Virtual Realities, Real Harms: Rethinking Violence in the Metaverse

Rungrot Tatiyawongwiwat

School of Environment, Resources, and Development
Asian Institute of Technology

Abstract: In a world where interconnected crises, ranging from climate change and digital warfare to forced migration, reveal mounting human vulnerability, the concept of *polycrisis* captures the overlapping nature of today's global risks. Within this context, the Metaverse has emerged as a new digital frontier offering refuge, simulation, and escapism. However, this environment also enables symbolically extreme yet normalized forms of violence, such as virtual harassment and aggressive communication, shaped by culturally embedded systems of value. As immersive tools like 3D interaction, Augmented Reality (AR), or Virtual Reality (VR) devices, and gadgets render the Metaverse increasingly tangible, they introduce new digital infrastructures, power relations, and social participation. This paper interrogates which dominant worldviews and hierarchies are embedded into virtual spaces, particularly concerning structural and historical violence. Building on Judith Butler's concept of *grievability*, the analysis identifies three dimensions of virtual violence: (1) unequal access to digital infrastructure and immersive technology; (2) cultural systems that shape meaning and representation; and (3) the influence of platform designers on digital norms and behavior. Case studies from platforms such as *VRChat*, *The Sandbox*, and *Meta Horizon Worlds* reveal how their design and algorithmic systems perpetuate discrimination, especially against marginalized users. The paper also explores how digital exclusion and algorithmic gatekeeping reflect *intellectual* and *institutional* violence, particularly through the commodification of data and the reproduction of biases. These phenomena disproportionately affect communities in the Global South, reinforcing digital inequality. This paper argues that virtual violence reproduces real-world power structures within emerging digital environments. It concludes by calling for co-developed digital justice frameworks between governments and developers to prevent the reproduction of gendered, cultural, economic, and racial violence. Ultimately, it situated virtual violence as a central issue within broader conversations on digital governance and the unfolding polycrisis.

Keywords: Metaverse, Digital Violence, Gender-Based Violence, Ethics

As the world enters a period of turbulence in which multiple crises are deeply intertwined, from climate change and displacement to conflict, the concept of *polycrisis* captures how one issue fuels another. At the same time, digital innovations have created alternate realities, expanding human action beyond physical limits. The Metaverse, introduced in Neal Stephenson's *Snow Crash* (1992),

exemplifies this shift, imagined as a realm of creative, social, and economic opportunities, not limited to 3D environments but also including digital spaces where users express themselves through avatars, virtual representations of self intentionally selected, shaped by media, norms, and cultural hierarchies, carrying deep socio-political implications (Boellstorff, 2015). Massive investment drives Metaverse growth, yet violence persists, beyond physical to structural, embedded in digital systems designed by developers shaped by their real-world social contexts. Ruha Benjamin defines this as the "New Jim Code," a direct reference to the Jim Crow laws that enforced racial segregation in the United States, but applied to the digital age where neutrality masks discrimination (2020). As an immersive, often unregulated space, the Metaverse reveals digital violence: racialized avatars (Groom et al., 2009), gender harassment (Peck et al., 2013), data commodification (Zuboff, 2017), and algorithmic bias (Eubanks, 2018), not as mere technical flaws, but as sociopolitical issues rooted in physical and virtual realities. This paper analyzes harassment, discrimination, exclusion, and cultural erasure through theories of power and Judith Butler's grievability (2016), with case studies from *VRChat, Meta Horizon Worlds,* and *The Sandbox*. It explores how social hierarchies persist and examines ethical challenges and opportunities for justice in virtual innovation. In this era of transformation, the Metaverse is not just technological progress; it reproduces structures of power where violence is embedded. Addressing it demands a deeper understanding of virtual environments as battlegrounds shaped by power, vulnerability, and resistance.

