Matthias Quent

Democratic Culture in the Metaverse

Interdisciplinary Perspectives on the Potentials and Perils

Nomos

Matthias Quent

Democratic Culture in the Metaverse

Interdisciplinary Perspectives on the Potentials and Perils

Nomos

Online Version
Nomos eLibrary

The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available on the Internet at http://dnb.d-nb.de

1st Edition 2025

Published by
Nomos Verlagsgesellschaft mbH & Co. KG
Waldseestraße 3–5 | 76530 Baden-Baden
www.nomos.de

Production of the printed version:
Nomos Verlagsgesellschaft mbH & Co. KG
Waldseestraße 3–5 | 76530 Baden-Baden

ISBN 978-3-7560-1919-9 (Print)
ISBN 978-3-7489-4811-7 (ePDF)

DOI https://doi.org/10.5771/9783748948117

Table of Contents

Democratic Culture, Virtual Spaces, and the Internet of Tomorrow

Matthias Quent

In the early 21st century, democracy finds itself confronted with profound changes and challenges. Worldwide, a backsliding of democratic cultures and structures can be observed: in many countries, democratic institutions and norms are coming under pressure while authoritarian tendencies gain traction. At the same time, the digital revolution is transforming the public sphere at a rapid pace and facilitating the rise of populist and illiberal forces. Societies have yet to find a satisfactory way to deal with the harmful side effects of the dynamic developments of social media in the Web 2.0 era. New platforms, artificial intelligence, and immersive technologies like the Metaverse are already becoming available. They offer unprecedented possibilities for communication and participation, but they also bring with them significant risks for user safety as well as for democratic culture.

Democratic culture generally refers to the totality of those beliefs, attitudes, and behaviors in a society that not only accept democratic principles like pluralism, human rights, equality, the rule of law, and participation, but actively support them. It manifests, for example, in a willingness to engage in open discourse, to tolerate divergent opinions and identities, and to actively participate in political processes. In the context of digital and immersive spaces in the Metaverse, the term Democratic culture takes on an expanded meaning: here, democratic culture encompasses not only the political values of individual users and how they interact with each other, but also the manner in which digital publics are designed, regulated, and experienced. A democratic culture in virtual environments requires that users are able to communicate as equals, that diversity is protected and made visible, and that discriminatory structures are recognized and curtailed. Ideally, platforms are not driven solely by commercial interests; instead, they enable user participation in decision-making and accept democratically legitimized regulations to balance freedom of expression and anti-discrimination. Democratic culture requires openness, transparency, and participation even beyond state institutions.

This book brings together transdisciplinary analyses and perspectives on the opportunities and risks of democratic culture in immersive digital environments, which are discussed under the term "Metaverse". These analyses were developed in the context of the "Immersive Democracy" project as part of the European Metaverse Research Network. The aim of the Immersive Democracy project is to empower social scientists, civil society actors, designers and creators of immersive environments, as well as political decision-makers, with theoretical foundations and empirical research to shape the internet of the future in an inclusive and democratic way.

In all regions of the world, democracy is under pressure – often from populist actors who reject political pluralism and seek unchecked power, while skillfully exploiting digital technologies. The background of this global trend is multifaceted – economic inequalities, nationalism as a reaction to the effects of globalization, the erosion of the middle class, cultural backlash against social change, and various crises all play a central role, as do changes in the media ecosystem. The latter directly links the crisis of democracy to the digital sphere, where a large part of political life, communication, and socialization takes place today.

Digital Public Spheres, Platform Structures, and the Concentration of Informational Power

Digitalization has fundamentally changed the foundations of public communication. While the internet in the 1990s was associated with hopeful visions of a free, decentralized information space, this image has shifted over the past two decades: the initial techno-utopian ideals have given way to the realization that legal frameworks, market forces, and social practices have transformed the web into a commercial platform economy. Today, a few digital platforms – above all global social networks, search engines, and messaging services – dominate the role of public forums. They function as new gatekeepers of opinion formation and use algorithmic curation to control which information becomes widely visible and which does not. This has been accompanied by an extreme concentration of informational power: only a handful of companies and their decision-makers have the ability to influence the online communication of billions of people worldwide to an unprecedented extent. These corporations collect and analyze vast quantities of user data and make decisions based on that – about, for

instance, newsfeed algorithms or content moderation – that have global societal and political impacts. Through their global economic and media power, they are able to thwart effective regulations by democratically legitimized governments.

These digital public spheres are privately organized and hierarchically controlled. Instead of transparent, accountable governance, control is concentrated in the hands of a few individuals – founders, CEOs, and major investors – who often act according to their own financial or ideological interests. As a result, private-sector actors with exclusive decision-making authority and sometimes a political agenda determine central aspects of freedom of expression and the order of discourse on the internet. This concentration of power raises grave questions in democratic theory: if public deliberation increasingly takes place in closed spaces rather than in open forums based on a consensual social framework of rules, principles of debate oriented toward the common good risk being pushed to the margins. Algorithms optimized for profit maximization and prolonged user engagement can, for example, reward sensationalism and outrage more than factual information – impairing the quality of public opinion formation. In this way, digital platforms become "private publics" with enormous power but without genuine accountability to society. The concentration of this power in a few hands – whether through algorithmic control over information flows or through surveillance practices for data analysis – jeopardizes the diversity and openness that a functioning democratic public sphere requires.

Democratic Risks of Monopolistic Platform Economies

Closely tied to the concentration of power in the digital realm are the quasi-monopolistic structures of today's platform economy. The global market for online communication, social media, and digital services is dominated by just a few large corporations – often referred to as "Big Tech" – such as Meta (Facebook, Instagram, WhatsApp), Alphabet (Google, YouTube), Apple, Amazon, Microsoft, or in Asia Tencent (WeChat, QQ) and ByteDance (TikTok). Each of these companies controls entire ecosystems of services and has attained a market position that presents traditional antitrust oversight with enormous challenges. This is democratically relevant primarily because extremely concentrated economic and informational power in private hands poses a potential danger to pluralistic societies and to the

principle of democracy itself. Monopolies tend to replace transparent rules with corporate interests and to convert public goods into private profits – while social costs are externalized to society. For example, the moderation of speech on Facebook or YouTube is ultimately subject to the guidelines and decisions of a single company – there is no democratic participation by the users or the general public in determining who may say what, or according to what criteria content is regulated.

The democratic risks of such monopolistic platform economies are manifold. First, a few companies can dominate the informational agenda of society – for instance, through selection of news or manipulation of reach – and thereby potentially influence political processes. Second, there is the danger of misuse of personal data and mass surveillance: Shoshana Zuboff (2019) coined the term "surveillance capitalism" for a regime in which corporations like Google or Meta analyze and influence citizens' behavior to generate profit – undermining individual autonomy and privacy. Third, public communication may become dependent on private infrastructure: if a dominant service fails or arbitrarily changes its terms of use, it can destabilize entire democracies and upend the relationship between truth and falsehood (think of the role of platform X under Elon Musk). Fourth, quasi-monopolistic digital corporations can convert their economic clout into political influence through lobbying, pressure on governments, or media agenda-setting.

A stark example outside of liberal democracies is Tencent in China: as the operator of WeChat, Tencent controls not only the social network but also payment transactions, everyday services, and news consumption for over a billion users – tightly interlinked with state surveillance. Here we see the dystopia of a corporate-provided every day and public space merging with authoritarian state power. Although liberal democracies differ significantly in political culture and legal context, the structural question remains: How can we prevent a few private platforms from becoming the quasi-infrastructure of democratic society without any democratic oversight mechanisms in place?

AI as a Challenge for the Political Public Sphere, Pluralism, and Participation

In parallel to platform concentration, another layer of the digital challenge is emerging: the rapid proliferation of Artificial Intelligence (AI) in the form of algorithmic decision-making systems, big data analytics, and more

recently generative AI (such as ChatGPT, DALL-E, or deepfake technologies). On one hand these developments promise efficiency gains and new tools – on the other hand, they bring significant dangers for the political public sphere, societal pluralism, and participation – a subject substantial enough for another important book in its own right. Above all, it should be noted here that developments in AI and the Metaverse are closely intertwined: the generation of environments and content in the Metaverse relies on vast amounts of data – AI is a prerequisite for the Metaverse, not its adversary.

Democratic Scope for Shaping Digital Spaces: Governance, Regulation, Innovation

In light of the risks outlined, the pressing question arises of how democratic principles and participation can be defended and shaped in the digital age. If digital platforms are becoming central arenas of political opinion-formation and AI systems increasingly intervene in societal processes, democracies need new strategies to enforce their values and rules in these spheres. In this context, several levels must be distinguished: first, state regulation and legal frameworks; second, governance innovations by the platforms themselves; and third, democratic innovations and alternative counter-models from civil society.

Regulation: In recent years, European democracies in particular have begun to develop comprehensive regulatory frameworks for the digital realm. A prominent example is the EU-wide Digital Services Act (DSA), which came into force at the end of 2023. With the DSA, the EU aims to rebalance the power relationship between users, platform companies, and public authorities. The legislation obliges online services to greater transparency (for example, through regular reports on content moderation and algorithmic systems), mandates the preservation of fundamental rights principles in their terms of service, and strengthens cooperation between platforms and supervisory authorities. Particularly large platforms (designated as Very Large Online Platforms, VLOPs) are subject to additional obligations. Among other things, they must conduct risk analyses regarding their societal impacts, submit to independent audits, offer users an option for non-personalized feeds, and provide data for research purposes. Overall, the DSA is designed to curb the information and economic power of tech corporations, give users more rights vis-à-vis platforms, and enable

authorities to act more effectively against disinformation, illegal content, and systemic risks. In addition, the Digital Markets Act (DMA) addresses monopolistic practices by imposing antitrust restrictions on gatekeeper platforms (for example, banning self-preferencing of their own services). These European initiatives are unprecedented globally and ultimately aim to structure digital spaces more in line with democratic ground rules. Likewise, in other democracies, such as Australia, Canada, or India, laws are being discussed to increase the accountability of large platforms (for example, regarding content moderation or the protection of personal data).

Governance and Self-Regulation: In parallel to state regulation, platforms and international policymakers are also searching for new governance models. One approach is the implementation of multi-stakeholder committees and oversight boards that incorporate a broader range of interests. For instance, Facebook – responding to public criticism – set up an "Oversight Board" that reviews content decisions in a limited capacity. Although its independence and influence remain disputed, this move indicates that platform rules need no longer be set in an entirely internal and opaque manner. Similarly, there are considerations to establish user councils or advisory boards to involve civil society in drafting moderation guidelines, algorithmic principles, or data practices. Digital platforms, as public spaces, should be a subject to democratic rules and oversight, while at the same time remain resilient against authoritarian and human-rights-violating co-optation. Politically, there is debate as to what extent competition policy (e.g. breaking up tech monopolies or stricter merger control) and transparency requirements (such as opening up recommendation algorithms) can improve the structural balance of power. All these measures aim at a technological order in which democratic value like transparency, accountability and participation, are anchored in digital systems.

Civil Society and Grassroots Innovations: At the level of civil society and grassroots initiatives, a variety of innovations are emerging to harness the opportunities of digitalization in a pro-democratic way. Under the banner of "civic tech," numerous digital participation platforms have been developed in recent years that open up new channels for citizens to have a voice. Examples include online petition platforms, participatory budgeting tools, e-voting pilot projects, and deliberative discussion platforms. These tools make it possible to involve broad segments of the population directly in political deliberation and decision-making processes – independent of time and place via the internet. For instance, citizen budgets can be voted on online, or legislative proposals can be collaboratively edited on the web.

From participatory budgeting to virtual citizen forums, digital participation platforms are already helping communities co-create policy, hold governments accountable, and amplify citizens' voices. Especially in times when traditional democratic institutions are losing public trust, such innovations offer a chance to make democracy more resilient and inclusive. Digital participation tools can build bridges between citizens and institutions and make political processes more transparent. Of course, these instruments also face challenges (e.g. the digital divide, abuse by trolls, data protection issues), but they reveal an important potential: Namely, to understand the internet not just as a threat, but as an expansion of the democratic realm.

Finally, the emergence of the Metaverse brings a new dimension of the digital sphere into focus, one that will increasingly need to be shaped in the future. In general, the Metaverse is understood as an immersive virtual environment in which a persistent, three-dimensional online world is created by means of virtual or augmented reality and interconnected platforms. Major technology corporations – especially Meta – are investing massively in this vision of a next-generation internet. For democracy, this means both opportunity and risk: on the one hand, novel spaces for social exchange, education, and even political participation could arise in Metaverse environments (for example, virtual town halls, simulations for citizen dialogues, and gatherings unconstrained by geography). On the other hand, everything discussed so far threatens to repeat itself in an even more acute form or even intensify. If the Metaverse infrastructure is provided by the same quasi-monopolistic firms, there is a danger that an immersive parallel public sphere under private control will emerge, in which the rules are dictated by a few platform operators. Governance and law enforcement are even harder to grasp in three-dimensional cyberspace – questions of accountability for virtual transgressions, hate speech, or manipulation in VR remain largely unresolved. The Metaverse could be misused to foment hatred, manipulate public opinion, and enable even more far-reaching surveillance, for example through the analysis of users' movement data. It is therefore crucial to embed democratic principles such as equality, transparency, data protection, and security into the architecture of the Metaverse and its AI systems from the outset. Moreover, open standards and alternative models should be promoted. Now is the time to set the course for an "immersive democracy," in which immersive virtual worlds do not become lawless zones or the monopolistic private realm of a few corporations, but are designed to be accessible, inclusive, and guided by fundamental rights.

The socio-political relevance of this topic could hardly be greater. While traditional democratic structures are coming under pressure worldwide, it is in the digital sphere – and soon in the Metaverse – that it will be decided whether democratic values and processes experience renewal and strengthening or are further hollowed out. The anthology *Immersive Democracy – Interdisciplinary Perspectives on the Potentials and Perils of Democratic Culture in the Metaverse* seeks to explore precisely this tension. The following contributions analyze, from a transdisciplinary perspective, how democratic culture can be preserved, fostered, and reimagined under the conditions of immersive technologies – to ensure that the next chapter of digital evolution does not become a step backward for democracy, but in the best case contributes to its deepening.

About this book

The contributions in this volume approach the relationship between democratic culture and immersive digital technologies from an interdisciplinary perspective and shed light on theoretical foundations as well as concrete fields of application.

In „Towards a Democratic Culture in the Metaverse?", Matthias Quent reflects on what it would take to cultivate democratic norms and structures in emerging virtual worlds. He discusses the risks of leaving metaverse development solely to corporate interests and highlights the importance of embedding principles like transparency, inclusivity, and human rights from the outset. Quent raises questions about governance models for the metaverse – for example, whether user communities can participate in rule-setting or if new institutions are needed to safeguard digital public spaces. He argues that without deliberate action, virtual environments could amplify existing social inequalities and undemocratic tendencies; however, with proactive engagement, the metaverse could become a sphere for enhanced civic participation and community empowerment. This chapter sets a forward-looking tone by challenging stakeholders in tech, government, and civil society to collaboratively steer the metaverse toward democratic ends.

Lenn Blaschke's article „A Theory of Immersive Democracy" proposes a conceptual framework for understanding how immersive technologies might transform democratic participation. Drawing from his background as a media artist, Blaschke explores the idea that immersion can deepen civic engagement by making democratic processes more experiential and

accessible. He discusses how virtual and augmented reality could enable citizens to "step into" political issues or community debates, potentially increasing empathy and understanding in polarized societies. At the same time, Blaschke critically examines challenges to this vision – such as ensuring broad access to immersive tech and guarding against the manipulation of virtual experiences for propaganda. Ultimately, A Theory of Immersive Democracy *serves* as an invitation to rethink traditional democratic practices in light of immersive media's potential to bring people closer to the issues and to each other.

Octavia Madeira and Georg Plattner („A Safe Space for Everyone – A Plea for a Democratic and Participative Metaverse) warn that as the metaverse evolves, extremists and other *malevolent actors* will creatively exploit it for propaganda and recruitment. In this chapter, they argue that these risks demand proactive measures to ensure the metaverse remains a democratic and participative space. The authors emphasize involving users, platforms, and democratic institutions in the governance of virtual worlds, urging that marginalized communities be empowered to counter hate and discrimination in immersive environments. They highlight that the metaverse is still in an early development phase – a critical window for building tools and norms that can make future virtual worlds safe and inclusive for all. This plea calls for collaboration across tech and civil society to develop a metaverse that prioritizes safety, participation, and equality as core features.

In "Navigating Democratic Challenges in the Age of Metaverses" Jean-François Lucas, Jessica Galissaire and Henri Isaac tackle the broad democratic dilemmas posed by the rise of multiple metaverse platforms. They observe that as we enter the "age of metaverses," questions of governance, accountability, and citizen rights become increasingly complex in these new virtual domains. The chapter discusses how the fragmentation of digital spaces among various private providers could lead to inconsistent norms and oversight, making it challenging to uphold public values across the metaverse. Key issues addressed include ensuring transparency in how immersive platforms are run, giving users a voice in platform policies, and preventing virtual worlds from becoming siloed echo chambers or zones of unchecked power. At the same time, Lucas, Galissaire, and Isaac highlight opportunities: metaverses might be used to *enhance* democratic engagement (for instance, through virtual town halls or collaborative world-building that involves citizens) if designed with inclusive principles. This chapter essentially navigates the tension between the exciting possibilities of immer-

sive technologies and the need for robust democratic safeguards as these technologies mature.

Esen K. Tütüncü and Danielle Shanley *(*"Hate Speech in the Metaverse"*)* explore how hateful content and harassment might manifest in immersive virtual environments, and what can be done about it. They note that defining "hate speech" is inherently challenging, but it remains crucial to address because virtual worlds can both reflect and *reinforce* real-world biases. The chapter shows that many forms of hate observed on traditional social media – racism, sexism, homophobia, and others – are already appearing in social VR platforms, often intensified by the immersive nature of these spaces. For example, users in VR can engage in nonverbal harassment (e.g. offensive gestures with avatars) and invade personal space, which can amplify the emotional impact of hate incidents. Tütüncü and Shanley emphasize that mitigation will require proactive measures: strong community standards, real-time content moderation tools, user education, and collaboration between platform developers, policymakers, and civil society. They also invoke concepts of "responsible innovation," arguing that the principles of deep democratic engagement should guide the design of the metaverse to prevent hate and bias from being "built in" to these emerging systems. In sum, this chapter highlights the imperative of early and concerted efforts to keep the metaverse's social spaces inclusive and respectful**.**

Julia Ebner examines in her article "Decentralized Autonomous Organisations (DAOs) in the Metaverse: The Future of Extremist Organisation?" *whether* decentralized, blockchain-based communities – known as DAOs – could become the next frontier for extremist groups operating in the metaverse. DAOs are online organizations without central leadership, governed by smart contracts, which allow members to coordinate and make decisions collectively across borders. Ebner warns that while DAOs hold promise for reimagining how we collaborate online, they might also be exploited by extremists to fundraise, spread propaganda, or even coordinate illegal activities in a leaderless, anonymous fashion. The chapter suggests that such decentralized structures could enable extremist networks to bypass traditional oversight, posing risks to minority rights and the rule of law if used for hate campaigns or plotting violence. Ebner discusses real and hypothetical scenarios where extremist or anti-democratic actors leverage DAOs to challenge democratic institutions for instance, using pooled funds for election interference or creating private governance systems that reject external authority. In response, she calls for foresight in monitoring these developments and adapting our counter-extremism and legal frameworks,

so that the metaverse economy (including DAOs) does not become a safe haven for anti-democratic organization.

Sandra Kero and Josephine B. Schmitt ("Narrative, Creative – Immersive? The Immersive Potential of Right-Wing Extremist Communication on Social Media") investigate how far-right actors leverage the immersive qualities of modern social media to spread their messages. They argue that "immersiveness" in this context isn't limited to VR headsets – it also refers to the deep emotional engagement and narrative techniques that draw users in on platforms like Instagram, TikTok, or YouTube. The chapter outlines mechanisms by which right-wing extremists make their content feel immersive and compelling. For example, blending political propaganda with lifestyle content, music, or interactive memes to create a sense of community and excitement around their ideology. This strategy normalizes extremist ideas by packaging them as part of everyday online culture, lowering the threshold for young or apolitical users to absorb these views. Kero and Schmitt also discuss how features like endless scrolling, Livestreaming, and algorithmic personalization can "pull" users deeper into a narrative bubble, effectively immersing them in an alternative reality curated by extremist influencers. A key contribution of this chapter is its interdisciplinary analysis (drawing on media and psychology theories) to explain why such content can be so emotionally resonant and persuasive. Finally, the authors offer recommendations on how to counter this immersive extremism, from media literacy efforts to platform design changes, aiming to disrupt the seductive storytelling that far-right communicators employ.

Arne Vogelgesang explores in this work „Platform Design, User Creativity, and Aesthetic Governance in Social VR" *how the design of* social VR platforms, combined with the creative agency of users, gives rise to what he calls "aesthetic governance." He examines two leading social VR environments – *Rec Room* and *VRChat* – as case studies to understand how different design philosophies can shape community culture and norms. For instance, *Rec Room* has a unified, game-like aesthetic with more controlled creation tools, whereas *VRChat* offers vast freedom for users to create avatars and worlds. Vogelgesang argues that these design differences lead to distinct forms of self-governance: platform rules and user behaviors evolve together as an *interaction between top-down design choices and bottom-up creative expression*. The chapter discusses how features like world-building tools, avatar customization, and moderation systems influence the development of social norms – what is considered acceptable or taboo – within these virtual communities. Vogelgesang's concept of aesthetic gover-

nance highlights that governance in the metaverse isn't only about formal rules; it's also about the implicit cultural standards that emerge from a platform's look, feel, and possibilities for user creativity. By comparing how *Rec Room* and *VRChat* communities differ, the chapter sheds light on the political implications of seemingly technical design decisions, suggesting that the way we build virtual spaces can encourage either more positive social interaction or, conversely, enable problematic subcultures. This analysis is crucial for understanding how we might foster *"immersive democracy"* – the idea that the design of virtual worlds should actively support democratic values and inclusive community-building.

Mick Prinz („Unpixelated Hate? Lessons from VR Gaming for a Digital Civil Society") draws on the world of video games to extract lessons for combating hate and extremism in the metaverse. He notes that video game environments have long been immersive social spaces where cultural values are both expressed and contested – virtual worlds are not just escapist playgrounds, but arenas where political debates occur and norms are formed. In toxic corners of gaming culture, anti-democratic actors have applied meta-political strategies, using games and gamer communities to normalize racist, sexist, or anti-Semitic narratives. Prinz points out that the games industry since the 1980s has been a testing ground for these dynamics, with some gaming platforms suffering from poor moderation that allows harassment and hate to flourish. At the same time, he acknowledges that gaming has also spawned positive movements – groups of players and developers working to promote diversity and inclusion in their communities. The chapter then asks what the burgeoning metaverse (which in many ways resembles a massively multiplayer game world) can learn from these experiences. Prinz discusses concrete examples, such as how lack of moderation in voice chats or VR lobbies leads to intimidation of marginalized users, and conversely how community-driven initiatives have successfully pushed back against toxic behavior in some games. He concludes that building a healthy *digital civil society* in the metaverse will require adapting effective strategies from gaming: robust moderation policies, user empowerment tools, and coalitions between industry and civil society to promote pluralistic values. In essence, *Unpixelated Hate?* calls for preemptive action so that the metaverse can be a space of playful democracy rather than a new wild west of unchallenged hate.

In the chapter „Identity and Safety in Social VR: Findings from an Experimental Avatar-Based Interview Study", Matthias Quent and Sara Lisa Vogl present results from an innovative study where they interviewed users

inside a social VR platform (using avatars) to understand issues of identity and safety. Their findings highlight several positive aspects of social VR. Many participants described how platforms like VRChat allow them to overcome real-life social anxieties and form meaningful friendships in a controlled virtual environment. For example, introverted users appreciated having tools to limit their exposure to overwhelming situations and to connect with others from the comfort of home. One interviewee noted that in VRChat people "get to know each other mainly based on their personality instead of physical appearance," which gave them a confidence boost and a safe space to practice social interaction. Quent and Vogl also delve into how users experiment with gender identity in VR. Some participants chose avatars of a different gender or presented in fluid ways; generally, they felt their choices were respected by others in the community, though instances of misgendering or harassment still occurred (mirroring real-world challenges). The ability to embody an avatar that aligns with one's felt identity or even to appear as fantastical as one wishes was seen as liberating and empowering by the interviewees, as long as the social environment remained supportive. On the flip side, the study acknowledges safety concerns: harassment does exist in social VR, and the chapter notes the importance of platform features (like blocking or muting tools) and community norms to protect users. Overall, Quent and Vogl's research provides an on-the-ground look at how identity expression and user safety play out in immersive social spaces, offering empirical insights to inform safer and more inclusive metaverse design.

Deborah Schnabel examines „Chances and Limits of Immersive Environments for Anti-Discrimination and (Historical-)Political Education". *She notes that each* new digital medium brings both problems and possibilities for civic education: on one hand, extremists quickly colonize new forums to spread racism and Antisemitism; on the other hand, those same forums offer educators innovative ways to reach people. Schnabel highlights the unique potential of VR to engage learners. For instance, virtual reality can preserve important historical testimonies (such as Holocaust survivor stories) by recreating sites and scenarios, enabling powerful experiential learning that would be impossible otherwise. She cites projects like the Anne Frank House VR tour and the "Inside Auschwitz" app, which give users immersive insights into history and empathy for victims. However, the chapter also underscores the ethical and pedagogical limits of such immersion. Schnabel discusses concerns about "artificial authenticity" – whether a simulated experience might distort history or provoke reactions

that educators can't properly handle. There is debate over how *realistic* these simulations should be: an overly realistic experience could be overwhelming or even inappropriate (for example, letting someone virtually "play" a victim of racism can cross ethical lines). The chapter describes how some projects deliberately include distancing elements (or even advise against using VR headsets in certain contexts) to avoid an "uncanny valley" of empathy that becomes counterproductive. In conclusion, Schnabel maintains that while immersive tech holds great promise for anti-discrimination and historical education by making lessons more engaging and visceral it must be used thoughtfully. The quality of outcomes will depend on aligning these new tools with sound educational principles, ensuring that empathy and critical thinking are fostered without trivializing or sensationalizing traumatic history.

In the chapter „Police Handling of Hate Crime: A Pilot Project to Use VR Technology for Professional Development in Sensitizing Police Officers to the Experiences of Victims of Bias Crime in Hamburg", the researchers Eva Groß, Ulrike Zähringer an Anabel Taefi report on a pioneering VR training program aimed at improving police responses to hate crimes. The authors describe how immersive technology was used to *sensitize officers* to the perspectives of victims of bias-motivated crimes. The rationale is that virtual reality can provide experience-based learning opportunities: by immersing police trainees in scenarios where they witness or feel the impact of racist or other bias-driven abuse, it can cultivate greater empathy and understanding. The chapter explains that trust in law enforcement, especially among minority communities, hinges on officers' ability to handle hate crime victims with professionalism and compassion. The VR pilot in Hamburg allowed officers to step into simulated environments that replicate common bias-crime situations and victim interactions. Preliminary results from the project showed promising benefits. Officers reported heightened awareness of victims' emotional states, and a better grasp of how their own behavior might affect victim trust. By *"walking a mile"* in a victim's shoes, even virtually, participants became more attuned to the subtleties of trauma and discrimination. Groß and colleagues also discuss the practical lessons learned: for example, how important it is to debrief participants after intense VR experiences, and how such training can be integrated into regular police education. They conclude that immersive training can be a powerful tool in building police capacity to combat hate crime, ultimately strengthening democratic resilience by improving the relationship between marginalized communities and law enforcement.

Jonas Fergert („Real Participation in Virtual Environments: Navigating Public Participation in the Metaverse") explores how immersive technologies could transform citizen participation in governance and public life, while also analyzing the obstacles that must be overcome. He begins by observing that the initial hype around the metaverse (sparked in 2021 by companies like Meta) has subsided, giving way to a more sober discussion about practical uses of these technologies. Fergert argues that this "hype hangover" period is an opportunity to proactively think about using virtual and augmented reality for civic purposes, not just commercial ones. One vision he discusses is integrating VR into public consultation processes – for example, virtual town halls or participatory urban planning where citizens can virtually experience proposed changes to their city and give feedback. Immersive systems can increase users' sense of involvement in democratic processes (a concept supported by research on how telepresence boosts engagement). The chapter notes that demand for digital participation tools was rising even before the metaverse, accelerated by the COVID-19 pandemic's push toward online engagement. However, Fergert also details significant challenges: ensuring equitable access to the necessary devices and internet connectivity, protecting personal data in virtual forums, and avoiding replicating the flaws of social media (such as algorithmic biases and commercialization) in the metaverse. He emphasizes that *deliberate design choices* are needed so that future virtual participation platforms encourage deliberation and inclusivity rather than division. In conclusion, this chapter paints a picture of both great potential and pitfalls in using VR/AR for public participation – with the core message that now is the time to shape these tools to serve democracy, before tech companies' agendas set the rules of the game.

Martin Müller and Matthias C. Kettemann („EU Platform Regulation and its Implications for the Metaverse: An Analysis of DSA, DMA, and Related Legal Acts") analyze the legal challenges posed by the metaverse and how existing law might apply to immersive digital spaces. They point out that, so far, no country or the EU has a comprehensive regulatory framework specifically for the metaverse, we are only at the stage of broad policy visions. Nevertheless, the authors show that many *pieces* of law will impact the metaverse: data protection rules, content moderation laws, product safety standards for VR hardware, and so on all form a patchwork of norms that govern aspects of virtual worlds. Key legal issues discussed include privacy and data security (given the vast personal data that devices and virtual interactions generate), user safety and content governance (how to deal

with illegal or harmful behavior in VR), questions of jurisdiction (which country's laws apply in a borderless virtual space), and interoperability and competition (ensuring open, non-monopolistic metaverse ecosystems). Müller and Kettemann highlight recent European regulatory approaches as especially relevant – for example, the EU's Digital Services Act (DSA), which sets obligations for online platforms and could enforce stricter content moderation and transparency in metaverse platforms. They also note the concept of the "constitutionalizing" of social media (embedding fundamental rights and constitutional principles into platform governance) and argue that a similar approach is needed for the metaverse. In essence, this chapter suggests that while the metaverse may feel like the Wild West, the rule of law *does* and *must* extend into it. The authors call for a combination of updated laws and proactive self-regulation to ensure that immersive digital environments develop in a way that protects users' rights and public interests.

In their contribution "Regulation of the Metaverse", Matthias C. Kettemann and Caroline Böck delve deeper into the normative foundations that should guide the metaverse's development, especially concerning the protection of fundamental rights. They build on the idea that social media underwent a "constitutionalizing" – where democratic societies asserted that platforms must uphold basic rights and values – and they argue that the metaverse now needs to undergo a similar process. Kettemann and Böck discuss how human rights (like freedom of expression, privacy, equality, and non-discrimination) can be translated and enforced in virtual environments that blur the line between physical and digital life. A core question in this chapter is how to prevent the metaverse from becoming a lawless space: the authors examine issues such as the lack of clear jurisdiction in virtual worlds, the challenge of holding platform owners accountable, and the potential role of international law or new institutions in governing the metaverse. They likely advocate for embedding public law values into the very architecture of the metaverse – for instance, through design choices that protect users by default and through governance models that include user representation. Kettemann and Böck emphasize that the rights and protections people enjoy "offline" should carry over to online immersive spaces, asserting that our rights must be "always on." They conclude with recommendations for policymakers and stakeholders on safeguarding democracy and the rule of law as we build these new worlds. This forward-looking chapter essentially serves as a legal and ethical com-

pass, insisting that innovation in the metaverse must go hand-in-hand with responsibility and respect for human dignity.

I would like to thank all the authors and contributors to this book, especially Susanne Haldrich and Lilli Walter for their support, and Sandra Frey at Nomos Publishing for her kind assistance.

Towards a Democratic Culture in the Metaverse: An Overview of Risks and Opportunities[1]

Matthias Quent[2]

Introduction

Whether intended or not: When Mark Zuckerberg, the founder and CEO of Facebook, rebranded his company as "Meta Platforms" in October 2021 and introduced the term "Metaverse" for the internet of the future, it was also a warning for democracy. The term "Metaverse" was first popularized in 1992 by Neal Stephenson's dystopian science fiction novel *Snow Crash*. In this novel, Stephenson describes a virtual parallel universe in future America, where one can dive into it through computer screens, interact socially and economically with avatars, and purchase virtual houses, properties, and objects. The digital avatars can—similar to leveling up in video games—improve their status positions in the virtual world. In the novel's Metaverse, avatars can attain a higher level of social recognition and status than the humans behind them possess in physical reality. However, there is a danger lurking in the Metaverse: *Snow Crash*. This experience, which can be interpreted as a drug, virus, or misinformation, spreads and infects users in front of their computer screens through the immersive experience. In *Snow Crash*, the real world is a right-wing libertarian dystopia: the USA is divided into regional zones controlled by corporations or the mafia. The wealthy are rulers with total decision-making power, including who can move where or live where and how. There is racist and social segregation, and the power of the ultra-rich prevails. Democracy and the environment are destroyed. In contrast, the Metaverse appears as an attractive escape point. Therefore, the literary concept of the Metaverse includes warnings for the social reality: warnings about the destruction of democracy and the destructive potential of a new generation of digital immersive technolo-

1 The article is an expanded and updated version of an article first published in German (Quent 2023).

2 With the support of Lilli Walter.

gies, which accompany new experiential spaces and freedoms of virtual representation.

Increasingly, civil society actors (e.g., Linux Foundation, 2023) and supranational institutions (e.g., European Commission, 2023; WEF no date) are trying to shape the Metaverse. Physical reality and Web 2.0 are complemented by the possibilities of virtual immersion, particularly through extended reality technologies, 360° environments, and both abstract and photorealistic avatars. It remains unclear and controversial how the Metaverse will look, whether it will even bear that name, whether the desired level of interoperability between different platforms within the Metaverse can be achieved, and how 'social' the Metaverse will be—in the sense of allowing all users to create content equally.

Discussions about the Metaverse – oscillating between hope and concern, hype and ignorance, economy and society – are particularly notable in light of the insufficient knowledge and research available. To be able to shape the internet of tomorrow – regardless of what name it will carry or what it will look like – it is essential to consider key developments and analyze the opportunities and risks. In this regard, the following will introduce important terms, outline key technological backgrounds, and sketch some aspects of the relationship between the Metaverse and democratic culture[3]. The summary findings of this contribution are based on an analysis of academic literature, background conversations with experts from the digital economy, (social) sciences, and civil society, professional symposia, as well as on exploratory and participatory observations in various virtual immersive environments within the framework of the *Immersive Democracy*[4] project. Finally, challenges for (digital) civil society and democracy research will be identified.

3 **Democratic culture** refers to the totality of attitudes, values, and behaviors in a society that support and further develop a democratic order. A democratic culture is characterized by pluralism and the respect for human rights.

4 The contribution is based on research in the *Immersive Democracy* project, led by Matthias Quent, within the framework of the independent *European Metaverse Research Network* (EMRN). The EMRN was founded in 2022 through an unrestricted donation from Meta. Further studies on the topic, as well as an overview of partners and symposia, can be found on the website www.Metaverse-forschung.de.

Immersive Experiences

The Metaverse does not yet exist. The tech industry speculates that by around 2030, a fully operational Metaverse could become a reality. Financial motivations, in particular, are driving the development of the Metaverse forward: The consulting firm McKinsey predicts that business models related to the Metaverse could reach a value of 5 trillion dollars by 2030 – especially around e-commerce (McKinsey & Company, 2022). Many large companies are actively working to offer both physical and virtual goods within digitally immersive environments. With the increasing performance and decreasing costs of various internet-enabled devices that allow access to immersive environments (smartphones, tablets, PCs, and especially AR and VR headsets), the technological trendsetting by large companies like Apple, Microsoft, Nvidia, Intel, Google, and Meta, as well as the transfer of immersive gaming experiences from younger generations into other areas of life, it is expected that more and more people will use immersive virtual environments. Since Zuckerberg introduced the term 'Metaverse,' the number of scientific publications on the topic has been increasing. Since many developments are still taking place, definitions are only provisional and subject to change. Park and Kim (2022) propose the following understanding, based on Wikipedia:

> *"Metaverse is a compound word of transcendence meta and universe and refers to a three-dimensional virtual world where avatars engage in political, economic, social, and cultural activities. It is widely used in the sense of a virtual world based on daily life where both the real and the unreal coexist" (p. 4221)*

The Metaverse is characterized by the idea of creating a permanently existing virtual universe in which a variety of decentralized immersive virtual environments are interconnected (interoperability). Users can, as avatars, engage in activities such as trading, playing, working, exercising, attending concerts, meeting friends, traveling, or participating in education. It is also possible to create one's own worlds and conduct election campaigns or demonstrations. A central distinguishing feature from Web 2.0 is the higher degree of immersion, meaning the stronger immersion into virtual environments. Already today, millions of people regularly navigate immersive virtual worlds, particularly through various gaming environments, which are considered central drivers of these developments.

The term *Metaverse* encompasses the totality of individual virtual immersive environments. The degree of immersion depends on a variety of factors, particularly:

- **Intensity and Quality**: Extended Reality: Augmented Reality, Mixed Reality, Virtual Reality, created by the respective technology and dependent on the performance of corresponding devices (e.g., VR headsets)
- **Realism/Intensity and Quality** of the representation of physical reality (pure fantasy world, abstract, photorealistic/digital twin)
- **Interactivity/Intensity and Quality** of interaction with the virtual world (no interaction to high interaction (Social VR))
- **(A)Synchronicity/degree of temporal presence** (Asynchronous, delayed, real-time presence)
- **•Sociality/degree of interaction and quality with other users** (none to audio-visual and haptic interactions)
- **Authenticity/degree of coherence and credibility of the immersive experience**

Relatively new application areas for immersive virtual experiences include, among others, trading virtual goods (primarily NFTs) and real-world products, virtual workplaces, sports, culture, education, therapy, tourism, urban planning, media, and Social Virtual Reality.

With the promise of more intense experiences in immersive environments, comes the risk that these experiences may not only be positive but can also—intended or not—have negative individual and societal consequences. Clearly, harassment and hate speech in the Metaverse can have particularly severe effects. In particular, real-time verbal communication and new non-verbal expression possibilities through avatars challenge existing (partially precarious) methods of dealing with archived toxic content in terms of regulation, law enforcement, and counter-speech. This applies both to law enforcement and the implementation of community standards, as well as to specific services aimed at supporting affected individuals.

State of Research

The founding and development of the Metaverse, as well as democracy-related questions, are explored in literature, particularly from the perspectives of ethics (Slater et al., 2020), responsibility and sustainability (Moro-Visconti, 2022), inclusion and diversity (Zallio & Clarkson, 2022), cli-

mate consequences (Palak et al., 2023), and regulation (Rosenberg, 2022). Dwivedi et al. (2023) use the term *"Darkverse"* to summarize the darker sides of the Metaverse – such as the threat to privacy, diminished reality, identity theft, invasive advertising, misinformation, propaganda, phishing, financial crime, terrorist activities, abuse, pornography, social inclusion, mental health, sexual harassment, and unintended negative consequences of the Metaverse. Some publications formulate critical positions and concerns about a new dimension of digital surveillance capitalism, raising the issue of data privacy (e.g., Bojic, 2022; Anderson & Rainie, 2022) or warning about scenarios of violent radicalization (Bajwa, 2022). Hine (2023) highlights critical challenges associated with content moderation in the Metaverse, particularly emphasizing the complex cross-border moderation conflicts arising from inconsistent international standards. The author argues that without clear global regulatory frameworks for acceptable content norms, the Metaverse may emerge as a significant new frontier for disputes over freedom of expression, amplifying existing tensions around digital governance and online speech.

