relational Vulnerabilities

The metaverse intends to be not only a space governing asymmetrical information exchange between individuals and corporate actors, but also one where interpersonal communication can thrive. The metaverse is expected to foster new levels of social presence in mediated exchanges — with humans and AI agents alike — as all manner of everyday real-world interactions, such as in-store shopping [24] or team meetings [25], are ported to XR. While social XR interactions are currently understudied [26], there are several plausible relational vulnerabilities posed by such settings which might guide considerations for research and design.

3.1 Relational Agency

The relational threat to agency in the metaverse may be caused by a lack of trust or knowledge about information flows. While anthropomorphic cues which engender social presence have been found to heighten trust [24, 27] and information control [28], they also produce feelings of surveillance in creating a direct gaze on participants [27]. With respect to information flows, when users perceive communication to be ephemeral, they are more motivated to disclose personal information [29]. However, it is unclear what perceptions of ephemerality will exist in the metaverse with prospective digital information exchanges that are meant to mirror F2F offline interactions.

3.2 Relational Privacy

Compared to current online communications, XR exchanges will present new contexts for sensitive self-disclosure and lateral surveillance, diminishing relational privacy and enabling further interpersonal and institutional context collapse. This is particularly dangerous within an immersive digital environment where users may feel as though they can behave and interact with others as they do in offline settings. Further, it remains unknown how users will be able to delineate human users from artificially-created bots, which is particularly important given that the latter can encourage the same degree of intimate self-disclosure as the former [30]. Thus, metaverse experiences pose a particular threat to privacy management, as the wider assortment of user data conveyed — explicitly or implicitly through embodied virtual interactions — may be shared beyond intended privacy boundaries.

3.3 Relational Safety

The anonymity and invisibility provided in online exchanges can lead to disinhibition, both benign and toxic [31]. In the metaverse, it is predicted that users will be similarly disinhibited, however, receivers and bystanders may experience the consequences of this disinhibition to a more realistic degree. For example, negative anti-social behaviors such as cyberbullying and harassment will continue to present psychological harm to users. It is possible that negative online disinhibition common to traditional online spaces will be mitigated by the richness of embodied XR representations and the resulting impressions of others as real, fully-formed persons; however, recent accounts of harassment and assault in XR settings suggests this may not be the case [32]. Thus, it is essential for designers of the metaverse to prepare for the prevention of toxic disinhibition in order to protect user safety in social settings.

4 Conclusion

Given recent accounts, including vision statements from firms invested in XR technologies, the metaverse is a near-term prospect. Amid the excitement and optimism, designers and users should duly attend to clear threats to both individual and relational vulnerabilities. The aforementioned typology is a starting point which maintains that these vulnerabilities can be conceptualized as threats to agency, privacy and safety. Notably, the current list of vulnerabilities is not exhaustive; rather, it serves as a launching point for continued discussion. We look forward to sharing this initial typology with the CHI community, drawing upon their insights and feedback as we refine this list as a tool for guiding future academic research, design, and platform policy considerations.

REFERENCES

[1] Jacquelyn Melinek. 2021. Microsoft CEO: The Metaverse Will Bring Real World into Any Digital Space. (November 2021). Retrieved Feburary 20, 2022 from https://blockworks.co/microsoft-ceo-the-metaverse-will-bring-real-world-into-any-digital-space/
[2] Tech@Facebook,. 2021. Connect 2021: Out vision for the metaverse. (October 2021). Retreived Feburary 20, 2022 from https://tech.fb.com/connect-2021-our-vision-for-the-metaverse/