Theoretical Lens

Understanding violence in digital spaces requires engaging with how power is constructed. Violence extends beyond physical harm to include control, coercion, and the restriction of thought and action. Michel Foucault stated that power is embedded in knowledge, language, and norms we accept as truth (1995). In regard to virtual spaces, Judith Butler introduces the concept of grievability in her book *Frames of War* (2016); this term can be used to understand how those in power exercise violence to govern, decide, and control. Grievability highlights how power determines whose lives are considered worth protecting and mourning. Some are erased or ignored, while others receive recognition, reflected in avatar design and user experiences in the Metaverse. Marginalized groups experience discrimination, neglect, or delayed responses, particularly women, ethnic minorities, and gender-diverse communities, further reinforcing digital hierarchies. Foucault's concept of knowledge as power applies to the Metaverse, where power and knowledge operate through user experience, platform design, and algorithmic visibility. These structures determine what users see and believe, shaping discourse and action, especially against vulnerable groups. Shoshana Zuboff extends Foucault's ideas by incorporating a concept of *surveillance capitalism*, which explains how corporations commodify human data, violating consent and privacy (2019). In the Metaverse, identity becomes a product, revealing new forms of economic violence and digital class distinctions. Critical race theory views race as embedded in digital realm systems. The Metaverse replicates racial bias via default system settings and white avatars, biases embedded even at the coding level, resulting in algorithmic exclusion. Ruha Benjamin notes that technology is never neutral; the design and usage of digital environments reinforce social, political, and economic hierarchies. Inequitable access to immersive tools further deepens digital inequality (2020). Postcolonial theory, with Gayatri Chakravorty Spivak (1988) as its leading figure, critiques how English-dominated digital design constitutes a form of violence by erasing local values and identities, despite much content originating in the Global South. Binary gender norms persist in avatar design; male avatars appear muscular, female avatars soft and pink, excluding queer and non-binary users. Avatars coded as feminine are often objectified. Judith Butler argues that denying recognition is itself a form of violence (2015).

Virtual violence is not incidental but systematically constructed, embedded in design, code, infrastructure, and visual systems. Through the insights from critical race, postcolonial, and gender theories, the Metaverse reveals patterns that reflect global class, racial, and gender hierarchies, perpetuating patriarchal norms in both the virtual and real world.

Forms of Violence in the Metaverse

While the Metaverse represents a new digital frontier for human experience, it is also a space of apparent freedom that is saturated with diverse forms of violence. These include symbolic violence, structural social violence, sexual violence, and intellectual oppression, all of which reflect deeper systems of inequality and domination embedded within the digital infrastructure.

1. Symbolic Violence: Racialized and Gendered Avatars, Body Norms

Pierre Bourdieu's concept of *symbolic violence* (1991) is key to Internet and critical race studies, explaining how representations of race, beauty standards, and gender norms shape digital violence, especially in the techno-utopian Metaverse. Originating in the physical world, traditional and digital media reinforce hegemonic standards of attractiveness and worth, normalizing hierarchical systems that privilege dominant groups (Bourdieu, 2002; Nakamura, 2013). A study by Cyan DeVeaux et al. (2025) on VRChat, an online virtual reality social platform, shows racial bias in through digital avatar designs. Users of color often report exclusion due to default system settings that fail to reflect their identities. Resorting to Do It Yourself (DIY) avatar creation, they face unstable systems that harm their well-being, inclusion, and immersion. Whiteness remains the default, and people of color must work harder to represent themselves authentically. Without physical force, symbolic violence becomes the dominant mode of harm in virtual spaces. Lisa Nakamura's *Cybertypes* (2013) describes racialized digital stereotypes that counter early internet ideals of liberation. Alongside Sarah Banet-Weiser (2018), Nakamura shows how digital aesthetics reproduce offline biases in design, advertising, and user experience based on race, gender, and age. This results in sexualized representations of women and hyper-masculinized males coded through tone, body size, and expression, limiting diversity and reinforcing class divides in a supposedly liberating space. The Metaverse thus becomes another site for constructing "us vs. them" dynamics. Through a gender lens, symbolic violence includes the rise of both feminism and misogyny in digital culture. Although feminism challenges bias, it is often co-opted by commercial interests, weakening its critique and undermining genuine efforts toward equity. Sarah Banet-Weiser warns that commercial feminism allows patriarchy to endure (2018). This is evident in practices like slut shaming and cyberbullying of outspoken women, where assertiveness is framed as immorality. Feminism must remain political to dismantle gendered power structures. Without this critical stance, symbolic violence will persist in digital spaces like the Metaverse.