The new density of data that can be collected through immersive technologies extends beyond haptic motion information, eye-tracking, microreactions in facial expressions, voice analysis, to camera-based capture of information about the physical spaces in which users are located. German-language publications addressing the Metaverse have also increased significantly since Mark Zuckerberg's announcement (e.g., Büchel & Klös, 2022). However, empirical reports, case studies, or analytical discussions of dimensions, social and democratic consequences of aspects of the Metaverse are still rare internationally, and empirical analyses of the impact on democratic culture are virtually nonexistent.

In an analysis for the Stiftung Zukunft Berlin and the Foundation Metaverse Europe, Hermann (2022) highlights the "Lock-In Effect" (p. 3) of a centralized Metaverse as a challenge. Negative effects already known from social media could be amplified in the Metaverse, particularly fake news and filter bubbles, hate speech and polarization, biases and discrimination, mental health, consumer and data privacy, surveillance, control, censorship, and targeted advertising (ibid., p. 4). A privately-run Metaverse appears problematic for fundamental rights, political public spheres, as well as democratic procedures and processes (ibid., p. 5). Therefore, long-term European structures and companies should be established, and democratic processes and regulations should be defined and adapted at the European level. Nehring (2023), in a policy paper for the Konrad Adenauer Foun-

dation, points to the necessity of regulation and law enforcement in the Metaverse and warns that disinformation in new virtual communication spaces will appear more realistic and intense, thus becoming even more dangerous—especially through deepfakes. He therefore recommends early media literacy regarding these new spaces.

Artificial intelligence represents key technology for the Metaverse in both industrial and social segments, such as for interactions with AI-supported avatars and environmental elements, personalized advertising, recognizing behavioral patterns, regulation, when it comes to combating hate messages and disinformation, and in automated worldbuilding. It is only with the support of AI that virtual environments can be designed on a large scale with a high degree of authenticity and realism, for example, using the Unreal graphics engine. Artificial Intelligence (AI) is rapidly shaping the development of the social consumer Metaverse, creating not only promising democratic potential but also significant risks for democratic culture (Quent, 2024).

Four key areas of AI application in immersive spaces are identified and analyzed: (1) content generation, including deepfake and misinformation scenarios; (2) moderation and regulation, examining the challenges of AI-driven content control; (3) avatars and interactions, highlighting risks of deception and manipulation through virtual personas; and (4) data collection and personalization, outlining potential abuses via micro-targeting and emotional profiling. Without strong ethical guidelines, transparency, and regulatory oversight, AI applications in the Metaverse could intensify political polarization, hate speech, and radicalization—posing a significant threat to democratic norms.

Journey Through the Metaverse

It is surprising to find a lack of ethnographic and democracy-related empirical studies in immersive virtual environments, given that there are multi-year experiences with individual technologies and application contexts. This is particularly true for the context of electronic games, gaming studies, and especially for Metaverse-like games such as *Second Life* (Boellstorf, 2015) or *Minecraft* (Nebel et al., 2015). The relatively well-researched sandbox game *Minecraft*, for example, was used by 140 million players monthly in 2021 (Bergert, 2021). Viral communities have formed around the game

on YouTube and Twitch. The game is also used in historical and political education, and with the "uncensored library" by Reporters Without Borders, users from different countries can access texts that are banned in their home country.[5]

The ADL (2022) notes that discrimination and far-right ideologies are increasing in online games – including in immersive environments like *Fortnite* or *Roblox*. For example, on *Roblox*, user-generated environments can not only enhance creativity, collaboration, or self-efficacy experiences but can also motivate group-based hostility and far-right extremism. This includes environments with social Darwinist quests such as running over homeless people. Also, Nazi concentration camps and (right-wing) terrorist attacks have been recreated and reenacted in the app, according to research by Prinz (2024), even though this violates the community standards.

Weimann and Dimant (2023) identify significant potential for terrorist exploitation of the Metaverse, characterizing it as a versatile toolbox for extremist activities. They outline how virtual immersive spaces might facilitate indoctrination and recruitment, the covert planning and coordination of attacks, sophisticated virtual training scenarios, and the dissemination of disinformation. Moreover, they express concerns about the financing of terrorism through virtual economies and cryptocurrencies within the Metaverse. To mitigate these risks, the authors stress the importance of fostering robust public-private partnerships (PPP) to establish comprehensive countermeasures.

Psychological effects of VR/XR technologies are well-researched, and these technologies have been successfully used for supporting psychotherapies for several years, particularly for treating anxiety disorders. The potential of Virtual Reality (VR), especially for the therapy of anxiety disorders and depression, promoting empathy through perspective-taking, and fostering participation and cultural inclusion, has also been explored (e.g., Herrera et al., 2018; van Loon et al., 2018). VR technologies can help develop empathy for marginalized social groups and reduce mechanisms of devaluation. For education in general, immersive learning offers new opportunities (e.g., Frehlich, 2020).

Hinduja and Patchin (2024) conducted an extensive quantitative study exploring risks and negative experiences among adolescents engaging with immersive virtual environments in the Metaverse. Drawing on survey data from a nationally representative sample of adolescents aged 13–17 in the

5 https://www.uncensoredlibrary.com/de [07.07.2023].

United States, the authors found that within one year, nearly half (44.1 %) of young users experienced hate speech or discriminatory slurs, while more than a third reported cyberbullying (37.6 %) and general harassment (35 %). Additionally, the study documented significant occurrences of trolling (43.3 %), malicious obstruction or restriction of movement within virtual spaces (31.6 %), threats of violence (29.5 %), doxing (18.2 %), catfishing (22.8 %), and exposure to unwanted sexual or violent content (20.8 %). The research identified important gender differences: while boys and girls faced similar levels of hate speech, harassment, and bullying, girls were significantly more likely to experience sexual harassment, grooming, or being targeted specifically due to their gender. Consequently, girls employed adaptive strategies such as selecting avatars less likely to attract harassment and utilizing platform-based protective tools to maintain distance from potentially abusive avatars. Hinduja and Patchin also emphasize the increasing importance of AI-based solutions, which are being deployed to automatically detect and mitigate toxic behaviors, underlining the relevance of algorithmic safeguards for youth safety in the Metaverse.

McIntosh and Allen (2024) explore the emerging challenges policymakers face concerning harassment in the Metaverse, emphasizing the critical need for governmental engagement in developing effective regulatory frameworks. They highlight that policymakers worldwide are actively evaluating whether existing laws adequately address the novel harms occurring within immersive virtual environments or if new legislative categories are required. In particular, the authors advocate recognizing a distinct category of harm associated with abusive behaviors and interactions uniquely enabled by the immersive and embodied nature of the Metaverse.

Dimensions of Democratic Culture in the Metaverse

Table 1 heuristically summarizes relevant questions about democratic culture, as well as opportunities and risks concerning the development of the Metaverse. Due to space constraints, the individual aspects cannot be elaborated in detail. Overall, on the individual level, there are many opportunities, especially through educational and perspective-shifting approaches, new freedoms in developing individual identities, virtual self-efficacy experiences, and the potential to gain recognition and status in the Metaverse. It is important to note that on the individual level, all three dimensions of the known digital divide are effective: 1) unequal access (e.g.,

when acquiring hardware/headsets), 2) unequal use (e.g., entertainment vs. education), and 3) unequal outcomes (e.g., socially valuable professional connections) (Matzat & Van Ingen, 2020).

On the micro and meso levels, the Metaverse offers enabling, experiential, and resonant spaces for interpersonal communication, small groups, as well as for orthodox (particularly those around the economic, socio-cultural, and political mainstream) and heterodox communities. A distinction should be made between positive forms of participation and what is known as dark participation (Quandt, 2018).

Level	**Questions in the Context of Democratic Culture**	**Opportunities**	**Risks (Darkverse)**
Individual Level (Avatars/Users)	**Users as consumers or individuals?** Rights of avatars, recognition and advancement, identity design, security, well-being, hate, manipulation, inclusion and diversity, participation opportunities, effects of immersion on individuals.	**Overcoming barriers and boundaries** (physical, psychological, social, cultural, economic, identity), engagement, participation, resonance/effectiveness experiences, perspective change, education and information, belonging, edutainment, self-efficacy.	Digital divide, data misuse, manipulation, isolation, disinformation, discrimination, hate, harassment and digital violence, desensitization and dehumanization, radicalization
Micro Level (Communities & Social Interactions)	How integrative and participatory are the communities?	Voting, activism, organization, (transformative) participation in platforms and society, solidarity	Dark participation/toxic and radicalizing communities, polarization, tribalism, and silo thinking
Meso Level (Immersive Environments with Various Technological Foundations)	How integrative and participatory are the communities?	(Political) education & edutainment, value communication, promotion of diversity, inclusion, and participation through design	Data misuse, digital divide, manipulation, structural discrimination, cyberattacks

Level	Questions in the Context of Democratic Culture	Opportunities	Risks (Darkverse)
Macro Level (Metaverse as a Universe of Various Connected Immersive Virtual Environments)	Ownership, governance, ethics, data protection, interoperability, design, regulation	(Political) education & edutainment, value communication, promotion of diversity, inclusion, and participation through design	Digital surveillance capitalism, monopolization and control of platforms, cyber attacks, unequal distribution of wealth and power, manipulation through disinformation, undermining of state order, loss of social and regional ties, rise of populist and nationalist counter-reactions, new fears and fear narratives (e.g., in connection with transhumanism), loss of shared reality

Table 1: Questions of Democratic Culture in the Metaverse at the Individual, Macro, Meso, and Micro Levels

While democratic participation includes aspects such as promoting engagement, empathy, solidarity, knowledge, critical public discourse, and (co-)creative collaboration, dark participation, as described by Kowert (2020) in the gaming context, particularly includes hate speech, (sexual) harassment, trolling, griefing, doxxing, fake news, cheating, trash talking, contrary play, and inappropriate role-playing. The latter aspects pose unique challenges to avatar-based environments that distinguish them from traditional social media.

On a macro level, the risks to democracy, as identified in literature and expert discussions, are particularly striking. A significant issue is the one-sided control and decision-making power of these technologies by globally operating private companies. Concerns include data protection, the dissolution of social communities and shared realities, and new possibilities for manipulation—especially in connection with artificial intelligence.

Furthermore, nationalist and conspiracy-theorist actors criticize the increasing fusion of real and virtual realities under the term "posthumanism", claiming that global liberal elites aim to gain total control over humanity and ultimately destroy the essence of being human (Dilger, 2022). Technological change may therefore lead to further political polarization, particularly between (right-wing) nationalist and populist movements and the socio-cultural and technological development of globalized capitalism.

The Metaverse is likely to accelerate individualization and the formation of sub-communities within digital environments. This is evident in the nearly unlimited possibilities for avatar customization and the formation of subcultural peer groups. Companies, through the design of virtual environments, can shape frameworks for defining normality, while democratically legitimized national and supranational institutions may lose significance and influence. The potential profits of corporations stand against unpredictable social consequences.

In this schematic comparison, it is important to recognize the interdependencies between the levels—the dynamics and directions of these interactions require future research. Communities and individual content creators can actively shape the development of immersive digital environments—even through protests or strikes. Through co-creative collaboration, social virtual environments are in a constant state of change and expansion, for example, through modifications or updates.

Approaches towards participation are visible but occasionally appear precarious or even counterproductive. Beyond the risk of pseudo-participatory processes, the design of participation mechanisms in immersive digital environments can also reinforce undemocratic, plutocratic developments. For instance, on the platform Decentraland, users can vote on the development and governance of this virtual world. However, voting power depends on the amount of digital currency (MANA tokens) a user holds (Decentraland, n.d.). This means that wealthier users have more influence, violating the democratic principle of "one person, one vote", a fundamental aspect of electoral equality. This structure mirrors dystopian visions such as those depicted in Snow Crash.

To counter this, Shapiro and Talmon (2022) propose a theoretical framework and architectural blueprint for a "grassroots democratic Metaverse," conceptualized as an interconnected network of autonomous digital communities that are collectively owned, operated, and governed by users themselves. Central to their approach is ensuring democratic equality through mechanisms designed to resist "sybil attacks"—the deceptive

creation of multiple false identities aimed at distorting democratic processes. The authors contrast their proposed democratic Decentralized Autonomous Organizations (DAOs) with existing, often plutocratic, DAO models. They advocate for DAOs built around digitally verified identities, democratic rather than monetary governance (rejecting the principle of "one coin, one vote"), and transparent, user-driven constitutional processes capable of democratically adapting every component, including underlying protocols. Their vision emphasizes fostering digital social movements, cooperatives, and political entities grounded in authentic democratic participation, highlighting a roadmap for a resilient, equitable Metaverse. On the other hand, Ebner (2024) sees dangers from extremist use of DAOs.

The Metaverse could even reduce political polarization: Shelley, Schmidt, and Ette (2025) discuss the potential role of the Metaverse in addressing political polarization, drawing particularly on the arguments proposed by Bruno Maçães. According to this perspective, immersive and compelling experiences within virtual environments may enable individuals to explore and live out their ideological beliefs without imposing these views on others in the physical world. The authors suggest that by facilitating diverse and personalized realities, the Metaverse could decrease individuals' drive to enforce a shared version of reality, potentially leading to more peaceful coexistence and reduced societal polarization.

Discussion

The overview shows that the emergence of the Metaverse may amplify existing challenges to democratic culture yet simultaneously reveals significant opportunities for fostering participation and democratic engagement. While individual and interpersonal levels offer substantial potential for identity development, empowerment, and overcoming traditional social barriers, critical risks become particularly evident at the societal and macro levels. Notably, the dominance of major technology corporations in controlling both infrastructural and normative frameworks within the Metaverse poses substantial challenges to democratic principles, emphasizing the urgent need for effective governance and regulation.

At the same time, core concepts such as "democratic culture" require further theoretical clarification and precise operationalization for the Metaverse. The lack of robust theoretical grounding generates ambiguity, limiting analytical precision and practical applicability. Moreover, despite

references to dystopian narratives and existing literature, there remains a notable gap in empirically grounded research. Specifically, systematic investigations into how social inequalities might be reproduced or transformed through immersive digital spaces remain sparse. Empirical scrutiny is equally lacking regarding the actual practices and mechanisms of political participation within virtual environments. Therefore, future research must prioritize methodologically robust empirical inquiries to clarify how digital inequalities persist or evolve and to empirically measure democratic engagement in immersive contexts. Interdisciplinary collaboration among researchers, policymakers, and civil society actors is necessary to develop comprehensive insights and effective strategies that safeguard democratic structures and maximize positive potentials in shaping the future of safe immersive digital environments.

References

ADL Center for Technology and Society. (2022). *Hate is no game: Hate and harassment in online games 2022.* Anti-Defamation League. https://www.adl.org/sites/default/files/documents/2022-12/Hate-and-Harassment-in-Online-Games-120622-v2.pdf

Anderson, J., & Rainie, L. (2022). *The metaverse in 2040.* Pew Research Center. https://www.pewresearch.org/internet/2022/06/30/the-Metaverse-in-2040/

Bajwa, A. (2022). Malevolent creativity & the Metaverse: How the immersive properties of the Metaverse may facilitate the spread of a mass shooter's culture. *The Journal of Intelligence, Conflict, and Warfare, 5*(2), 32–52. https://doi.org/10.21810/jicw.v5i2.5038

Bergert, D. (2021). *Minecraft zählt mittlerweile 140 Millionen Nutzer monatlich.* PC-Welt. https://www.pcwelt.de/article/1195175/minecraft-zaehlt-mittlerweile-140-mio-nutzer-monatlich.html

Boellstorff, T. (2015). *Coming of age in Second Life: An anthropologist explores the virtually human* (2nd ed.). Princeton University Press. https://doi.org/10.2307/j.ctvc77h1s

Bojic, L. (2022). Metaverse through the prism of power and addiction: What will happen when the virtual world becomes more attractive than reality? *European Journal of Futures Research, 10*(1), Article 22. https://doi.org/10.1186/s40309-022-00208-4

Büchel, J., & Klös, H.-P. (2022). *Metaverse: Hype or "next big thing"?* Institute of the German Economy Cologne e.V. https://www.iwkoeln.de/studien/jan-buechel-hans-peter-kloes-hype-oder-next-big-thing.html

Decentraland. (n.d.). *DAO user guide.* Retrieved June 28, 2023, from https://docs.decentraland.org/player/general/dao/dao-userguide/

Dilger, F. (2022). *#Faktenfuchs: Transhumanists are not planning a global conspiracy.* Bayerischer Rundfunk. https://www.br.de/nachrichten/deutschland-welt/faktenfuchs-transhumanisten-planen-keine-globale-verschwoerung,TQBAPEm

Dwivedi, Y. K., Kshetri, N., & Hughes, L. (2023). Exploring the Darkverse: A multi-perspective analysis of the negative societal impacts of the Metaverse. *Information Systems Frontiers.* Advance online publication. https://doi.org/10.1007/s10796-023-10400-x

Ebner, J. (2024). *Decentralized Autonomous Organizations (DAOs) in the Metaverse – Opportunities and risks in the context of democratic culture.* Project Immersive Democracy. Retrieved March 12, 2025, from https://www.metaverse-forschung.de/en/2023/09/06/decentralized-autonomous-organizations-daos-in-the-Metaverse-opportunities-and-risks-in-the-context-of-democratic-culture

European Commission. (2022). *Call: Culture, the arts and cultural spaces for democratic participation and political expression, online and offline.* Horizon Europe Framework Programme. https://ec.europa.eu/info/funding-tenders/opportunities/docs/2021-2027/horizon/wp-call/2023-2024/wp-5-culture-creativity-and-inclusive-society_horizon-2023-2024_en.pdf

Frehlich, C. (2020). *Immersive learning: A practical guide to virtual reality's superpowers in education.* Rowman & Littlefield.

Hermann, I. (2022). *Democratic values in the Metaverse from a European perspective.* Foundation Metaverse Europe.

Herrera, F., Bailenson, J., Weisz, E., Ogle, E., & Zaki, J. (2018). Building long-term empathy: A large-scale comparison of traditional and virtual reality perspective-taking. *PLOS ONE, 13*(10), e0204494. https://doi.org/10.1371/journal.pone.0204494

Hinduja, S., & Patchin, J. W. (2024). Metaverse risks and harms among US youth: Experiences, gender differences, and prevention and response measures. *New Media & Society.* Advance online publication. https://doi.org/10.1177/14614448241284413

Hine, E. (2023). Content moderation in the Metaverse could be a new frontier to attack freedom of expression. *Philosophy & Technology, 36*(3), Article 43. https://doi.org/10.1007/s13347-023-00645-4

Kowert, R. (2020). Dark participation in games. *Frontiers in Psychology, 11,* Article 598947. https://doi.org/10.3389/fpsyg.2020.598947

Linux Foundation. (2023). *Linux Foundation announces launch of the Open Metaverse Foundation.* https://www.linuxfoundation.org/press/open-Metaverse-foundation

Matzat, U., & Van Ingen, E. (2020). Social inequality and the digital transformation of Western society. In S. Maasen & J.-H. Passoth (Eds.), *Soziologie des Digitalen—Digitale Soziologie?* (pp. 379–397). Nomos. https://doi.org/10.5771/9783845295008-379

McIntosh, V., & Allen, C. (2024). What do policymakers need to know about harassment in the Metaverse? *Frontiers in Virtual Reality, 5,* Article 1443384. https://doi.org/10.3389/frvir.2024.1443384

McKinsey & Company. (2022). *Value creation in the Metaverse: The real business of the virtual world.* https://www.mckinsey.com/capabilities/growth-marketing-and-sales/our-insights/value-creation-in-the-Metaverse

Moro-Visconti, R. (2022). ESG-compliant Metaverse ecosystems: Clues for impact investing. *SSRN.* https://doi.org/10.2139/ssrn.4193294

Nebel, S., Schneider, S., & Rey, G. D. (2016). Mining learning and crafting scientific experiments: A literature review on the use of Minecraft in education and research. *Journal of Educational Technology & Society, 19*(2), 355–366.

Nehring, C. (2023). *Manipulation and disinformation in the Metaverse.* Konrad-Adenauer-Stiftung.

Palak, S., Gulia, P., Gill, N. S., & Chatterjee, J. M. (2023). Metaverse and its impact on climate change. In A. E. Hassanien, A. Darwish, & M. Torky (Eds.), *The future of Metaverse in the virtual era and physical world* (pp. 211–222). Springer. https://doi.org/10.1007/978-3-031-29132-6_12

Park, J., & Kim, N. L. (2022). From Metaverse to the real world: The role of avatar identification in consumer's virtual purchasing behavior. *Innovate to Elevate*, 4218–4222. Iowa State University Digital Press. https://doi.org/10.31274/itaa.15732

Prinz, M. (2024). *Unpixelated hate: Lessons from VR gaming for a digital civil society.* Metaverse Research Lab. Project Immersive Democracy. Retrieved March 12, 2025, from https://www.metaverse-forschung.de/en/2024/02/22/unpixelated-hate-lessons-from-vr-gaming-for-a-digital-civil-society/

Quandt, T. (2018). Dark participation. *Media and Communication, 6*(4), 36–48. https://doi.org/10.17645/mac.v6i4.1519

Quent, M. (2023). Democratic culture and the next internet: Opportunities and risks of virtual immersive experience spaces in the Metaverse. In Institute for Democracy and Civil Society (Ed.), *Knowledge creates democracy: Focus on internet cultures and platform politics* (Vol. 14, pp. 30–43). Jena.

Rosenberg, L. (2022). Regulation of the Metaverse: A roadmap—The risks and regulatory solutions for large-scale consumer platforms. In *Proceedings of the 6th International Conference on Virtual and Augmented Reality Simulations*(pp. 21–26). ACM Digital Library. https://doi.org/10.1145/3546607.3546611

Shapiro, E., & Talmon, N. (2022). Foundations for grassroots democratic Metaverse. *arXiv.* https://doi.org/10.48550/arXiv.2203.04090

Shelley, T., Schmidt, E. W., & Ette, F. (2025). Can the Metaverse remedy political polarization? *Society.* Advance online publication. https://doi.org/10.1007/s12115-024-01056-7

Slater, M., Gonzalez-Liencres, C., Haggard, P., Vinkers, C., Gregory-Clarke, R., Jelley, S., Watson, Z., Breen, G., Schwarz, R., Steptoe, W., Szostak, D., Halan, S., Fox, D., & Silver, J. (2020). The ethics of realism in virtual and augmented reality. *Frontiers in Virtual Reality, 1,* Article 1. https://doi.org/10.3389/frvir.2020.00001

Stephenson, N. (2021). *Snow Crash.* Fischer.

van Loon, A., Bailenson, J., Zaki, J., Bostick, J., & Willer, R. (2018). Virtual reality perspective-taking increases cognitive empathy for specific others. *PLOS ONE, 13*(8), e0202442. https://doi.org/10.1371/journal.pone.0202442

World Economic Forum (WEF). (n.d.). *The Metaverse.* Retrieved May 7, 2023, from https://www.weforum.org/topics/the-Metaverse

Zallio, M., & Clarkson, P. J. (2022). Designing the Metaverse: A study on inclusion, diversity, equity, accessibility, and safety for digital immersive environments. *Telematics and Informatics, 75,* Article 101909. https://doi.org/10.1016/j.tele.2022.101909

A Theory of Immersive Democracy

Autonomy, Participation, and Agency in a Digitalized Society

Lenn Blaschke

"The miracle that saves the world, the realm of human affairs, from its normal, 'natural' ruin is ultimately the fact of natality, in which the faculty of action is ontologically rooted. It is, in other words, the birth of new men and the new beginning, the action they are capable of by virtue of being born."

Hannah Arendt (Vita activa, 1958)

A Theory of Immersive Democracy[1]

Immersive Democracy refers to a political and social practice that does not understand democratic participation solely as a rational-discursive process but explicitly integrates the affective, embodied, and existential dimensions of participation. This perspective builds on Hannah Arendt's concept of *natality* — the human capacity to bring something fundamentally new into the world.

According to Arendt, democracy should be seen less as a static condition and more as a dynamic space of political action, where people experience themselves collectively and create political realities through joint action. Immersive Democracy can also be understood as a space of collective beginnings, where citizens actively engage in political processes through sensory-emotional experiences, technological mediation, and symbolic orders.

The symbolic order, a concept from Jacques Lacan's psychoanalysis, offers an important framework for understanding how political subjectivation might occur within the context of Immersive Democracy. Lacan describes the symbolic order as a dense network of language, norms, and social rules that structure human thought, emotion, and behavior. By entering

1 This text was created with the support of an AI language model (ChatGPT 4o) and subsequently revised by the author. Thanks to: Prof. Dr. Matthias Quent, Alina Mönig and *Radikale Töchter.*

this order — for instance, through language acquisition — the individual becomes embedded in social reality and begins to perceive themselve as part of broader social contexts (Widmer, 2018, p. 43).

In Immersive Democracy, this symbolic order is not understood as a rigid set of rules but rather as a dynamic structure that enables affective and symbolic experiences of political participation. Yet this is precisely where a key challenge emerges: How can it be ensured that the symbolic order, which mediates political experiences, does not function solely in an affirmative or manipulative way, but instead unfolds emancipatory potential?

Put differently: Can Lacan's symbolic order truly function democratically in digital spaces, or does it risk being deformed by algorithmic selection and capitalist interests?

Immersive Democracy goes beyond purely deliberative or representative models by emphasizing the active design of environments that affectively engage individuals and connect their being-in-the-world with political actions. While deliberative models understand political processes as being based on consultation and thorough consideration (Bundeszentrale für politische Bildung, n.d.), immersion in this context does not merely refer to immersion in virtual or media-based experience spaces. Rather, it signifies the active embodiment of democratic principles in everyday life, allowing political action to be experienced as a creative, processual enactment of *natality*.

In an era where digital technologies increasingly shape the perception of reality, Immersive Democracy raises the question of how affective mechanisms and immersion-enhancing structures can be employed not merely for manipulation but for emancipatory self-empowerment. By emphasizing the concept of *natality* — the human potential to change the world through the new — it offers a perspective that envisions democratic processes as open, emergent, and radically participatory spaces of experience.

1 The Foundation: Immersion and Democracy as Lived Reality

The term *immersion* originates from the Latin word *immergere*, meaning "to dive in" or "to be embedded." In its original sense, it refers to the physical act of submersion in a substance — such as diving into water or during ritual baptisms. In a figurative sense, however, *immersion* describes not only a physical state but also a mental, emotional, or social absorption/embeddedness in a particular environment, reality, or structure.

In media and cultural studies, *immersion* is described as a state in which individuals become so deeply engaged in an environment or experience that the boundary between their own self and external reality becomes blurred. This can be triggered by narrative fiction, audiovisual media, performative art, or social dynamics. In immersive environments, the perceiving individual no longer feels like a detached observer but rather as an active part of the unfolding events — whether in a novel, a theatrical performance, a virtual reality experience, or a social movement (Schütz, 2015, p. 7).

This definition of *immersion* highlights that individuals can be involved in experiences not only cognitively but also affectively and physically. People experience reality not solely through abstract reflection but also through an embodied being-in-the-world shaped by emotions, sensory perceptions, and social interactions. Applying immersion to democratic practice opens the possibility of understanding political participation not exclusively as a rational-deliberative or representative process, but as a holistic, affective, and embodied reality.

An Immersive Democracy would thus not merely be a space for debate but, above all, a space of sensory-emotional, collective experience in which democratic self-empowerment becomes tangibly perceptible. At the same time, however, this expansion reveals fundamental challenges — particularly the question of whether affective involvement genuinely promotes democratic emancipation or rather heightens the risk of affective manipulation.

Especially in digital environments, which are heavily shaped by commercial and technological infrastructures, there is a risk that immersive experience spaces may not provide the anticipated autonomy but instead create new dependencies. Consequently, for Immersive Democracy, the critical question arises regarding the independence and freedom of the platforms on which immersion takes place, and how these spaces must be consciously and critically designed to unfold democratic potential without succumbing to manipulative practices.

2 Connection to Natality in Hannah Arendt's Thought

In *The Human Condition* (1958), Hannah Arendt introduces the concept of *natality* as a fundamental characteristic of human existence. While philosophers — from Plato to Heidegger — have often emphasized mortality as the

central feature of human existence in ontology, focusing on the awareness of finitude and life's orientation toward death, Arendt offers an alternative perspective: Human existence is not primarily defined by its mortality but by the potential to initiate something new.

This capacity is linked to the fact of birth — every human life begins with a unique entry into the world and carries with it the possibility of creating something original (Arendt, 2020, pp. 25–26).

Natality thus represents an expression of existential openness that distinguishes the human condition. Humans are not only born into existing structures but also possess the ability to change these structures through their actions. For Arendt, this potential to bring something new into the world constitutes the essence of political action.

While labor and work in Arendt's thought are often characterized by repetition and purpose-driven activity, *action* is the realm in which individuals reveal themselves as unique beings and, through their interaction with others, contribute to shaping a shared world (Arendt, p. 23–24).

Natality is therefore not merely a biological concept but also a political category: it points to the possibility of creating a world that is not defined solely by repetition, but by unpredictable, creative, and collective action. For Arendt, politics is thus not merely the administration of what already exists but rather a space of appearance, where individuals constitute themselves through action and actively shape the conditions of their world.

Politics as a Space of Appearance for New Beginnings

For Arendt, democracy is not merely an institutional order based on representation and legality but rather a vibrant space of political practice. She criticizes modern mass societies and bureaucracies for restricting opportunities for political action and reducing people to passive spectators rather than recognizing them as active co-creators of their shared world (Arendt, 1974, p. 198).

A vibrant democracy must therefore be more than a system of elections, rules, and procedures — it must be a space where political *natality* can genuinely be lived. This occurs not merely through casting votes or managing what already exists, but through collective, public action that creates new realities. The world is not simply "administered"; rather, it is continuously renewed through human action (Arendt, 2020, pp. 42ff).

Here, the close connection between *natality* and democracy becomes evident: Democracy is not a stable order but an ongoing process of emergence and becoming. It is not a finalized state but a field of possibility in which people continuously discover new political forms and modes of expression.

However, this raises an important theoretical question that Arendt herself does not directly address: How can this existential openness and the creative potential of new beginnings be meaningfully translated into digital, immersive environments? This presents a potential tension, as Arendt primarily conceived political spaces as physical, real spaces of encounter, where individuals experience and engage with one another in embodied presence.

Transposing this concept to digital environments therefore raises the question of whether and to what extent the concept of *natality* can be authentically realized in digital contexts, or whether immersive democratic spaces may instead represent a mere simulation of new beginnings — one that, rather than opening up new possibilities for action, may in fact prove alienating.

It thus remains uncertain whether Arendt's concept can be straightforwardly transferred to immersive practice spaces, or whether this very translation into digital and affectively designed platforms requires a critical reassessment.

3 Affects, Affectivity, and Immersive Power

Affects as Political Forces: Expanding the Rational Citizen Concept

A central aspect of Immersive Democracy is its affectivity. While traditional democratic theories often assume a rational image of the citizen — for example, in John Rawls' liberal theories (1971), which conceive justice as a rationally calculated principle, or in Jürgen Habermas' deliberative democracy (1981), which emphasizes democratic decision-making primarily through rational argumentation and discursive processes — practice clearly shows that political processes are not driven solely by rational reflection but are equally shaped by affective, emotional, and situational dynamics.

This classical conception of democracy is largely based on the assumption that political decisions are primarily the result of rational deliberations,

arguments, and discursive processes. Yet, as previously highlighted, political reality is often far more complex: Political movements, social protests, and mobilizations frequently arise not primarily from argumentative processes but from intense affective experiences.

Voting decisions and political engagement are often driven by emotional states and moods, such as a deep sense of injustice, outrage, fear, or hope (cf. Frevert, 2022). These affective dimensions of political practice are precisely what lie at the heart of Immersive Democracy, which acknowledges and shapes political processes as affective and physically experienced realities.

Immersive Democracy draws a key conclusion from this insight: it seeks to promote political participation not only as a rational-deliberative process but also by actively incorporating the affective and emotional dimensions of democratic engagement. Rather than viewing affective forces as irrational or disruptive, Immersive Democracy deliberately asks how democracy can be designed when affective involvement is recognized as a productive component of political processes.

However, this requires careful theoretical reflection on the role and potential of affects, as explored particularly in *Affect Theory*.

Affect Theory: Affects as Pre-Reflective Forces in Politics

The significance of affects in political processes can be explained through *Affect Theory*, which builds on the works of Baruch de Spinoza, Gilles Deleuze, and Brian Massumi. In this framework, affects are not merely individual emotions or subjective feelings; rather, they are interpersonal (trans-individual) and bodily forces that emerge prior to conscious reflection. They move between bodies and subjects before these impulses are processed cognitively or linguistically (Mühlhoff, 2018, pp. 14f).

Spinoza describes affects as the capacity of bodies to be affected and to affect others. Affects arise within relational dynamics and are not isolated states of closed-off individuals but expressions of the causal interconnectedness of all things. In social and political contexts, they can be amplified, altered, or channeled (Mühlhoff, 2018, pp. 20f).

Massumi emphasizes that affects manifest on a pre-reflective level — they often operate before conscious opinion formation and shape how people perceive and respond to political events (Massumi, 1995, pp. 83–109). An example of this is visual or narrative staging in politics, which

impacts audiences not primarily through arguments but through feelings of fear, anger, or hope. A populist politician, for instance, may not require a coherent political agenda if they succeed in mobilizing followers through affective stimulation, evoking states of outrage, pride, or fear (Massumi, 2010, pp. 105f).

This demonstrates that democracy inherently encompasses both discursive-rational and affective dimensions. Precisely this insight underscores the need to critically reflect on the role of affects within Immersive Democracy, taking into account both their potential and the possible risks they may entail.

4 The Role of Mediality and Digital Structures

From Physical to Digital Democracy: A Shift in Spaces of Experience

Historically, democracy has often been understood as a space of collective, rational, and physically experienced encounters. From the ancient Greek *agora* — a village gathering place where festivals, assemblies, and markets took place, serving as a central institution of the *polis* (Höcker, 2008, pp. 2–4) — to modern parliaments or street demonstrations, political action has always been closely tied to embodied co-presence.

In an increasingly digitalized world, however, this democratic space is undergoing a fundamental transformation: Political participation is shifting more and more into virtual, algorithmically structured spaces, which has direct implications for how people perceive, experience, and shape political reality.

Digital media offer new opportunities for expanded and location-independent participation; at the same time, however, they carry the risk of strategically channeling affects, amplifying opinions through algorithms, or even rendering critical discourses invisible. This presents a key challenge for Immersive Democracy: How can digital spaces be designed to enable affective engagement without manipulating political processes or endangering democratic autonomy through algorithmic or commercial interests?

In particular, this raises the question of whether the affective and resonant dimension that is evident in physical political action can be authentically experienced on digital platforms — or whether what emerges is instead a simulation that ultimately limits or distorts political participation.

Digital Technologies as Mechanisms for Shaping Political Experience

The proliferation of digital technologies such as Virtual Reality (VR), Augmented Reality (AR), social media, and algorithmically curated platforms not only expands political spaces of experience but also significantly shapes and influences them. This development produces two parallel and interwoven effects:

On the one hand, digital platforms enable new forms of political participation by allowing people to connect across national borders, engage in deliberative processes, and collaboratively shape political decisions in virtual spaces.

On the other hand, these new political spaces are subject to control and regulation that is far from neutral; instead, they are shaped by platforms, algorithms, and economic interests. In her analysis of *surveillance capitalism*, Shoshana Zuboff (2019) has demonstrated that digital technologies are not merely neutral communication tools. Rather, through data extraction, algorithmic control, and personalized manipulation, they actively steer affects and perceptions in targeted ways (Zuboff, 2019).

Platform capitalism has thus created a new form of political power: While classical democratic systems were based on deliberative negotiation and public discourse, today's political debates are increasingly shaped by algorithmic selection and commercial interests. Who sees which information, which topics are made visible, and which remain hidden is no longer determined solely by political institutions but is now heavily influenced by technological infrastructures.

This development leads to a paradoxical situation: On the one hand, digitalization has indeed made democratic processes more accessible, interactive, and diverse. On the other hand, these very processes are often invisibly influenced by commercial, political, or ideological interests, potentially endangering fundamental democratic values such as autonomy and transparency.

A striking example of digital technologies' influence on political processes is the *Facebook-Cambridge Analytica data scandal.* In 2018, it was revealed that the British consulting firm Cambridge Analytica had, without users' knowledge or consent, harvested personal data from up to 87 million Facebook profiles. These data were used to analyze voter behavior and target political advertisements, notably during the 2016 U.S. presidential election and the Brexit referendum (AP News, 2025).

This incident highlights how digital platforms and the data they collect can be used to influence political decision-making processes, often without the knowledge or consent of the individuals affected. It raises critical questions about transparency, ethical responsibility, and the regulation of digital technologies in political contexts.

Despite the revelations and ensuing criticism, similar practices have persisted, underscoring the need for stricter data protection laws and greater oversight of the use of personal data in political campaigns (Wikipedia, 2025).

Immersive Democracy as an Alternative: Designing Digital Spaces Democratically

Immersive Democracy raises the question of how digital structures can be employed not for control or passive consumption but as active democratic spaces of experience. The aim is to understand digital technologies not merely as channels for political communication but as spaces of experience and *spaces of appearance*, where people can engage with democracy affectively, bodily, and interactively. This approach envisions digital environments that foster democratic engagement as something that can be felt, embodied, and collectively shaped.

One possible strategy for achieving this lies in the creation of interactive citizen forums that are designed not to generate algorithmic attention but to enable genuine political participation. Such platforms open up new forms of democratic decision-making by encouraging deeper political engagement and collective deliberation. Rather than relying on social media platforms shaped by engagement algorithms and clickbait logic, these forums provide dedicated spaces that focus on meaningful discourse and thoughtful exchange. Examples of such initiatives can be found in platforms like *Decidim*, an open-source tool developed by the city of Barcelona to involve citizens in political decision-making processes (Decidim, n.d.); *vTaiwan*, a digital platform used by the Taiwanese government to conduct public debates on legislative proposals (vTaiwan, n.d.); and *LiquidFeedback*, a deliberative online tool employed by the *Piraten Partei* in Germany that applies principles of liquid democracy (LiquidFeedback, n.d.). Each of these platforms illustrates how digital spaces can be intentionally designed to foster political participation in ways that are immersive, affective, and

interactive, rather than reinforcing passive consumption or manipulative practices.

Immersive Simulations: Experiencing Political Processes through Narrative and Interactive Design

Immersive Democracy explores how political processes can be made tangible through narrative and interactive simulations. Rather than conveying political education solely through texts or debates, immersive simulations invite people into dynamic experiential spaces where they can engage with different perspectives — whether through role-playing, interactive storytelling formats, or multisensory stagings.

Examples of this approach include *Democracy* (Positech Games, n.d.), a political simulation game in which players assume the role of a head of government and must make political decisions, and *100 % City* (Rimini Protokoll, n.d.) by the research and theater collective Rimini Protokoll. Using documentary methods, Rimini Protokoll makes political structures tangible by bringing 100 representative citizens on stage to demonstrate democracy in action. Such formats foster empathy, reflection, and active political participation by engaging individuals not just intellectually but also emotionally and physically.

Participatory Digital Narratives: Engaging Citizens in Shaping Political Processes

Participatory digital narratives empower citizens to actively shape political processes rather than merely consuming them passively. Through interactive digital storytelling, complex social issues can be experienced from multiple perspectives, fostering deeper understanding and encouraging stronger engagement. An example of this approach is the *INDCOR* project, which explores how interactive digital narratives can be used to address societal challenges such as racism, war, and disinformation (INDCOR, n.d.).