[3] James J. Cummings and Jeremy N. Bailenson. 2015. How immersive is enough? A meta-analysis of the effect of immersive technology on user presence. Media Psychology 19,2 (May 2015), 272-309. https://doi.org/10.1080/15213269.2015.1015740
[4] Catherine S. Oh, Jeremy N. Bailenson, and Greg F. Welch. 2018. A systematic review of social presence: Definition, antecedents, and implications. Frontiers in Robotics and AI, 114 (October 2018), https://doi.org/10.3389/frobt.2018.00114
[5] Yongquiang Sun, Nan Wang, and Xiao-Liang Shen. 2021. Calculus interdependency, personality contingency, and causal asymmetry: Toward a configurational privacy calculus model of information disclosure. Information & Management, 58, 8 (December 2021), 103556. https://doi.org/10.1016/j.im.2021.103556
[6] James J. Cummings, Mina Tsay-Vogel, Tiernan J. Cahill and Li Zhang. 2021. Effects of immersive storytelling on affective, cognitive, and associative empathy: The mediating role of presence. New Media & Society (February 2021) 1-24, https://doi.org/10.1177/1461444820986816
[7] Fernanda Herrera and Jeremy N. Bailenson. 2021. Virtual reality perspective-taking at scale: Effect of avatar representation, choice, and head movement on prosocial behaviors. New Media & Society, 23, 8 (August 2021), 2189–2209. https://doi.org/10.1177/1461444821993121
[8] Beatrice S. Hasler, Daniel H. Landau, Yossi Hasson, Noa Schori-Eyal, Jonathan Giron, Jonathan Levy, Eran Halperin and Doron Friedman. 2021. Virtual reality-based conflict resolution: The impact of immersive 360° video on changing view points and moral judgment in the context of violent intergroup conflict. *New* Media & Society, 23, 8 (August 2021), 2255–2278. https://doi.org/10.1177/1461444821993133
[9] Mel Slater. 2009. Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. Philosophical Transactions of the Royal Society B: Biological Sciences, 364, 1535 (December 2009), 3549-3557. https://doi.org/10.1098/rstb.2009.0138
[10] Shyam S. Sundar, Jin Kang, and Danielle Oprean. 2017. Being there in the midst of the story: How immersive journalism affects our perceptions and cognitions. Cyberpsychology, Behavior, and Social Networking, 20, 11 (November 2017), 672-682. https://doi.org/10.1089/cyber.2017.0271
[11] Priska Breves. 2021. Biased by being there: The persuasive impact of spatial presence on cognitive processing. Computers in Human Behavior, 119 (June 2021) 106723. https://doi.org/10.1016/j.chb.2021.106723
[12] Florencia Luna. 2009. Elucidating the concept of vulnerability: Layers not labels. IJFAB: International Journal of Feminist Approaches to Bioethics, 2, 1 (Spring 2009), 121-139. https://doi.org/10.3138/ijfab.2.1.121
[13] Natalya N. Bazarova and Philipp Masur. 2020. Towards an integration of individualistic, networked, and institutional approaches to online disclosure and privacy in a networked ecology. Current Opinion in Psychology, 36 (December 2020), 118-123. https://doi.org/10.1016/j.copsyc.2020.05.004
[14] Frank Rudy Cooper. 2014. Always already suspect: Revising vulnerability theory. NCL Rev., 93, 1339.
[15] Ryan Calo. 2016. Privacy, vulnerability, and affordance. DePaul L. Rev., 66, 591.
[16] Fred Turner. 2015. The politics of virtual reality. The American Prospect, 26, 3, 25-29.
[17] Jon Penney. 2021. Understanding Chilling Effects. 106 Minnesota Law Review 101 (2021, Forthcoming). Available at SSRN: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3855619
[18] Ben Egliston. 2020. Surveillance technicity: affect, retention and videogame analytics. Media, Culture & Society, 42, 6 (January, 2020), 915-931.
[19] Jeremy Bailenson. 2018. Protecting nonverbal data tracked in virtual reality. JAMA pediatrics, 172, 10 (January, 2020) 905-906. https://doi.org/10.1177/0163443719880139
[20] Divine Maloney, Guo Freeman, and Donghee Yvette Wohn. 2020. "Talking without a Voice" Understanding Non-verbal Communication in Social Virtual Reality. Proceedings of the ACM on Human-Computer Interaction, 4 (CSCW2), 1-25. https://doi.org/10.1145/3415246
[21] Daniel J. Solove. 2007. I've got nothing to hide and other misunderstandings of privacy. San Diego L. Rev., 44, 745.
[22] Federal Trade Commission. Complying with COPPA: Frequently Asked Questions. Retrieved February 18, 2022 from https://www.ftc.gov/tips-advice/business-center/guidance/complying-coppa-frequently-asked-questions-0
[23] Colin M. Gray, Yubo Kou, Bryan Battles, Joseph Hoggatt and Austin L. Toombs. 2018. The dark (patterns) side of UX design." In Proceedings of the 2018 CHI conference on human factors in computing systems, Montreal, QC, Canada, 1-14. https://doi.org/10.1145/3173574.3174108
[24] Yinshu Zhao, Nilufar Baghaei, Alexander Schnack and Lehan Stemmet. 2021. Assessing Telepresence, Social Presence and Stress Response in a Virtual Reality Store. In 2021 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct). IEEE, Bari, Italy, 52-56 doi: 10.1109/ISMAR-Adjunct54149.2021.00020
[25] Mark Manuel, Poorvesh Dongre, Abulaziz Alhamadani and Denis Gracanin. 2021. Supporting Embodied and Remote Collaboration in Shared Virtual Environments. In International Conference on Human-Computer Interaction, 639-652. https://doi.org/10.1007/978-3-030-77599-5_44
[26] Amal Yassien, Passant ElAgroudy, Elhassan Makled, and Slim Abdennadher. 2020. A design space for social presence in VR. In Proceedings of the 11th Nordic Conference on Human-Computer Interaction: Shaping Experiences, Shaping Society. Tallin, Estonia, 1-12. https://doi.org/10.1145/3419249.3420112
[27] Nuria Rodriguez-Priego, Rene van Bavel, R. and Shara Monteleone. 2021. Nudging online privacy behaviour with anthropomorphic cues. Journal of Behavioral Economics for Policy, 5, 1, 45-52.
[28] Wenxi Pu, Siyuan Li, Gregory J. Bott, Marie Esposito and Jason Bennett Thatcher. 2022. To Disclose or Not to Disclose: An Evaluation of the Effects of Information Control and Social Network Transparency. Computers & Security, 112 (January 2022), 102509. https://doi.org/10.1016/j.cose.2021.102509
[29] Xiaofen Ma, Yuren Qin, Zhuo Chen, and Hichang Cho. 2021. Perceived ephemerality, privacy calculus, and the privacy settings of an ephemeral social media site. Computers in Human Behavior, 124 (November 2021) 106928. https://doi.org/10.1016/j.chb.2021.106928
[30] Yi-Chieh Lee, Naomi Yamashita, Yun Huang and Wai Fu. 2020. "I Hear You, I Feel You": Encouraging Deep Self-disclosure through a Chatbot. In Proceedings of the 2020 CHI conference on human factors in computing systems. (CHI '20). 1-12. Honolulu, HI. https://doi.org/10.1145/3313831.3376175
[31] John R. Suler. 1999. To get what you need: healthy and pathological Internet use. CyberPsychology & Behavior, 2, 5 (January, 2009). 385-393. https://doi.org/10.1089/cpb.1999.2.385
[32] Sheera Frankel and Kellen Browning. 2021. The Metaverse's Dark Side: Here Comes Harassment and Assaults. (December 2021). Retrieved February 18, 2022 from https://www.nytimes.com/2021/12/30/technology/metaverse-harassment-assaults.html