2. Intellectual Oppression and Violence: Limited involvement of Non-Western Languages and Identities

While earlier discussions focused on visible harms like symbolic violence, intellectual oppression demands a deeper analysis of digital systems. This occurs when the knowledge of certain communities is rendered silent or insignificant through structural power arrangements. In digital spaces, such erasure is embedded in systemic default settings. Academic critiques highlight how narrative production is predominantly controlled by Western creators, embedding intellectual violence into digital environments. While some see cultural context as natural in design, Gayatri

Chakravorty Spivak (1988) and Ruha Benjamin (2020) identify intellectual violence in English defaults, Western worldviews, and Euro-American storylines that shape digital narratives and perspectives. Moderation systems often prioritize Western norms while neglecting linguistic, cultural, and spiritual diversity, reinforcing intellectual marginalization and weakening global pluralism. Spivak's subaltern theory defines intellectual violence as exclusion from dominant power structures. In virtual worlds, minority identities are rendered inaudible and absent as leaders, replaced by tall, white, masculine avatars rooted in European history, filtered out by dominant systems. This is not exclusive to the West; state power also erases voices. Indigenous mountain communities are misrepresented as ecological threats under assumptions of ignorance, ignoring their historical role as environmental stewards. Intellectual violence intersects with racial discrimination, expressed through exclusion from power structures due to differences in race, culture, and belief. Ruha Benjamin shows how data, algorithms, and automation embed inequality under claims of neutrality (2020). Her "New Jim Code" reveals racism encoded in facial recognition, credit scores, and predictive systems; discrimination masked as objective efficiency. These systems reinforce bias through hidden structures. Benjamin calls for abolitionist, community-centered technology, ethical design, and digital social movements to resist systemic inequality. As the Metaverse evolves, it must confront these embedded hierarchies. Without intentional intervention, colonial knowledge systems will continue shaping digital architecture in invisible yet pervasive ways.

3. Economic Violence: Access Barriers and Data Colonization

Economic conditions drive inequality and violence, especially in capitalist systems where power and capital control access. Virtual worlds, constructed atop real-world socio-economic foundations, replicate structural violence through scarcity and unequal distribution of resources. This includes inequitable access to virtual resources and the commodification of user data. Zuboff calls this surveillance capitalism, where companies extract data without consent, not only to predict but also to shape behavior (2019). A key example is digital land speculation in *the Sandbox*. Corporations like Adidas, PwC, and Gucci invest in high-traffic zones, where value is tied to location, social status, and income potential, benefiting capital-rich early adopters and reinforcing wealth disparities. These markets favor capital-rich actors, often from the Global North, excluding many from the Global South. Though platforms promote decentralization, control remains with developer-backed investors who originally secured key land, reproducing hierarchy and power concentration. Axie Infinity, a blockchain-based online game as part of the broader Metaverse ecosystem, once provided income to low-income users, especially in the Philippines, but market crashes left many in debt after entering with limited knowledge and high risk, driven primarily by hopes for economic mobility. Economic violence in virtual spaces mirrors real-world inequalities through unequal access to land, tools, and internet infrastructure. In the Global South, poor access to reliable, high-speed internet limits knowledge and participation. User data is extracted and monetized without true consent, disproportionately affecting marginalized users who often remain unaware of the commodification. This phenomenon, known as data colonization, masks exploitation under the guise of free digital services. Nick Couldry and Ulises A. Mejias call this data colonialism: the systemic appropriation of life through data extraction for capitalist gain, where data is the new gold to be seized and commodified (2019). Profit motives and speculative land grabs reproduce corporate violence. Despite the rhetoric of decentralization, Metaverse platforms remain dominated by capital-backed developers and early investors, reinforcing exploitative structures under the illusion of freedom and equal access.