The goal of these approaches is to extend the affective and embodied dimension of democracy into (digital) spaces. While traditional political participation often relies on rational argumentation and text-based communication, immersive (digital) environments can convey political experience through bodily involvement, emotional resonance, and symbolic spaces for action. By engaging people on multiple sensory and emotional levels, these formats aim to deepen participation and foster a more experiential understanding of democracy.

Despite the potential of these immersive democratic approaches, some fundamental limitations and risks must be critically considered. In partic-

ular, the issue of the *digital divide* poses a significant challenge: Not all citizens have equal access to the necessary technologies or possess sufficient digital literacy to effectively participate in immersive formats of political engagement. As a result, rather than reducing existing social and economic inequalities, these disparities could potentially be reinforced.

Beyond issues of accessibility, a critical challenge lies in the question of the actual autonomy and independence of the platforms being used. Even open-source projects like Decidim are potentially reliant on technological infrastructures that may themselves be subject to commercial or political interests. The very possibility of *natality* in the sense described by Hannah Arendt — the freedom to initiate something new and radically reshape political spaces — could be restricted in digital environments if platforms and technologies are influenced by external actors or if algorithmic mechanisms subtly regulate user behavior.

Another significant limitation concerns the sustainability and long-term impact of immersive political experiences. While immersive simulations and participatory narratives can indeed stimulate empathy and political interest in the short term, it remains uncertain whether they can consistently strengthen political action over time or whether they risk becoming temporary, affective experiences without lasting engagement. These aspects require ongoing reflection and critical examination to ensure that Immersive Democracy does not become a superficial experience but instead fosters enduring democratic self-empowerment.

5 The Transition from Subject to Community: Socio-Psychological Aspects

The Political Subject Between Autonomy and Embeddedness

A central aspect of Immersive Democracy is the question of how the political subject is shaped and how it moves within a collective political space of experience. While classical democratic theories often conceive of the subject as a rational, autonomous entity, social-psychological and philosophical theories demonstrate that identity and political action are always embedded in social, symbolic, and affective structures.

The French theorists Jacques Lacan and Michel Foucault have made significant contributions to understanding the subject not as a closed, sovereign unit, but as a product of social relations and discursive power structures.

In his psychoanalysis, Lacan describes the subject as a being that is always situated within a web of language, symbols, and imaginary self-images. The subject does not perceive itself as an isolated individual but always in relation to others — whether within the symbolic order of language or through affective relationships with an *Other*. Political subjectivity, therefore, is not simply given; rather, it emerges through a process of identification and differentiation within a symbolic structure (Pagel, 2012, pp. 30f).

Foucault, in turn, describes the subject as shaped by power relations and discursive structures. Political identities do not arise from individual rationality but are produced through the conditions established by social institutions, norms, and rules (Breite, 2023, pp. 33f). Subjectivity is thus not autonomous but is formed by social dispositifs and mechanisms of governmentality, which not only regulate individuals but also guide them toward self-governance (Breite, p. 35).

These perspectives are crucial for Immersive Democracy because they reveal that political participation is not solely a matter of rational will formation but is also shaped by affective, discursive, and symbolic practices. People participate in democracy not just as individuals but always as part of social spaces in which they are positioned through experiences, affects, and power relations.

Immersive Democracy as the Design of Resonance Spaces

If political identity is understood not as isolated autonomy but as a relational practice, Immersive Democracy must ask how political spaces can be designed to engage people not only rationally but also affectively and physically. Hartmut Rosa's concept of *resonance* offers a particularly valuable framework in this regard. Resonance describes a vibrant relationship with the world in which individuals do not merely absorb information but are affectively moved and experience themselves as capable agents within their surroundings. In this understanding, democratic practice can be designed in a way that citizens no longer perceive themselves merely as passive listeners or spectators of political processes but actively recognize their own involvement and agency (Rosa, 2019, pp. 362–380).

Immersive Democracy builds on this idea by asking how political experiences can shift from being passively consumed to being actively co-created. The focus here is particularly on designing political processes in ways that

are not only intellectually comprehensible but also sensually and emotionally accessible. The goal is to create symbolic and affective spaces that genuinely empower people to act politically. Resonance in this sense is not merely an emotional reaction but rather a dynamic relationship between the subject and the political environment — one in which political agency emerges. A space of political resonance enables individuals to perceive themselves not as passive recipients of political measures but as active co-creators, whose emotions, bodies, and perceptions are integral to political processes.

Politics as Embodied Experience

Immersive Democracy must therefore go beyond classical participation models and understand engagement not only as a cognitive act but explicitly as an embodied experience. This requires consistently conceptualizing political subjectivation as a process: individuals are not born as "ready-made" citizens but develop their political identity through experience, perception, and action. Democratic participation thus unfolds not only through the expression of opinions but also through symbolic actions, collective experiences, and affective involvement.

Political processes are therefore never purely rational reflections; they are always also embodied by presence, social interaction, and sensory experience, as seen in demonstrations, assemblies, protests, or performative political actions. These forms of political participation extend beyond mere argumentation and use the body as a medium of politics.

Ultimately, this means that political spaces should not be conceived solely as neutral arenas for discourse but intentionally designed as affectively charged experiential worlds in which people do not merely understand democracy but directly experience and live it. Only in such spaces can political power structures be immediately recognized, critically questioned, and actively transformed.

From Subject to Community – Immersive Democracy as Collective Practice

Immersive Democracy departs from the classical notion of the isolated, rationally acting political subject and instead conceives political identity as a

processual, affective, and socially embedded formation. In this understanding, democracy is not merely a space for purely rational decision-making processes but an embodied, interactive, and affectively experienced practice in which individuals directly perceive themselves as part of a community.

Central to this perspective is the transition from the individual subject to collective experience: democracy is not primarily a mechanism of individual will-formation but a *resonance space* in which people perceive themselves as active political agents through their relationships with others and collectively create political reality through affective, bodily, and symbolic practices.

Immersive Democracy thus opens up a perspective that extends significantly beyond deliberative and representative models of democracy. It envisions democracy not only as a discursive or institutionalized practice but as a vibrant field of communal experience that makes political participation tangible in affective, embodied, and collective ways.

6 An Example of Immersive Democracy: Radikale Töchter and Action Art

The preceding theoretical reflections on Immersive Democracy — particularly regarding *natality* (Arendt), the symbolic order (Lacan), and *resonance* (Rosa) — demonstrate that political participation encompasses far more than rational argumentation or institutionalized procedures. In order for political processes to enable genuine participation, they must be affectively and physically tangible. A concrete example that embodies these theoretical concepts can be found in the action art workshops of *Radikale Töchter* (Radikale Töchter, n.d.).

Radikale Töchter employ methods of action art to facilitate political participation within immersive experiential spaces. Their workshops consist of performative, artistic, and activist interventions in which participants not only discuss political situations theoretically but also experience them directly on an affective and bodily level. During the workshops, participants independently develop their own action art concepts on issues that are personally meaningful to them. These concepts are then collaboratively translated into staged political interventions, where participants address political crises, make collective decisions, or develop performative forms of protest. Through the independent and creative development of such immersive scenarios, participants gain direct insight into how power structures, social dynamics, and affects can shape and influence political realities.

A central element of the workshops is the intentional integration of digital media and spaces. Digital research plays a crucial role: participants explore their political topics not only through analog sources but also via digital archives, online research, and social media. This digital investigation is not merely supplementary but often essential to the substantive depth of their work. Here, a clear connection to Immersive Democracy becomes visible — a model that understands both analog and digital spaces as equal sites of political insight and participation.

Moreover, the workshops deliberately incorporate digital media such as smartphones, social media, AI-powered tools, and other digital platforms into their creative processes. These digital tools function as stages, means of communication, or symbolic elements within the performative enactments, enabling participants to convey their political messages in innovative ways.

A particularly innovative concept within the workshops is the method of "Digital Stages." In this approach, digital spaces are explicitly understood as arenas for political action art. Participants are encouraged to design their political interventions with digital publics in mind, thereby opening up new spaces for political participation. Many of the action art concepts developed in the *Radikale Töchter* workshops intentionally use digital platforms to generate visibility and mobilize political concerns.

This creative engagement with both digital and analog spaces illustrates the versatility with which Immersive Democracy can be shaped. By actively integrating digital media into immersive experiences, *Radikale Töchter* expand the traditional notion of action art to include a digital dimension, while simultaneously fostering initiative and self-efficacy among participants.

The approach of *Radikale Töchter* exemplifies what Immersive Democracy might look like in practice: through methods rooted in action art, they succeed in raising awareness for political processes, involving participants emotionally, and making it directly tangible that they themselves can be powerful agents within democracy. This suggests considerable democratic potential, as political experience is not merely consumed but actively co-created.

At the same time, it is important to critically reflect on whether and to what extent these workshops truly foster sustainable political change. While action art can create experiences that are immediate and emotionally powerful, it remains unclear how lasting and impactful the resulting political insights and impulses for action are. Can action art methods lead to long-term democratic subjectivation and resonance, or do these

experiences ultimately remain confined to the short-lived moment of performance? Furthermore, it is worth questioning whether affective and immersive strategies may unintentionally become manipulative — particularly when political content is primarily aestheticized and rendered consumable.

Radikale Töchter thus reveal both the potential and the limits of Immersive Democracy: they create vibrant spaces of political experience while also highlighting the challenges associated with affective, symbolic, and immersive methods. Especially the integration of digital spaces opens new opportunities for activating political engagement and strengthening participants' sense of empowerment — yet it also carries the risk that political processes may become flattened in the aesthetics of digital media and devolve into symbolic activism without long-term political impact.

7 Concluding Reflections and Discussion of Limitations

Immersive Democracy, by integrating affective, embodied, and symbolic experiences, offers a promising and expanded understanding of democratic practice—one that goes well beyond deliberative and representative models. By drawing on the human capacity for *natality* as conceptualized by Hannah Arendt and the idea of *resonance* as developed by Hartmut Rosa, it provides theoretical foundations for reimagining political spaces as radically participatory. It emphasizes that political participation should not be understood solely as a cognitive act but also as an affective and bodily experience.

At the same time, significant theoretical and practical challenges emerge that must not be underestimated. A central area of tension lies in the transferability of these theoretical concepts to digital and immersive spaces of experience. Rosa's concept of resonance, which describes a vibrant, responsive relationship with the world, may encounter limitations in digital contexts. The critical question remains whether immersive experiences can truly generate authentic resonance, or whether they risk producing a sense of alienation by simulating participation and agency rather than enabling them.

The appropriation of Jacques Lacan's concept of the symbolic order also raises critical questions: If political subjectivity is always embedded in social and discursive structures, one must ask to what extent digital, immersive structures can truly be autonomous and emancipatory — or

whether they inevitably reproduce symbolic orders that reinforce existing power relations and dependencies.

Another essential aspect concerns the technological and social infrastructure of immersive democratic forms. Digital platforms are often shaped by commercial and algorithmic logics, increasing the risk that political experiences may be affectively manipulated rather than fostering genuine participation. In light of existing social and digital divides, it also remains uncertain whether immersive democratic formats are truly accessible to all, or whether they might even exacerbate existing inequalities.

The action art workshops by *Radikale Töchter* vividly illustrate both the potential and the limitations of an immersive democratic practice. Their workshops, centered on performative and aesthetic interventions, make political participation directly tangible. However, it remains an open question whether — and how — they genuinely strengthen long-term political agency, or whether their impact is more temporary and episodic. This question marks a crucial task for future theoretical and empirical research.

It must also be acknowledged that *Radikale Töchter* represent just one example among many possible forms of immersive democratic practice. To develop a broader understanding of Immersive Democracy, it would be necessary to extend its principles with additional approaches and ideas — such as Philip Dingeldey's proposals for the establishment of a citizens' chamber (Dingeldey, 2025), or other concepts that emphasize autonomy, participation, and agency.

Ultimately, Immersive Democracy revolves around the core democratic goals of autonomy, participation, and empowerment. The various theoretical and practical challenges that emerge from this show that Immersive Democracy must not be understood as a fixed or finalized concept. Rather, it should be regarded as an open field of theoretical reflection and practical experimentation — one in which the affective, symbolic, and technological dimensions of democratic practice must be continuously reconfigured, critically examined, and empirically evaluated.

References

Arendt, H. (2020). *Vita activa oder Vom tätigen Leben*. Piper Verlag.

Arendt, H. (1974). *Über die Revolution*. Piper Verlag.

Breite, E. (2023). *Die Verfügbarmachung der Literatur: Neoliberale Logiken und ihr Einfluss auf literarische Bildung*. transcript Verlag.

Bundeszentrale für politische Bildung. (o.D.). *Deliberation*. https://www.bpb.de/kurz-knapp/lexika/das-europalexikon/176777/deliberation/

Decidim. (o.D.). *Open-Source-Plattform für deliberative Demokratie*. https://decidim.org

Dingeldey, P. (2025). Repräsentation und Partizipation zusammendenken: Bürger:innenkammer und parlamentarische Demokratie. *Bipar*. https://bipar.de/wp-content/uploads/2025/03/Praesentation-Dingeldey.pdf

Frevert, U. (o.D.). Wohldosierte Emotionen. *Max-Planck-Gesellschaft*. https://www.mpg.de/18101870/regieren-mit-gefuehl.

Höcker, C. (2008). *Metzler Lexikon der antiken Architektur*. Metzler.

INDCOR. (o.D.). *Interactive Digital Narratives for Conflict Resolution*. https://indcor.eu

LiquidFeedback. (o.D.). *Open Source Plattform für deliberative Demokratie*. https://liquidfeedback.com/de

Massumi, B. (2010). *Ontomacht – Kunst, Affekt und das Ergebnis des Politischen*. Merve Verlag.

Massumi, B. (1995). The autonomy of affect. *Cultural Critique*, 31, 83–109.

Mühlhoff, R. (2018). *Immersive Macht – Affekttheorie nach Spinoza und Foucault*. Campus Verlag.

Radikale Töchter. (o.D.). *Radikale Töchter*. https://radikaletoechter.de/

Rimini Protokoll. (o.D.). *100 % Stadt*. https://www.rimini-protokoll.de/website/de/projects/100-stadt-7-1

Rosa, H. (2019). *Resonanz – Eine Soziologie der Weltbeziehung*. Suhrkamp.

Pagel, G. (2012). *Jacques Lacan – zur Einführung*. Junius.

Positech Games. (o.D.). *Democracy 4*. http://www.positech.co.uk/democracy4.

Schütz, T. (2015). *Gestaltete Erfahrungsräume und die Teilhabe an der „totalen, radikalen Fiktion": Dimensionen immersiver Erfahrung im Gegenwartstheater am Beispiel von „Ventestedet" von SIGNA und „John Gabriel Borkman" von Vinge/Müller*. https://signa.dk/files/editor/Bibliography/MA_Skript_Schutz.pdf

vTaiwan. (o.D.). *Digital deliberation platform*. https://info.vtaiwan.tw

Widmer, P. (2018). *Subversion des Begehrens*. Turia + Kant Verlag.

Wikipedia. (o.D.). *Facebook – Cambridge Analytica data scandal*. https://en.wikipedia.org/wiki/Facebook%E2%80%93Cambridge_Analytica_data_scandal

Zuboff, S. (2019). *Das Zeitalter des Überwachungskapitalismus: Der Kampf für eine menschliche Zukunft im neuen Zeitalter der Macht.* Campus Verlag.

A Safe Space for Everyone – A Plea for a Democratic and Participative Metaverse

Octavia Madeira & Georg Plattner

Malevolent Actors in the Metaverse

The vision of a metaverse presented by Meta CEO Mark Zuckerberg in October 2021 was a watershed moment for society and for the tech world. Although the concept of a second digital life, including a digital identity, is neither new nor exclusive to Meta (see, for example, Second Life), this was the first time that almost all the possible functionalities of the metaverse had been presented up to that point. Here, there is a special emphasis on immersion via virtual reality which, as an extension of today's Internet applications, is meant to give users a completely new sense of participation and let them experience the metaverse in a multimodal and multi-sensory way. The presentation of the vision also focused on technological permeability, on the diffusion of social media in all areas of human life and therefore on the displacement of social media as a purely entertainment platform.

Should the metaverse turn out to be as Zuckerberg and other proponents envision it, this would mean a radical transformation of social interaction with the digital space, and also a radical change in our everyday lives. Shopping could increasingly shift to the metaverse as an immersive experience, sports classes could take place in a virtual environment and virtual church services could be held with believers from all over the world. The world of work has already permanently changed, partly due to the coronavirus pandemic – and we could soon move from working from home to working in the meta-office.

But these innovations will not only change our everyday lives – they will also cause extremism and radicalisation to strike out in new directions and transform to adapt to new environments. Extremists use technologies that are cheap, readily available, easy to use and widely accessible for their purposes, like propaganda, communication and recruitment. Using technology for a function other than that intended by the developers with the intention of doing harm to others is an inherently creative process.

Cropley, Kaufman and Cropley (2008) call this "malevolent creativity". They define it as a form of creativity that "is deemed necessary by some society, group, or individual to fulfil goals they regard as desirable, but has serious negative consequences for some other group, these negative consequences being fully intended by the first group" (p. 106). We describe actors who display malevolent creativity (such as extremists or spreaders of fake news) as malevolent actors.

In the past, malevolent actors were very creative especially when it came to realigning their own organisation and distributing their own ideology. The digital revolution has equipped them with an unprecedented number of tools with which to further their cause: from (encrypted and instant) mass communication for propaganda and recruitment to alternative instruments for financing operations and logistics through to new means of destruction and terror. Recent technological advances have opened up a wide range of new opportunities for malevolent actors. For example, Web 2.0, the rise of social media and the availability of nearly all content on the Internet have enabled these actors to easily connect with other like-minded individuals and form almost entirely closed communities that reinforce their own views.

Research into the metaverse as the successor to social media and the mobile Internet can provide important insights into how malevolent actors could creatively use the metaverse. While we generally agree with Joe Whittaker and others (Whittaker, 2022; Valentini, Lorusso & Stephan, 2020) that distinguishing between offline and online radicalisation does not make sense from an analytical perspective, the way in which malevolent actors currently use social media could give an idea of the metaverse of the future.

It is generally recognised that malevolent actors (with different ideological backgrounds) began to make use of the Internet and its possibilities at an early stage (Feldman, 2020; Fisher, 2015; Stewart, 2021; Lehmann & Schröder, 2021). They used new technologies in creative ways in order to evade monitoring and detection and also to improve their own operations. As an anonymous place of countless possibilities where one can find a wealth of information tailored to one's own interests, the Internet is a gold mine for extremists (Bertram, 2016, p. 232).

While research on the radicalisation patterns of convicted jihadi terrorists has shown that offline networks played a much greater role in their radicalisation than online networks (Hamid & Ariza, 2022), other research indicates that the Internet has a more important role for right-wing extremists. This applies especially to the planning of their attacks and actions (von

Behr et al., 2013; Gill et al., 2017). "The Internet is largely a facilitative tool that affords greater opportunities for violent radicalization and attack planning. Nevertheless, radicalization and attack planning are not dependent on the Internet […]." (Gill et al. 2017, p. 113).

Social media has been used by malevolent groups to create, target and distribute self- generated content without the traditional processes of vetting used by traditional media companies and while avoiding policing and censorship from nation states (Droogan et al., 2018, p. 171). Furthermore, social media has also become an instrument of social interaction for those who are already radicalised and those whom they want to convince or who are interested in their activities (Conway, 2017).

The introduction of the metaverse could further reinforce this momentum. By further bridging the gap between offline and online, it could be even more difficult to maintain the distinction between the two spheres of radicalisation (and extremism and terrorism). At present, offline networks provide familiarity and a close environment and are more likely to evade security services than online extremists (Hamid & Ariza, 2022). The future metaverse could bring together these advantages of the offline world in an extensive and immersive digital experience. Combined with the advantages of the online world – instant mass communication and propaganda – the metaverse could become an even bigger game-changer than the Internet and social media were.

The Metaverse as a Democratic Space

The metaverse is still in the early stages of development and has a long way to go before it reaches a certain stage of maturity in which promises, and actual functionalities are implemented. It is already apparent that the risks of the metaverse are comparable to those of social media and, in the past, a response often came too late. Freedom and security will probably be the decisive variables in this technology of the future, which makes engaging with malevolent actors all the more crucial (Neuberger, 2023). In the initial phase of the metaverse, it is already becoming clear that malevolent actors are finding fertile ground – as illustrated, for example, by incidents of sexual harassment that have occurred in the current test versions of the metaverse (Bazu, 2021; Bovermann, 2022; Diaz, 2022; Wiederhold, 2022).

How can these developments be tackled? How can they be prevented before they cause harm? It will be important to ensure the democratic

involvement of actors and marginalised groups in decision-making and development processes. While this would now be a genuinely reactive process in the case of social media, the developers of the metaverse still could build beneficial structures. Democratisation of social media is desirable from a sociopolitical perspective because it is a powerful tool due to its widespread use and its economic and cultural importance. This power should be democratically legitimised and controlled (Engelmann et al., 2020). However, democratic safeguarding should not follow a party-political pattern.

In the development of the metaverse, social media should be informative in various ways – from the creativity with which malevolent actors use new media and technologies (see above) through to the democratic involvement of users. Social media operators have already tried to take account of the aspect of participation:

- META conceived the idea of an Oversight Board[1] in 2018 as a body whose independent judgement could help the company make tough content decisions. This board is committed to being independent, accessible, and transparent. META has granted it the authority to decide whether content should be allowed or removed.
- Twitter has been advised by a Trust and Safety Council in the past. This consisted of various NGOs and researchers who advised the company on online security issues. Elon Musk dissolved the Council after taking over the company (The Associated Press, 2022).
- On its YouTube video platform, Google has introduced the Priority Flagger Programme[2]. This enables NGOs and public authorities to use highly effective tools to report content that violates the Community Guidelines. This flagged content is then reviewed by moderators as a priority. However, the deletion criteria are the same as for any other reports. The programme was revised by YouTube in 2021, which led to major criticism from the community (Meineck, 2021).

In general, there seems to be a worrying trend on social media to cut back on these participative models of moderation and security in favour of artificial intelligence (AI) applications (Gorwa et al., 2020; Llansó, 2020). However, AI solutions cannot and should not replace the involvement of civil society in decision-making processes and questions of democratic culture, not least because AI-supported content moderation solutions are

1 https://www.oversightboard.com/.
2 https://support.google.com/youtube/answer/7554338?hl=de.

still prone to error and lack transparency (Gillespie, 2020; Gorwa et al., 2020).

Encouraging Participation and Democratic Involvement

In social media research and particularly in platform governance research, important approaches can be found that may help to enable a democratic and inclusive metaverse. In addition to essential cooperation between operators and governmental and non-governmental actors on issues of transparency and research, there is an emphasis on actively strengthening democratic actors and narratives (Bundtzen & Schwieter, 2023; Engelmann et al., 2020; Rau et al., 2022).

This strategy is crucial to ensure that a state's repressive apparatus is actually only used as a measure of last resort to stop malevolent actors. Democratic argument and discourse must be possible in an inclusive metaverse without people constantly having to fear repression and restriction. Instead, platform operators can also take steps in the metaverse to consciously and actively promote democratic actors and narratives and thus build democratic resilience in the metaverse.

Here too, the metaverse can take inspiration from existing approaches in the social media field, such as YouTube's trusted flagging programme. Democratic actors, e.g. NGOs and government organisations, specialising in areas such as hate speech, group-focused enmity or strengthening democracy could have access to special reporting tools. They could also be given extended powers to contextualise questionable content.

However, as well as reinforcing democratic narratives, the democratisation of the platform itself is a crucial factor for inclusivity. Involving users in decision-making and design processes can have enormous added value for a platform that is interested in democratic interaction. Marginalised groups and their representatives know exactly where hate and harassment may be lurking in the digital space. By involving such stakeholders at an early stage, some of the mistakes that were made on social media could be minimised from the outset.

In political practice, mini publics have already proved effective as an instrument of user participation (Escobar & Elstub, 2017; Smith & Setälä, 2018). Mini publics are groups of (randomly or systematically) selected citizens who work together over an extended period to examine socially relevant issues, with the inclusion of external sources, e.g., scientific exper-

tise. Topics are examined, discussed and assessed from a broad range of perspectives, and the resulting recommendations are forwarded to political decision-makers (Escobar & Elstub, 2017; Pek et al., 2023). One example of this is the virtual citizens' assembly in Germany. In June 2022, its members debated the consequences of using artificial intelligence (Buergerrat.de, 2022). These types of assemblies allow platform-specific topics to be discussed with the aim of ensuring that decision making is more democratic.

Although quite controversial (see above), platform councils can also develop potential for promoting democracy if they are able to operate independently, objectively and transparently (Haggart & Keller, 2021; Rau et al., 2022). To ensure this, platform councils of this type could be based on the press and broadcasting councils that are already established in Germany, in line with the recommendations of Kettemann and Fertmann (2021). It should be noted, however, that responsibilities (geographical, practical), participants (citizens, experts, NGOs, political decision-makers), and not least powers (quasi-judicial, advisory) must be part of the social discourse and cannot yet be conclusively clarified (Cowls et al., 2022; Kettemann & Fertmann, 2021). Furthermore, such councils could boost public confidence – the more diverse and transparent their line-up is and the more publicly visible the effects of their recommendations are.

Finally, the aim must also be to strengthen media literacy and policy competence by means of various training opportunities. These should be designed in such a way that individuals who are not (or no longer) associated with the education system are also able to benefit from them. Here it is vital to provide the necessary tools for dealing with fake news, other manipulated or extremist content and hate speech on the Internet. One example to mention is the Good Gaming – Well Played Democracy project[3] directed by the Amadeu Antonio Foundation, which aims to raise the gaming community's awareness of extremist content, among other things.

In addition, it must be noted that building a democratic metaverse is not solely a task for citizens. The creation of a digital twin in the sense of a well-fortified democracy is also important. However, according to Rau et al. (2022), this does not exclusively mean the use of repressive measures such as deletion or suppression of problematic content (see, for example, Bellanova & De Goede, 2022) but also, coupled with this, the strengthening of democratic actors, e.g. through algorithmically increased visibility. In

3 https://www.amadeu-antonio-stiftung.de/projekte/good-gaming-well-played-democracy/.

this context, the empowerment of marginalised democratic actor groups becomes especially important to adequately represent social diversity. They are properly trained to recognise problematic content at an early stage, for example, and can thus also be consulted for advice (Rau et al., 2022). The use of counter speech could also be another strategy for tackling extremist content in the metaverse (Clever et al., 2022; Hangartner et al., 2021; Kunst et al., 2021; Morten et al., 2020). The term (digital) counter speech refers to comments or other content posted as a response to hate speech in order to minimise and weaken the impact of it or to support potential victims (Ernst et al., 2022; Garland et al., 2022). In this regard, studies have shown that counter speech can be an effective means of tackling extremist content and reducing it effectively (Garland et al., 2022; Hangartner et al., 2021). In the context of newer technological complexes, e.g. AI, consideration is currently being given to implementing counter speech automatically in certain circumstances, although final concepts and responsibilities are still the subject of intensive discussion (Clever et al., 2022).

In addition to participatory methods, legislation can also be used to prevent extremist content. In Germany, the dissemination of unconstitutional symbols and signs is forbidden, and perpetrators can be prosecuted. Germany's Network Enforcement Act (Netzwerkdurchsetzungsgesetz, NetzDG) also provides a legal framework for dealing with hate crime on social media. Accordingly, the Terrorist Content Online Regulation (European Union, 2021) requires platform operators offering services in the EU to remove or block reported terrorist content within one hour. Recent results of extremism research indicate, however, that so-called legal but harmful content is already proving to be a major challenge and is likely to be of significance in the metaverse as well (Jiang et al., 2021; Rau et al., 2022). This includes, for example, digital content that may have a subtle radicalising effect but is not unlawful. However, it should be noted in this regard that content moderation must comply with the constitutional principle of free speech. Consequently, it is to be assumed that the ongoing discussion on the relationship between freedom and security will also significantly influence the design of the metaverse and will, or must, be the result of a negotiation process involving society as a whole in order to guarantee the democratic dimension.

Discussion

If the immersiveness of the metaverse lives up to Mark Zuckerberg's vision, it is very likely to have a huge impact on our everyday lives and on social interaction. This immersiveness would mean that the operators of the metaverse (or metaverses) would need to deal intensively with questions of democratisation. Not only would the state probably play a (yet to be defined) role in a metaverse, but its users must also be enabled to participate democratically in it. This would help to make the platform inclusive and as safe as possible from malevolent actors.

Building on the social media research of recent decades, there are many common points of reference which can support and steer the design of a democratic metaverse. As mentioned above, the metaverse is still at an early stage of development. However, given the rapid pace of advancement, it is vital to support this process, stay on the ball and take an active role in discussions. A multi-perspective approach from all stakeholders involved is also relevant to ensuring a balance between security and freedom for all users. The possibilities outlined here for building a metaverse present some solutions for implementing democratic pillars. In summary, the following solutions should be deployed by platform operators:

- early implementation of methods for user participation, e.g. mini-publics or independent platform councils
- strengthening of democratic actors and inclusion of marginalised groups
- reference to existing scientific research findings on social media, hate speech and (digital) extremism, as well as open cooperation with research institutions
- offer of educational opportunities in cooperation with democratic actors

Final and concrete implementation is still currently the subject of lively discussion. However, the status of the early development phase of the metaverse is encouraging active participation, which is also reflected in this Immersive Democracy Project and can be understood as an invitation to this process. Participation is not a panacea for the dangers lurking in the digital space. But it is an important source of support that can help to empower marginalised groups or individuals in specific ways and thus give them the tools to work together with operators against discrimination and hate in the metaverse. Now is the time to develop these tools and to make sure that a future metaverse is as safe and secure as possible for everyone.

References

Bazu, T. (2021). The metaverse has a groping problem already. *MIT Technology Review*. https://www.technologyreview.com/2021/12/16/1042516/the-metaverse-has-a-groping-problem/

von Behr, I., Reding, A., Edwards, C., & Gribbon, L. (2013). *Radicalisation in the digital era: The use of the internet in 15 cases of terrorism and extremism*. RAND Europe.

Bellanova, R., & De Goede, M. (2022). Co-producing security: Platform content moderation and European security integration. *JCMS: Journal of Common Market Studies*, *60*(5), 1316–1334. https://doi.org/10.1111/jcms.13306

Bertram, L. (2016). Terrorism, the internet and the social media advantage: Exploring how terrorist organizations exploit aspects of the internet, social media and how these same platforms could be used to counter-violent extremism. *Journal for Deradicalization*, *2016*(7), 225–252.

Bovermann, P. (2022). Online-Belästigungen im Metaverse – Am eigenen Leib. *Süddeutsche Zeitung*. https://www.sueddeutsche.de/kultur/metaverse-vr-virtual-reality-microsoft-sexuelle-belaestigung-1.5519527?print=true

Buergerrat.de. (2022). Bürgerrat diskutierte über künstliche Intelligenz. *Buergerrat.de*. https://www.buergerrat.de/aktuelles/buergerrat-diskutierte-ueber-kuenstliche-intelligenz/

Bundtzen, S., & Schwieter, C. (2023). *Datenzugang zu Social-Media-Plattformen für die Forschung: Lehren aus bisherigen Maßnahmen und Empfehlungen zur Stärkung von Initiativen inner- und außerhalb der EU*. Institute for Strategic Dialogue (ISD).

Clever, L., Klapproth, J., & Frischlich, L. (2022). Automatisierte (Gegen-)Rede? Social Bots als digitales Sprachrohr ihrer Nutzer*innen. In J. Ernst, M. Trompeta, & H.-J. Roth (Eds.), *Gegenrede digital: Neue und alte Herausforderungen interkultureller Bildungsarbeit in Zeiten der Digitalisierung* (pp. 11–26). Springer Fachmedien. https://doi.org/10.1007/978-3-658-36540-0_2

Conway, M. (2017). Determining the role of the internet in violent extremism and terrorism: Six suggestions for progressing research. *Studies in Conflict & Terrorism*, *40*(1), 77–98. https://doi.org/10.1080/1057610X.2016.1157408

Cowls, J., Darius, P., Santistevan, D., & Schramm, M. (2022). Constitutional metaphors: Facebook's "supreme court" and the legitimation of platform governance. *New Media & Society*. https://doi.org/10.1177/14614448221085559

Cropley, D. H., Kaufman, J. C., & Cropley, A. J. (2008). Malevolent creativity: A functional model of creativity in terrorism and crime. *Creativity Research Journal*, *20*(2), 105–115. https://doi.org/10.1080/10400410802059424

Diaz, A. (2022). Disturbing reports of sexual assaults in the metaverse: 'It's a free show'. *New York Post*. https://nypost.com/2022/05/27/women-are-being-sexually-assaulted-in-the-metaverse/

Droogan, J., Waldek, L., & Blackhall, R. (2018). Innovation and terror: An analysis of the use of social media by terror-related groups in the Asia Pacific. *Journal of Policing, Intelligence and Counter Terrorism*, *13*(2), 170–184. https://doi.org/10.1080/18335330.2018.1476773

Engelmann, S., Grossklags, J., & Herzog, L. (2020). Should users participate in governing social media? Philosophical and technical considerations of democratic social media. *First Monday*, *25*(12). https://doi.org/10.5210/fm.v25i12.10525

Ernst, J., Trompeta, M., & Roth, H.-J. (2022). Gegenrede digital – Einleitung in den Band. In J. Ernst, M. Trompeta, & H.-J. Roth (Eds.), *Gegenrede digital* (pp. 1–7). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-36540-0_1

Escobar, O., & Elstub, S. (2017). Forms of mini-publics: An introduction to deliberative innovations in democratic practice [Research and Development Note]. *New Democracy*.

European Union. (2021). Regulation (EU) 2021/784 of the European Parliament and of the Council of 29 April 2021 on addressing the dissemination of terrorist content online (Text with EEA relevance). *Official Journal of the European Union*.

Garland, J., Ghazi-Zahedi, K., Young, J.-G., Hébert-Dufresne, L., & Galesic, M. (2022). Impact and dynamics of hate and counter speech online. *EPJ Data Science*, 11(1), Article 3. https://doi.org/10.1140/epjds/s13688-021-00314-6

Gill, P., Corner, E., Conway, M., Thornton, A., Bloom, M., & Horgan, J. (2017). Terrorist use of the internet by the numbers: Quantifying behaviors, patterns, and processes. *Criminology & Public Policy*, 16(1), 99–117. https://doi.org/10.1111/1745-9133.12249

Gillespie, T. (2020). Content moderation, AI, and the question of scale. *Big Data & Society*, *7*(2). https://doi.org/10.1177/2053951720943234

Gorwa, R., Binns, R., & Katzenbach, C. (2020). Algorithmic content moderation: Technical and political challenges in the automation of platform governance. *Big Data & Society*, *7*(1). https://doi.org/10.1177/2053951719897945

Haggart, B., & Keller, C. I. (2021). Democratic legitimacy in global platform governance. *Telecommunications Policy*, *45*(6). https://doi.org/10.1016/j.telpol.2021.102152

Hamid, N., & Ariza, C. (2022). *Offline versus online radicalisation: Which is the bigger threat?* London: Global Network on Extremism & Technology.

Hangartner, D., Schmid, L., Helbling, M., & Zingg, A. (2021). Empathy-based counterspeech can reduce racist hate speech in a social media field experiment. *Proceedings of the National Academy of Sciences*, 118(50). https://doi.org/10.1073/pnas.2116310118

Jiang, J. A., Scheuerman, M. K., Fiesler, C., Brubaker, J. R., & Bovet, A. (2021). Understanding international perceptions of the severity of harmful content online. *PLOS ONE*, 16(8). https://doi.org/10.1371/journal.pone.0256762

Kettemann, M. C., & Fertmann, M. (2021). *Die Demokratie Plattformfest Machen: Social Media Councils als Werkzeug zur gesellschaftlichen Rückbindung der privaten Ordnungen digitaler Plattformen*. Friedrich-Naumann-Stiftung.

Kunst, M., Porten-Cheé, P., Emmer, M., & Eilders, C. (2021). Do "Good Citizens" fight hate speech online? Effects of solidarity citizenship norms on user responses to hate comments. *Journal of Information Technology & Politics*, 18(3), 258–273. https://doi.org/10.1080/19331681.2020.1871149

Llansó, E. J. (2020). No amount of "AI" in content moderation will solve filtering's prior-restraint problem. *Big Data & Society*, *7*(1). https://doi.org/10.1177/2053951720920686

Meineck, S. (2021). Trusted Flagger: YouTube serviert freiwillige Helfer:innen ab. *netzpolitik.org*. https://netzpolitik.org/2021/trusted-flagger-youtube-serviert-freiwillige-helferinnen-ab/

Morten, A., Frischlich, L., & Rieger, D. (2020). Gegenbotschaften als Baustein der Extremismusprävention. In J. B. Schmitt, J. Ernst, D. Rieger, & H.-J. Roth (Eds.), *Propaganda und Prävention* (pp. 581–589). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-28538-8_32

Neuberger, C. (2023). Sicherheit und Freiheit in der digitalen Öffentlichkeit. In N. J. Saam & H. Bielefeldt (Eds.), *Sozialtheorie* (pp. 297–308). transcript Verlag.

Pek, S., Mena, S., & Lyons, B. (2023). The role of deliberative mini-publics in improving the deliberative capacity of multi-stakeholder initiatives. *Business Ethics Quarterly*, 33(1), 102–145. https://doi.org/10.1017/beq.2022.20

Rau, J., Kero, S., Hofmann, V., Dinar, C., & Heldt, A. P. (2022). Rechtsextreme Online-Kommunikation in Krisenzeiten: Herausforderungen und Interventionsmöglichkeiten aus Sicht der Rechtsextremismus- und Platform-Governance-Forschung. *Arbeitspapiere des Hans-Bredow-Instituts*. https://doi.org/10.21241/SSOAR.78072

Smith, G., & Setälä, M. (2018). Mini-publics and deliberative democracy. In A. Bächtiger, J. S. Dryzek, J. Mansbridge, & M. Warren (Eds.), *The Oxford Handbook of Deliberative Democracy* (pp. 299–314). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780198747369.013.27

The Associated Press. (2022, December 12). Musk's Twitter has dissolved its Trust and Safety Council. *National Public Radio (NPR)*.

Wiederhold, B. K. (2022). Sexual harassment in the metaverse. *Cyberpsychology, Behavior, and Social Networking*, 25(8), 479–480. https://doi.org/10.1089/cyber.2022.29253. editorial

Navigating Democratic Challenges in the Age of Metaverses

Jean-François Lucas, Jessica Galissaire & Henri Isaac

The Metaverse is an opportunity to bring together "all the questions surrounding the development of digital technology in the years ahead" (LINC, 2023, pp. 25), in particular "all of the digital world's usual ethical issues: network neutrality, protection of personal data, digital identity, online harassment, addiction, isolation, and also exclusion" (Basdevant et al., pp. 81).

This ecosystem of interconnected, immersive digital worlds, while promising unparalleled opportunities for innovation and engagement, also embodies critical democratic concerns. As digital spaces become more immersive and ubiquitous, the traditional boundaries between what's virtual and what's real are getting blurred, raising complex questions about governance, participation, and individual rights. The concept of the "Metaverse" suggests an ideal of fully interoperable and synchronized digital spaces; however, the reality often presents a constellation of "metaverses" managed by private corporations, public authorities, or other stakeholders. This fragmentation brings with it the risk of reproducing, or even intensifying, the governance limitations seen in contemporary digital platforms, such as the monopolistic control exerted by a few dominant players. For example, centralised governance models, typical of many current digital platforms, can impose restrictions on user autonomy, transparency, and democratic engagement. These concerns are compounded by the rapid regionalisation of digital spaces, as seen in the diverging regulatory frameworks of the European Union, the United States, and China, suggesting that democratic principles in the Metaverse will be inherently context-dependent and highly variable.