4. *Sexual and Psychological Violence: Harassment in VR Spaces*

The Metaverse enables users to present alternate selves through avatars. Embodiment in VR allows users to feel and perceive actions as if they are happening in a physical world. This proximity to reality intensifies emotional and psychological engagement. Although VR doesn't replicate physical reality, its psychological impact is profound. In late 2021, SumOfUs (current Ekō) reported a sexual harassment incident in *Horizon Worlds*, where a researcher's avatar was subjected to unwanted contact. At the time, there were no protective measures. This led to the 2022 launch of a personal boundary feature, a design mechanism that creates virtual space between users to prevent unwanted proximity (Ekō, 2022). Reports show that users who identify as female or gender-diverse disproportionately experience virtual violence, including assault, sexualized comments, and non-consensual touching driven by online misogyny and what Emma Jane terms *e-bile*, or digitally-expressed sexual hostility (2017). This contradicts the Metaverse's inclusive vision. Jane calls upon platform regulators to prioritize safety and urges society to recognize the structural nature of virtual violence. Supporting this, Jesse Fox and Wai Yen Tang found that repeated exposure to online harassment normalizes harmful beliefs, including rape myths and victim-blaming, especially among youth in unsafe digital spaces (2014). These patterns expose the unpreparedness of immersive platforms to regulate harm. Existing frameworks lack preventive systems and legal protections tailored to VR, placing responsibility on users instead of developers and overlooking the psychological depth of virtual experiences. Safety and belonging must be built through ethical design. Without it, the Metaverse risks becoming a digital extension of real-world sexual violence, replicating harmful dynamics rather than transforming them. In VR, psychological harm can be as real and damaging as physical harassment.

Violence, Real-World Inequality, and the Reordering of Society

Violence in virtual spaces is not spontaneous: it is imported by creators and users, reproducing real-world inequality. The illusion of freedom hides boundaries set by capitalist systems that commodify personal data without consent (Zuboff, 2019). Structural violence, rooted in inequality, spans both physical and digital realms (Galtung, 1969). Symbolic violence and intellectual oppression arise from developers and users reinforcing systemic norms in avatars, narratives, and governance. English interfaces and Eurocentric design dominate, marginalizing non-Western users and deepening Global North–South divides (Benjamin, 2020; Nakamura, 2013). Sexual harassment of female and gender-diverse avatars normalizes violence in VR, with real-world mental and physical consequences (Jane, 2017). Economic violence also appears; high hardware costs and weak internet exclude low-income users. This reflects *digital colonialism*, where Global South users are exploited for data by Global North entities (Couldry & Mejias, 2019). Though still developing, the Metaverse, originating from Neal Stephenson's *Snow Crash* (1992), shows clear immersive potential. Despite claims of decentralization, economic capital continues to determine influence (van Dijck, 2013). Privatized governance dominates, with surveillance capitalism manipulating user behavior via data extraction under coercive consent (Zuboff, 2019). Governments also regulate virtual spaces, turning them into geopolitical arenas. Sovereign Metaverses hint at rising authoritarianism and eroding digital rights (Morozov, 2011). To evolve as a public space, the Metaverse must address ethical design and accountability. Current platforms rely on reactive safety, placing responsibility on users, rarely enforcing preemptive protections (Gillespie, 2018). This failure harms behavior, mental health, and thought patterns. Corporate-state control threatens digital civil liberties. Users become passive, losing voice and agency, weakening democratic engagement.

Virtual violence will persist, evolving through layers of reproduced injustice. Studying this is vital; it reveals how normalized harm becomes when users are promised freedom but constrained by unequal systems. The Metaverse mirrors real-world norms. Rather than an escape, it becomes a space to question and challenge dominant narratives, showing how even appearance shapes collective well-being in often overlooked ways.

The Metaverse as a Space for Resisting Dominant Discourses and Real-World Hierarchies

While often seen as a space of digital violence, the Metaverse also offers potential for resistance. Redesigning platforms to let users embody perspectives beyond their own (e.g., men using female avatars) can subvert stereotypes and biases rooted in real-world social structures. Scholars and digital communities are exploring ways to resist harm and build ethical digital spaces. This resistance is not merely about safety; it lays the groundwork for digital justice rooted in equity, access, and dignity. The Metaverse can also bridge online and offline activism. Groups like Feminist Internet and Access Now lead this work. Launched in Malaysia in 2014, Take Back the Tech campaigns against violence toward women and promotes digital safety across countries like the Philippines, Bangladesh, Indonesia, and Cambodia (Association for Progressive Communications, 2021). It empowers women to document harm and build safer tech ecosystems inclusive of LGBTQ+ communities. Feminist Internet challenges patriarchal norms in algorithms, governance, and interfaces (Internet, 2020). These movements empower women and LGBTQ+ communities to share testimonies and build safer online environments. Developers have introduced safety features. Meta's Horizon Worlds, for example, added a four-step personal boundary function. However, privacy tools and protective policies alone remain insufficient for creating a truly just Metaverse. Jean-Marc Seigneur and Mohamed-Amine Choukou call for accountability structures and inclusive governance, letting users, especially marginalized groups, actively shape the systems they inhabit (2022). Momentum for inclusion is growing. Ethical design is key, not just a technical fix, but a tool that opens pathways for social, political, and economic participation. Research highlights five dimensions of resistance in the Metaverse:

1. Representation and Cultural Diversity: Designing avatars and environments that reflect diverse cultures, ethnicities, gender identities, and linguistic expressions (Benjamin, 2020).
2. Genuine Ownership and Consent: Developing tools that grant users true ownership of their virtual presence, informed data consent, and participatory control over how their information is used (Zuboff, 2019).
3. Digital Social Accountability: Enforcing harm reporting and moderation with transparent oversight, including third-party mechanisms.
4. Equity and Access: Supporting infrastructure and education to bridge digital divides, especially in the Global South (Couldry & Mejias, 2019).
5. Participatory System Design: Involving survivors of violence and marginalized communities in platform governance to ensure inclusive system design (Costanza-Chock, 2020)

Conclusion

It is no longer feasible to dismiss the Metaverse as fantasy or speculative fiction. With digital technologies generating immersive experiences where even the tactile sensation of a device can lead to doom scrolling, the trajectory toward a future shaped by virtual and extended reality is clear. As daily life intertwines with digital realms, this shift brings emerging violence, power

reconfigurations, and entrenched inequalities. This paper shows that immersive platforms are not neutral or harmless but environments that reproduce, reinforce, and normalize violence rooted in social, economic, gendered, and racial hierarchies. From symbolic violence to intellectual oppression, these harms are not system failures but outcomes of designs reflecting real-world inequities. Pain in virtual worlds mirrors offline injustices, carrying psychological and physical consequences. Drawing on Foucault's power-knowledge, Butler's grievability, and Zuboff's surveillance capitalism, this paper argues that Metaverse violence is no accident but a repetition of real-world values, ideologies, and economic structures embedded by creators and users. The promise of being "anything you want" is constrained by design from avatar customization to default settings, reinforcing dominant ideals. Case studies from *VRChat, The Sandbox,* and *MetaHorizon Worlds* show how user freedom is limited by poorly regulated systems, allowing violence to persist. Virtual violence is not isolated. When immersive experiences reward class behaviors or social dominance, they replicate hierarchies offline. Digital class systems, algorithmic bias, generational divides, and commodification of users contribute to a culture that treats differences in ethnicity, gender identity, geography, or class as grounds for unequal recognition. Governance of the Metaverse will shape power distribution. Without transparent, accountable, participatory structures, dominant hierarchies may persist unless digital justice and inclusion are prioritized. In a polycrisis era of conflict, migration, instability, and disaster, this new frontier cannot be ignored. The Metaverse, fueled by billions in digital transformation, is becoming integral to human life. It must be designed not to reproduce violence but to inspire equity, empathy, and ethical pluralism. Understanding violence in the Metaverse, from symbolic to systemic, is essential to developing technologies that prioritize justice, equality, and diversity. The future of digital spaces must be generative, grounded in care, accountability, and collective imagination.

References

Association for Progressive Communications. (2021). *Take Back the Tech: 2021 Campaign Summary.* APC. https://www.apc.org/en/project/take-back-tech

Banet-Weiser, S. (2018). *Empowered: Popular Feminism and Popular Misogyny.* Duke University Press. https://doi.org/10.2307/j.ctv11316rx

Benjamin, R. (2020). *Race after technology: Abolitionist tools for the new Jim code* (Reprinted). Polity.