This chapter aims to examine some democratic challenges posed by the emergent Metaverse. It is structured into three sections. The first section explores the concept of the Metaverse, distinguishing between its idealised form as a single, interconnected space and the reality of multiple, fragmented "metaverses" governed by varied and often conflicting interests. The second section delves into the governance implications of these metaverses, drawing parallels with the existing digital platforms ecosystem and highlighting risks associated with centralised control. Finally, the chapter

outlines some stages of a roadmap for the promotion of democratic governance in metaverses, including recommendations for multi-stakeholder processes, user representation, and agile regulatory frameworks that can adapt to evolving technological contexts. Through this analysis, the chapter contributes to ongoing debates on ensuring that the Metaverse can develop as an inclusive, democratic digital space, responsive to the needs and rights of its diverse user base.

1 The "Metaverse" as a Concept, Metaverses as Embodiments

The “Metaverse” has a multitude of definitions, both in scientific literature and in the media. While "the concept of the ‘virtual world’ underlying the Metaverse has been assimilated by the public [in particular] thanks to the imaginary universe deployed in the entertainment industries, [...] its definition is still unclear to consumers, who do not grasp its significance or the scope of its use" (Galienni & Truphème, 2023, pp. 116).

The Metaverse embodies the idea of a universe of interoperable digital worlds and environments, enabling an unlimited number of people to enjoy immersive, collective, and synchronised experiences. It's a concept, an ideal, that seems hardly feasible today. In this respect, it is interesting to draw a distinction between "the abstract concept of the Metaverse 'with a capital M' (as we refer to the Internet 'with a capital letter'), which describes a concept of immersion, [and] metaverses in lower case. The term metaverse in lower case is used to describe instantiations or implementations of the Metaverse principles. In other words, the Metaverse concept holds within it a multitude of possibilities, services, and spaces, which are open, to a greater or lesser extent, and which will hereafter be referred to as 'metaverses' with a lower-case 'm'" (Basdevant et al., 2022, pp. 9). The use of the term “metaverse” reflects the reality of today's digital worlds and simulated universes: they are diverse, heterogeneous and plural.

We can therefore assume that there will not be just one large Metaverse, but rather several metaverses, each constituting a sort of galaxy of worlds governed by large companies, public authorities, and various communities, or even users. As a result, “some leaders, like Tim Sweeney, are convinced that, in the end, each company will have to run its own virtual world, both as individual planets and as stakeholders in the main virtual world platforms, such as Fortnite and Minecraft. As Sweeney put it, in the same

way that, a few decades ago, every company created its website, and then, after a while, every company created a Facebook page'" (Ball, 2023, pp. 56).

Moreover, adds the author, "over the past 15 years, what we call 'the internet' has become increasingly regionalised. All countries use the internet Protocol suite, but each market's platforms, services, technologies, and agreements have diverged, in part due to the emergence of non-American tech giants. [...] If the metaverse is to play a greater role in human society and the workplace, then it is likely that its emergence will also lead to more and stronger regional players" (Ball, 2023, pp. 322). This perspective needs to be taken seriously, at a time when the European Commission, as well as countries such as the United Kingdom and China, among many others, are communicating about projects and strategies linked to the development of their own Metaverse; even if these are often "digital" strategies that are closer to industrial digital twins (industrial metaverses) than to the concept of Metaverse that we have just described. The lack of a consensual and shared definition is a primary obstacle to the existence of a democratic debate on this object.

2 The Platforms Legacy

The debate about the democratic nature of metaverses is part of an ongoing debate about digital platforms. These often intense and polemical debates have highlighted the weak democratic nature of digital platforms. Social networks and other types of platforms, such as user-generated content platforms or electronic marketplaces, are characterised by a very undemocratic governance. As the Metaverse seeks to articulate these different logics in a single immersive space, it is to be feared that it will simply reproduce the current governance mechanisms without really amending them. Most of the existing platforms operate on a centralised model.

Centralised platforms rely on a single authority for decision-making. Decision rights lie with a single person or an entity representing the platform owner (Baldwin & Woodard, 2009). Centralised platforms typically exhibit no or very little transparency. Since transparency and user participation are related to platform control and commercialisation, centralised platforms do not provide the platform user with information on what governance processes take place and why (Staykova & Damsgaard, 2015). Although certain units within the company may be involved and, therefore, be given

access to information, this does not account for all units and participants of a platform.

Furthermore, platforms with centralised governance mechanisms rely on strict participation regulations and poor accessibility. Centralised platform governance is typically accompanied by strict participation boundaries, resulting in a low level of accessibility. Working on or improving the platform is, therefore, only possible based on a working order being given by the governing entity, or by another organisational unit. For instance, developers who want to cooperate or improve the platform have neither the possibility to get information about the areas to be developed nor the code available (Trang et al., 2014).

The main source of trust in centralised platforms are established Terms and Conditions, enforced by the platform management. These Terms and Conditions are the basis for dispute resolution. They specify in which cases and in which way the company has to settle disputes between its users (Parker & Van Alstyne, 2017). In short, trust is based on top-down processes and mechanisms (Söllner et al., 2016).

Finally, incentives for engagement in platform improvement focus on order-based improvements of the platform with no or low (pecuniary) incentives. Related to the accessibility dimension, central authorities might give monetary incentives to users who contribute to developing the platform.

As it stands, there is a strong fear that metaverses will only replicate the centralised platform model. Meta's Horizon Worlds, for instance, illustrates a continuity from digital platforms' centralised model. It is therefore fundamental that the regulation of metaverses provides a framework that constrains their governance, particularly on the issue of users' rights over these spaces and the transparency of the rules governing them.

3 Some Governance Principles to Promote Democratic Spaces in Metaverses

Towards More Harmonised Criminal Law Frameworks

In today's democracies, legal frameworks exist to protect individuals from criminal offences and harmful behaviour, both in the physical world and online. The EU Charter of Fundamental Rights and the French Penal Code, for example, apply to the Metaverse (as they apply to the internet

in general). In France, cyberbullying is considered an offence, in the same way as moral or sexual harassment, and is punishable under article 222–33–2–2 of the Penal Code. Furthermore, there are some standards that apply to everyone, both online and offline, and which form the basis of democracies: fundamental rights. In fact, they constitute the highest standards that apply to metaverses. At supranational level, several texts are dedicated to the fundamental rights that are likely to be impacted by the mass adoption of virtual worlds[1].

Despite the existence of such a framework, it can be difficult, in the age of the Metaverse and social networking platforms, to clearly define criminal offences and therefore to counter them. How can harassment be defined in an immersive virtual space? Can rape take place in the Metaverse? Should a distinction be made between private and public spaces? How can the perpetrators of criminal offences and harmful behaviour be held accountable? How can such behaviour be proven, and victims compensated? It is all quite unclear.

Moreover, there are major gaps in this regard in our legal systems, starting with the lack of definitions. For example, it is still difficult to establish a common legal definition of what is illegal. At European level, there is no harmonised definition of "hate speech", for instance. Prior to 15 March 2017, and the coming into force of the Counter-Terrorism Directive, even "terrorism" did not have a common definition across EU Member States. Nor is there a harmonised Penal Code in Europe – each Member State has its own. What constitutes illegal content or behaviour therefore varies from country to country. In Denmark and Germany, for example, Holocaust denial is not punishable by law, whereas it is in France. Therefore, how can we identify and fight against illegal behaviour in the Metaverse? Illegal according to whom? Illegal where? This remains an open question.

One key example of how this lack of common definition can be an obstacle to the prevention and punishment of illegal behaviour in the Metaverse is the criminalisation of virtual rape, which has been under discussion for over thirty years and is well documented[2]. In France, under Article 222–23 of the Penal Code, "[a]ny act of sexual penetration, whatever its nature,

1 These include: the 1948 Universal Declaration of Human Rights; the European Convention for the Protection of Human Rights and Fundamental Freedoms, enforced by the European Court of Human Rights, which is based in Strasbourg; the Charter of Fundamental Rights of the European Union, which has been enforceable by the Member States since 2009: any citizen can refer to it if their rights are not respected.

2 On this subject, see for example: Horne, 2023.

committed against another person by violence, coercion, threat, or surprise is rape". For rape to be considered a criminal offence, there must be physical contact. In other words, in France, rape is not currently recognised as such in the cyberspace, as it is not a physical space in which there can be a physical penetration. However, at a time when the boundaries between physical online experiences and physical offline experiences are becoming increasingly blurred, we should keep a close eye on developments in immersive technologies and how people feel about them. Some legal adjustments may indeed be needed.

Despite the lack of common definitions for key concepts such as "hate" or "rape", efforts are being undertaken by European legislators to come up with a shared framework of what is "illegal" or "harmful" online. For instance, the EU's Digital Services Act (DSA)[3] is based on the principle that what is illegal offline is also illegal online. It lays down a series of rules to make digital platforms more accountable and to combat the spread of illegal or harmful content or products: racist attacks, child pornography, disinformation, the sale of drugs or counterfeit goods, etc. The aim is to better protect European internet users and their fundamental rights (freedom of expression, consumer protection, etc.) and to strengthen the democratic control and supervision of the very large platforms and reduce their systemic risks (manipulation of information, etc.). It's a first (and big) step, but it doesn't solve the problem of harmonising laws from the point of view of freedom of expression at international level, for example (which is probably a utopia). Thus, we urge the new European Parliament and Commission to make legal consistency and enforceability a priority in the coming years. The possible advent of the Metaverse appears to be a perfect laboratory in this respect.

A Collective Approach

Multi-Stakeholder Processes

At the start of 2022, the European Commission took up the subject of metaverses and virtual worlds, to analyse if there was a need to regulate their development and, if so, the potential role of regulators in doing so. In

3 European Commission, "The Digital Services Act": https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/europe-fit-digital-age/digital-services-act_en

line with this, the President of the European Commission, Ursula von der Leyen, announced the launch of an initiative on virtual worlds in 2023, as part of the "A Europe fit for the digital age"[4] programme.

Questions are being raised as to whether the existing legal framework is sufficient to protect users from certain practices that could be detrimental to them in metaverses, to protect their rights in these spaces, and to guarantee an "open, secure, trustworthy, fair, and inclusive digital environment"[5] as called for by the European executive. But we must also question how this framework is being decided. Who formalises these frameworks? In what way? According to what criteria? What role do users play in decision-making processes?

In this regard, we recommend a multi-stakeholder approach, bringing together metaverse operators, terminal suppliers, users, the relevant regulatory authorities, legislators, researchers and civil society, in order to have a holistic and structured approach to Metaverse governance.

Thierry Breton, former European Commissioner for the Internal Market, was calling for the launch, "similarly to the European Bauhaus", of "a creative and interdisciplinary movement, aiming to develop standards, increase interoperability, maximising impact with the help of IT experts, regulatory experts citizens' organisations and youth" (European Commission, 2022). With this in mind, between February and April 2023 the European Commission convened a "European citizens' panel on virtual worlds". Bringing together 140 citizens from the 27 Member States, the panel published 23 recommendations on the values and actions needed to create attractive and equitable European virtual worlds. These recommendations are now feeding into the Commission's work on virtual worlds and tomorrow's internet.

In our report "Governing the Metaverse and tomorrow's internet", we propose the development of experimental multi-stakeholder processes (e.g. via regulatory sandboxes and policy prototyping), in order to analyse the relevance of the existing legal framework in relation to the Metaverse, and

4 This work culminated in the presentation, on 11 July 2023, of a strategy to "place the EU at the forefront of Web 4.0 and virtual worlds". See: European Commission. (2023). Towards the next technological transition: Commission presents EU strategy to lead on Web 4.0 and virtual worlds. https://ec.europa.eu/commission/presscorner/detail/en/ip_23_3718

5 European Commission. (2023). Towards the next technological transition: Commission presents EU strategy to lead on Web 4.0 and virtual worlds. https://ec.europa.eu/commission/presscorner/detail/en/ip_23_3718

put forward recommendations relating to the technical operationalisation of concepts such as privacy, the protection of personal data, or the prevention of cyber-bullying.

Involving Users in Content Moderation

Users of digital platforms contribute significantly to the creation of value on these platforms (Renaissance Numérique, 2020a). Therefore, they should be integrated in their regulation. One way to do this would be to impose user representation in their governance and decision-making bodies. The role of a user-representative body should focus on several elements: the definition of moderation rules and their evolution, the definition of a collaborative platform-specific approach to moderation, and the development of a moderation culture specific to each platform (Renaissance Numérique, 2020b).

For metaverses, like for platforms, content moderation requires defining the right balance and processes, hand in hand with public authorities, civil society and the end users. The notion of value co-creation is inherent to platforms that host content generated by their users. The same will apply to metaverses. This substantial input from end-users should be reflected in the moderation efforts of metaverses, for example by encouraging online service providers to involve users in the moderation of content and behaviour, and more generally in the regulation of metaverses. However, this requires the establishment of real discursive processes between platforms and their users, and cannot be limited to the outsourcing of moderation tasks (Renaissance Numérique, 2020c).

An Agile Approach

Without wiping the slate clean of existing restrictive legal frameworks, which are necessary, the challenge is to simultaneously establish multi-stakeholder mechanisms that are as agile as possible, with feedback loops to adapt to technologies as they evolve. This aspect is all the more important with regard to the Metaverse, as the underlying technologies are not yet fully mature, and the uses to which they will be put, and therefore the business models that will be linked to them, remain largely undefined. In this sense, we propose to put in place agile, multi-stakeholder governance mechanisms to structure every person's right and duty in tomorrow's internet. The new

European Commission should encourage and develop the implementation of this kind of approach at European level.

A Holistic Approach

It is also highly likely that the responsibilities of the various players involved in the governance of the Metaverse will evolve. Nowadays, in the Web 2.0 era, we are faced with highly centralised systems. It is technically the online services providers, also known as "intermediaries", via their moderation and Trust & Safety policies, and via their Terms & Conditions, who decide, within the limits set by the law, what is acceptable or not on their platforms. However, the Metaverse is not destined to become a space controlled solely by a few dominant players. On the contrary, a multitude of metaverses and hence owners of immersive spaces should be able to emerge.

In Meta's Horizon Worlds, for example, a third-party company could create its own space, in which it sets its own rules. As is the case in immersive worlds that have been in use for several decades now, such as Second Life, there would be several layers of responsibilities and rules: a technical layer, covering what the source code allows or does not allow in terms of actions; a layer managed by the operator of the metaverse; a layer managed by the owner of the specific world in that metaverse; and above all that, the law (Lucas, 2013).

The ambition of a collective, agile, holistic approach should be to move towards a more effective allocation of responsibilities across various layers, so that they can be implemented more effectively. Renaissance Numérique encourages the new European Commission to embrace this approach and to facilitate the establishment of agile, multi-stakeholder governance mechanisms to organise the rights and duties of all stakeholders in tomorrow's internet.

3 Conclusion

The rise of the Metaverse encapsulates a unique convergence of social, economic, environmental, cognitive, ethical, and legal considerations that define today's digital era. As it takes root in our existing digital landscape, the Metaverse compels us to reimagine how we want to communicate,

connect, learn, entertain, and trade in a more immersive and interactive internet. Rather than merely replicating past models, it offers us an unprecedented chance to craft an internet that is more egalitarian, inclusive, ethical, and sustainable—one that aligns with the aspirations of a responsible, democratic society.

The governance of the Metaverse calls for a multi-stakeholder approach that prioritises agility, transparency, and inclusivity. This would empower not only institutions and corporations but also individual users and citizens, facilitating a collective re-evaluation of our shared values. Legal frameworks currently provide baseline protections against harmful conduct in digital environments, yet inconsistencies remain, especially at the international level. Harmonising these definitions and aligning policies with the realities of immersive experiences is essential to ensure users' safety and rights across borders.

However, the rapid evolution of metaverse technologies, driven by intense competition and undefined business models, poses the risk of recreating the centralised, platform-dominated structures of the past. To avoid this, alternative governance mechanisms must be considered, such as decentralised models enabled by blockchain or emerging cooperative frameworks, that could uphold users' rights and trust without relying solely on a central authority.

As we face these challenges, the Metaverse should be viewed not only as a technological innovation but as an opportunity to redefine digital governance and rethink the ways we interact socially online. By treating the Metaverse as a living, evolving project, we can build a digital future that reflects the democratic values we strive for, setting the stage for a truly inclusive, participatory, and resilient internet. The journey is complex, but the potential to shape a Metaverse that serves society as a whole is a compelling invitation to reimagine the foundations of our digital lives.

References

Baldwin, C. Y., & Woodard, J. (2009). The architecture of platforms: A unified view. In A. Gawer (Ed.), *Platforms, markets, and innovation*. Edward Elgar Publishing.

Ball, M. (2023). *Le Metavers: Comment va-t-il tout révolutionner*. De Boeck Supérieur.

Basdevant, A., François, C., & Ronfard, R. (2022). *Mission exploratoire sur les métavers*. Ministère de l'Économie, des Finances et de la Relance. https://www.economie.gouv.fr/files/files/2022/Rapport-interministeriel-metavers.pdf

European Commission. (2022, September 14). *Europe's plan to thrive in the metaverse* [Blog post by Commissioner Thierry Breton]. https://ec.europa.eu/commission/presscorner/detail/en/STATEMENT_22_5525

Galienni, S., & Truphème, S. (2023). *La vague Web3: Maîtriser les nouveaux codes du marketing 3.0. Blockchain, NFT, DAO, Métavers…*. Dunod.

Horne, C. (2023). Regulating rape within the virtual world. *Lincoln Memorial University Law Review*, *10*(2). https://digitalcommons.lmunet.edu/lmulrev/vol10/iss2/8

LINC. (2023). *Data, footprint and freedoms: Exploring the overlaps between data protection, freedoms and the environment* (IP Reports, Innovation and foresight, No. 9). CNIL. https://linc.cnil.fr/sites/linc/files/2023-09/cnil_ip9_data_footprint_and_freedoms.pdf

Lucas, J.-F. (2013). *De l'immersion à l'habiter dans les mondes virtuels: Le cas des villes dans Second Life* [Doctoral dissertation, European University of Brittany, Rennes 2].

Trang, S. T. N., Zander, S., & Kolbe, L. M. (2014). Examining its alignment in inter-organizational networks: Development and application of a measurement instrument. In *PACIS 2014 Proceedings* (p. 183). https://aisel.aisnet.org/pacis2014/183

Parker, G., & Van Alstyne, M. (2017). Innovation, openness, and platform control. *Management Science*, *64*(7). https://doi.org/10.1287/mnsc.2017.2757

Renaissance Numérique. (2020a). *Regulating digital platforms: Why and how?* https://www.renaissancenumerique.org/wp-content/uploads/2022/06/renaissancenumerique_note_platformsregulation.pdf

Renaissance Numérique. (2020b). *Digital platforms: For a real-time and collaborative regulation*. https://www.renaissancenumerique.org/en/publications/digital-platforms-for-a-real-time-and-collaborative-regulation/

Renaissance Numérique. (2020c). *Moderating our (dis)content: Renewing the regulatory approach*. https://www.renaissancenumerique.org/en/publications/moderating-our-discontent-renewing-the-regulatory-approach/

Söllner, M., Pavlou, P. A., & Leimeister, J. M. (2016). Understanding the development of trust: Comparing trust in the IT artifact and trust in the provider. In *Academy of Management Proceedings* (Vol. 2016, No. 1, p. 10167). Academy of Management.

Staykova, K. S., & Damsgaard, J. (2015). The race to dominate the mobile payments platform: Entry and expansion strategies. *Electronic Commerce Research and Applications*, *14*(5), 319–330.

Hate Speech in the Metaverse

Esen K. Tütüncü & Danielle Shanley

Hate speech, as defined by the United Nations (UN), refers to any form of communication, gesture, or conduct that may incite violence, discrimination, hostility, or prejudicial action against individuals or groups based on attributes such as race, ethnicity, religion, gender, sexual orientation, or other characteristics. It encompasses expressions that demean, dehumanize, or stereotype individuals or communities, perpetuating harmful stereotypes and promoting intolerance (UN, n.d.).

Despite this definition appearing relatively clear and concise, defining hate speech is no easy task. Any attempt at defining hate speech effectively has to try to strike a balance between safeguarding freedom of expression – as a fundamental human right – and protecting individuals and groups from harm. This is further complicated given the fact that different individuals may perceive the same speech act differently, and what one person considers offensive or harmful, another may view as a legitimate expression of opinion. The subjective nature of determining whether specific expressions cross the line into hate speech makes it challenging to establish universally applicable definitions.

Despite these definitional difficulties, it is crucial to engage with the topic, especially in online settings like the metaverse. As we learned from early research into virtual worlds, what happens within these worlds can shape the values that influence individuals' real-world lives, and vice versa. For example, writing in the early 1990s, Christine Ward Gailey argued that video games reflect dominant cultural values in society, reinforcing and promoting behaviors that align with the dominant ideology (Gailey, 1993). In essence, commercially successful games often replicate and reinforce the values and activities associated with prevailing societal norms.

Biases in Virtual Worlds

As a result, it is perhaps unsurprising that the portrayal of different social groups in the virtual worlds of video games tends to reflect existing biases

held by individuals in the real world, including sexism and racism. For example, the gender stereotyping of video game characters' appearances, where women are often depicted as thin with large breasts and emphasized sexual features (a trend epitomized by the character Lara Croft in the Tomb Raider series) and men are often portrayed with muscular physiques and aggressive attitudes, has been well documented. Psychological studies among adolescents and college students indicate that exposure to these stereotypical portrayals can desensitize individuals to real-world sexism, making instances of sexism seem less shocking, potentially perpetuating harmful beliefs such as "rape myths" that blame victims of sexual assault (Breuer et al., 2015).

Of course, sexism in video games is not only an issue when it comes to content. User behavior, particularly within multiplayer online games, is also often deeply problematic. For example, female players are often targeted and harassed and have to develop their own coping mechanisms and strategies in order to safeguard themselves against undesirable behavior.

Racist imagery has also been a persistent issue in video games since the early 1990s. Games like Duke Nukem 3D and Shadow Warrior exemplify how troubling racial stereotypes have shaped the narratives of video games. Duke Nukem 3D's storyline revolves around eugenic panic concerning race mixing between invading aliens and white women in a future Los Angeles depicted as mono-ethnic. The main character embarks on a mission to stop the alien invaders in order to preserve the genetic purity of the human species. In Shadow Warrior meanwhile, the protagonist's stereotypical and generic "Asian" identity is accompanied by a skill set portrayed as biologically determined, perpetuating ideas about the character's perceived deficient masculinity (Weise, 2021).

Sexist and Racist Ideas Have Real World Consequences

The sexist and racist ideas and representations depicted in video games can transcend the virtual world and have real-world consequences. Psychological research suggests that playing violent video games can increase ethnocentrism and trigger heightened aggression when individuals encounter someone who is different from themselves (Ewoldsen et al., 2012). Despite activists like Anita Sarkeesian taking up these issues within the gaming

industry,[1] hate and discriminatory attitudes within video games have become normalized over time, which is likely to influence the design and development of other virtual worlds too. In 2022, media outlets were quick to notice the metaverse's "groping problem", which only goes to show how important it is that we address and challenge these issues in order to create inclusive and safe online environments.

Hate Speech on Social VR

On social VR platforms, such as VRChat, AltSpace, and Meta Horizons, hate speech often takes similar forms to that of more traditional social media. Users can engage in discriminatory or offensive behavior targeting individuals or groups based on their characteristics or identities.

- Racist or derogatory remarks: Users may verbally express racial slurs, engage in racial stereotyping, or make discriminatory comments based on a person's race or ethnicity.
- Homophobic or transphobic behavior: Hate speech can manifest as verbal harassment, bullying, or exclusion, targeting individuals based on their sexual orientation or gender identity.
- Religious or cultural intolerance: Users may engage in hate speech by expressing discriminatory attitudes or insulting remarks against specific religions or cultural groups.
- Cyberbullying and harassment: Like traditional social media, social VR platforms can become spaces for targeted harassment, where individuals are subjected to online bullying, threats, or offensive behavior.

While there are important lessons to be learned from how these forms of behavior take place on social media, there are a number of important differences between them and social VR platforms, particularly given their immersive nature.

Social VR platforms introduce additional elements to user experiences, which can affect the nature and extent of hate speech, bullying, and discrimination. These platforms offer users the opportunity to embody avatars and engage in more immersive interactions. As a result, new or different forms of behavior may emerge. Some examples of this are:

1 Feminist Frequency (2013). Damsel in Distress: Part 1 – Tropes vs Women in Video Games. Online: https://youtu.be/X6p5AZp7r_Q (Accessed on June 27th 2023).

- Nonverbal expressions: Hate speech can extend beyond verbal communication. Users can utilize avatars to engage in offensive or discriminatory gestures, actions, or visual representations, which can amplify the impact of hate speech.
- Spatial proximity and presence: In VR environments, users can physically navigate and interact with others in close proximity. This physical presence can intensify the emotional impact of hate speech, leading to increased feelings of harassment or discrimination.
- Immersive experiences and anonymity: The immersive nature of VR can provide a heightened sense of anonymity and disinhibition, potentially leading to more extreme or offensive behavior compared to traditional social media platforms.

Mitigating Hate Speech on Social VR

To mitigate the negative impact of hate speech in social VR platforms, it is important for platform providers to prioritize proactive moderation, establish clear community standards, and promote user empowerment through reporting tools and educational initiatives. Collaborative efforts involving platform developers, users, and relevant stakeholders can help create safe and welcoming virtual environments where users can freely express themselves without fear of harassment or discrimination.

Recent literature on technology ethics emphasizes the value of deliberative engagement for shaping technological development.[2] These works widely – both implicitly and explicitly – draw on concepts of "deep democracy" (Buhmann and Fieseler, 2021, p. 101475) by highlighting the epistemic potential of open engagement processes. Broadly speaking, this literature proposes that innovators, as proactive participants of a wider public debate and discourse, can contribute to responsible processes of innovation. So essentially, the goal is to harness the potential of different forms of engagement in order to help find optimal solutions.

Around 2010, the term "Responsible Innovation" (RI) became a popular way of talking about responsibility-related issues for academics and policy-makers alike. It is used to refer to a way of organizing research and innovation so that its impacts are safe, equitable, and aligned with societal needs.

2 See, for example, several articles published in the Journal of Responsible Innovation and the Journal of Responsible Technology.

New and emerging technologies, like AI and VR, are going to shape our future in powerful new ways. As we can already see, the result of this is that we will need to confront new questions about risks, ethics, justice, and equity. Responsible innovation essentially provides us with the concepts and practices that are required to address these sorts of questions, helping us to think about things like hype, scale, power, and inclusion in research and innovation.

Drawing on the work of Buhmann and Fieseler, responsible innovation can be seen as encompassing three main dimensions (which are often reflected by the public discourse surrounding new and emerging technologies). First, the *responsibility* to *avoid harm*, which refers, for example, to risk management approaches supposed to control for potentially harmful consequences. Second, the *responsibility to do good,* which refers to the improvement of living conditions, such as are set out in the sustainable development goals. Finally, *governance responsibility,* which refers to the responsibility to create and support global governance structures that can facilitate the former two responsibilities (Ibid).

Examples of the sorts of tools and concepts that fall under the umbrella of responsible innovation are:

- Value Sensitive Design: A framework for exploring stakeholder's values in order to then translate those values into operational design criteria, through iterative conceptual, empirical, and technical investigations. VSD asks questions such as: What values to include in design? How to make these values bear on the design process? How to make choices and trade-offs between conflicting values? How to verify whether the designed system embodies the intended values?
- Scenario Planning Workshops: Narratives, or scenarios, are essentially hypothetical sequences of events constructed for focusing attention on causal processes and decision points. In this sense, the development of scenarios can be used for learning and deliberation, producing decision-making processes that are based on the involvement and interaction of different stakeholders.
- Envisioning Cards: Combines both VSD and scenario planning, the Envisioning Cards are built upon a set of envisioning criteria that are intended to raise awareness of long-term and systemic issues in design. Cards provide prompts for thinking through various implications and value tensions and can be used within workshops or team meetings to trigger discussion and reflection.

It is ultimately in and through these sorts of approaches to technology development that we can try to confront and mitigate potential harms, such as hate speech and bias, before they become locked in.

Discussion

With regards to social VR applications, we can already see the emergence of several key issues that need to be addressed sooner rather than later. For example, managing user behavior in online platforms presents a serious challenge, due to the fragmented domain area and multiple systems involved. Each platform may have its own rules and guidelines, leading to conflicting values and standards. Furthermore, enforcing age restrictions becomes difficult, as users can misrepresent their age or simply bypass the restriction. Another issue is that the speed of development and deployment in the digital space often outpaces the ability to implement effective moderation measures. It is also difficult to legally mandate industry-wide standards, making it hard to establish consistent guidelines for content regulation. Testing these platforms under real-world conditions is also complicated, as the virtual environment is still relatively uncharted territory. In light of these factors, we must ensure that virtual worlds are developed responsibly, which will require continuous adaptation and collaboration between platform developers, users, and regulatory bodies.

Until now, efforts to combat hate speech have largely involved fostering dialogue, promoting education, and encouraging media literacy to enhance understanding, empathy, and respect among individuals and communities. In the case of the development of the metaverse, collaborations between governments, civil society organizations, and technology companies will also play a crucial role in developing guidelines, policies, and tools to address hate speech effectively while respecting diverse perspectives and cultural sensitivities.

It is important to acknowledge that while social VR platforms undoubtedly offer opportunities for socialization and creativity, they also face significant challenges ahead when it comes to addressing hate speech, bullying, and discrimination. As discussed, moderating content and enforcing policies are both complex tasks due to the dynamic and immersive nature of VR environments. However, drawing upon ideas and concepts from responsible innovation, implementing reporting mechanisms, educating

users, and developing community guidelines all can and more importantly *should* play a role in fostering inclusive and respectful virtual communities.

References

Breuer, J., Kowert, R., Festl, R., & Quandt, T. (2015). Sexist games= sexist gamers? A longitudinal study on the relationship between video game use and sexist attitudes. *Cyberpsychology, Behavior, and Social Networking,* 18(4), 197–202.

Buhmann, A. & Fieseler, C. (2021). Towards a deliberative framework for responsible innovation in artificial intelligence. *Technology in Society,* 64, 101475.

Ewoldsen, D. R., Eno, C. A., Okdie, B. M., Velez, J. A., Guadagno, R. E., & DeCoster, J. (2012). Effect of playing violent video games cooperatively or competitively on subsequent cooperative behavior. *Cyberpsychology, Behavior, and Social Networking,* 15(5), 277–280.

Gailey, C. W. (1993). Mediated messages: Gender, class, and cosmos in home video games. *Journal of Popular Culture,* 27(1), 81.

UN (n.d.). *What is hate speech?* https://www.un.org/en/hate-speech/understanding-hate-speech/what-is-hate-speech

Weise, M. (2021*). The Hidden, Destructive Legacy of 'Duke Nukem'.* https://www.vice.com/en/article/pkdvxb/the-hidden-destructive-legacy-of-duke-nukem

Decentralized Autonomous Organisations (DAOS) in the Metaverse: The Future of Extremist Organisation?

Julia Ebner

The transition from Web 2.0 to Web 3.0 is likely to leverage the latest technological advances including AI, machine learning and blockchain technology. This might not only bring about a new interplay of the physical, virtual and augmented reality but could also fundamentally change the ways in which corporate entities, social communities and political movements organise themselves.

Decentralisation is a key property of the Metaverse. Policies, legal contracts and financial transactions that were traditionally the domain of governments, courts and banks might be replaced with smart contracts and financial transactions in blockchain. Cryptocurrencies can be used as a new medium of exchange outside of established banking systems, while non-fungible token (NFT) might serve as unique digital identifiers to certify ownership and authenticity. Taken together, these new forms of self-governance could lead to the explosion of so-called Decentralized Autonomous Organisations (DAOs) in the Metaverse.

Decentralized Autonomous Organisations (DAOs) are digital entities that are collaboratively governed without central leadership and operate based on blockchain (Jentzsch, 2016). As such, DAOs allow internet users to establish their own organisational structures, which no longer require the involvement of a third party in financial transactions and rulemaking. DAOs allow online communities to simplify their transactions and use a community-based approach to establish rules (World Economic Forum, 2023). However, as this study will explore, they might also give rise to new threats emerging from decentralised extremist mobilisation, pose a risk to minority rights, challenge the rule of law, and disrupt institutions that are currently considered fundamental pillars of our democratic systems.

Research Aim and Methods

The aim of this chapter is to explore potential ways in which DAOs could be exploited by extremist and anti-democratic actors. To what extent are extremist movements incentivised and able to make use of such new forms of self-governance? What are the types of threats that might emerge from anti-democratic or anti-minority DAOs?

To make progress on these research questions, the chapter reviews existing literature and summarises the findings from expert interviews and digital ethnographic research. More specifically, qualitative interviews were carried out with two leading experts on DAOs, a manual review of roughly 350 existing DAOs was performed, and exploratory digital ethnographic research was carried out across the extremist fringe platforms Odysee, Bitchute and Gab as well as the encrypted apps Telegram and Discord.[1]

The chapter will first provide an overview of DAOs, their relationship with the Metaverse and current use cases. It will then assess potential areas of misuse before providing an outlook of future trends, including key challenges and opportunities, and ideas for future research areas.

DAOs and the Metaverse

The next iteration of the internet, the Metaverse, is inherently tied to decentralised forms of communication, collaboration and finance. The metaverse is built on blockchain technology. Researchers have pointed out that collective governance and decentralised finance will likely be a key characteristic of the Metaverse (Chao et al., 2022; Goldberg & Schär, 2023; Tole & Aisyahrani, 2023). Ownership over an asset or "a piece of land" in the Metaverse would be governed by NFTs (Laeeq, 2022).

The World Economic Forum described DAOs as "an experiment to reimagine how we connect, collaborate and create" (World Economic Forum, 2023, p. 27). DAOs operate differently from traditional organisations in their allocation of resources, coordination of activities and decision-making processes. Code-driven and community-oriented structures allow stakeholders to be directly involved in governance and operations via voting systems (World Economic Forum, 2023). Decentraland is the

1 All quoted and referenced primary source content was documented and archived in the form of screenshots and can be made available to fellow researchers upon request.

first large-scale virtual world based on the architecture and premises of a DAO. With the aim of distributing power among its participating users, it operates on the Ethereum blockchain. The "residents" of Decentraland can initiate policy updates based on decentralised voting systems that are embedded in Decentraland DAO governance structure (Goldberg & Schär, 2023, p. 3).

There is an active ongoing debate about the legal status of DAOs. Proponents of DAOs have argued that smart contracts are self-executing codes on blockchain that can operate independently of legal systems. They point to lower transaction costs and reduce the need for intermediaries. However, more skeptical commentators have suggested that DAOs should be owned and/or operated by humans (Jentzsch, 2016, Mondoh et al., 2022). A more detailed discussion of existing efforts to regulate DAOs can be found in the final section of this chapter.

To date, research into DAOs and the Metaverse remains scarce. However, in recent years there has been a notable rise in publications that investigate the trends and implications of DAOs for various sectors. For example, Chao et al. (2022) discuss the advantages and disadvantages of DAOs for non-profit organisations. While they argue that DAOs provide a "strong democratic system", they also caution that they "are not subject to legal, physical, or economic constraints" and are therefore capable of operating "outside the control of a single central authority or a single governing body" (Chao et al., 2022).

Mondoh et al. (2022) discuss whether DAOs might be the future of corporate governance, while Goldberg and Schär (2023) write that DAOs could disrupt the monopoly market structures in the technology sector. Their assessment is that "open standards and blockchain-based governance are a necessary but not a sufficient condition for a decentralized and neutral platform" (Goldberg & Schär, 2023). Meanwhile, Tole and Aiyahrani (2023) predict that the Metaverse and DAOs have the potential to revolutionise the education system. Through DAOs, they claim, "technology courses, certificates, and more can become automated and authenticated on the blockchain" (Tole & Aisyahrani, 2023, p. 1). DAOs could provide the basis for what they call "Metaversity", whereby the necessary infrastructure is provided for decentralised learning centers that have incentive structures and tailored courses in place (Tole & Aisyahrani, 2023, p. 3).

Most recently, computer scientists provided an in-depth analysis of the genesis and evolution of DAOs, as well as their classification and ethical implications (Amhaz et al., 2024). Another large-scale study published in 2024

argued that DAOs "resemble early online communities", in particular open-source projects. The authors assessed the impact of several DAO properties, including voting mechanisms, on levels of decentralization (Sharma et al., 2024). Moreover, the United Kingdom's Law Commission published a scoping paper on DAOs, discussing the legal characterization of DAOs, liability of participants as well as questions around financial regulation and tax (HM Law Commission, 2024).

Current Use of DAOs

The current DAO ecosystem is highly diverse and growing at a fast pace. By December 2024, DAOs collectively manage over 40 billion dollars and count more than 11 million token holders across at least 13,000 entities (DeepDAO, 2024). DAOs exist across a range of industries such as finance, philanthropy and politics. A manual review of the 356 DAOs listed on the website Decentralist.com at the time of the analysis (July 2023), illustrates the highly diverse nature and purpose of DAOs. The descriptions and stipulated mission statements include social, activist, investment, gaming, media and collector aims (Decentralist, 2023). Some DAOs are clearly satirical in nature, including the Café DAO which aims "to replace Starbucks", the Doge DAO which wants to "make the Doge meme the most recognizable piece of art in the world" and the Hair DAO, "a decentralised asset manager solving hair loss".

By 2023, DAOs have primarily attracted entrepreneurs, libertarians, activists, pranksters and hobbyists. However, there are also activist and political DAOs. For example, Decentralist.com lists DAOs that are tied to the climate movement and Black Lives Matter community as well as social justice and anti-banking DAOs that want to tackle social inequality. Figure 1 provides an overview of the most common DAOs found on Decentralist:

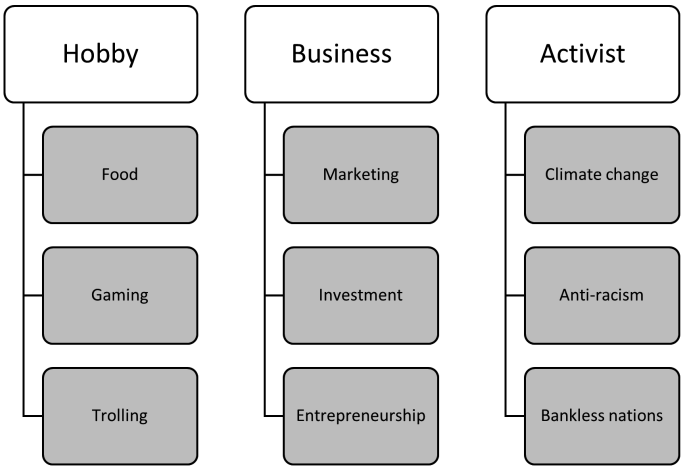

Figure 1: Overview of most common existing DAOs

While the review of existing DAOs did not identify any openly anti-democratic or anti-minority DAOs, some DAOs made use of alt-right codes and conspiracy myth references, or expressed support for far-right populist leaders. For example, the "Redacted Club DAO" claimed to be a secret network with the aim of "slaying" the "evil Meta Lizard King". The official Discord chat of the "Redacted Club DAO", which counted 850 members at the time of analysis, contained Pepe the Frog memes and references to "evil lizards" and "good rabbits". Another DAO, which was not listed on Decentralist, was called "Free Trump DAO". Its Twitter account described it as being "for Patriots" and serving as "a powerful tool that fights for freedom and liberty around the world." Its Telegram channel, which counted 474 members, was filled with Trump-glorifying memes, MAGA symbols and announcements such as "We're supporting Trump. Bring back Trump". There was also the "Trump DAO – the 47th U.S. President", which claimed to want to "raise money through the crypto to support Trump for 2024". The aim was to shape political campaigns by fundraising and giving complete anonymity to people who want to support Trump.

Extremist Incentives to Exploit DAOs

This sub-chapter used ethnographic research in extremist forums and chat groups to assess their conversations about DAOs. The analysis was

performed on the far-right fringe platforms Bitchute, Odysee and Gab as well as the encrypted apps Discord and Telegram. An exploratory discourse analysis was used, as the volume of content that is specifically focused on DAOs is still limited in extremist communities and therefore does not provide enough data points for statistical analysis.

A total of 85 pieces of content were identified across the analysed far-right fringe platforms, 46 on Odysee, 23 on Bitchute and 16 on Gab. Additionally, 100+ channels related to DAOs were detected on Telegram and Discord, including far-right extremist channels. While many of the posts on Odysee, Bitchute and Gab simply explain the characteristics and advantages of DAOs and how they might shape the Web 3.0, some of the detected conversations also reveal that there is much appetite for decentralised alternative forms of collaboration, communication and crowdfunding. Broadly speaking, the intent of far-right activists to use DAOs can be divided into two overarching types of incentives: practical and ideological motivations.

Important practical reasons are that DAOs can help users to circumvent monitoring, regulatory mechanisms and traditional institutions. For example, extremists may view them as useful to escape surveillance by security services, avoid perceived censorship by tech firms and find an alternative to frozen bank accounts. Many extremist and even terrorist movements already created their own cryptocurrencies and make use of anonymous bitcoin wallets. In particular, non-transparent cryptocurrencies such as Monero served extremists whose bank accounts have been frozen. Jihadists used cryptocurrencies as early as 2016 to fund violent activities (Irwin & Milad, 2016). In 2019, Treasury Secretary Steve Mnuchin warned that cryptocurrencies pose a national security threat, allowing malicious actors to fund criminal activities (Rappaport & Popper, 2019). While decentralised finance is already a trend among extremists, a shift to entirely decentralised forms of self-governance could be the next step (Krishnan, 2020). For example, Gab users highlighted that DAOs can help organisations to "unlock their full potential and usher in a new era of decentralized governance." The post continued: "Embrace the power of automation and embark on a journey of efficient and transparent decision-making within your DAO." Another one shared that DAOs may not be "capable of being subject to sanctions".

There are also ideological incentives that might lead extremists to use DAOs for their purposes. In particular, fundamental distrust in the establishment means that DAOs can be an appealing alternative. Users who believe that the "deep state" or the "global Jewish elites" control everything

from governments and big tech to the global banking system might prefer to set up their own technological, financial and logistical networks. For example, QAnon-related groups on Telegram were discussing the future of decentralized finance and how this might be an escape route to evade the U.S. federal banking system. DAOs also fit ultra-libertarian utopian visions of online worlds that are entirely unregulated, in which speech and actions are "truly free". "FREEDOM Meta-DAO" declared in its mission statement on Discord: "We believe in freedom of speech, privacy, and protection from cancel-culture bullying. Using a DAO as a service platform, we bring society together as a whole by removing borders and adaptation blockages." Finally, some users may be motivated by the outlook of creating their own digital state that operates based on their own rules, values and ideologies.

Practical Incentives	Ideological Incentives
Safe haven from surveillance by security services or journalists	Conspiracy myths about the establishment and financial institutions
Escape route from social media removal policies	Ultra-libertarianism and desire for unlimited free speech
Financial solution to frozen bank accounts	Creation of digital state or corporation according to own rules

Table 1: Overview of practical and ideological incentives

Most of the posts about DAOs on far-right websites were positive, however there were also a few warnings among the analysed pieces of content. For example, one user on Gab wrote that "though Blockchain technologies make traditional authoritarianism less likely, they make a new kind of authoritarianism, born of decentralized autonomous organizations (DAOs), more likely." The user continued: "By design and accident, DAOs will tend to develop into the computational equivalents of eusocial colony animals such as ants, bees, and termites. Once formed into such superorganisms, DAOs will exhibit emergent behaviors like swarming and collective intelligence." Indeed, the user pointed out that one of the risks is that "humans venturing into the DAOs' native habitat would then find themselves forced to live under the arbitrary will of not another human, but instead of a vast, mysterious hoard of nonhuman, and perhaps inhuman, entities."

Extremist Capabilities to Exploit DAOs

In this section, an analysis of potential areas for exploitation was performed based on a review of literature and interviews with two leading technology experts. The first interviewee was Carl Miller, the Director of the Centre for the Analysis of Social Media (CASM) at the think tank Demos who has long warned of potential threats emerging from the misuse of DAOs. The second interviewee was Christoph Jentzsch, a leading developer of the blockchain Ethereum and head behind "The DAO", which became one of the largest crowdfunding campaigns in history, raising over $150 million after launching in April 2016 via token sale, but was hacked soon later.

This sub-chapter addresses questions such as: Could trolling armies start cooperating via DAOs to launch election interference campaigns? What happens if anti-minority groups establish their own digital states in which they impose their own governing structures? Finally, how might terrorists leverage DAOs to fund and plot criminal activities and violent attacks?

DAOs can be used by movements to further their social, political and criminal objectives. Three potential threats were identified related to extremist and terrorist use of DAOs: 1.) influence campaigns, 2.) radicalisation of sympathisers, and 3.) attacks on political opponents. In all three potential threat areas, DAOs can be exploited in at least three ways on a tactical level: a.) by coordination and planning activities, b.) by crowdfunding and purchasing activities, and c.) by radicalisation and training activities. As such, they could change the nature of rebellion movements as well as their relative position of power, effectiveness and resilience to governmental countermeasures (Krishnan, 2020). Figures 2 and 3 summarise the potential threats as well as potential tactical exploitation areas associated with the use of DAOs by extremists and terrorists:

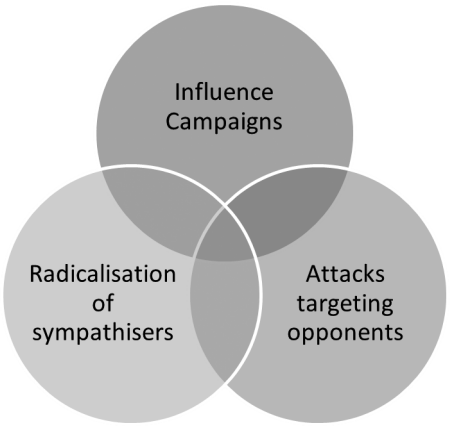

Figure 2: Overview of potential threats

Figure 3: Overview of potential tactical areas for exploitation

1 Influence Campaigns

One risk is that DAOs might be used in the future to carry out large-scale influence and election interference campaigns. Carl Miller told me that "beyond the speculative activities around crypto and NFTs, the deeper simmering experimentation around governance provided more fundamental challenges to the state and security." DAOs allow extremist groups to engage in deliberative and collective governance and decision-making in

ways where, according to Carl, it is a.) very unclear who is doing it, and b) hard for statutory security agencies to do anything about it. "As we gear up for the next U.S. presidential election, we might see DAOs being used for election campaigns," he warns. For example, Trump DAO could incentivise people to donate and participate in campaigns. "One could imagine a campaign being funded by tens of thousands of dark wallets that do not have clear links to real-world identities." Would this be recognised as electoral campaigning? Carl noted that the regulatory space in U.S. currently regulates at wallet level, not at DAO levels. However, these wallets can be fully anonymous. "There are already crypto-based gig sites where small jobs can be done anonymously based on crypto remuneration," Carl said. "In the future we might see gig workers being paid to run disinformation campaigns or stage protests."

2 Radicalisation of Sympathisers

Another risk is that DAOs could facilitate radicalisation efforts undertaken by extremist groups. Carl noted that "the basic problem of extremists is that they are being denied the conventional, easiest ways of reaching people". For example, they find it hard to rent halls, to have Facebook groups, to raise money or get the word out. He continued "You could easily see a world where there are protocol-based alternatives or replacements for conventional organizational structures", allowing extremist movements to surmount collective action problems, create international coalitions and fundraise for their activities.

Extremist movements could potentially even create their own digital states, for example in the form of digital white ethnostates or cyber caliphates. In 2022, the former chief technology officer at Coinbase Balaji Srinivasan described in his book *The Network State* how DAOs could soon give rise to new forms of digital statehood. Any group of online users could decide to start their own country, with their own laws, social services and financial transactions (Srinivasan, 2022).

3 Attacking Political Opponents

Finally, DAOs might also serve as safe havens for the planning and plotting of violent terrorist attacks and cybercrimes. The circumvention of govern-

ment regulation and monitoring activities could make them particularly useful for violent extremist and criminal organisations. Moreover, white nationalist movements have long been advocating for decentralised structures and so-called "leaderless resistance" (Michael, 2012; Malone et al., 2022). Increasingly, jihadist organisations such as Islamic State and Al Qaeda have adopted similarly decentralised models of leadership. Researchers have argued that leaderless organisations are more resilient to disruptions and interventions than hierarchical organisations. They compared hierarchical organisations to spiders who will die when you cut off their head, while the equated leaderless organisations with starfish whose legs will regrow when you cut them off (Brafman & Beckstrom, pp. 19–29).

According to Christoph Jentzsch, the most notable difference between DAOs and traditional forms of organisations is that, unlike associations or co-operatives, DAOs cannot be outlawed and their assets or bank accounts cannot be frozen. "If you can get organised without relying on traditional infrastructure, this also means that you won't be controlled by traditional institutions. It would be difficult to intervene on a statutory level". Jentzsch argues that cash already provides an alternative route for clandestine groups to fundraise, hold or spend money. Yet, DAOs might make it easier to conduct international, anonymous operations.

DAOs could give rise to new forms of the dark market. It might even be possible for DAOs to facilitate an anonymous assassination market (Krishnan, 2020). Carl noted that there would probably be law enforcement responses to terrorist plots and other serious criminal activities, which could include "a mix of infiltration and subversion, and perhaps direct cyber offensive activity." DAOs would be resilient in some ways because they are decentralised and rely on smart contracts. However, they might also be more vulnerable to hacking offenses. "It's difficult for creators to know when they are safe from a hacking attack due to their complex structures," Carl explained.

Conclusions, Future Research and Challenges on the Horizon

DAOs bring both a range of challenges and opportunities for democratic culture in cyberspace. Advocates of DAOs have argued that these blockchain-based forms of self-governance promise enhanced security, transparency and trust and reduce transaction costs arising through inter-

mediaries. Meanwhile, this chapter has explored some of the key challenges and risk areas for national security and democracy.

The potential exploitation of DAOs for extremist or criminal purposes has not received enough attention in the research and policy communities. This chapter identified a range of ways in which DAOs might be misused by extremist movements in the future, which could challenge the rule of law, pose a threat to minority groups, and disrupt institutions that are currently considered fundamental pillars of democratic systems. More specifically, the study explored how extremist movements might tap into DAOs to plan, coordinate and launch influence and interference operations, radicalisation campaigns and violent attacks.

Risks associated with the misuse of DAOs for extremist and criminal purposes has not been on the radar of global policymakers. Many governments have started to develop or pass legal frameworks to regulate AI. However, few countries or regions have even recognised the existence of DAOs or considered regulating them. Technology expert Carl Miller said that "even though DAOs behave like companies, they are not registered as legal entities". As of 2024, there are only a few exceptions: The U.S. States Wyoming, Utah, Vermont and Tennessee, as well as Switzerland, Estonia, Malta, Gibraltar, and the Republic of the Marshall Islands have passed laws to legally recognise DAOs.

This chapter understands the potential risk areas of DAOs in the Metaverse. While this study focused on non-state actors, exploring emerging threats from hostile state actors might be another important area for future research. Studies could also use experimental methods and interviews with policymakers and law enforcement to map the threats landscape. "DAOs can be used by anyone: by charities, investment funds, but yes, they can also be used by terrorist organisations," Christoph Jentzsch argues. However, he believes that the positive cases of use significantly outweigh the negative ones. Future studies should investigate both the potential opportunities and challenges related to collective decision-making and self-governance, diversity and the protection of minorities, radicalisation and extremism in decentralized communities in the Metaverse. The positive ways in which DAOs can shape future democratic culture are just as poorly understood as the negative impact they could have on politics and society.

References

Amhaz, R., Bobenrieth, C., & Marz, M. (2014). The impact of decentralised autonomous organisations (DAO) on Society 5.0. In *5th International Conference on Machine Learning, IOT and Blockchain*. https://aircconline.com/csit/papers/vol14/csit140301.pdf

Brafman, O., & Beckstrom, R. A. (2007). *The starfish and the spider: The unstoppable power of leaderless organizations*. Penguin.

Chao, C.-H., Ting, I.-H., Tseng, Y.-J., Wang, B.-W., Wang, S.-H., Wang, Y.-Q., & Chen, M.-C. (2022). The study of decentralized autonomous organizations (DAO) in social network. In *Proceedings of the 9th Multidisciplinary International Social Networks Conference* (pp. 59–65).

Decentralist. (2023). The list of DAOs. *Decentralist*. https://www.decentralist.com/list-of-daos

DeepDAO. (2024). Organizations. *DeepDAO*. https://deepdao.io/organizations

Goldberg, M., & Schär, F. (2023). Metaverse governance: An empirical analysis of voting within decentralized autonomous organizations. *Journal of Business Research*, 160.

HM Government, Law Commission. (2024, July). *Decentralised autonomous organisations (DAOs): A scoping paper*. https://lawcom.gov.uk/project/decentralised-autonomous-organisations-daos/#related

Irwin, A. S. M., & Milad, G. (2016). The use of crypto-currency in funding violent jihad. *Journal of Money Laundering Control*, 19(4), 411.

Jentzsch, C. (2016). *Decentralized autonomous organization to automate governance* [White paper].

Krishnan, A. (2020). Blockchain empowers social resistance and terrorism through decentralized autonomous organizations. *Journal of Strategic Security*, *13*(1).

Laeeq, K. (2022, February). *Metaverse: Why, How and What* [Presentation]. https://www.researchgate.net/publication/358505001_Metaverse_Why_How_and_What

Malone, I., Blasco, L., & Robinson, K. (2022, September). Fighting the Hydra: Combatting vulnerabilities in online leaderless resistance networks. *National Counterterrorism, Innovation, Technology and Education Center, U.S. Department of Homeland Security*. https://digitalcommons.unomaha.edu/nctereportsresearch/24/

Michael, G. (2012). *Lone wolf terror and the rise of leaderless resistance*. Vanderbilt University Press.

Mondoh, B. S., Johnson, S. M., Green, M., & Georgopoulos, A. (2022, June 23). Decentralised autonomous organisations: The future of corporate governance or an illusion? http://dx.doi.org/10.2139/ssrn.4144753

Rappaport, A., & Popper, N. (2019, July 15). Cryptocurrencies pose national security threat, Mnuchin said. *New York Times*. https://www.nytimes.com/2019/07/15/us/politics/mnuchin-facebook-libra-risk.html

Sharma, T., Potter, Y., Pongmala, K., Wang, H., Miller, A., Song, D., & Wang, Y. (2024). Future of algorithmic organization: Large scale analysis of decentralized autonomous organizations (DAOs). In *Proceedings of ACM* (New York, US). https://arxiv.org/pdf/2410.13095

Srinivasan, B. (2022). *The network state: How to start a new country*. Self-published.

Sutikno, T., & Aisyahrani, A. I. B. (2023). Non-fungible tokens, decentralized autonomous organizations, Web 3.0, and the metaverse in education: From university to metaversity. *Journal of Education and Learning*, 17(1), 1–5.

Venis, S. (2022, July 5). Could new countries be founded – on the internet? *The Guardian*. https://www.theguardian.com/commentisfree/2022/jul/05/could-new-countries-be-founded-on-the-internet

World Economic Forum. (2023, January). *Decentralized Autonomous Organizations Toolkit* [Insight Report]. https://www3.weforum.org/docs/WEF_Decentralized_Autonomous_Organization_Toolkit_2023.pdf

Narrative, Creative – Immersive? The Immersive Potential of Right-Wing Extremist Communication on Social Media

Sandra Kero & Josephine B. Schmitt

Introduction

Social media platforms like Instagram or TikTok are popular tools of right-wing actors. As political outsiders, they exploit the opportunities that social media offers to establish themselves as independent voices outside journalistic media for the purpose of orchestrating their narratives, channelling their political opinions and ideological attitudes and recruiting new members (Fielitz & Marcks, 2020; Rau et al., 2022; Schmitt, Harles, et al., 2020; Schwarz, 2020). On Instagram, right-wing groups deliberately rely on creators, lifestyle and a connection to nature to make their ideology more easily digestible (Echtermann et al., 2020). As a result, ideological content is not only subtly presented – often, the corresponding ideological perspectives are explicitly stated (Kero, 2023). But also TikTok, which is particularly popular among young users (mpfs, 2022), has become increasingly important as a platform for the far-right to interact with target groups that have not yet formed firm political views (pre:bunk, 2023; Hohner et al., 2024). The range of content provided is broad here too: alongside supposedly humorous and musical content, there is also material from AfD politicians who showcase their beliefs. Extreme right-wing ideas are made accessible through normalisation strategies, e.g. pseudo-scientific misinformation and emotionalisation (Müller, 2022).

Social media platforms not only give everyone the opportunity to enter the public political discourse – the disseminated content also allows boundaries between political information, entertainment and social issues to become blurred. At the same time, various functions of social media facilitate users' *immersive* experiences – something that far-right extremists in particular can benefit from in their communication practices.

This article aims to contextualise the immersiveness of extreme right-wing content on social media and present selected immersive mechanisms within these environments as examples. On the basis of this, preventive measures are identified.

What Do we Understand by Immersion?

Immersion essentially means being deeply absorbed or engrossed in a particular environment, activity or experience (Murray, 1998). Immersion is often mentioned as a feature of digital technologies such as virtual reality (VR) or augmented reality (AR), or in the gaming industry (Mühlhoff & Schütz, 2019; Nilsson et al., 2016). From a media theory and media psychology perspective, the term primarily describes the state of dissolution of physical and fictitious boundaries, i.e. users' subjective experiences of plunging into an imaginary world or narrative environment and being emotionally involved in it (e.g. E. Brown & Cairns, 2004; Haywood & Cairns, 2006). This is usually associated with a strong feeling of presence in which perception and attention are focused on the specific immersive content. Conversely, attention to time and (real) space decreases for the period of immersion (Cairns et al., 2014; Curran, 2018). Immersive environments and mechanisms can encourage engagement and learning (Dede, 2009), but can also promote the ideological effect of content among users (Braddock & Dillard, 2016).

Moving away from an understanding of the term that is limited to a technically or culturally versed or psychological phenomenon, Mühlhoff and Schütz (2019) consider immersion from an affect theory and social theory perspective. Here, immersion is described as a dynamic interaction between (non)human individuals and between them and their environment, i.e. as "a specific mode of emotional and affective involvement in a current or mediatised social event" (p. 20). Affects are spontaneous, instinctive reactions that can be influenced both individually and socially and which shape a person's behaviour (Strick, 2021, among others). The design of certain (media) environments can influence affective reactions in specific ways and thus also modulate power relations and social dynamics (Mühlhoff & Schütz, 2019). In this sense, certain immersion offerings can help to control and regulate behaviour, thus creating "immersive power" (p. 30). Immersion can therefore be considered a form of situational influence exerted on the thoughts and feelings of the individual. It represents an indirect exercise of power that arises without hierarchies and through social contexts.

Immersion is considered in a wide range of scientific disciplines (e.g. psychology, computer science, cognitive science, design); therefore, the synonyms and closely related concepts are just as diverse. The literature thus includes terms such as *presence, flow, involvement* and *engagement*,

which describe similar phenomena (e.g. Curran, 2018; Nilsson et al., 2016). To describe psychological immersion in the context of the reception of narrative media content, the term *transportation* is used in media psychology and communication science research (e.g. Moyer-Gusé, 2008). Immersion as a subjective experience of users is not limited to a specific medium or technology. Learning situations, books, films and social media content can therefore also have an immersive effect, as well as computer games or VR applications.

Against the backdrop of this integrated understanding of the term, we focus in this article on immersive mechanisms in and through social media and their content. We are therefore primarily concerned with the social and psychological dimension of immersiveness.

Immersive Mechanisms in Social Media

Social media can contribute to immersive mechanisms and effects in various ways. With an understanding of immersion from a media theory and media psychology perspective, these experiences relate to the constant presence of smartphones and, with them, social media applications in our everyday lives, resulting in a merging of virtual, physical and social spaces. Users' emotional involvement is also intensified through the primary type of use of social media platforms – such as direct sharing of everyday experiences, interests and opinions – and the (psychological) connection with individual producers of this content, hereinafter referred to as creators, and their posts. From a social theory perspective, a power structure is created here in which affective dynamics can be used in a targeted way to influence the behaviour, attitudes and perceptions of individuals.

In the following, we aim to take a closer look at the underlying mechanisms in order to thoroughly examine the immersive potential of social media in the communication of extreme right-wing creators. A distinction is made between two contexts: the platform environment as an environment of social events, and the interaction between individuals and content in this environment.

Presenting the Persona and the Story

In the competition for the undivided attention of users, various far-right extremists are also taking to social media. The amount of extreme right-

wing content has been steadily increasing for years (e.g. Munger & Phillips, 2022). Opinions are becoming things that can be sold. Social media marketing strategies are being used to attract and retain the attention of users (Cotter, 2019) and shape opinions in line with right-wing ideology. How a particular media persona is presented and the way in which the stories are told play a crucial role when it comes to the immersiveness and impact of the content provided.

Several creators and influencers have emerged from the far-right scene in recent years. They are an essential prerequisite for reaching users, especially those who are otherwise not interested in political content. Generally speaking, influencers are previously unknown social media users who become well-known personalities through careful self-promotion and regular posting of content on social media, and who cover a range of topics on their channels (Bauer, 2016).

Female activists in particular use popular social media platforms to present extreme right-wing ideology to a wide audience in a personal and emotive way through recipes, beauty tips and inspiring landscape photos (Ayyadi, 2021; Echtermann et al., 2020; Kero, 2023); the boundaries between entertainment, lifestyle and politics are fluid. The creators act as role models (Zimmermann et al., 2022), are opinion leaders as regards content that threatens democracy (Harff et al., 2022), can attract users to content and channels, and encourage interaction via personal connections (Leite et al., 2022). The greater the trustworthiness and centrality of the creators, the more the recipients report immersive experiences with the media offering (Jung & Im, 2021).

Figure 1: TikTok channel "victoria" – a mix of nature, homeland and historical scenes

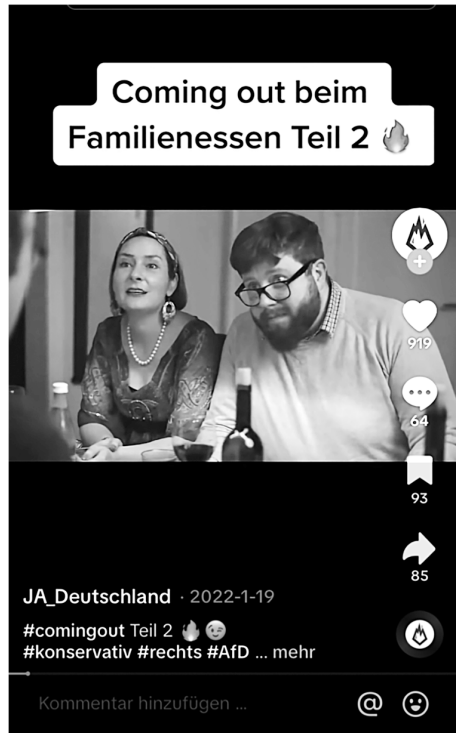

Figure 2: The youth organisation of the AfD – classified as an extremist group that threatens the constitution – posts satirical clips on TikTok

Above all, it is the communicative presentation styles of the content producers that let recipients really immerse themselves in their everyday world. They tell stories of nature, homeland and a sovereign German people, and provide simple answers to complex political questions. Concepts of enemies as those responsible for social problems are soon found and quickly identified based on supposedly distinct characteristics (see Figure 1). Through their narratives, they not only engage to create identity and meaning, but also offer recipients a strong group that provides interested individuals with a framework and (political) orientation. Formats with an emphasis on humour, pop culture and youth culture are used to make content easy

to digest and build recipients' loyalty to channels and creators (Schmitt, Harles, et al., 2020, see also Figure 2). The network of right-wing actors on social media is dense; they refer to each other's stories and narratives – including across platforms (Chadwick & Stanyer, 2022). An apparent consistency of narratives, i.e. the same story is told by several actors, makes a story appear even more credible.

Mechanisms of Narrative Persuasion Are Used for Specific Purposes

Research in the fields of media psychology and communication science identifies three mechanisms that make stories convincing and thus make right-wing extremist communication particularly immersive and effective: transportation, identification and parasocial interactions (e.g. Braddock, 2020).

Transportation describes the process whereby; in order to understand the narrative presented in a story, people must shift their attention from the real world around them to the world constructed in the narrative. Ideally, because of this psychological immersion, they lose their awareness of the real world and completely immerse themselves in the fictional world (Braddock, 2020). More engagement with the narrative world means less questioning of persuasive information (Igartua & Cachón-Ramón, 2023; Moyer-Gusé, 2008).

Identification means the adoption of the media character's viewpoint or perspective by recipients. This happens, for example, where the media figure is perceived as particularly similar.

Parasocial interaction is a psychological process in which media users like and/or trust a media figure so much that they feel as if they are connected to them. Where this happens over a longer period, e.g. several videos or articles, it is also called a parasocial relationship. A parasocial interaction or relationship with a media figure reduces users' reactance to the content presented and the desire to object to content. This facilitates the adoption of beliefs, attitudes and behaviours (Braddock, 2020).

Through narrative storytelling, the direct presentation of everyday situations, self-disclosure and directly addressing recipients, right-wing creators can increase the level of their authenticity and reinforce parasocial relationships with their followers. The more socially attractive a media figure is perceived to be (Masuda et al., 2022) and the greater the recipients' identification with the media figure (Eyal & Dailey, 2012), the stronger the

parasocial relationship. Social media creators offer young users in particular great potential for identification by presenting themselves as similar to them and approachable (Farivar & Wang, 2022; Schouten et al., 2020). An example of this is the account belonging to Freya Rosi, a young, right-wing creator. With tips on beauty, baking and cooking, and photos of nature, she orchestrates the image of an approachable, traditional young far-right supporter who loves her homeland (a detailed analysis of the account is also found in Rösch, 2023). A particularly close and enduring parasocial relationship with social media protagonists can even lead to addiction-like behaviours among users (de Bérail et al., 2019). In other words: users have difficulty escaping the fictional world manufactured by the social media creators – in this context their content has a highly immersive effect.

Figure 3: Screenshots of Freya Rosi's Instagram account. With tips on beauty, baking and cooking and photos of nature, the creator orchestrates the image of an approachable, traditional young far-right supporter who loves her homeland. A detailed analysis of the presentation strategies can also be found in Rösch, 2023

Self-disclosure by creators can lead to a higher level of parasocial interactions and relationships among followers via the feeling of social presence,

i.e. perceiving the media figure as a natural person (Kim et al., 2019; Kim & Song, 2016), and to greater engagement with and commitment to communicators and their content (Osei-Frimpong & McLean, 2018). In turn, the feeling of social presence makes it less likely that people will check the truth of information (Jun et al., 2017). This is particularly problematic when dealing with extreme right-wing content, e.g. disinformation or conspiracy theories.

Emotionalisation of Content

In the context of their communication, radicalised communicators talk about fears, uncertainties and developmental tasks; they also use emotionalising content and images to make their content relatable and convince followers of their world view (Frischlich, 2021; Frischlich et al., 2021; Schmitt, Harles, et al., 2020; Schneider et al., 2019). In the so-called post-truth era, emotions, as well as perceived truths, are becoming a key manipulation tool of extreme right-wing creators. An experimental study by Lillie and colleagues (2021) indicates that narrative content which triggers *fear* among recipients leads to more flow experiences – i.e. the complete absorption of users in the reception of the content – and encourages behavioural intentions; on the other hand, fear reduces the willingness to contradict content. This makes it appear more convincing.

The development of social media into platforms for creating, publishing and interacting with (audio-)visual content supports the communication and effect of this sort of content. In particular, (audio-)visual content is suitable for conveying emotional messages (Houwer & Hermans, 1994) and thus generating corresponding emotional responses. In turn, this affects information processing, as the brain pays more attention to processing the emotional reaction than to the information (Yang et al., 2023). Tritt et al. (2016) found that politically conservative people appear to respond more easily to emotional stimuli than liberals.

The immersive effects outlined so far may result in followers becoming more intensely involved in the world portrayed by creators, forging a deeper emotional bond and identifying more strongly with the content presented (Feng et al., 2021; Hu et al., 2020; Slater & Rouner, 2002; Wunderlich, 2023).

The Role of Platform Functions

Time plays an important role on social media. Users' attention must be captured quickly to attract them to the platform. It is fair to assume that immersive effects of communication on social media are intensified through the design and functional logic of the platforms. Platform-specific algorithms and dark patterns play a special role here.

Algorithms Determine Relevance and Importance

In addition to highlighting the relevance and importance of certain content, algorithms encourage the emotional involvement of users by preferentially showing them content that they or similar users potentially prefer. This is designed to attract users to the platform and content and retain them for as long as possible; it is part of the platforms' business model. In this way, platforms let users plunge deeper and deeper into the virtual worlds. This makes them an immersive and effective recruitment tool. The algorithmic selection and sorting of content depends on various factors including the type of content, its placement, the language used, the degree of networking between those disseminating it and the responses from users (Peters & Puschmann, 2017).

Commercial platforms prioritise posts that trigger emotional reactions, polarise opinions or promote intensive communication between members (Grandinetti & Bruinsma, 2023; Huszár et al., 2022; Morris, 2021). In particular, visual presentation modes such as photos, videos and memes trigger affective emotions and reactions among recipients and draw them in (Maschewski & Nosthoff, 2019). Algorithmic networking provides technical reinforcement of users' emotional connection and involvement with the content. This also applies to contact with anti-democratic content. Studies suggest that social media algorithms can facilitate the spread of extreme right-wing activities (Whittaker et al., 2021; Yesilada & Lewandowsky, 2022) – sometimes similar keywords are enough, as with unproblematic content (Schmitt et al., 2018). In particular, people who are interested in niche topics (e.g. conspiracy theories) (Ledwich et al., 2022), and who follow platform recommendations for longer (M. A. Brown et al., 2022) are manoeuvred into corresponding filter bubbles by the algorithms. At the same time, however, most users are shown moderate content (M. A. Brown et al., 2022; Ledwich et al., 2022).

In recent years, TikTok has faced scrutiny particularly because of its strong algorithmic control. This makes it even more likely that users will unintentionally encounter radicalised content (Weimann & Masri, 2021). In addition, the stream of content continues to run if users do not actively interrupt it, thus creating a much more immersive experience on the platform compared to other social media (Su, Zhou, Gong, et al., 2021; Su, Zhou, Wang, et al., 2021). Sometimes the platform is even said to have addictive potential due to its functions (e.g., Qin et al., 2022). It is difficult to break away from, and this poses a great danger for users, especially given the increase in extremist communication on TikTok.

Dark Patterns as Mechanisms to Enhance Immersion

Dark patterns also play a role in discussions around the immersive effect of social media. These describe platforms' design decisions that use deceptive or manipulative tactics to keep users on the platforms and encourage them to make (uninformed) usage decisions (Gray et al., 2023). On social media, dark patterns are primarily aimed at retaining users' attention for long periods, ensuring excessive and immersive use of the platforms and getting users to perform actions (e.g. likes). Examples of dark patterns are: activity notifications, the major challenge on Facebook of *really* logging out of the platform or even deleting an account, or the often counter-intuitive colours of cookie notifications which mean that users accept all of them, including unnecessary ones. Nearly all popular platforms and applications have dark patterns; many users do not notice them (Bongard-Blanchy et al., 2021; Di Geronimo et al., 2020). Creators also aim to influence user behaviour through dark patterns. For example, they manipulate popularity metrics (e.g. likes, number of followers) and images to try to artificially boost their credibility and generate attention for their content (Luo et al., 2022).

What Is an Effective Preventive Response?

The immersiveness of extremist social media communication poses a major challenge in terms of preventive measures. This relates to the many different levels that must be considered in this context. In the following, the preventive measures are specified at a) content level, b) platform level and c) media level. To clearly define the measures from the users' point of

view, we will use the three prevention levels of *awareness, reflection and empowerment* (Schmitt, Ernst, et al., 2020). Awareness includes a general consideration or a general awareness (e.g. of one's own usage behaviour, extremist narratives, platform functioning), while reflection refers to critical reflection on content and the functioning of the media. Empowerment, however, means the ability of recipients to position themselves as regards certain media content or functions and be capable of acting.

At content level (a) awareness must be raised about extremist content and its specific digital forms of communication and representation on social media. This is therefore about creating awareness of how right-wing extremist creators communicate in order to gain the attention of social media users, generate interest in their narratives and ultimately motivate people to act in accordance with the far-right world view (Schmitt, Ernst, et al., 2020). Furthermore, users should be enabled to take a critical stance as regards extremist content so they can position themselves accordingly in social discourse. This positioning does not only have to take place in the context of political discussions – the fact of reporting content to the platform itself or to one of the common reporting platforms (e.g. Jugendschutz.net) also indicates a stance. As well as a strong ability to take a critical view of media, it is particularly important here to have historic, intercultural and political knowledge. Tolerance of ambiguity, which enables people to tolerate complex and possibly contradictory information, is also important.

Platform-related prevention measures (b) should enable users to be aware, understand and reflect on algorithmic functional logic (e.g. what am I being shown and why?) and raise awareness of the mechanisms of action of the platform design (e.g. how is content being shown to me? How is that affecting my actions?) (Di Geronimo et al., 2020; Silva et al., 2022; Taylor & Brisini, 2023). For a school context, there is, for example, the second learning package in the CONTRA lesson series (Ernst et al., 2020) which has been explicitly designed to facilitate the ability to criticise media with regard to algorithms.

On the platform side, it is important to create transparency around algorithms. Disclosure of the platform design, such as clarifying the functional logic of the recommendation algorithm, can help people identify dark patterns – and thus also understand the potential spread dynamics of extremist messages and counteract these (Rau et al., 2022). In a social media context, dark patterns have been relatively poorly researched to date (Mildner et al., 2023). In this field, information from research activities is provided, for

example, by the Dark Pattern Detection Project, which also enables users to report dark patterns.

Political regulatory measures, such as those in the Network Enforcement Act (NetzDG) and the Digital Services Act (DSA), can also help as regards the disclosure of platform design and require transparency reports to be provided, for example. With a view to potential new virtual media environments – such as the metaverse – it is essential to enforce transparency rules and consider new forms of design. In what circumstances does the virtual interaction take place? How are the digital activity spaces designed? What criteria are used for the visual representation of the actors, e.g., avatars, and how inclusive is their design?

On the media level (c), it is crucial to raise users' awareness of their own media behaviour and encourage them to reflect on it. Among other things, this involves questions about the time spent using certain applications (e.g. how long have I spent on TikTok today?) and also the type of use[1] (e.g. what content and mechanisms are drawing me in?). Help could be provided by apps that set time limits or send warnings about the usage times of other apps or block them (e.g. StayFree or Forest). The digital well-being function, which can measure the usage time of individual apps, is already established in various smartphone brands.

Summary and Outlook

The statements above illustrate the levels on which the social media communication of far-right creators can have an immersive effect from a social and psychological perspective. This occurs both as a result of the communication styles they choose and through the platforms themselves. When these are combined, a parallel world is created – a metaverse – that attracts users in a variety of ways and can therefore have an effect.

From a research perspective, several unanswered questions arise. For example, TikTok, as an extremely popular platform, should be examined more thoroughly regarding its immersive potential. The new possibilities offered by AI-based image generation tools also raise questions about the effect and immersiveness of synthetic imagery produced by extremist communicators. The topic of gaming and right-wing extremism has also gained a lot more attention in recent years (e.g. Amadeu Antonio Stiftung, 2022;

1 Tools such as Dataskop can show TikTok use.

Schlegel, no date; Schlegel, 2021). Becoming absorbed in virtual worlds is a key motivation for gamers (for an overview see, for example, Cairns et al., 2014). Currently, there are few reliable findings regarding exactly how far-right actors use gaming for their own purposes, what role avatars, 3D worlds and VR play here and what impact this can have in the context of radicalisation processes. However, some initial findings are available on how interactive games and VR can be used for prevention purposes (e.g. Bachen et al., 2015; Dishon & Kafai, 2022). Through a fun and immersive approach, more abstract topics like politics and democracy – including their controversial aspects – can be explored in an engaging and cooperative way. This makes it possible for people to take different points of view and learn and practise interactions that are also relevant in everyday life. Games can also promote knowledge, empathy and critical thinking. This potential should be explored more fully by researchers and those working in prevention. Although we have outlined a selection of considerations for the prevention of immersive extremist communication, those in the field of prevention are asked to address the various facets of immersiveness in greater detail and possibly use comparable mechanisms for their own goals, taking into account Germany's prohibition on overwhelming students (*Überwältigungsverbot*) (bpb, 2011; for criticism of the Beutelsbach Consensus see also Widmaier & Zorn, 2016).