Boellstorff, T. (2015). *Coming of age in Second Life: An anthropologist explores the virtually human* (First new edition paperback). Princeton University Press.

Bourdieu, P. (2002). *Distinction: A social critique of the judgement of taste* (11. print). Harvard Univ. Press.

Butler, J. (2015). *Gender Trouble: Feminism and the Subversion of Identity* (First issued in hardback). Routledge, Taylor & Francis Group.

Butler, J. (2016). *Frames of War: When Is Life Grievable?* Verso.

Costanza-Chock, S. (2020). *Design Justice: Community-Led Practices to Build the Worlds We Need.* The MIT Press. https://doi.org/10.7551/mitpress/12255.001.0001

Couldry, N., & Mejias, U. A. (2019). *The Costs of Connection: How Data is Colonizing Human Life and Appropriating it for Capitalism.* Stanford University Press.

DeVeaux, C., Han, E., Hudson, Z., Egelman, J., Landay, J. A., & Bailenson, J. N. (2025). Black immersive virtuality: Racialized experiences of avatar embodiment and customization among Black users in social VR. *Computers in Human Behavior, 168,* 108639. https://doi.org/10.1016/j.chb.2025.108639

Dijck, J. V. (2013). *The Culture of Connectivity: A Critical History of Social Media.* Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199970773.001.0001

Ekō. (2022). *Metaverse: Another Cesspool of Toxic Content.* Ekō. https://www.eko.org/images/Metaverse_report_May_2022.pdf

Foucault, M. (1995). *Discipline and punish: The birth of the prison* (A. Sheridan, Trans.; Second Vintage Books edition). Vintage Books, a division of Random House, Inc.

Fox, J., & Tang, W. Y. (2014). Sexism in online video games: The role of conformity to masculine norms and social dominance orientation. *Computers in Human Behavior, 33,* 314–320. https://doi.org/10.1016/j.chb.2013.07.014

Galtung, J. (1969). Violence, Peace, and Peace Research. *Journal of Peace Research, 6*(3), 167–191. https://doi.org/10.1177/002234336900600301

Gillespie, T. (2018). *Custodians of the internet: Platforms, content moderation, and the hidden decisions that shape social media.* Yale University Press.

Groom, V., Bailenson, J. N., & Nass, C. (2009). The Influence of Racial Embodiment on Racial Bias in Immersive Virtual Environments. *Social Influence, 4*(3), 231–248. https://doi.org/10.1080/15534510802643750

Internet, F. (2020). *Feminist Internet Research Manifesto.* Feminist Internet. https://feministinternet.org

Jane, E. (2017). *Misogyny Online: A Short (and Brutish) History.* SAGE Publications Ltd. https://doi.org/10.4135/9781473916029

Morozov, E. (2011). *The Net delusion: The dark side of internet freedom.* PublicAffairs.

Nakamura, L. (2013). *Cybertypes: Race, Ethnicity, and Identity on the Internet.* Taylor and Francis.

Peck, T. C., Seinfeld, S., Aglioti, S. M., & Slater, M. (2013). Putting yourself in the skin of a black avatar reduces implicit racial bias. *Consciousness and Cognition, 22*(3), 779–787. https://doi.org/10.1016/j.concog.2013.04.016

Seigneur, J.-M., & Choukou, M.-A. (2022). How should metaverse augment humans with disabilities? *13th Augmented Human International Conference,* 1–6. https://doi.org/10.1145/3532525.3532534

Spivak, G. C. (1988). Can the Subaltern Speak? In C. Nelson & L. Grossberg (Eds.), *Marxism and the Interpretation of Culture.* University of Illinois Press.

Stephenson, N. (1992). *Snow Crash.* Bantam Books.

Zuboff, S. (2019). *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power.* PublicAffairs.
https://www.hachettebookgroup.com/titles/shoshana-zuboff/the-age-of-surveillance-capitalism/9781610395694/?lens=publicaffairs