References

Amadeu Antonio Stiftung. (2022). *Unverpixelter Hass—Gaming zwischen Massenphänomen und rechtsextremen Radikalisierungsraum?*. https://www.amadeu-antonio-stiftung.de/neue-handreichung-unverpixelter-hass-gaming-zwischen-massenphaenomen-und-rechtsextremen-radikalisierungsraum-81173/

Ayyadi, K. (2021, August 25). Rechte Influencerinnen: Rechtsextreme Inhalte schön verpackt. *Belltower.News*. https://www.belltower.news/rechte-influencerinnen-rechtsextreme-inhalte-schoen-verpackt-120301/

Bachen, C. M., Hernández-Ramos, P. F., Raphael, C., & Waldron, A. (2015). Civic play and civic gaps: Can life simulation games advance educational equity? *Journal of Information Technology & Politics*, 12(4), 378–395. https://doi.org/10.1080/19331681.2015.1101038

Bongard-Blanchy, K., Rossi, A., Rivas, S., Doublet, S., Koenig, V., & Lenzini, G. (2021). "I am definitely manipulated, even when I am aware of it. It's ridiculous!"—Dark patterns from the end-user perspective. In *Proceedings of the 2021 ACM Designing Interactive Systems Conference* (pp. 763–776). https://doi.org/10.1145/3461778.3462086

bpb. (2011, April 7). Beutelsbacher Konsens. *bpb.de*. https://www.bpb.de/die-bpb/ueber-uns/auftrag/51310/beutelsbacher-konsens/

Braddock, K. (2020). Narrative persuasion and violent extremism: Foundations and implications. In J. B. Schmitt, J. Ernst, D. Rieger, & H.-J. Roth (Eds.), *Propaganda und Prävention: Forschungsergebnisse, didaktische Ansätze, interdisziplinäre Perspektiven zur pädagogischen Arbeit zu extremistischer Internetpropaganda* (pp. 527–538). Springer Fachmedien. https://doi.org/10.1007/978-3-658-28538-8_28

Braddock, K., & Dillard, J. P. (2016). Meta-analytic evidence for the persuasive effect of narratives on beliefs, attitudes, intentions, and behaviors. *Communication Monographs*, 83(4), 446–467. https://doi.org/10.1080/03637751.2015.1128555

Brown, E., & Cairns, P. (2004). A grounded investigation of game immersion. In *CHI '04 Extended Abstracts on Human Factors in Computing Systems* (pp. 1297–1300). https://doi.org/10.1145/985921.986048

Brown, M. A., Bisbee, J., Lai, A., Bonneau, R., Nagler, J., & Tucker, J. A. (2022). Echo chambers, rabbit holes, and algorithmic bias: How YouTube recommends content to real users (SSRN Scholarly Paper 4114905). https://doi.org/10.2139/ssrn.4114905

Cairns, P., Cox, A., & Nordin, A. I. (2014). Immersion in digital games: Review of gaming experience research. In *Handbook of Digital Games* (pp. 337–361). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781118796443.ch12

Chadwick, A., & Stanyer, J. (2022). Deception as a bridging concept in the study of disinformation, misinformation, and misperceptions: Toward a holistic framework. *Communication Theory*, 32(1), 1–24. https://doi.org/10.1093/ct/qtab019

Cotter, K. (2019). Playing the visibility game: How digital influencers and algorithms negotiate influence on Instagram. *New Media & Society*, 21(4), 895–913. https://doi.org/10.1177/1461444818815684

Curran, N. (2018). Factors of immersion. In *The Wiley Handbook of Human Computer Interaction* (pp. 239–254). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781118976005.ch13

de Bérail, P., Guillon, M., & Bungener, C. (2019). The relations between YouTube addiction, social anxiety and parasocial relationships with YouTubers: A moderated-mediation model based on a cognitive-behavioral framework. *Computers in Human Behavior*, 99, 190–204. https://doi.org/10.1016/j.chb.2019.05.007

Dede, C. (2009). Immersive interfaces for engagement and learning. *Science*, 323(5910), 66–69. https://doi.org/10.1126/science.1167311

Di Geronimo, L., Braz, L., Fregnan, E., Palomba, F., & Bacchelli, A. (2020). UI dark patterns and where to find them: A study on mobile applications and user perception. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (pp. 1–14). https://doi.org/10.1145/3313831.3376600

Dishon, G., & Kafai, Y. B. (2022). Connected civic gaming: Rethinking the role of video games in civic education. *Interactive Learning Environments*, *30*(6), 999–1010. https://doi.org/10.1080/10494820.2019.1704791

Echtermann, A., Steinberg, A., Diaz, C., Kommerell, C., & Eckert, T. (2020). Kein Filter für Rechts. *correctiv.org*. https://correctiv.org/top-stories/2020/10/06/kein-filter-fuer-rechts-instagram-rechtsextremismus-frauen-der-rechten-szene/?lang=de

Ernst, J., Schmitt, J. B., Rieger, D., & Roth, H.-J. (2020). #weARE – Drei Lernarrangements zur Förderung von Medienkritikfähigkeit im Umgang mit Online-Propaganda in der Schule. In J. B. Schmitt, J. Ernst, D. Rieger, & H.-J. Roth (Eds.), *Propaganda und Prävention: Forschungsergebnisse, didaktische Ansätze, interdisziplinäre Perspektiven zur pädagogischen Arbeit zu extremistischer Internetpropaganda* (pp. 361–393). Springer Fachmedien. https://doi.org/10.1007/978-3-658-28538-8_17

Eyal, K., & Dailey, R. M. (2012). Examining relational maintenance in parasocial relationships. *Mass Communication and Society*, 15(5), 758–781. https://doi.org/10.1080/15205436.2011.616276

Farivar, S., & Wang, F. (2022). Effective influencer marketing: A social identity perspective. *Journal of Retailing and Consumer Services*, 67, 103026. https://doi.org/10.1016/j.jretconser.2022.103026

Feng, Y., Chen, H., & Kong, Q. (2021). An expert with whom I can identify: The role of narratives in influencer marketing. *International Journal of Advertising*, 40(7), 972–993. https://doi.org/10.1080/02650487.2020.1824751

Fielitz, M., & Marcks, H. (2020). *Digitaler Faschismus: Die sozialen Medien als Motor des Rechtsextremismus*. Dudenverlag.

Frischlich, L., Hahn, L., & Rieger, D. (2021). The promises and pitfalls of inspirational media: What do we know, and where do we go from here? *Media and Communication*, 9(2), 162–166. https://doi.org/10.17645/mac.v9i2.4271

Grandinetti, J., & Bruinsma, J. (2023). The affective algorithms of conspiracy TikTok. *Journal of Broadcasting & Electronic Media*, 67(3), 274–293. https://doi.org/10.1080/08838151.2022.2140806

Gray, C. M., Sanchez Chamorro, L., Obi, I., & Duane, J.-N. (2023). Mapping the landscape of dark patterns scholarship: A systematic literature review. In *Companion Publication of the 2023 ACM Designing Interactive Systems Conference* (pp. 188–193). https://doi.org/10.1145/3563703.3596635

Harff, D., Bollen, C., & Schmuck, D. (2022). Responses to social media influencers' misinformation about COVID-19: A pre-registered multiple-exposure experiment. *Media Psychology*, 25(6), 831–850. https://doi.org/10.1080/15213269.2022.2080711

Haywood, N., & Cairns, P. (2006). Engagement with an interactive museum exhibit. In T. McEwan, J. Gulliksen, & D. Benyon (Eds.), *People and Computers XIX — The Bigger Picture* (pp. 113–129). Springer. https://doi.org/10.1007/1-84628-249-7_8

Hohner, J., Kakavand, A. E., & Rothut, S. (2024). Analyzing radical visuals at scale: How far-right groups mobilize on TikTok. *Journal of Digital Social Research*, 6(1). https://doi.org/10.33621/jdsr.v6i1.200

Houwer, J. D., & Hermans, D. (1994). Differences in the affective processing of words and pictures. *Cognition and Emotion*, 8(1), 1–20. https://doi.org/10.1080/02699939408408925

Hu, L., Min, Q., Han, S., & Liu, Z. (2020). Understanding followers' stickiness to digital influencers: The effect of psychological responses. *International Journal of Information Management*, 54, 102169. https://doi.org/10.1016/j.ijinfomgt.2020.102169

Huszár, F., Ktena, S. I., O'Brien, C., Belli, L., Schlaikjer, A., & Hardt, M. (2022). Algorithmic amplification of politics on Twitter. *Proceedings of the National Academy of Sciences*, 119(1), e2025334119. https://doi.org/10.1073/pnas.2025334119

Igartua, J.-J., & Cachón-Ramón, D. (2023). Personal narratives to improve attitudes towards stigmatized immigrants: A parallel-serial mediation model. *Group Processes & Intergroup Relations*, 26(1), 96–119. https://doi.org/10.1177/13684302211052511

Jun, Y., Meng, R., & Johar, G. V. (2017). Perceived social presence reduces fact-checking. *Proceedings of the National Academy of Sciences*, 114(23), 5976–5981. https://doi.org/10.1073/pnas.1700175114

Jung, N., & Im, S. (2021). The mechanism of social media marketing: Influencer characteristics, consumer empathy, immersion, and sponsorship disclosure. *International Journal of Advertising*, *40*(8), 1265–1293. https://doi.org/10.1080/02650487.2021.1991107

Kero, S. (2023). Jung, weiblich und extrem rechts. Die narrative Kommunikation weiblicher Akteurinnen auf der Plattform Instagram. *easy_social sciences*, *69*, 2836. https://doi.org/10.15464/EASY.2023.10

Kim, J., Kim, J., & Yang, H. (2019). Loneliness and the use of social media to follow celebrities: A moderating role of social presence. *The Social Science Journal*, 56(1), 21–29. https://doi.org/10.1016/j.soscij.2018.12.007

Kim, J., & Song, H. (2016). Celebrity's self-disclosure on Twitter and parasocial relationships: A mediating role of social presence. *Computers in Human Behavior*, 62, 570–577. https://doi.org/10.1016/j.chb.2016.03.083

Ledwich, M., Zaitsev, A., & Laukemper, A. (2022). Radical bubbles on YouTube? Revisiting algorithmic extremism with personalised recommendations. *First Monday*. https://doi.org/10.5210/fm.v27i12.12552

Leite, F. P., Pontes, N., & Baptista, P. de P. (2022). Oops, I've overshared! When social media influencers' self-disclosure damage perceptions of source credibility. *Computers in Human Behavior*, 133, 107274. https://doi.org/10.1016/j.chb.2022.107274

Lillie, H. M., Jensen, J. D., Pokharel, M., & Upshaw, S. J. (2021). Death narratives, negative emotion, and counterarguing: Testing fear, anger, and sadness as mechanisms of effect. *Journal of Health Communication*, 26(8), 586–595. https://doi.org/10.1080/10810730.2021.1981495

Luo, M., Hancock, J. T., & Markowitz, D. M. (2022). Credibility perceptions and detection accuracy of fake news headlines on social media: Effects of truth-bias and endorsement cues. *Communication Research*, 49(2), 171–195. https://doi.org/10.1177/0093650220921321

Maschewski, F., & Nosthoff, A.-V. (2019). Netzwerkaffekte. Über Facebook als kybernetische Regierungsmaschine und das Verschwinden des Subjekts. In R. Mühlhoff, A. Breljak, & J. Slaby (Eds.), *Affekt Macht Netz: Auf dem Weg zu einer Sozialtheorie der Digitalen Gesellschaft* (1st ed., Vol. 22, pp. 55–80). transcript Verlag. https://doi.org/10.14361/9783839444399

Masuda, H., Han, S. H., & Lee, J. (2022). Impacts of influencer attributes on purchase intentions in social media influencer marketing: Mediating roles of characterizations. *Technological Forecasting and Social Change*, 174, 121246. https://doi.org/10.1016/j.techfore.2021.121246

Mildner, T., Savino, G.-L., Doyle, P. R., Cowan, B. R., & Malaka, R. (2023). About engaging and governing strategies: A thematic analysis of dark patterns in social networking services. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (pp. 1–15). https://doi.org/10.1145/3544548.3580695

Morris, L. (2021, October 27). In Poland's politics, a 'social civil war' brewed as Facebook rewarded online anger. *The Washington Post*. https://www.washingtonpost.com/world/2021/10/27/poland-facebook-algorithm/

Moyer-Gusé, E. (2008). Toward a theory of entertainment persuasion: Explaining the persuasive effects of entertainment-education messages. *Communication Theory*, 18(3), 407–425. https://doi.org/10.1111/j.1468-2885.2008.00328.x

mpfs. (2022). *JIM-Studie 2022: Jugend, Information, Medien. Basisuntersuchung zum Medienumgang 12- bis 19-Jähriger*. https://www.mpfs.de/fileadmin/files/Studien/JIM/2022/JIM_2022_Web_final.pdf

Mühlhoff, R., & Schütz, T. (2019). Die Macht der Immersion. Eine affekttheoretische Perspektive. https://doi.org/10.25969/MEDIAREP/12593

Müller, P. (2022). Extrem rechte Influencer*innen auf Telegram: Normalisierungsstrategien in der Corona-Pandemie. *ZRex – Zeitschrift für Rechtsextremismusforschung*, 2(1), Art. 1. https://www.budrich-journals.de/index.php/zrex/article/view/39600

Munger, K., & Phillips, J. (2022). Right-wing YouTube: A supply and demand perspective. *The International Journal of Press/Politics*, 27(1), 186–219. https://doi.org/10.1177/1940161220964767

Murray, J. H. (1998). *Hamlet on the Holodeck: The future of narrative in cyberspace*. The MIT Press.

Nilsson, N. Chr., Nordahl, R., & Serafin, S. (2016). Immersion revisited: A review of existing definitions of immersion and their relation to different theories of presence. *Human Technology*, 12(2), 108–134.

Osei-Frimpong, K., & McLean, G. (2018). Examining online social brand engagement: A social presence theory perspective. *Technological Forecasting and Social Change*, *128*, 10–21. https://doi.org/10.1016/j.techfore.2017.10.010

Peters, I., & Puschmann, C. (2017). Informationsverbreitung in sozialen Medien. In J.-H. Schmidt & M. Taddicken (Eds.), *Handbuch soziale Medien* (pp. 211–232). Springer VS.

pre:bunk. (2023). Rechtsextremismus und TikTok, Teil 1: Von Hatefluencer*innen über AfD und bis rechter Terror. *Belltower.News*. https://www.belltower.news/rechtsextremismus-und-tiktok-teil-1-147835/

Qin, Y., Omar, B., & Musetti, A. (2022). The addiction behavior of short-form video app TikTok: The information quality and system quality perspective. *Frontiers in Psychology*, 13. https://www.frontiersin.org/articles/10.3389/fpsyg.2022.932805

Rau, J., Kero, S., Hofmann, V., Dinar, C., & Heldt, A. P. (2022). Rechtsextreme Online-Kommunikation in Krisenzeiten: Herausforderungen und Interventionsmöglichkeiten aus Sicht der Rechtsextremismus- und Platform-Governance-Forschung. *Arbeitspapiere des Hans-Bredow-Instituts*. https://doi.org/10.21241/SSOAR.78072

Rösch, V. (2023). Heimatromantik und rechter Lifestyle. Die rechte Influencerin zwischen Self-Branding und ideologischem Traditionalismus. *GENDER – Zeitschrift für Geschlecht, Kultur und Gesellschaft*, *15*(2), Art. 2. https://www.budrich-journals.de/index.php/gender/article/view/41966

Schlegel, L. (n.d.). Super Mario Brothers Extreme. Gaming und Rechtsextremismus. *Gaming-Rechtsextremismus.de*. https://gaming-rechtsextremismus.de/themen/super-mario-brothers-extreme/

Schlegel, L. (2021). The role of gamification in radicalization processes. *modus | zad*. https://modus-zad.de/publikation/report/working-paper-1-2021-the-role-of-gamification-in-radicalization-processes/

Schmitt, J. B., Ernst, J., Rieger, D., & Roth, H.-J. (2020). Die Förderung von Medienkritikfähigkeit zur Prävention der Wirkung extremistischer Online-Propaganda. In J. B. Schmitt, J. Ernst, D. Rieger, & H.-J. Roth (Eds.), *Propaganda und Prävention: Forschungsergebnisse, didaktische Ansätze, interdisziplinäre Perspektiven zur pädagogischen Arbeit zu extremistischer Internetpropaganda* (pp. 29–44). Springer Fachmedien. https://doi.org/10.1007/978-3-658-28538-8_2

Schmitt, J. B., Harles, D., & Rieger, D. (2020). Themen, Motive und Mainstreaming in rechtsextremen Online-Memes. *Medien & Kommunikationswissenschaft*, *68*(1–2), 73–93. https://doi.org/10.5771/1615-634X-2020-1-2-73

Schmitt, J. B., Rieger, D., Rutkowski, O., & Ernst, J. (2018). Counter-messages as prevention or promotion of extremism?! The potential role of YouTube recommendation algorithms. *Journal of Communication*, *68*(4), 780–808. https://doi.org/10.1093/joc/jqy029

Schneider, J., Schmitt, J. B., Ernst, J., & Rieger, D. (2019). Verschwörungstheorien und Kriminalitätsfurcht in rechtsextremen und islamistischen YouTube-Videos. *Praxis der Rechtspsychologie*, *1*, Art. 1.

Schouten, A. P., Janssen, L., & Verspaget, M. (2020). Celebrity vs. influencer endorsements in advertising: The role of identification, credibility, and product-endorser fit. *International Journal of Advertising*, *39*(2), 258–281. https://doi.org/10.1080/02650487.2019.1634898

Schwarz, K. (2020). *Hasskrieger: Der neue globale Rechtsextremismus*. Herder.

Silva, D. E., Chen, C., & Zhu, Y. (2022). Facets of algorithmic literacy: Information, experience, and individual factors predict attitudes toward algorithmic systems. *New Media & Society*. https://doi.org/10.1177/14614448221098042

Slater, M. D., & Rouner, D. (2002). Entertainment—education and elaboration likelihood: Understanding the processing of narrative persuasion. *Communication Theory*, 12(2), 173–191. https://doi.org/10.1111/j.1468-2885.2002.tb00265.x

Strick, S. (2021). *Rechte Gefühle: Affekte und Strategien des digitalen Faschismus*. Transcript.

Su, C., Zhou, H., Gong, L., Teng, B., Geng, F., & Hu, Y. (2021). Viewing personalized video clips recommended by TikTok activates default mode network and ventral tegmental area. *NeuroImage*, 237, 118136. https://doi.org/10.1016/j.neuroimage.2021.118136

Su, C., Zhou, H., Wang, C., Geng, F., & Hu, Y. (2021). Individualized video recommendation modulates functional connectivity between large scale networks. *Human Brain Mapping*, 42(16), 5288–5299. https://doi.org/10.1002/hbm.25616

Taylor, S. H., & Brisini, K. S. C. (2023). Parenting the TikTok algorithm: An algorithm awareness as process approach to online risks and opportunities for teens on TikTok (SSRN Scholarly Paper 4446575). https://doi.org/10.2139/ssrn.4446575

Tritt, S. M., Peterson, J. B., Page-Gould, E., & Inzlicht, M. (2016). Ideological reactivity: Political conservatism and brain responsivity to emotional and neutral stimuli. *Emotion*, 16(8), 1172–1185. https://doi.org/10.1037/emo0000150

Weimann, G., & Masri, N. (2021). TikTok's spiral of antisemitism. *Journalism and Media*, 2(4), Art. 4. https://doi.org/10.3390/journalmedia2040041

Whittaker, J., Looney, S., Reed, A., & Votta, F. (2021). Recommender systems and the amplification of extremist content. https://doi.org/10.14763/2021.2.1565

Widmaier, B., & Zorn, P. (Eds.). (2016). *Brauchen wir den Beutelsbacher Konsens? Eine Debatte der politischen Bildung*. Bundeszentrale für politische Bildung. https://www.lpb-mv.de/nc/publikationen/detail/brauchen-wir-den-beutelsbacher-konsens-eine-debatte-der-politischen-bildung/

Wunderlich, L. (2023). Parasoziale Meinungsführer? Eine qualitative Untersuchung zur Rolle von Social Media Influencer*innen im Informationsverhalten und in Meinungsbildungsprozessen junger Menschen. *Medien & Kommunikationswissenschaft*, 71(1–2), 37–60. https://doi.org/10.5771/1615-634X-2023-1-2-37

Yang, Y., Xiu, L., Chen, X., & Yu, G. (2023). Do emotions conquer facts? A CCME model for the impact of emotional information on implicit attitudes in the post-truth era. *Humanities and Social Sciences Communications*, 10(1), Art. 1. https://doi.org/10.1057/s41599-023-01861-1

Yesilada, M., & Lewandowsky, S. (2022). Systematic review: YouTube recommendations and problematic content. *Internet Policy Review*, 11(1), 1652. https://doi.org/10.14763/2022.1.1652

Zimmermann, D., Noll, C., Gräßer, L., Hugger, K.-U., Braun, L. M., Nowak, T., & Kaspar, K. (2022). Influencers on YouTube: A quantitative study on young people's use and perception of videos about political and societal topics. *Current Psychology*, 41(10), 6808–6824. https://doi.org/10.1007/s12144-020-01164-7

Platform Design, User Creativity, and Aesthetic Governance in Social VR

Arne Vogelgesang

Introduction

While the term "metaverse" is often employed to gesticulate towards a larger paradigm of digitized sociality, immersive applications of virtual reality technology (VR)[1], commonly represented as humanoid figures communing in 3D-generated landscapes, are central to its vision. Currently, it is so-called Social VR (SVR) platforms that most approximate this vision socially and aesthetically and therefore allow for an evaluation of current practices and possible trajectories of digital embodiment as a basis of "metaversed" online cultures. The following text is not a thorough empirical investigation of existing SVR platform culture, but an exemplary sketch of parts of the landscape delineating the conditions and possible effects of aesthetic governance in VR.

The purpose of this overview is to close a gap in the discussion about SVR and future "metaverse" spaces at large. In terms of structural and aesthetic design choices, the main focus of inquiry has been on usability and safety – the latter arguably because of extensive bad experience with and public discourse on forms of digital(ized) violence in online virtual spaces. Less attention has been given to the powerful interplay of emerging SVR platforms' reliance on user creativity and their control over the conditions and limits of this creative labor. Thinking of this only in terms of "content creation" analog to the current social media platform paradigm risks missing an important factor: Other than in non-VR virtual social spaces, where users produce social meaning by co-creating and co-consuming the media environment they then come to metaphorically inhabit, in VR physically experienced habitation is a central issue right from the start. The "content" SVR communities produce is first and foremost the virtual bodies their

1 In this text, the term "VR" denotes technologically mediated immersive digital 3D environments, while the word "virtual" may in a wider sense also refer to other non-physical/online spaces, communities, practices, or phenomena.

members are present through and present themselves as, and the virtual spaces they sensually experience these bodies in. Controlling what such bodies and spaces can do and look like also means profoundly controlling the scope of users' embodied experience. The emergence of digital aesthetic governance over embodiment is what this text is meant to draw attention to.

Since the development and deployment of the *Oculus Rift* started the consumer VR mainstreaming phase around 2013, numerous social worlds and platforms centering the technology have been created. From weblogger Ryan Schultz's more than 160 entries long list of VR-capable social virtual worlds (Schultz, 2024), only a few have garnered a 4-or-more-digit user count, however. The two most prominent ones as of 2024 are *Rec Room* and *VRChat*. A comparison of these two protagonists helps understand how different concepts of aesthetic worldbuilding and user creation can influence community development in terms of culture and politics – which is crucial when thinking about what "immersive democracy" might mean or come to be.

This text is divided into several parts. After a general introduction to Social VR, I will give an overview of the genesis and characteristics of the two chosen platforms. This overview is followed by a rough description of the communities that have formed on each platform during the past years. The concluding part will discuss aesthetic governance as a process of negotiation between design paradigms and community culture(s). For my descriptions, I largely rely on observations during unstructured preliminary field research on various SVR platforms 2020–2022. Full-fledged ethnographic research through participatory observation on Social VR is still lacking, but I have taken into account some literature using qualitative methods like interviews (Freeman et al., 2020; McVeigh-Schultz et al., 2019; Shriram & Schwartz, 2017), guided group walkthroughs (Liu & Steed, 2021) or social media discourse analysis (Zheng et al., 2022), as well as primary and secondary online sources.

Social VR

Social VR platforms are "immersive systems that prioritize and focus on the in-environment communication" (Liu & Steed, 2021). In earlier decades, such systems have been discussed as "collaborative social environments"

(CVEs), but the arrival of mass consumer VR hardware has shifted terminology (Jonas et al., 2019). SVR practice can be described as embodied social role-playing within a system of connected and confined virtual spaces inhabited by avatars[2] tethered[3] to human users. The terms "space", "room", "world" or sometimes "map" are often used interchangeably when talking about places inside SVR. In this text, I follow that mode of usage and reserve the term "platform" for speaking about the whole system of spaces, the infrastructure of which is most often run and owned by one company. While in a general sense, socializing in virtual networked environments has been a foundational element in digital culture for decades, it is the possibility of first-person embodiment through VR technology that makes Social VR distinct, introducing the bodily effects of place illusion and presence (Slater, 2009; Slater et al., 2009) into digital socialization.

Most, though not all, SVR platforms focus on meeting and connecting with strangers and have implemented functions to build user networks, like friends lists or groups. Users communicate verbally via microphone and/or non-verbally with their avatar bodies, be it via live VR body tracking or through prerecorded movements. Other established means of social online communication like emojis and written chat are common as well, though less central to the experience. Almost all SVR platforms allow usage without a dedicated head-mounted device (HMD) to lower entrance barriers and enable user growth – in fact, the majority of people using the bigger SVR platforms currently are non-VR users, because VR hardware is still relatively pricey, quick to become technically obsolete, and of limited everyday utility. Nevertheless, since immersive VR usage is core to the platforms' appeal and affordances, most insight can be gained from concentrating on this part of the user experience.

Due to technical limitations, a single room on an SVR platform can currently usually host no more than about 50 people simultaneously, which structurally encourages the dynamic creation and dissolving of social groups. Virtual rooms can be instanced multiple times in different social states: public, private, open only for friends, through invitation, or only for people using a specific link or token. Since digital assets can be copied, uniqueness in virtual environments is a rare good. Consequently, it has –

2 For a closer look at avatars within SVR see Kolesnichenko et al. (2019).

3 Turkle (2006) discusses the psycho-social implications of such "tethering" with respect to mobile phones in a way that might be even more relevant for virtual and mixed reality.

like presence – become decoupled from (virtual) materiality to mainly exist as a transient psycho-social fact: as an experience. But before discussing this, I will first delineate the properties of the two platforms examined here.

Rec Room

History and Availability

In the spring of 2016, a group of six men – partly *Microsoft* employees who had formerly worked on the mixed reality device *HoloLens* – founded the company *Against Gravity* to release *Rec Room*. The application was marketed as a "virtual reality social club where you play active games against competitors from all around the world"[4] and featured different virtual spaces for users to play and socialize in. *Rec Room*'s name refers to its central social metaphor, which is also the source of its unified aesthetics: "a prototypical rec[reation] center from the year 1987" (McVeigh-Schultz et al., 2019). Most of the provided games were and are indeed competitive, but over the years simulations of typical physical sports games like dodgeball were surpassed in popularity by more martial ones like laser tag, embedding the common "first-person shooter" experience of online gaming into the social sandbox.

Rec Room was initially released to meet the market entry of the new *HTC Vive* HMD, while also being available for the *Oculus Rift* and expanding availability to *Playstation 4*'s VR system in late 2016. Since then, the platform has become accessible to a fairly large number[5] of devices and operating systems: *Windows* PC desktops (either as a downloadable standalone application or via the digital distribution platform *Steam)*, *SteamVR* compatible as well as *Oculus Rift* and (*Meta*) *Quest* HMDs[6], mobile *iOS*

4 Cited from the original press release accompanying the application's launch, archived under https://web.archive.org/web/20160620140618/http://www.againstgrav.com/press [accessed 2024, September 23].

5 "Fairly large" should be understood in comparison to other SVR apps. While technically, browser-based platforms like the recently discontinued *Mozilla Hubs* are accessible from any device with a compatible browser and thus have the lowest threshold for entry and widest possible adoption, in practice companies controlling access to VR applications via their stores have been reluctant to include and sometimes actively excluded WebXR compatible browsers, to the effect of restricting access to non-proprietary platforms.

6 Support for *Quest* 1 devices was discontinued in the first half of 2023 when *Meta* deprecated the relevant SDK.

and *Android* devices, *Xbox* and *PlayStation*. Like most SVR applications, the platform so far does not support *Linux* and *macOS* desktop devices.

Economy and Adoption

Likely due to its founders already being well-connected in the industry, *Against Gravity* started with substantial seed funding from multi-billion-dollar venture capital firm *Sequoia Capital* in 2016 and was able to raise investments to almost $300 million until late 2021 – the bulk of which poured into the firm during the Covid-19 pandemic[7]. The company has since changed its name to *Rec Room Inc*. Like almost all current SVR platforms (and most social media platforms more generally), *Rec Room* is free to use, with some advanced features only accessible to paying customers. An in-game economy of tokens to spend on virtual items and clothing was included from the start, and custom creations of users made with the platform tools can be traded via those tokens inside the app's ecosystem. In 2020, the ability to purchase tokens with "real" money for an exchange rate set by the company was added, as well as a monthly subscription feature called "Rec Room Plus" that allows creators of in-game assets to cash out their earnings after reaching a threshold of 250,000 tokens (currently converting to $100). On the virtualization side of economics, room creators can also create their own sub-currencies, which may then be traded against *Rec Room* tokens[8], allowing for a potentially unlimited number of virtual micro-economies. The company calls its meta economy "Community Commerce" – a term that has gained popularity in recent years with TikTok's growing success – and promotes it to users as a potential way of "making a sustainable income"[9].

On its website, *Rec Room* boasts more than 100 million lifetime users. While an impressive figure, this amount does not reflect the number of people actually using the platform, since it includes abandoned, multiple, and otherwise inactive accounts. Occasionally, the company publishes numbers of its monthly active user count (MAU) at peak times. In 2022, this number was reported as 3 million accounts that had logged into the

7 Numbers cited from https://www.crunchbase.com/organization/against-gravity/companyfinancials [accessed 2023, December 5].

8 See https://recroom.com/roomcurrencies [accessed 2024, September 23].

9 Cited from https://blog.recroom.com/posts/2021/10/12/community-commerce-report [accessed 2024, September 23].

platform at any given time over the course of a month (Au, 2022b). Meant to demonstrate growing adoption, this number still does not convey much information about the amount of time people spend on the platform and what they actually do there. Independent numbers are not available and would be hard to obtain from the outside without access to a usage data API, because users are spread over thousands of rooms at any given time.

Aesthetic Concept

Rec Room's visual concept is a virtual youth nostalgia – not only regarding the choice of its metaphorical location but also in the sense that most of its users are too young to have any personal memories of a US college or university recreation center in 1987. The virtual spaces provided by the platform itself, called "Rec Room Originals", are dominated by warm colors and rounded shapes creating a family-friendly[10] nostalgic vibe. Simple materials and "low-poly"[11] 3D objects ensure fluid rendering and interoperability across different devices and add to the overall retro aesthetics[12].

Fluid playability on mobile devices is also a major reason for the stylized humanoid user avatars on the platform, which for the longest time did not feature any legs[13]. Platform users are represented through torsos floating above ground, with aligned but unconnected hands and heads. These avatars, called "floating beans", can be customized individually inside the app with regard to their facial features, hairstyle, skin color, gender

10 For these and the following descriptions, compare McVeigh-Schultz et al. (2019), who have interviewed *Rec Room* designers about their aesthetic and functional decisions.

11 The number of simple geometric polygons a 3D object consists of correlates with the computational power needed for its visual rendering. Since technological advancement in graphics computation is accompanied by a drive for higher fidelity 3D realism, simpler "low poly[gon]" aesthetics have become associated with a nostalgic vibe.

12 YouTuber Retr0's video "The Evolution of Rec Room (Release, 2016 and 2017)" gives an impression of the aesthetic development, but also consistency over the years ([Retr0], 2021).

13 Most consumer VR hardware only provides movement tracking of three points – head and hands. Leg movement and positioning usually have to be inferred computationally. The company describes the rationale of the original avatar design in a blog post as follows: "We avoided showing untracked legs and arms because it could break the feeling of presence; we kept facial features cute and minimal to avoid the uncanny valley effect; and we chose simplicity over visual detail so the game ran smoothly" (https://blog.recroom.com/posts/avatars) [accessed 2024, September 23].

attributes, clothing, and accessories. Stylized mouths with animations synchronized to the user's microphone input make social interactions feel more "alive" and have been designed to predominantly convey friendly facial expressions. This design decision is a form of aesthetic nudging towards a more "positive" social atmosphere where, as one *Rec Room* developer put it, "everyone looks happy all the time" (McVeigh-Schultz et al., 2019).

Besides the "Rec Room Original" spaces/games developed by the company itself, users can build their own rooms from an assortment of basic 3D elements and materials, as well as design custom avatar "costumes" and thereby body shapes. This is done with an in-game tool called "Maker Pen" – a stylized hot-glue pistol – and a visual scripting system called "Circuits" for interactive functionalities like buttons, dynamic architecture, collision detection, or scoring systems. In 2023, an additional development kit called "Rec Room Studio" was rolled out in beta state and made accessible to a small portion of the user base. The kit allows the import of environments and elements created with the game engine *Unity3D*, thereby significantly expanding the 3D design options. If widely adopted, this is likely to break up the fairly unified aesthetics of Rec Room in the future. *Rec Room Studio* is on the one hand targeting companies that want to be present on the platform with their corporate visual designs[14]. On the other hand, it can also be understood as a reaction to the success of *Rec Room*'s direct competitor *VRChat*, which follows a different logic of aesthetic creation and has spawned a culturally much more influential creator community.

VRChat

History and Availability

VRChat was released by software engineer Graham Gaylor for the then-new *Oculus Rift* HMD in early 2014. Alongside the later discontinued platform *Riftmax*, the app quickly assembled a small community of VR enthusiasts using it for socializing, exploration, development, and discussion in the early years of consumer VR. At the point of release, *VRChat* was in a very basic state, and it has retained the status of being an "early access" product in development until now. Its core functionality was, and still is,

14 A paragraph on the feature webpage addresses readers that "are a company or brand" (https://recroom.com/studio) [accessed 2024, September 23].

the hosting and mediating of networked virtual co-presence through 3D avatars, leaving most everything else to its users. Contrary to *Rec Room*, *VRChat* never had a unified aesthetic design concept: user-created content is hugely important to the platform and has been the main reason for its popularity.

Like most SVR platforms, *VRChat* does not limit accessibility exclusively to users with VR hardware. Desktop clients for *Windows* and *macOS* were deployed early, though the latter was discontinued in the first half of 2016 when support for the newly released *HTC Vive* HMD via *SteamVR* was added. Direct client downloads from the *VRChat* homepage were phased out in the following years in favor of distribution through the different app stores tied to the disjunct and competing VR hardware device ecosystems. Most importantly, a combined PC desktop and VR version accessible via the *Steam* software platform's early access program in mid-2017 drew in a lot of users who approached the application from a video-gaming perspective. There is no native support for *Linux* or *macOS*. The beta version of a mobile app for *Android* was released in December 2023 but has so far found less adoption than the mobile versions of direct competitors like *Rec Room* or *Roblox*.

Economy and Adoption

Since its inception, the initial two-person LLC (Gaylor teamed up with programmer and game designer Jesse Joudrey shortly after the initial release to launch the company) has evolved into a business with several dozen reported full-time employees. *VRChat Inc*. has been financed through several funding rounds with about $95 million[15]. To the author's knowledge, the company has never disclosed revenue or valuation figures or even a business model. The application is largely free to use, with a subscription service called "VRChat Plus" offering exclusive or early access to select features, but the revenue from subscriptions is unlikely to support a significant part of the cost of infrastructure, support, and development. The latest – and by far largest – funding round in 2021, providing the company with an $80 million backing led by US venture investment firm *Anthos Capital,* was linked in a company blog post on behalf of the "VRChat Team & Investors"

15 Numbers from https://www.crunchbase.com/organization/vrchat/company_financials [accessed 2023, December 5].

to the ambition of further growing the user base and implementing a "creator-driven economy" ([Tupper], 2021b), i.e. mechanisms allowing users to pay each other inside the platform. Such a payment infrastructure similar to the *Rec Room* "Community Commerce" would enable the company to profit off transaction fees that have so far been taken in by external platforms like *Booth*, *Gumroad*, or *Patreon*, which have become hosts to the community's lively informal content market economy (Au, 2021).

Similarly to *Rec Room Inc.*, the company is not interested in making its adoption and usage data fully transparent. Instead, it occasionally publishes new record highs of concurrent users, i.e. the maximum number of accounts logged in simultaneously at a select moment. Those were reported to be about 40,000 on New Year's Eve 2020 ([Tupper], 2021a) and more than twice that number one year later (Au, 2022a). There is, though, a community-driven documentation of *VRChat*'s API that runs a comprehensive usage dashboard allowing continuous insight[16]. Generally, *VRChat*'s total user base is often assumed to be lower than that of *Rec Room* but with a higher percentage of actual VR hardware users due to its advanced motion tracking support. *Steam* usage statistics of PC desktop and VR users usually rank *VRChat* significantly higher than *Rec Room*[17], but do not represent mobile or any other users not connecting via the service, the former being a significant part of *Rec Room* users according to the company[18].

On the technological side, *VRChat* supports more advanced VR hardware technology than most of its competitors, like up to 11-point full-body tracking[19], and features a generous scripting API. Despite prominent claims that "legs are hard"[20] in VR, *VRChat* avatars have long been able to accommodate not only legs with inverse kinematics and tracking but also

16 See https://metrics.vrchat.community [accessed 2024, September 23].

17 For at least the last year, the *Steam* user count for *VRChat* has been roughly 20 times the one of *Rec Room* as per https://steamdb.info/charts/?category=53&select=1&compare=438100%2C471710 [accessed 2024, September 23].

18 A *Rec Room* representative reported in 2022 that "at this point VR is a pretty low percentage of our monthly players" and then referred to the bulk of users coming from various ecosystems not represented on *Steam* (Lang, 2022).

19 Tracking accuracy of a user's physical body can be increased by adding more capture points at the feet or between key joints like hips, knees, or elbows. *VRChat* supports tracking devices that interface with *Valve*'s optical "lighthouse" system but can also be expanded by solutions compatible with *SteamVR*'s protocols. See https://docs.vrchat.com/docs/full-body-tracking [accessed 2023, December 5] and compare FN 13.

20 "Seriously, legs are hard" was famously proclaimed by *Meta*'s Mark Zuckerberg at the "Meta Connect VR" conference in 2022 when announcing full-body avatars for

dynamically moving tails/hair/costume parts, advanced custom shaders, prerecorded movement animations and a wide range of avatar sizes. This has led to the platform garnering a power-user base of people willing and able to invest in VR hardware allowing for higher degrees of embodiment. Consequently, users with VR hardware and "deskie" users without it can have very different experiences when using the platform, which sometimes leads to differing social behavior and contributes to cultural stratification along hardware ownership lines.

Aesthetic Concept

VRChat's significant informal community content market, with users selling, trading, and commissioning avatars and virtual rooms among each other, is a result of its aesthetic production paradigm. The platform has encouraged and relied on user-created content pretty much from the start by providing a software development kit (SDK) plugging into the free-to-use *Unity3D* game engine. Early on, *VRChat* founder Graham Gaylor expressed his belief that custom content creation was key to evolving metaverse applications, as it had been for social web platforms[21] – virtual environments and avatars being the equivalent of user-generated text and image content on "web 2.0" social media. Like with these previous platforms, Social VR's appeal and worth would come to depend on its users' creative labor.

The "look and feel" as well as the social dynamics on *VRChat* today are a direct consequence of the decision to have almost all content generated[22] by users. The first *VRChat* application had been quickly assembled in *Unity3D* by Gaylor, using a scene from the *Unity Asset Store* and a simple humanoid (male) avatar as readymades for testing the functionality of networked VR[23]. Since there were no aesthetic guidelines but only technical limitations, interested users soon began experimenting with possibilities and lim-

Horizon Worlds, followed by the erroneous statement "[…] which is why other virtual reality systems don't have them either" (Hern, 2022).

21 See Thompson (2014) at minutes 17:22 & 48:56.

22 "Generated" may at the most basic not mean much more than "uploaded" – "stealing"/copying content from other creators is not uncommon, although frowned upon in the community if it is outside the scope of fan art. See e.g. the sentiment expressed by one interviewee of Quent and Vogl (2025) in this book.

23 An impression can be gained from the release thread on Reddit and links to the first run's documentation in the comments: https://www.reddit.com/r/oculus/comments/1vcbwk/vrchat_v01_released/ [accessed 2024, October 19].

itations for creating avatars and spaces. An influx of "very online" users in the following years brought recreations of video games, pop-culture figures, and memes. Since the knowledge threshold and learning curve for user creation in 3D spaces is significantly higher than on "classical" social media, with a wider gap between content production and social practice, avatars in particular have become a desirable commodity on *VRChat* – a kind of social trading good in the community, sometimes spreading very fast and creating memetic phenomena spilling into the wider online culture. Over time, *VRChat* users thus developed a deliberate aesthetic eclecticism that made the platform increasingly attractive for content creators on video and streaming platforms like YouTube and Twitch, who became part of the developing informal cultural economy.

Virtual Communities

VRChat's eclecticism and avatar affordances have become a breeding ground for distinctive and overlapping communities around identities and practices with a high emphasis on embodied aesthetics. It is beyond the scope of this text to describe these communities in detail. They notably include

- a long-standing club/party scene as well as a dedicated dance community holding competitions supported by full-body tracking gear, both mainly focused on, but not limited to, e-girl & e-boy avatar styles,
- a transgender community using the affordances of virtual morphological freedom[24] and sharing advice on gendered body movement and voice training through socializing as well as workshops and community events,
- a virtual furry community enjoying the low entry threshold of VR avatars as opposed to the high prices of physical fursuits, with the last

24 "Morphological Freedom" is an idea from transhumanist enhancement discourse. In the present context, compare founder of the *VRChat* "Trans Academy" Tizzy in an interview with VTuber Phia: "In 2016, when I was looking to have facial feminization surgery, I brought a screenshot of my second life avatar because it was the person that I felt the most comfortable and happy as. That might seem a little bit taboo now but I think that as Social VR and the metaverse become more of an integral part of our society in the future, we're going to see a lot more people prototyping their identity in these spaces and embracing the idea of having morphological freedom" (Bollinger, 2023).

convention on the platform according to the organizers having had more than 21,000 participants[25],
- a diverse role-playing community with different game worlds and stories as well as meta-role-playing troupes with high mobility on the platform like the "Loli Police Department",
- and a meme community that strongly influenced the platform's public image because of its attractiveness to live streamers.

The latter's appeal to underage users and people close to online trolling and "shitposting" culture and their often dominating and abrasive behavior in public *VRChat* rooms has driven other local communities to largely avoid public worlds and rely on non-public rooms and invitation mechanics, operating their own events and social spaces somewhat shielded from the platform's wider ecosystem. This dynamic has begun to create something akin to a *VRChat* society, where interest groups negotiate their sometimes aligned, sometimes conflicting interests through different channels and carve out virtual space for themselves.

VRChat is also frequently referred to as having been instrumental in developing distinct virtual socio-physical practices and conventions: "head-patting" as a gesture of affection, silent rooms where users can doze or sleep while wearing their HMDs, and a growing number of users engaging in erotic role-play (ERP) in VR. The latter has been met with concern by longer-term users because it amplifies or contributes to a growing sexualization of avatars[26] and exacerbates existing disputes about the status of minors on the platform.

All these practices and their exemplary sub-communities share a strong connection with corporeality, at least and foremost for users of dedicated VR hardware. They are about negotiating and transforming the relationship of physical bodies, self-images, and forms of expressive movement on the one hand to the possibilities of virtual bodies, self-images, and mediated relationships on the other – tethering and expanding embodiment. Thanks to its advanced body tracking support *VRChat* has become one of few plat-

25 https://furality.org/ [accessed 2024, September 23].

26 Arguably, sexualization is part of the complex intercultural history of anime aesthetics at large, so this tendency was prevalent in a community relying heavily on those aesthetics for their avatars and virtual cross-dressing practices pretty much from the start. It only seems to have become problematic for this community when combined and thus increasingly identified with publicly performed socio-sexual practices – an example of the differential value judgments at play in communities forming around visual representations of bodies, identities, and desire.

forms that can accommodate the aesthetic realization of this relationship with and desire for virtual embodiment, where "physical bodies [are] the immediate and sole interface between [users] and their avatars" (Freeman et al., 2020).

The relatively large degree of technical freedom in the creation especially of avatars has also given *VRChat* a long history of hacks and so-called "crashers" – code-based modifications that can be employed as a weapon to freeze or kick other users out of the game, sometimes in quite elaborate and aesthetically overwhelming ways. Especially crashers weaponizing shader code combine the affective experience of being forcefully ejected from a (virtual) social reality with an intense aesthetic overload likely to provoke strong physical reactions in HMD users: they not only crash the software, but they also attack the sensory system of its corporeal users, forcing them to embody the severance of their own tethering.

Like avatars in general, such crashers have long been traded among *VRChat* users, be it for offensive or defensive purposes. The technical affordances allowing for such virtual weapons as well as the comparably weak content moderation on the platform have made many community members somewhat resilient to attacks, insults, flaming, etc., making them regard harassment as an annoying, yet not truly avoidable social phenomenon at least in public worlds. The danger of being attacked or insulted is seen as a trade-off for the power of forming, defining, and developing community and community aesthetics "from the ground up". The aesthetic sandbox is a social sandbox as well, where too many preventive restrictions are undesired even by users experiencing harassment, "as they might prevent the open dialogues that drew users to the technology in the first place" (Shriram & Schwartz, 2017).

Contrastingly, *Rec Room* communities, with their limitations in avatar design, have developed less around virtual corporeality and more around competitive playful practices. Many users are heavily invested in the games the platform offers – not least because especially the "Rec Room Original" PvP games like paintball or laser tag work well from a vsports[27] perspective. But there is also a creative community focused on building worlds and costumes or painting in *Rec Room,* as well as sub-communities based on such aesthetic creation, like for (military) role-playing or pop-cultural fan-

27 While the term "vsports" seems to be not in use yet, it makes a lot of sense to distinguish virtual sports activities with their emphasis on whole body movement from egaming/esports that require more isolated hand-eye coordination.

doms. For creators, being confined by the narrower aesthetic limits of the platform is a creative challenge balanced by the entanglement of attention and token economy. Lastly, like in *VRChat*, there are also identity-centered communities/servers for LGBTQ or furry users on *Rec Room*, although they appear to be less prominent. Generally, most users seem to follow the central metaphor and conceptual idea of *Rec Room* as a "social club" around sports activities, and also partake in the regular special events the company designs around tasks and token/item collection, which also try to build and develop a narrative around the fictional platform universe.

When, in 2023, *Rec Room* announced the upcoming integration of full-body avatars (i.e. bodies with legs) and single-finger movement, a significant portion of users seemed rather wary of such changes[28]. Especially longer-time users seem to identify with the stylized aesthetics of the platform and take a rather conservative stance towards changing the simplified look. When discussing such changes, users regularly invoke *VRChat* as the aesthetic negative to their appreciation of *Rec Room,* emphatically describing the dread they feel when confronted with *VRChat*'s radical aesthetic inconsistency of avatars and worlds. In contrast, it seems they specifically value the stable and defined aesthetic normality across the *Rec Room* universe, for it allows them to concentrate on the core activities of gaming and socializing. In response, the company places a lot of emphasis on explaining upcoming changes in terms of aesthetic consistency. A recent developer blog post on the full-body avatar feature, while praising the aesthetic change of "adding noses, eyebrows, fingers, arms, legs, and feet" as "more ways for all of us to make money", ended with a section titled "Our Commitment to the Floating Bean Avatars"[29].

Governance

As the introduction already made clear, I assume a connection between aesthetic and social regulation of Social VR platforms. If that connection existed as a correlation – however complicated by the fuzziness of cultural processes –, then we would assume spending time in *Rec Room* to be

28 For an exemplary discussion among *Rec Room* users that focuses on the aesthetics of single fingers, see https://www.reddit.com/r/RecRoom/comments/143hytj/what_are_your_opinions_on_rec_room_having_hand/ [accessed 2024, September 23].

29 https://blog.recroom.com/posts/avatardeepdive [accessed 2024, September 23].

an experience significantly less likely to be disruptive or stressful. Indeed, the platform is not only aesthetically and economically more coherent, it also has more developed moderation/policing features than *VRChat*. A system of appointing and rewarding community moderators, a third-party algorithm constantly surveilling users' speech for forbidden words[30], and features like an embodied gesture for quickly blocking other users in threatening situations speak of user safety being considered on a variety of levels. It is no wonder then that in the academic literature on Social VR, *Rec Room* is being discussed more prominently and also more positively than *VRChat* when it comes to questions of safety and harassment[31], with the latter platform usually being characterized as a form of "wild west" (McVeigh-Schultz et al., 2019) "known for non-normative social interactions" (Zheng et al., 2022).

While this is intuitively plausible, there might also be some bias at play. Academic research on Social VR, when more than pure literature review, has so far concentrated on design features, and on harassment as a potential design problem. Skimming through papers and their methodologies shows that researchers spend surprisingly little time actively using the platforms they write about[32]. There is a serious lack of ethnographies about and on SVR that would enable outsiders to learn how those platforms' users make sense of and navigate the social space(s) they inhabit and, for the most part, create. Harassment is one part of this social space and users respond to it within the frame of the general community politics, explicit policies, and tacit rules of their specific platform – their response is part of the "attendant literacies, interaction conventions, and common practices that exist in a feedback loop between the (top-down) designed affordances of various online social platforms and the (bottom-up) practices of virtually embodied players seeking to communicate" (Tanenbaum et al., 2020).

30 See company blog posts at https://recroom.zendesk.com/hc/en-us/articles/4419902650135-Applying-for-Moderator-Volunteer-Mod and https://blog.recroom.com/posts/voice-moderation-updates [accessed 2024, September 23].

31 One literature review conspicuously lists *VRChat* as "known for harassment and unpredictable social encounters" in a long table of otherwise neutral or even advertisement-like descriptions of different platforms' functionalities/USPs (Handley et al., 2022).

32 In addition, *Rec Room* company staff seem to be much more accessible for interviews with researchers, which also leads to a certain representational bias.

In fact, hate speech is as much a problem in both *Rec Room* and *VRChat* as bullying of certain user groups like e.g. furries[33] is – despite the different grades of moderation and implementation of safety features. On both platforms, it does not take long to encounter nazi roleplaying or discriminatory talk. On both platforms, sexual harassment is a problem evolving from its already prevalent and well-described occurrences in virtual social spaces in the wider sense into the new embodiment and immersion affordances of VR technology – a problem made even more pressing by the significant presence of minors. Additionally, underage users themselves form, on both platforms, a group that many older members see more as annoying than as vulnerable, denigrating them as "squeakers" for especially younger boys' high-pitched frequent yelling.

As has often been established for all sorts of virtual environments, such social problems will keep appearing and shape-shifting in online spaces as long as they exist in the so-called "real world". While design-relevant, they are not design-solvable problems. "[I]ntensified old concerns in the new world" (Zheng et al., 2022), they now appear in a context with new conditions and possible complications. This new context is on the one hand defined by the intensified bodily experience of interactions in virtual reality, which as a "social" medium paradoxically also bears the characteristics of disembodied online interactions – with the consequence of "less boundaries [...] that can rule and determine what are reasonable, psychologically safe and permissible ways for other people to behave around self and how self will respond when someone steps outside those limits" (Zheng et al., 2022). But this context on the other hand also carries the vectors and effects of the platforms' differing creative/aesthetic paradigms. How can those paradigms be described when thinking about governance in these new and ambiguous spaces of virtual embodiment and embodied virtuality?

Of the two example platforms regarded in this text, *Rec Room* seems to fit the top-down model of a benevolent ruler. "Rooms are behavior", as one of its developers put it in an interview (McVeigh-Schultz et al., 2019), and the company retains relatively much control over the social cues they allow virtual spaces to give users on their platform. Communitization takes place around competitive playful activities, mediated by a ubiquitous gamified economy and within a unifying aesthetic atmosphere

33 Searching for "furries rec room" on YouTube yields plenty of videos with titles like "trolling furries on rec room", "Killing furries in Rec Room", "Making furries cry in Rec Room", "Infiltrating Furry Rec Room Servers" etc.

regulating the expressions users are able and allowed to perform. It is, in short, a virtual model of happy capitalism. On a platform of thousands of parallel and synchronous bounded virtual rooms, centralized control over the conditions of social experience within these rooms hedges boundary testing experiments as much as violations of the social contract. In that, *Rec Room* policy follows Blackwell et al.'s recommendation that "designers could directly influence the norms of individual communities and groups through design ‚nudges'" (Blackwell et al., 2019) – a socio-aesthetic technology of governance that has implications far beyond the scope of dealing with harassment. This is ever more true because *Rec Room's* vision of "democratization" has from its inception been linked closely to monetization through community commerce[34]: it is at its core an economic experiment. In consequence, community politics "on the ground" appear to develop between the poles of an aesthetic conservatism shying from "too much" diversity and a growing consciousness about the stratification effects and exploits of the platform's token economy[35].

In contrast, *VRChat*'s focus on embodiment effects and a very liberal user-driven asset production ecology have given birth to a multiplicity of partly overlapping, partly averse sub-communities that have made the platform something like the *Reddit* of VR. In an equally liberal low-moderation environment, many members of those communities have developed platform-specific resilience against equally platform-specific threats. The lively and sometimes unhinged creativity of community members has influenced the pop-cultural image of Social VR more than existing research has acknowledged, and *VRChat* communities politicize mainly around the conditions for this appeal – especially when they find them endangered. The company was forced to acknowledge this in mid-2022 when users became enraged about a new anti-cheat function that was meant to prevent tampering with the client software but effectively barred an entire modding

34 *Rec Room*'s General Partner at main investor *Sequoia Capital* describes the platform's vision of building community around games "both for fun and to earn money" in a blog post like this: "Rec Room's vision is to democratize access for anyone to create [...]. The team is also excited to launch P2P monetization to enable creators to monetize their own creations — enabling the new side hustle for kids" (Zhan, 2020).

35 Community Vtuber BVR proposed a system of upper, middle, and lower classes depending on users' token wealth in a video titled "Why is Everything SO EXPENSIVE in Rec Room?" ([BVR], 2022), assigning content creators to the wealthiest class. *Road to VR* editor Scott Hayden pointed to the risk of "gambling, money laundering, and other illicit behavior" within *Rec Room* in 2020 already (Hayden, 2020).

community that had also taken responsibility for providing users with impaired eyesight or hearing access to the platform; a rage that manifested in large scale review bombing[36] and numerous active and creative users leaving for smaller competitors like *Neos VR* or *Chillout VR*.

While methodically robust ethnographic research is yet to be desired, it seems plausible that the less safe and less regulated environment of *VRChat* has led to a higher degree and valuation of self-governance amongst most of the platform's multiple communities[37]. This form of community self-governance aims at protecting the peer group against the dangers of the platform's evolving social ecosystem. As a consequence, it tends to produce entry barriers and exclusion mechanics around sub-communities. There is a "dark social" VR space on the platform consisting of disparate systems of non-public rooms only accessible through invitations from community members who control the health and safety of their virtual social spaces. If this tendency becomes too strong, neglect of the open social space between sub-communities – the "metaverse" equivalent to the democratic concept of public space – might become a problem for social reproduction as well as for user and company growth at large, because it is those liminal communal rooms where onboarding of new users commonly happens but which can only sustain the platform's appeal when they are lively social spaces themselves.

Both differing platform cultures and models of governance thus provide starting points for thinking about how democratic structures might develop and be stabilized in virtual worlds employing VR technology. While the two platforms' development trajectories seem to converge – *Rec Room* opening up aesthetically with a new *Unity3D* SDK, *VRChat* working towards integrated community commerce –, it remains to be seen what role their different community cultures will play in said conversion. This is of interest especially because whatever social technologies are developed in Social VR in the strict sense, they will have wider implications for an increasingly virtualized social reality at large as envisioned by "metaverse" evangelists: If

36 Thousands of furious reviews by users temporarily lowered *VRChat*'s *Steam* rating to "mostly negative", prompting gaming and technology magazines to conjure apocalyptic imagery of the platform "being absolutely nuked into the ground" (Taylor, 2022).

37 Common theory about the effects of VR technology, like place/plausibility illusion (Slater, 2009) and body ownership illusion (Slater et al., 2010), indicates higher vulnerability of immersed players, but this seems to be balanced at least for highly invested users by the desire to experience those very effects and thus accept higher social risk.

VR technology finds more users, SVR ecology will likely become the model of further digital community politics to come.

New forms of governing through design have been conceptualized for urban planning (Ghertner, 2015) or (social) media studies (A. Elias et al., 2017) before, but they become even more relevant when and where the virtual production of space, bodies, and sociality merge. Reflecting on the structural role of possibilities and limits of aesthetic creation in VR, of how it forms the basis for making sense of, representing and sensually experiencing bodies and worlds, of its entanglement with economic flows and the production of social order acknowledges the intuition that "the affordances that designers and other practitioners deem important will inevitably shape an extensive portion of human social interactions today and in the future" (Kolesnichenko et al., 2019). Design decisions for social worlds are always political decisions, and aesthetic governance is an important part of intersectional affective biopolitics in a mediatized world.

If we regard the current two largest SVR platforms as for how their different paradigms of worldbuilding and aesthetic creation relate to democratic culture, we cannot ignore the fact that both platforms are proprietary infrastructures run by competing private companies – spawning and harboring social communities is their mode of redeeming the venture capital invested in them. It is the economic allure of the "metaverse" that redesigning and virtualizing the social might allow increased extraction of value from the very basic human need of being with others. Both platforms are examples of possible pathways towards the likely conflictual realization of this goal, while also affording new ways of relating to and mediating the complexities of corporeal existence. These pathways differ from the start – one beginning as an integrated business concept with thoughtful planning, the other as an experiment growing out of a VR tech enthusiast community trying and often struggling to keep up with its own development –, but whether they also lead to different outcomes depends on the politics negotiated between the respective companies and their users. Even more and most importantly, it depends on whether this provider-client relationship can evolve into something resembling democratic citizenship, instead of just being the virtual model for the ongoing real-world transformation of citizens into customers.

Paradoxically, while *Rec Room* takes on the "classical" role of a governing state much more than *VRChat* – setting and enforcing social policies, controlling the economic infrastructure, regulating the possible and impossible relations of what is "normal" and what is not, ensuring fairly equal access

for different (hardware) populations –, its users seem to regard it more as a regular online game provider than those of *VRChat* treat "their" platform. This might for one be because the libertarian plurality of *VRChat* indeed resembles the current image of a neoliberal (post)democracy more than the "all fun and games" uniformity of *Rec Room* does, from its unregulated laissez-faire economy and identity-based cultures down to the rituals of partaking in mass demonstrations (like the review bombing mentioned earlier) or performing the disgruntled citizen alienated from "the powers that be". The more powerful element charging this relationship, though, might be the higher grade of embodiment afforded by the platform, tethering its core user base much more intensely to the experience of having a second body living a second social life in a second reality. Many invested *VRChat* users choose the platform not primarily for leisure or monetary gain, but because it allows them to realize themselves – to become, and to become with others. If the claim to diversity and plurality of current (liberal) democracies is to be taken seriously, then this indicates that these concepts will need to mean more in SVR than just choosing the skin color and gender attributes of an otherwise standardized 3D comic character or even embodying a "realistic" 3D scanned copy of ones own physical body: it rather means the ability and possibility to access the "morphological freedom" the technology promises in the first place.

On another note, the economic aspects of this freedom have only begun being tested. Who controls the infrastructures facilitating the production and trade of virtual bodies? What does body ownership in VR mean not as a psychological effect, but as a social question stretching between the communicative practice of fast-swapping dozens of freely copyable avatars during a single conversation on the one hand and identifying with a unique virtual body, demanding structural protection of its integrity and uniqueness, on the other hand? Who will profit off the existence of virtual bodies to start with? Will certain ways of looking be valued and prized higher than others, as is true for much of the physical world, or will beauty and its valorization become subject to a radical re-negotiation amongst bodies-as-humans, bodies-as-animals, bodies-as-objects, bodies-as-rooms and other yet unimaginable forms of being or being-experienced?

Companies invested in building a "metaverse" extending or even supplanting the "real world" as the primary realm of the social[38] are quick to acknowledge that platforms that "enable anybody to create and share their own social virtual worlds [...] shouldn't be built privately, but rather alongside a passionate community who can help shape the future"[39]. While it stands to reason that platforms are eager to enlist their users' labor for building their virtual realities – especially when they do not have to pay for them for their work –, it is yet another question who will own these realities. The more the actual fabric of a platform consists of the results of its users' creative labor, the more contested this question will and should be. Asking for the distribution and implementation of aesthetic governance can give us hints on how it could or should be answered.

References

Au, W. J. (2021, September 22). Inside VRChat's Informal Economy: An Estimated 350+ People Make Real Money from that Metaverse—Many Likely Earning a Full-Time Income! UPDATE: More Like Over 1000 [Weblog]. *New World Notes*. https://nwn.blogs.com/nwn/2021/09/vrchat-community-economy-gumroad-patreon-youtube-1.html

Au, W. J. (2022a, January 4). VRChat User Concurrency Hit Nearly 90,000 Last New Year's Eve! [Weblog]. *New World Notes*. https://nwn.blogs.com/nwn/2022/01/vrchat-concurrency-2021.html

Au, W. J. (2022b, April 13). BREAKING: Rec Room Has Over 3 Million Monthly Active VR Users, Most on Meta's Quest 2; That's 10x More Users Than Meta's Horizon Worlds & Venues! [Weblog]. *New World Notes*. https://nwn.blogs.com/nwn/2022/04/rec-room-vr-mau-2022-install-base.html

Blackwell, L., Ellison, N., Elliott-Deflo, N., & Schwartz, R. (2019). Harassment in Social Virtual Reality: Challenges for Platform Governance. *Proceedings of the ACM on Human-Computer Interaction*, *3*(CSCW), 1–25. https://doi.org/10.1145/3359202

Bollinger, S. (2023, November 4). *Why are there SO many Trans people in VRChat? Gender, Identity, and Self Discovery.* [Video]. YouTube. https://www.youtube.com/watch?v=Mh3HTTa5NFU

[BVR]. (2022, September 29). *Why is Everything SO EXPENSIVE in Rec Room?* [Video]. YouTube. https://www.youtube.com/watch?v=dnkq0m2lEeA

38 *Rec Room* cofounder Nick Fajt's vision as relayed by their main provider of investment capital: "[W]hereas the the [sic!] last era of social centered on sharing real world experiences online, the next era of social would be centered on both creating and sharing these moments online" (Zhan, 2020).

39 Quoted from the developer statement about *VRChat*'s "early access" status on *Steam:* https://store.steampowered.com/app/438100/VRChat/ [accessed 2023, December 5].

Elias, A., Gill, R., & Scharff, C. (2017). Aesthetic Labour: Beauty Politics in Neoliberalism. In A. S. Elias, R. Gill, & C. Scharff (Hrsg.), *Aesthetic Labour* (S. 3–49). Palgrave Macmillan UK. https://doi.org/10.1057/978-1-137-47765-1_1

Freeman, G., Zamanifard, S., Maloney, D., & Adkins, A. (2020). My Body, My Avatar: How People Perceive Their Avatars in Social Virtual Reality. *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*, 1–8. https://doi.org/10.1145/3334480.3382923

Ghertner, D. A. (2015). *Rule by aesthetics: World-class city making in Delhi*. Oxford University Press.

Handley, R., Guerra, B., Goli, R., & Zytko, D. (2022). *Designing Social VR: A Collection of Design Choices Across Commercial and Research Applications*. https://doi.org/10.48550/ARXIV.2201.02253

Hayden, S. (2020, December 1). 'Rec Room' Now Lets Premium Users Sell Creations for In-game Currency [Weblog]. *Road to VR*. https://www.roadtovr.com/rec-room-sell-items-tokens-economy/

Hern, A. (2022, October 12). Meta's virtual reality project will finally have legs – literally. *The Guardian*. https://www.theguardian.com/technology/2022/oct/12/meta-virtual-reality-legs-mark-zuckerberg-facebook

Jonas, M., Said, S., Yu, D., Aiello, C., Furlo, N., & Zytko, D. (2019). Towards a Taxonomy of Social VR Application Design. *Extended Abstracts of the Annual Symposium on Computer-Human Interaction in Play Companion Extended Abstracts*, 437–444. https://doi.org/10.1145/3341215.3356271

Kolesnichenko, A., McVeigh-Schultz, J., & Isbister, K. (2019). Understanding Emerging Design Practices for Avatar Systems in the Commercial Social VR Ecology. *Proceedings of the 2019 on Designing Interactive Systems Conference*, 241–252. https://doi.org/10.1145/3322276.3322352

Lang, B. (2022, April 13). Virtual Social Platform 'Rec Room' Hits 3 Million Monthly Active VR Users [Weblog]. *Road to VR*. https://www.roadtovr.com/rec-room-monthly-active-vr-users-3-million-peak/

Liu, Q., & Steed, A. (2021). Social Virtual Reality Platform Comparison and Evaluation Using a Guided Group Walkthrough Method. *Frontiers in Virtual Reality*, *2*, 668181. https://doi.org/10.3389/frvir.2021.668181

McVeigh-Schultz, J., Kolesnichenko, A., & Isbister, K. (2019). Shaping Pro-Social Interaction in VR: An Emerging Design Framework. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 1–12. https://doi.org/10.1145/3290605.3300794

Quent, M., Vogl, S. L. (2025). Identity and Safety in Social VR: Findings from an Experimental Avatar-Based Interview Study. Interdisciplinary Perspectives on the Potentials and Perils of Democratic Culture in the Metaverse, 167–179. Nomos.

[Retr0]. (2021, March 28). *The Evolution of Rec Room (Release, 2016 and 2017)* [Video]. YouTube. https://www.youtube.com/watch?v=PH7Wsa2V1Nk

Schultz, R. (2024, April 29). Welcome to the Metaverse: A Comprehensive List of Social VR/AR Platforms and Virtual Worlds (Including a List of Blockchain Metaverse Platforms) [Weblog]. *News and Views on Social VR, Virtual Worlds, and the Metaverse*. https://ryanschultz.com/list-of-social-vr-virtual-worlds/

Shriram, K., & Schwartz, R. (2017). All are welcome: Using VR ethnography to explore harassment behavior in immersive social virtual reality. *2017 IEEE Virtual Reality (VR)*, 225–226. https://doi.org/10.1109/VR.2017.7892258

Slater, M. (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), 3549–3557. https://doi.org/10.1098/rstb.2009.0138

Slater, M., Lotto, B., Arnold, M. M., & Sanchez-Vives, M. V. (2009). How we experience immersive virtual environments: The concept of presence and its measurement. *Anuario de Psicología*, *40*.

Slater, M., Spanlang, B., Sanchez-Vives, M. V., & Blanke, O. (2010). First Person Experience of Body Transfer in Virtual Reality. *PLoS ONE*, *5*(5), e10564. https://doi.org/10.1371/journal.pone.0010564

Tanenbaum, T. J., Hartoonian, N., & Bryan, J. (2020). „How do I make this thing smile? ": An Inventory of Expressive Nonverbal Communication in Commercial Social Virtual Reality Platforms. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 1–13. https://doi.org/10.1145/3313831.3376606

Taylor, M. (2022, July 26). VRChat is being review bombed to hell and back after anti-cheat announcement. *PC Gamer*. https://www.pcgamer.com/vrchat-is-being-review-bombed-to-hell-and-back-after-anti-cheat-announcement/

Thompson, G. S. (2014, May 19). *Virtually Incorrect with Gunter—Episode #5* [Video]. YouTube. https://www.youtube.com/watch?v=jxQRzOckQu0

[Tupper]. (2021a, January 5). *VRChat's New Years 2021—Or, what the $%@& was that?* [Weblog]. https://medium.com/vrchat/vrchats-new-years-2021-or-what-the-was-that-d84334789f77

[Tupper]. (2021b, June 25). *VRChat Partners with Anthos Capital to Close $80M Series D* [Weblog]. https://medium.com/vrchat/vrchat-partners-with-anthos-capital-to-close-80m-series-d-d706a1ab481b

Turkle, S. (2006). Tethering. In C. A. Jones & B. Arning (Hrsg.), *Sensorium: Embodied experience, technology, and contemporary art* (1st MIT Press ed, S. 220–226). MIT Press : The MIT List Visual Arts Center.

Zhan, S. (2020, December 18). *RecRoom: A digital third place for kids at heart* [Weblog]. https://medium.com/sequoia-capital/a-digital-third-place-for-kids-at-heart-2646dfd55e00

Zheng, Q., Wang, L., Do, T. N., & 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, New York, NY, USA. https://doi.org/10.48550/arXiv.2204.07121

Unpixelated Hate? Lessons from VR Gaming for a Digital Civil Society

Mick Prinz

Introduction

In Germany, over half of 6–69-year-olds play computer and video games (Game – Verband der deutschen Games-Branche e. V., 2023). While many game titles are increasingly breaking stereotypes, some video games incorporate anti-democratic narratives and promote discrimination in their plots and character design. It is not just the game settings that have political implications, but also the gaming communities that are becoming increasingly political. Discussions on topics such as Black Lives Matter (GameStar Redaktion, 2022), the invasion of Ukraine, or the situation in the Middle East are also taking place in in-game chats and on gaming platforms. These debates are often characterized by toxic arguments and discrimination against marginalized positions – a climate that anti-democratic actors exploit. Gaming spaces with poor moderation are used to spread misanthropic narratives, intimidate people, and strive for international networking. Players experience toxic agitation on gaming platforms such as Twitch, Steam, or Roblox, as well as in in-game spaces, voice or text chats, and virtual realities. The intensity of hate and discrimination varies within immersive gaming realities.

In this article's first step, the nature of video game culture will be examined, and a definition of immersive gaming realities will be provided. Various examples will be outlined, and the cornerstones of VR gaming will be traced. The topic of dark participation and toxicity in video game contexts will then take center stage. Different levels of discrimination will be outlined, and examples of reactionary agitation within different gaming communities will be discussed. The extreme right in existing immersive realities will also be the focus. Three forms of instrumentalization will be examined here: networking, mobilization, and the 'meta-politics' of the extreme right in video game contexts. These examples emphasize the challenges that metaverse-like structures face and the tasks they will have to deal with in the future. The text introduces the idea that immersive realities

have great potential for promoting democratic values. Both titles created for entertainment purposes and serious games (games with an educational focus) are discussed here. These types of video games can effectively convey values through virtual realities. The passage concludes with a summary of the lessons and challenges for designing metaverse-like structures based on the cultural aspect of gaming.

What do we Mean by Immersive Realities in Video Games?

The term „immersion" initially refers to being fully immersed in a specific environment or setting (Murray, 1998). Digital technologies such as virtual reality, augmented reality (AR), mixed reality (XR), and the entire ecosystem of digital games are all categorized as immersive realities (Mühlhoff & Schütz, 2019; Nilsson et al., 2016). Video games often contain elements of immersive realities. They have a virtual environment that is partially disconnected from the physical world but can still be based on reality. A prime example of this is the popular game Minecraft, which has over 180 million users each month in 2023 (Redaktion G-Portal, 2023). Players can create and design entire worlds, establish ecosystems, and work towards cooperative or competitive goals. Along with countless fictional worlds, Minecraft also has servers where players can recreate real German cities, while elsewhere, digital depictions of Nobel Peace Prize winners educate players on basic democratic values (Diedrich, 2022). Additionally, the project „Reporters Without Borders" created the „Uncensored Library" in a digital Minecraft environment to showcase works that are banned in autocratic regimes. Thanks to the free accessibility of Minecraft's immersive reality, progressive literature remains accessible even in dictatorships (Redaktion Reporter ohne Grenzen e. V., 2023).

Figure 1: Screenshot of the Saudi Arabia section in the "Uncensored Library" (source: Reporters Without Borders, 2023)

Another important aspect of immersive realities can be seen in video games: avatarization. Players often have the opportunity to take on the role of a fictional character, create their own storyline, or follow a predefined narrative. Designing a unique character is also a central component, particularly in (online) role-playing games. Many modern titles implement a wide range of options for customizing your own digital image closely to real-life conditions if desired. Titles such as the space role-playing game "Starfield" allow for a free choice of pronouns, while other examples such as the fantasy role-playing game "Baldur's Gate 3" allow players to choose their voice, gender attribution, and external characteristics independently. While many players welcome these diversity options, allowing players to correctly integrate their own identities into a game, other gamers react with reactionary rejection and toxic criticism (Seng, 2023).

It is also noticeable that many online video games enable multi-layered interaction options with other entities. One example of this is the escapist simulation game „Animal Crossing: New Horizons". A game mechanic here makes it possible to visit neighboring islands of other players. In addition to social interaction opportunities, a visit can also be economically motivated. Different islands mean potentially different beet prices – a rare commodity for which prices vary daily and are different on different players' islands. The focus here is less on social motives and more on the capitalist maximization of one's own money within the game cosmos (Zeitz, 2020). In addition to these obvious forms of interaction dictated by the game, Animal

Crossing has also attracted attention through political protest. Players join forces and demonstrate in favor of a free Hong Kong, the anti-racist Black Lives Matter movement, or against toxic corporate structures in gaming companies (Mack, 2020).

Figure 2: Black Lives Matter protest in The Sims 4 (source: BlackSimsMatter, 2020)

Digital protests with references to real-life conflicts take place in immersive realities in many places, while other titles focus on co-operative or competitive scenarios. Potential characteristics of an immersive gaming reality therefore include the creation of separate ecosystems, the ability to live through fictional or realistic scenarios with an avatar, shaping stories, and interacting with other players.

Gaming and the Changing Relationship with Virtual Reality

Debates about the use of VR technology are nothing new. Decades before Mark Zuckerberg and Meta publicized the idea of a metaverse-like, digital parallel world, the first attempts at VR were already being made in the game industry. In the 1980s, projects tried their hand at the first VR versions, but the VR arcade machines suitable for the masses were not created until 1991 by the company "W-Industries" (Redaktion TechMonitor, 1991). Based on the Amiga 3000, W-Industries presented the one-kilogram arcade machine.

Two years later, the SEGA VR was announced, which was reminiscent of today's VR glasses for the first time (Bezmalinovic, 2017). However, due to immense production costs caused by expensive motion sensors and the then still unsolvable problem of motion sickness, SEGA's VR goggles were never released. In the years that followed, there were repeated attempts to use VR-like technology in the gaming world, such as Nintendo's "Virtual Boy" in 1995 (Norman, 2013). Ultimately, however, insufficient processors, low frame rates, high latency, high production costs, and extensive reports of eye and headaches put a stop to these first steps in VR gaming.

It was not until 2012 that the newly founded company "Oculus VR" once again attempted to integrate VR into the gaming market. The VR goggles "Oculus Rift" were financed via crowdfunding (Harris, 2019). Meta discovered the potential of the VR market and bought Oculus VR. Many other companies, such as HTC and Microsoft, signalled their interest and tried their hand at their own VR hardware. Once again, the virtual reality boom failed to materialize, with VR headsets from various manufacturers selling too poorly. This was mainly due to the lack of major game titles that would have justified a purchase for gamers. It wasn't until "Meta Quest" and "Meta Quest 2" and games such as the rhythm lightsaber game "Beat Saber" (2018) and "Half-Life-Alyx" (2020) that VR technology achieved higher sales figures and the interest of many video gamers. However, there is currently disagreement in the gaming community. With the PlayStation VR2 released by Sony, the company is demonstrating how powerful current video game hardware is, and high-resolution titles such as "Horizon Call of the Mountain" are impressing the trade press and gaming public. However, there is still no broad catalogue of games for the VR glasses from Sony, Meta, or Apple that would anchor the technology more firmly among gamers. As of 2023, VR remains a niche in gaming culture that is perceived by the broad mass of gamers but is far from being favored by them.

Dark Participation, Toxicity and Gaming

Despite the current scepticism towards VR hardware, immersive realities and virtual reality have long been an integral part of gaming culture. Video gamers are usually part of various gaming communities that form around different genres, influencers, or game titles. The majority of gamers have already experienced hate in gaming spaces – whether in-game chats, on platforms, or through private messages. According to the "Hate is no

Game" study by the Anti-Defamation League, 77 % of over-18s have experienced hate speech in video game contexts. Alarmingly, even 20 % of adult gamers have encountered right-wing extremists within their communities. Primarily, female gamers, non-white gamers, Jewish gamers, and parts of the LGBTQIA+ community are affected by hate in gaming. These groups are defamed because they belong to specific groups (e.g. Jews or African Americans) (Anti-Defamation League, 2022).

"Dark participation" or toxic behavior primarily excludes marginalized groups from various fields of interaction in immersive realities. The term "dark participation" encompasses insults, hate speech, sexual harassment, discriminatory profiles/nicknames, doxxing, and the spread of disinformation. In gaming practice, there are various levels at which dark participation becomes visible. Firstly, there is the "in-game" level. Whether via text message or voice chat, almost every online video game must deal with insulting, patronizing, and even misanthropic statements within player communication. Studios and game makers try to reduce hate speech with moderation tools and word filters, but toxic gamers adapt systems and circumvent restrictions, for example with leetspeak (i.e. mixing letters and similar-looking numbers). For example, many e-athletes and gamers report on toxic everyday life in the video game "League of Legends" (Gogoll, 2021) or female gamers report sexism in the voice chat of the shooter "Valorant" (eSports-Redaktion, 2021). VR video games, in particular, harbor the risk of ableist, racist, or misogynistic discrimination reaching those affected unfiltered. There are examples of "Among Us VR" or VR sports games in which, for example, the N-word is uttered by an avatar in the immediate vicinity. It is not just gaming experiences, but also misanthropy that is made more immersive by VR glasses (MayoMonkey Bowblax[1], 2022).

Another level at which dark participation in gaming takes place is the platform level. Toxic communication is also noticeable on various gaming platforms outside of the actual game. For example, when Black Lives Matter (BLM) is discussed on Steam, a platform popular with PC gamers. In addition to solidarity with George Floyd and the BLM movement, gamers reproduce racism and spread disinformation (Amadeu Antonio Foundation, 2022). Elsewhere, toxic gamers postulate that diversity in video games leads to a boycott and that political attitudes must be kept out of normative works.

1 Content warning about racism.

Figure 3: Toxic user review of the Horizon Forbidden West expansion on Metacritic. The narrative: Children are "homosexualized" by the game story (source: Metacritic, 2023)

The comments on the game "Starfield," where players can freely choose pronouns for their avatars, are characterized by such subjective criticism (Steam, 2023a). Other interactive platforms like Twitch are also not safe spaces for marginalized groups. Apart from discrimination in livestream chats, there are countless talk shows where far-right opinions are given a platform (Belltower.News, 2021).

Immersive Realities and the Far Right

Dark participation is gaining ground in various areas of the gaming world with little resistance. This is largely due to players ignoring problematic content and reporting functions being underutilized. In addition, platform operators are inconsistent and lacking in their moderation efforts. As a result, extreme right-wing actors are also active in immersive realities, alongside toxic voices. There are three ways in which the extreme right utilizes immersive realities.

The first is networking (1) within digital worlds. In the online mode of the popular action game "Grand Theft Auto V," users can create their own avatars. On their own role-playing servers, players establish their own objectives and follow a daily routine. Within the digital world, players who follow the SHAEF conspiracy narrative also gather. In GTA Online, they

even go on digital patrol. SHAEF refers to the "Supreme Headquarters Allied Expeditionary Force," a term used by supporters of this conspiracy theory claiming that Germany is an occupied state, a "BRD GmbH" (De:hate, 2021). The platform Roblox allows players to create their own mini games/"experiences" and offer them for free download, but it also features numerous right-wing extremist worlds and experiences. On one map, players can assume the role of an SS officer and march through a fictionalized 1940s Berlin, greeting others with the Hitler salute. In other areas, players navigate through an extermination camp where the Shoah is trivialized. Digital avatars can also participate in target practice on swastika emblems. Known for its popularity among 9–12-year-olds, Roblox has over 200 million active users. Many of the far-right "experiences" can be found with just a few clicks (Prinz, 2023). Another title on Steam, however, is more subtle. The description states that "[game name] is a truly immersive and uncensored social experience where you can engage with like-minded people in a mutually constructive way." At first glance, this may sound promising. However, the game allows for the expression of any political opinion in a digital environment. It incorporates symbols from the Canadian Covid-19 protest movement and the "QANON" conspiracy ideology (Steam, 2023). This use of immersive realities as alternative spaces for interaction is a tactic employed by extreme right-wing actors to build their own world of experience with like-minded individuals, without encountering resistance.

This is also connected to another motivation: gaming and immersive realities are seen as part of a cultural struggle for the right, known as "metapolitics" (2). The "new right" considers "metapolitics" as a strategy for shifting social discourse to the right by occupying pre-political spaces. In practice, this can involve the dissemination of right-wing extremist symbols and codes or the glorification of extreme right-wing actors within immersive gaming realities and on video game platforms. For instance, there are numerous profiles on Steam that use the far-right slogan "Read Siege", inspired by the instructions for "race war" and "leaderless resistance" followed by the right-wing terrorist group "Atomwaffen Division" (Ayyadi, 2020). The perpetrators of Christchurch, Halle and Buffalo were also praised on Steam through their profile pictures and descriptions (Gensing, 2020). Roblox even had its own video game worlds where players could assume the role of right-wing terrorists and re-enact these attacks. For example, one downloadable experience allows players to play as the right-wing extremist from Buffalo and replay the attack on the "Tops" supermarket.

Figure 4: Right-wing extremist "experience" on Roblox. Here, players can re-enact the Buffalo attack (source: Roblox, 2023)

The attacks in Christchurch and Halle were available on the platform until the moderators reacted months later (Graf & Marx 2022). The far right primarily uses gaming areas where they can create and design their own content. Sandbox-like game concepts such as "Roblox" and "Minecraft" are used here, as well as the workshop area of Steam. Modifications (mods) for existing games can be developed here. This includes mods in which players can play as the Waffen SS, a legitimate faction in "Company of Heroes", or select Adolf Hitler as head of state in "Civilization VI" (Steam, 2023b). Therefore, it primarily utilizes existing gaming infrastructures in which right-wing extremists push their "meta-politics". Occasionally, neo-Nazis also try to develop their own propaganda games, such as a right-wing developer studio belonging to the „Identitäre Bewegung". In this game, players can choose a cadre figure from the so-called "new right" and fight against right-wing enemies as Martin Sellner or Alexander "Malenki" Kleine in a dystopian 2D pixel platformer (Jugendschutz.net, 2020). Developing independent VR games does not seem to be a priority for right-wing extremist groups.

The final motive for the extreme right to use immersive gaming spaces is mobilization (3). Extreme right-wing actors use gaming evenings or patriotic video game tournaments to appeal to potentially interested parties and reactivate their own in-group. Roblox's immersive realities have been used to target young people. Young players were invited to join right-wing extremist groups and motivated to participate in joint shooting exercises (Vollbild, 2023). Elsewhere, merchandise from conspiracy ideology influencers is distributed in the game "Second Life". Player avatars act as digital billboards for disinformation. Right-wing extremist groups also appear in

existing VR games like "VRChat" to spread and provoke inhumane narratives, but also to arouse interest. For example, players in Ku Klux Klan outfits have appeared in the VRChat game and given racist monologues (ScaryToaster, 2022).

Figure 5: Screenshot from VR Chat. A person in a KKK outfit harassed other players (source: VRChat, 2022)

It is clear that the far right uses various areas of immersive realities to spread racist and anti-Semitic ideological elements, (re)mobilize people, and network at the same time. Creating their own avatars, implementing gestures, and creating their own immersive realities (such as in the Roblox sandbox or Minecraft) harbors potential dangers. As a result, parts of the digital world are becoming places where those affected by discrimination, in particular, cannot move freely and safely.

Democracy in VR and Immersive Realities?

Gaming culture highlights the challenges faced by immersive realities and metaverse-like spaces. At the same time, experiences from various video game communities emphasize that digital worlds harbor democratic potential. More and more video games convey a canon of values that emphasize plurality and tolerance. This also applies to video games that primarily serve entertainment purposes and have a commercial interest. For example, the game "The Last of Us 2" implements a queer love triangle and deals with experiences of discrimination against trans people. "Spider Man 2", on the other hand, depicts anti-racist and diverse narratives in its setting, while modern role-playing games, such as "Baldur's Gate 3", break with

binary gender stereotypes in the character editor (Hart, 2022). Elsewhere, the action game "Assassin's Creed Mirage" provides insights into Islamic beliefs. So-called "serious games", i.e. video games with educational claims, are also significantly more immersive than the static educational games of the early 2000s. "The Darkest Files", for example, is a historical investigation and courtroom game. Players are part of Fritz Bauer's team of public prosecutors and have to bring Nazi perpetrators to justice (Steam, 2023c). The game "Hidden Codes" by the Anne Frank Educational Centre, on the other hand, provides information about current right-wing extremist symbolism. Thanks to a classroom version, the title can be used in school lessons without much effort (Bildungsstätte Anne Frank, 2023). Democratic values can thus be played and experienced.

Democratic movements are also articulated within immersive realities. For example, hundreds of players come together every year in the world of Tyria in the online role-playing game Guild Wars 2 to organize a protest march for the LGBQIA+ community. In addition to its symbolic effect, donations were also collected for initiatives that take a stand against misanthropy (Guild Wars 2 Forum, 2023). Elsewhere, players in the worlds of "The Sims" and "Animal Crossing" show their solidarity with the Black Lives Matter movement by designing cosmetic clothing with BLM lettering and highlighting everyday racism and discrimination in digital demonstrations (Grayson, 2020). In response to the Russian invasion of Ukraine, players in the world of "The Elder Scrolls Online" took a stand and organized a march for peace for those affected in Ukraine (Marasigan, 2022). Democratic protests are no longer a rarity, especially in online video games where gamers create their own avatars and organize themselves into player communities.

Figure 6: Pride Parade in the game Guild Wars 2; once a year, players organize a solidarity demonstration for the LGBQIA+ community (source: Guild Wars 2, 2022)

Within VR worlds, on the other hand, there are only a few pilot projects that promote greater awareness of discrimination. On one hand, this is due to the limited distribution of VR headsets compared to other gaming hardware. On the other hand, it is also due to the educational and didactic challenges that immersive experiences in VR worlds present. The depiction of discriminatory and hateful statements in virtual realities can be traumatic, especially for those affected, due to the immersion. The "Debug" project therefore accompanies users in their VR experience away from the headset. Players are followed in VR by a swarm of mosquitoes that continuously utter racist stigmas. Debug conveys the omnipresence of everyday racism for those affected and the intensity of stigmatization (Prinz, 2022). In the title "1,000 Cut Journey", on the other hand, players experience different life cycles and perspectives of a POC protagonist. Whether during disciplinary measures in the classroom, as a teenager by the police, or as a young adult through discrimination in the workplace – racism runs through the life of the character (Games for Change, 2022).

Another VR experience entitled "I Am A Man" makes the history of the American civil rights movement of African Americans tangible. The VR presentation aims to provide personal understanding of the struggles of marginalized groups. The title uses historical film footage, photographs, and voice recordings of actual participants in the civil rights movement. "I

Am a Man" sees itself as an interactive documentary experience to provide a deeper awareness of anti-racist struggles (Meta, 2018).

Figure 7: Screenshot of the VR game "I AM A MAN" (source: Meta, 2018)

Virtual realities cannot fully depict the suffering and pain of those affected. However, they attempt to make the perspectives of those affected accessible to those who have not yet had to deal with discrimination from privileged perspectives or who simply did not want to. They can also generate a sense of closeness and tangibility to historical events through the integration of historical sources and make processes more tangible.

Conclusion and Appeal for an Democratic Civil Society

The experiences of immersive realities and VR experiences in the context of gaming culture highlight opportunities, possibilities, and challenges that future metaverse-like ecosystems will face. Extreme right-wing actors utilize various segments of immersive reality, while democratic movements also make use of these spaces. In conclusion, clear recommendations for participation in future virtual realities are formulated.

- *1) Avatarisation*: If VR spaces allow users to create and design their own avatar, diversity options must be considered. An editor that only thinks in terms of binary gender models or ignores skin or hair types excludes broad sections of society. It is also important to ensure that discriminatory and/or criminally relevant symbolism cannot be reproduced.
- *2) Moderation of inhuman content*: In addition to comprehensive reporting and notification functions, proactive moderation is necessary

to identify hateful symbolism and recognize anti-democratic narratives. It is also important to make it more difficult for extremist groups to participate in VR. Anti-hate language programs are already being tested in video game rooms and criminal content should ideally be recognized in its early stages.

- *3) Promote democratic participation*: Immersive realities can be used to draw attention to problematic aspects in the analog world and to show solidarity with those affected. Ideally, VR spaces should provide an infrastructure that not only enables but also promotes democratic participation. For example, through moderated discussion forums, lectures, or workshops in immersive realities. Positive examples such as the "Uncensored Library" (Minecraft) should be made accessible to users.
- *4) Protect marginalized groups/security settings:* The so-called "metapolitics" and mobilization attempts in games such as Roblox or VRChat make it clear that extreme right-wing actors use immersive realities strategically. People affected by discrimination in particular, but also all other users, must be able to make comprehensive security settings and be informed about them. It should be possible to mute text and voice chats, hide gestures, and activate a security area around your own avatar. This is essential to guarantee physical proximity if desired.

In order to ensure democratic participation and not exclude marginalized positions, these four points are a minimum requirement for immersive realities.

The article emphasized the potential of VR spaces and digital worlds for a progressive understanding of values. At the same time, there are many pitfalls and anti-democratic endeavors that need to be taken into account in the potential technology of tomorrow. The problem becomes particularly clear with increasingly realistic graphics in immersive realities: if right-wing terrorist attacks, such as those in Christchurch or Halle, switch from a static block Roblox graphic to a hyper-realistic guise, the pull effect for extremist actors will increase. Developers, companies, authorities, and digital civil society must all play their part in a democratic metaverse. The lessons learned from game culture should not be forgotten.

References

Anti-Defamation League. (2022). *Hate no game: Hate and harassment in online games*. https://www.adl.org/resources/report/hate-no-game-hate-and-harassment-online-games-2022

Ayyadi, K. (2020, August 6). James Masons Anleitung zum „Rassenkrieg" für die „Atomwaffen Division". *Belltower.News*. https://www.belltower.news/siege-james-masons-anleitung-zum-rassenkrieg-fuer-die-atomwaffen-division-94551/

Belltower.News – Redaktion. (2021, November 25). Gaming und Rechtsextremismus im Livestream: Was passiert auf Twitch? *Belltower.News*. https://www.belltower.news/gaming-rechtextremismus-im-livestream-was-passiert-auf-twitch-111417/

Bezmalinovic, T. (2017, August 10). Sega VR: Vom Aufstieg und Fall der legendären VR-Brille. *MIXED*. https://mixed.de/sega-vr-vom-aufstieg-und-fall-der-legendaeren-vr-brillle/

Bildungsstätte Anne Frank. (2023). *Hidden Codes*. https://www.hidden-codes.de/

De:hate. (2021, January 27). „Commander" Jansen verbreitet Trumps Willen auf Telegram. *Belltower.News*. https://www.belltower.news/shaef-verschwoerung-commander-jansen-verbreitet-trumps-willen-auf-telegram-123213/

Diedrich, S. (2022, December 10). Minecraft und der Dalai Lama. *GameStar*. https://www.gamestar.de/artikel/minecraft-nobel-dalai-lama,3378349.html

eSports-Redaktion. (2021, May 3). Streamerinnen ranten auf Twitter: So groß ist das Sexismus-Problem in Valorant. *esports.com*. https://www.esports.com/de/streamerinnen-rant-auf-twitter-so-gross-ist-das-sexismus-problem-in-valorant-242179

Game – Verband der deutschen Games-Branche e.V. (2023). *Jahresreport 2023*. https://www.game.de/wp-content/uploads/2023/08/230809GME_Jahresreport_2023_168x240_DE_Web.pdf

GameStar Redaktion. (2022, June 10). Live-Stream mit der Chefredaktion: Black Lives Matter und der Einfluss auf die Gaming-Branche. *GameStar*. https://www.gamestar.de/artikel/live-stream-mit-der-chefredaktion-blacklivesmatter-und-der-einfluss-auf-die-gaming-branche,3358410.html

Games for Change. (2022). *1000 Cut Journey*. https://www.gamesforchange.org/games/1000-cut-journey/

Gensing, P. (2020, May 18). Steam-Nutzer glorifizieren Attentäter. *Tagesschau*. https://www.tagesschau.de/investigativ/steam-christchurch-terrorismus-101.html

Gogoll, M. (2021, July 14). Trainer über Diskriminierung im E-Sport. *taz*. https://taz.de/Trainer-ueber-Diskriminierung-im-E-Sport/!5812767/

Graf, M., & Marx, M. (2022, October 27). Hat das Gaming ein Rechtsextremismus-Problem? *GameStar*. https://www.gamestar.de/artikel/podcast-gaming-rechtsextremismus,3387301.html

Grayson, N. (2020, June 9). Sims players hold virtual Black Lives Matter rally. *Kotaku*. https://kotaku.com/sims-players-hold-virtual-black-lives-matter-rally-1843998733

Guild Wars 2 Forum. (2023, June 24). Tyria Pride 2023. *Guild Wars 2 Forum*. https://en-forum.guildwars2.com/topic/129869-tyria-pride-2023/

Harris, B. J. (2019). *The history of the future: Oculus, Facebook, and the revolution that swept virtual reality*. Dey Street Books.

Hart, A. (2022, December 8). Baldur's Gate 3 adds non-binary option. *Gayming*. https://gaymingmag.com/2022/12/baldurs-gate-3-adds-non-binary-option/

Jugendschutz.net. (2020, September 29). Praxisinfo: Computerspiel Heimat Defender Rebellion. *Jugendschutz.net*. https://www.jugendschutz.net/mediathek/artikel/praxisinfo-computerspiel-heimat-defender-rebellion

Mack, M. (2020, April 10). China verbietet Animal Crossing nach virtuellen Hongkong-Protesten. *zeit.de*. https://www.zeit.de/zett/politik/2020-04/china-verbannt-animal-crossing-nach-virtuellen-hongkong-protesten

Marasigan, M. (2022, March 1). Elder Scrolls Online protest Russia's invasion of Ukraine with peace marches in Cyrodiil. *MMOs.com*. https://mmos.com/news/elder-scrolls-online-protest-russias-invasion-of-ukraine-with-peace-marches-in-cyrodiil

MayoMonkey Bowblax. (2022, November 14). RACIST Father & Son Call BLACK Man N WORD in Among Us VR! [Video]. *YouTube*. https://www.youtube.com/watch?v=qbuAtWkr46s

Meta. (2018). *I Am A MAN*. https://www.meta.com/de-de/experiences/pcvr/1558748774146820/

Mühlhoff, R., & Schütz, A. (2019). Immersive Medien, Virtual Reality und die Ethik der Simulation. In S. Nagenborg & J. Nink (Ed.), *Handbuch Maschinenethik* (S. 1–18). Springer. https://doi.org/10.1007/978-3-658-17484-2_37-1

Norman, J. M. (2013). Nintendo's "Virtual Boy," the first mass-produced virtual reality game system. *History of Information*. https://www.historyofinformation.com/detail.php?id=3637

Prinz, M. (2022, July 12). Videospiel gegen Rassismus: Virtual Reality ermöglicht einen Perspektivenwechsel. *Belltower.News*. https://www.belltower.news/videospiel-gegen-rassismus-virtual-reality-ermoeglicht-einen-perspektivenwechsel-141205/

Prinz, M. (2023, May 13). Rechte Erlebnisse. *taz*. https://taz.de/!5909022/

Redaktion G-Portal. (2023, August 7). Wie viele Leute spielen Minecraft? *G-Portal*. https://www.g-portal.com/wiki/de/wie-viele-leute-spielen-minecraft/

Redaktion Reporter ohne Grenzen e.V. (2023). *Uncensored Library*. https://www.uncensoredlibrary.com/de

Redaktion TechMonitor. (1991, July 16). W INDUSTRIES MAKES VIRTUAL REALITY A REALITY AT UKP20,000. *TechMonitor*. https://techmonitor.ai/technology/w_industries_makes_virtual_reality_a_reality_at_ukp20000

ScaryToaster. (2022, December 11). Meeting the KKK in VRchat #shorts [Video]. *YouTube*. https://www.youtube.com/watch?v=j6Oxzjt2oxg

Seng, M. (2023, August 18). Pronomen in Games: Wokeness im Weltall. *Frankfurter Allgemeine Zeitung*. https://www.faz.net/aktuell/feuilleton/medien/pronomen-in-games-wokeness-im-weltall-19288095.html

Steam. (2023a). *Starfield*. https://store.steampowered.com/agecheck/app/1716740/

Steam. (2023b). *Sid Meier's Civilization VI*. https://store.steampowered.com/app/2890070/Sid_Meiers_Civilization_VI/?l=german

Steam. (2023c). *The Darkest Files*. https://store.steampowered.com/app/2058730/The_Darkest_Files/

Vollbild. (2023, March 31). Hakenkreuz & Bombenbau: Neonazis unterwandern Kinderspiele [Video]. *YouTube*. https://www.youtube.com/watch?v=-01FPD-SYsc

Zeitz, S. (2020, March 30). Animal Crossing: New Horizons – Rübenpreise vorhersagen. *GamePro*. https://www.gamepro.de/artikel/animal-crossing-new-horizons-ruebenpreise-vorhersagen,3357127.html

Identity and Safety in Social VR: Findings from an Experimental Avatar-Based Interview Study

Matthias Quent & Sara Lisa Vogl[1]

Introduction

Virtual Reality (VR) has evolved in recent years from a niche technology to a widely adopted medium with numerous applications. The modern use of the term VR was coined by the computer scientist Jaron Lanier in the 1980s (LaValle, 2020, p. 5). According to Steven M. LaValle, "The most important idea of VR is that the user's perception of reality has been altered through engineering, rather than whether the environment they believe they are in seems more 'real ' or 'virtual'. A perceptual illusion has been engineered." (LaValle, 2020, p. 6) This immersive technology enables users to enter worlds vastly different from their physical surroundings, opening new dimensions of experience and interaction. Applications of VR systems range from video games and immersive cinema to telepresence (e.g. Google's Street View) and educational settings (LaValle, 2020, pp. 9–22).

Among the most significant and increasingly relevant applications of VR systems are Social VR platforms. Social VR refers to virtual reality platforms that emphasize interpersonal interactions, enabling users to meet, communicate, and engage in shared activities within virtual environments. One of the most popular and pioneering platforms in this domain is VRChat. Gaining broader popularity since its public release in 2017, VRChat offers a diverse array of user-generated worlds and avatars, allowing participants to interact in an almost limitless range of settings. Beyond real-time voice and text-based conversations, users can partake in activities such as games, karaoke, movie nights, and educational sessions. As of 2023, VRChat has over four million registered users, exemplifying the growing appeal of virtual spaces for social connection and collaborative experiences. Users can access the platform via PC, VR headsets and a mobile app on smartphones. High-quality VR experiences require VR headsets connected to powerful computers**.** The technology is particularly useful for displaying high-quali-

1 With assistance of Lilly Walter.

ty, high-resolution avatars. The visual level of detail can be reduced through software settings. Fully immersive experiences require expensive hardware.

Arne Vogelgesang highlights, that „Social VR as a practice can be understood as embodied social role-playing in a system of networked and bounded virtual 3D spaces inhabited by avatars connected to human users." (Vogelgesang, 2024, pp. 1–2) Krell and Wettmann (2023) explore corporeal interactions within the Social VR platform VRChat, emphasizing that users experience intense bodily closeness and intimacy despite spatial separation. Central to their analysis is the concept of "body synchronization", which describes the alignment of physical movements with virtual avatars, leading to the simultaneous perception of physical and digital embodiment in a "coextensive space." This hybrid embodiment results in phenomena such as "phantom sense," where virtual touches trigger physical sensations. Routine interactions such as hugging, head-patting, and dancing, as well as intimate sexual practices (Erotic Roleplay, ERP), illustrate how mediated corporeality facilitates strong affective and social bonds. The authors argue for a theoretical expansion of interaction theory, stressing that mediatised situations should be understood as situated on a continuum of interactional intensity and corporeal involvement, approaching physical co-presence. While VRChat offers vast opportunities for creativity and social interaction, it also presents significant challenges. Sabri et al. (2023) discuss these issues in their study 'Challenges of Moderating Social Virtual Reality' highlighting the difficulties associated with moderating such platforms. The anonymity and deep immersion provided by VR can foster both positive and negative behaviors, posing a challenge for moderators who must ensure a safe and enjoyable environment without stifling user freedom and creativity. These challenges range from combating harassment and abuse to enforcing age restrictions and ensuring privacy. These challenges are not unique to VRChat but reflect broader societal concerns about immersive technologies. Matthias Quent (2023) highlights in his analysis of democratic cultures in the 'next Internet' that while VR offers significant opportunities at an individual level, it also confronts major social challenges especially at a social level (Quent, 2023, p. 30). He sees immersive technologies like VR as offering significant opportunities, such as supporting psychotherapy for anxiety and depression, fostering empathy and inclusion through perspective-taking, and enabling innovative educational experiences. However, they also pose risks, as seen in platforms like Roblox, where user-generated content can promote discrimination, far-right ideologies, and group-focused enmity, violating community standards (Quent, 2023, p. 38). A study

by Hinduja and Patchin (2024) investigates risks and harms experienced by adolescents within metaverse and social VR environments. Based on a representative survey of 5.005 US adolescents aged 13–17, the authors examine the prevalence of various negative experiences, including bullying, hate speech, sexual harassment, and grooming. Their findings reveal notable gender differences, with girls being more frequently targeted for sexual harassment and predatory behaviors, primarily due to their gender. Despite the availability of safety mechanisms within VR platforms, these are underutilized, although girls employ them slightly more than boys.

Reference to the Metaverse

The "Metaverse" refers to a collective virtual shared space. It is a vision of a fully immersive digital universe that people can inhabit, work, play, and socialize in, akin to a massive, interconnected digital society. Platforms like VRChat and other Social VR experiences serve as early indicators or prototypes of what the metaverse might look like in the future. They allow users to transcend physical limitations, enabling them to interact with others in a virtual universe with diverse worlds and activities. In VRChat, for instance, the idea of user-generated content—wherein participants can create their own avatars, design their worlds, and curate their experiences—aligns closely with the broader metaverse vision. Current social VR platforms pave the way for a comprehensive and integrated digital ecosystem, where physical and virtual worlds become increasingly interconnected. As technologies advance and more platforms emerge, the vision of a unified metaverse becomes progressively tangible, with current social VR platforms serving as its foundational building blocks and environments for learning and experimentation.

Research Design and Data Collection Methods

As part of our experimental research, we aimed to find out why people use Social VR, what experiences they have had there, particularly regarding aspects of identity, safety, and hate speech, and how they protect themselves from harassment. Questions asked in the interviews conducted inside the social VR platform VRChat included aspects of gender, harassment, hate speech, and safety.

To approach these questions, we conducted four qualitative guided interviews as avatars in VRChat in autumn 2023. The anonymized participants come from different countries and were recruited as a self-selected sample from the virtual network of VR designer Sara Lisa Vogl. The interviews were transcribed, coded according to the key questions and analyzed. The results of this explorative pilot research are limited and cannot be generalized. However, they provide first-hand insights into user behavior and experiences in Social VR and thus fill an empirical gap in the primarily theoretical and technological research literature. Further research with more interviewees in virtual worlds can build on this. Through the unique ability of VR to completely alter appearances and virtual embodiment, the interviewees can show up in various identities that are unrelated to their real-life bodies and identities regarding age, gender and abilities.

Key Findings from the Interviews

Social VR Usage

To gain insights into the patterns and diversity of VR usage among our participants, and how they relate to their identity and safety issues inside and outside of VR, we asked interviewees when and for what purpose they have been using Social VR. Duration of usage between participants is 2–5 years in Social VR, mainly VRChat. For some of the participants, the experience of Social VR is an alternative to real-life social interaction, while still performing shared activities and feeling connected to friends. Especially for self-described introverts, meetings in Social VR have the benefit of added layers of control over the environment. People can adjust the number of users visible to them as well as colors and audio of virtual worlds and avatars. They can also completely blend out other users they do not want to see or interact with. They describe how they can adjust their VR experience to their preferences and avoid having to be exposed to people or situations they have difficulties coping with: “I do not need to interact with people that I don't want to, and having that extra bit of control over the situation is just amazing for me, especially because I'm an introvert.”

Additionally, they value the control VRChat offers over their social environment. They can adjust the volume of individual users, hide overly bright avatars, or block individuals they do not wish to interact with. They further highlight that in VRChat, people get to know each other mainly based on

their personality instead of their physical appearance. This aspect helps overcome insecurities related to physical appearance, boosting confidence by focusing interactions on the personality projected through the virtual avatar.

For example, Interviewee 2 discusses using VRChat as a means of coping with social anxiety. They explain that the social interaction within VRChat functions as a form of training in a controlled environment where positive feedback encourages them to engage in social exchanges. Interviewee 3 uses Virtual Reality primarily to detach from the stresses of reality and to relax. They express satisfaction with this approach as it allows them to meet new people, establish friendships, and practice their English skills.

Gender-Identity and Embodiment of Diverse Genders

Another aim of our research was exploring how the interviewees experienced and expressed their gender identity in Social VR. We asked them about their gender identity in real life and how it matched or differed from their virtual avatar appearance and gender identification. We also inquired about their feelings of embodiment and agency in Social VR, and how they were perceived and interacted with by other users. Through this set of questions, we expected to gain insights into the challenges and opportunities of fundamentally free and fluid gender expression and embodiment in Social VR, as well as if and how the according experience relates to their real-life identity and gender expression.

Interviewees describe voice characteristics as the main element of gender identification by other users in VR. They also report that, similarly to reality, some people do not respect their chosen pronouns: "Of course you've run into some people that don't respect that, but you'll run into that anywhere, unfortunately." One Interviewee states that their pronouns and gender identification in VR change with the Avatar they wear and the way they present in VR: "In VR I usually go with the avatar, so when I am a female, I prefer people use she. When I am a male, I prefer he, and when I am an object so like a lamp or stuff like that, it doesn't matter."

This differs from their real life, where they represent their birth gender and use corresponding pronouns only: "No, it's something I just used to do in VRChat – in real life I am not into this stuff and people call me for what is it."

Experienced Harassment or Hate Speech

Harassment – especially sexual harassment – as well as anti-Semitic, racist, transphobic hate comments, or extremist content not only have harmful individual and societal consequences but can also lead people to avoid certain virtual platforms and spaces. Protecting users from harassment and creating an inclusive, human dignity-oriented culture is also particularly important in immersive virtual environments, especially since it must be assumed that the intense experiences through embodiment and virtual reality are experienced more intensively by users than in text-based media. Few studies or figures are yet available on the spread of hate messages in immersive environments, partly because the platform providers (so far) – unlike the particularly large social media platforms – are not obliged to provide public and transparent information on measures against hate. Due to real-time presence, real-time verbal communication, and the absence of recordings, the documentation of hate messages in Social VR is particularly difficult. In the interviews, respondents reported varying degrees of experience with harassment or hate speech in VR.

One interviewee observes “definitely a lot of racism in this game, like in a lot of unfiltered language that is very racist or just offensive to people and genders and all that stuff. It's very offensive." This can be encountered by users in very normal situations. Influenced by political discourse in the world, the interviewee observes mostly hate-filled messages against the LGTBIQ+ community: "Hateful comments and stuff towards the entirety of LGBTQ [community]. [...] Because it's become pretty prominent outside of VR Chat as well [...].”

In Social VR, harassment and hatred can be found primarily in certain groups or spaces. The manner or characteristics to which the experienced hate messages refer are closely related to the virtual appearance as an avatar, so that the user is attacked for their avatar. However, the user relates this to personal identity and does not say “[...] my avatar was attacked for being a furry.” but "I've been hated on for being a furry.". This speaks to a great deal of identification with the avatar and that attacks are perceived personally, rather than as avatar related. In addition, there is gender-based discrimination: "I've been hated on for being a female." and the association of both characteristics with weight-based discrimination: "I've been hated on for being a female furry because apparently that immediately makes you 500 pounds." In such situations, the user makes short work of it: "It's just the block button and then you don't need to talk to those people anymore."

This raises the question of whether avatars and sub-communities in virtual realities give rise to new forms of digital hate communities that are directed against certain classes of avatar identities.

Experiences of harassment are reported by one user mainly in public spaces, which they therefore avoid and their own circle of friends in VR: "Harassment typically comes in whenever I go to public worlds [...] that's one of the main reasons why I don't go to public worlds anymore."

Most notably, the interviewee experiences harassment in verbal form because the user has hidden other avatars by default in the settings: "It's mostly speech, especially because I personally have everyone's avatars turned off by default, and then I have to manually go in and show people's avatars." However, there are also particular glitches related to the specific implementation of the embodiment in the metaverse. For example, offensive images can be used as avatars or high-resolution avatars can cause the hardware to crash due to overload. The user protects themselves against this by disabling avatars by default.

The statements reveal the high technical level of the early VR experts, who developed various tricks to deal with unintended, sometimes malicious interventions in the environments. This addresses key technological challenges to the goal of photorealistic real-time presences or highly authentic avatars: The consumer hardware available today is predominantly not yet capable of achieving this reliably. The varying quality of technological equipment can lead to this being used to the disadvantage of people – consciously or unconsciously – with less powerful hardware – a new form of digital divide in immersive environments. For the development of innovative technologies on the one hand and the establishment of interoperability on the other, this represents a major challenge that is directly relevant to users' sense of security and stability in immersive virtual environments.

Overall, the interviewees agree that while social media and VRChat share some certain similarities regarding harassment, the immersive nature of VRChat intensifies negative interactions, making them feel more personal and intrusive, requiring more proactive measures to manage and mitigate. The small survey indicates that harassment in VRChat is particularly prevalent in public spaces, prompting users to seek refuge in private groups. Common forms of misconduct include verbal harassment and technical exploits, with users frequently needing advanced technical skills to protect themselves. The digital divide exacerbates this issue, with less powerful hardware leaving users vulnerable.

Although VRChat provides various protective mechanisms, their effectiveness remains limited. Users adopt personal strategies, such as avoiding public spaces and employing advanced blocking techniques, to enhance their safety and cultivate a more positive experience in virtual environments.

Blocking as a Personal and Structural Security Strategy

We specifically examined how the interviewees used and perceived the blocking system in VRChat. We asked them about the number and characteristics of people they had blocked, the reasons, outcomes, and the challenges and drawbacks they experienced. Through further investigation of this popular, but profoundly different to real-life protection mechanism, we aim to gain insights into the role and impact of blocking as a personal and structural security strategy in Social VR.

One interviewee states, "I don't think there's really any counter speech, it's just a straighter blocking." The person observes varying levels of sensitivity in dealing with hate speech and in utilizing the blocking function. The interviewee makes a direct comparison to traditional social media, noting that in VR chat there is "[…] definitely not as much censorship in chat as there is on social media [...] But via chat, it's probably the worst. The difference is that you have to be unlucky enough to end up in the same room as those people, whereas with social media, everyone has access to it." Subsequently, the technical barrier to accessing Social VR still provides some protection from being flooded with hate messages compared to traditional social media. At the same time, the interviewee points out that such negative content has a more intense effect in Social VR and requires more active countermeasures: "It's just like I feel like via chat is worse in person [...] But I'd say via chat is way more in your face about and then social media is." Regarding their view of potentially getting blocked by other users, they stated: "I mean, I don't know if I've ever been blocked. It doesn't tell you like when you do. Honestly, it doesn't make a difference to me."

Another person reports about their experiences with blocking as something that can be confusing and even hurtful:

> *"So, I have gotten blocked before [..] for absolutely no reason at all. I know I didn't do anything or say anything hurtful. […] And when I get blocked, if it's from a random person […] sometimes I'll just laugh it off and I'll be like, OK, whatever. They have their reason. Now, if it's someone that's*

a friend who has blocked me before, I'll feel kind of hurt and I'll be like, well, what did I do? What did I do to hurt you? What can I do to make this better?"

One interviewee describes how the blocking feature affects the visibility of avatars of users:

"The biggest thing is that if you block someone, it blocks it for you, but not for everyone else in the instance. So, I've run into things where I blocked someone and then like the poor person standing next to me didn't want to block them for whatever reason. And then they started harassing the person next to me trying to get them to talk to me […]."

Overall, the Interviewees describe blocking as a commonly used tool for managing harassment in VRChat, however, its effectiveness is limited to the user who initiates the block. But the feature can lead to social complications and emotional reactions, depending on the context and the nature of the relationship between users.

Potential Legal Consequences of Actions in Social VR

A central focus of our study was to examine the legal implications of user actions in Social VR, specifically within VRChat. This is a challenging topic because anonymity remains still highly valued by most users, and currently, no Social VR platform implements a mandatory legal identity verification system. We asked users about their awareness and opinions regarding existing and emerging laws applicable to Social VR platforms to understand the challenges and opportunities of legal governance in these environments. We also explored their perspectives on ethical and moral concerns arising from Social VR usage, including issues related to identity, impersonation, consent, and harassment.

One interviewee suggests that there should be certain crimes that are relevant to be prosecuted in the metaverse and ruled out at the user's physical location. Specifically, they highlight that avatar theft and unauthorized appropriation of 3D files as major yet unprosecuted issues, emphasizing: "Now when it comes to the stealing avatars [...] that is a breach of copyright. You are stealing somebody's intellectual property."

Another interviewee sees the safety of minors and the presence of pedophiles on VRChat as the most urgent issues to be tackled on both platform and policy side:

> "*Yes, there are so many pedophiles in this game it is actually ridiculous. […] And they will kind of prowl upon anyone who has some sort of need or wanting for attention, regardless of age. And there are a lot of really social, angsty teenagers in this game aged like 13 to 17, […] some of these guys or girls are really like just sinking their teeth into because they can give them attention that they need in exchange for [...] certain types of like child pornography and like abuse towards minors happening on the platform.*"

The interviewees highlighting critical issues such as harassment, avatar theft, and the presence of pedophiles within VRChat. They emphasize the necessity for legal accountability concerning criminal activities in VR, stressing the importance of protecting minors and managing intellectual property rights within the platform.

What Should Policy Makers, Regulators or Designers Know about VR Experiences Like VRChat?

The final question of our study aimed to gather recommendations and insights for policymakers, regulators or designers who are interested in or responsible for Social VR experiences such as VRChat. We asked participants about their expectations and suggestions for enhancing the quality, safety and accessibility of these platforms. Interviewees highlighted several critical issues related to user safety, policy implications, and social inclusion. One participant expressed significant concerns about minors' vulnerability to sexual exploitation by pedophilic users, as well as potential legal repercussions for adults inadvertently interacting with minors. They underline the practical challenges of prosecuting online crimes due to the anonymity and inadequate identification mechanisms prevalent on these platforms.

One interviewee advocated the introduction of mandatory identity verification measures, such as digital ID checks and security recordings, to effectively hold users accountable and facilitate the legal prosecution for online offenses. At the same time, they emphasized the necessity of ensuring cost-free verification processes to prevent the exclusion of economically disadvantaged individuals or users from low-income countries, thus fostering equitable access and social inclusion. Another interviewee stressed the necessity for policymakers and regulatory authorities to gain direct, immersive experiences with Social VR platforms to create informed, user-centric regulatory frameworks. Additionally, this participant pointed out that the heightened sense of immersion characteristic of virtual environments in-

tensifies the urgency for comprehensive safety regulations and policy interventions, acknowledging that virtual interactions carry tangible emotional and psychological impacts for users behind virtual avatars. Interviewees recommended that policymakers and platform designers enhance safety features within Social VR, such as implementing effective vote-kick systems, strengthening protection for minors, and considering mandatory ID verification for improved user accountability. VRChat has now implemented an age verification system through IDs, at least for VRC+ users. The interviewees emphasized the necessity for policymakers to gain first-hand experience to better understand the platform's dynamics and stressed the importance of making security measures accessible to all users, regardless of their economic situation. Additionally, they emphasized the importance of parental supervision to safeguard children from inappropriate content and interactions. They also highlight the unique benefits of VR for improving social interaction skills, particularly for introverts, due to the controlled environment and customizable features available. Furthermore, they advocated higher age ratings for Social VR spaces given the elevated immersion and potential risks, suggesting that regulatory framework should be adjusted to the advanced nature of VR technology.

Discussion

The experimental study is limited by the relatively small sample size of only four participants. Additionally, the sample is restricted in terms of age and gender diversity, although it includes male, female, and diverse gender identities. Geographically, participants represent predominantly the United States and Europe, further limiting the generalizability of the findings. Despite these limitations, the study provides first-hand information and experiences from users and demonstrates that it is relatively straightforward to conduct avatar-based interviews as a researcher in the metaverse. This is especially true when following a community-based approach involving gatekeepers in virtual communities.

Experiencing and expressing personalities independently of physical bodies gives users the opportunity to experiment with social interactions and learn more about them. Friendships, relationships and Smalltalk with strangers are embraced in VR for various reasons, such as users' location or social anxiety. While anonymity in online spaces has significantly impact users' behavior positively and negatively, protective mechanisms are still

underdeveloped, buggy and not the main priorities for platforms. Ensuring minors 'safety through robust identity verification emerges as the most pressing area requiring improvement.

Key concerns include limited understanding among policy-, law- and decision makers regarding the immersive dimension of Social VR. Users emphasize the uniquely immersive nature of VR and its profound social implications, such as fostering friendships, relationships and personal development through prolonged interaction in digital environments.

The surveys clearly indicate that Social VR allows users to experience and express personalities independently of their physical bodies, facilitating experimentation and learning about social interactions. Friendships, relationships, and casual interactions with strangers are highly valued, especially by those geographically isolated or socially anxious. Social VR offers opportunities such as global social interactions with like-minded people, geographical logistical and environmental independence, accelerated learning and collaboration through high levels of immersion, and creative self-expression through interactive tools and accessible audiences. Nonetheless, the medium faces challenges, such as fraud and harassment enabled by anonymity, the particularly dangerous nature of harassment due to VR's immersive quality, and the risks posed by anonymity and insufficient age verification, which exposes minors to age-inappropriate behavior and abuse. Subsequent research could explore socializing in VR compared to physical reality and its impacts on users across different age groups and gender identities.

References

Hinduja, S., & Patchin, J. W. (2024). Metaverse risks and harms among US youth: Experiences, gender differences, and prevention and response measures. *New Media & Society, 26*(1), 1–22. https://doi.org/10.1177/14614448241284413

Krell, F., & Wettmann, N. (2023). Corporeal interactions in VRChat: Situational intensity and body synchronization. *Symbolic Interaction, 46*(2), 159–181. https://doi.org/10.1002/symb.629

LaValle, S. M. (2020). *Virtual reality*. Cambridge University Press.

Quent, M. (2023). Demokratische Kultur und das nächste Internet: Chancen und Risiken virtueller immersiver Erfahrungsräume im Metaverse. In Institut für Demokratie und Zivilgesellschaft (Ed.), *Wissen schafft Demokratie: Schwerpunkt Netzkulturen und Plattformpolitiken* (Vol. 14, pp. 30–43). Institut für Demokratie und Zivilgesellschaft.

Sabri, N., Chen, B., Teoh, A., Dow, S. P., Vaccaro, K., & ElSherief, M. (2023). Challenges of Moderating Social Virtual Reality. In *CHI '23: Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (pp. 1–20). ACM. https://doi.org/10.1145/3544548.358132

Vogelgesang, A. (2024). *Social VR platform design, Nutzerkreativität und ästhetische Kontrolle* (Project Immersive Democracy Working Paper). European Metaverse Research Network.

Chances and Limits of Immersive Environments for Anti-Discrimination and (Historical-)Political Education

Deborah Schnabel

The digitalization of the public sphere puts political educators under considerable pressure. With every new technology and app, a new forum emerges that can become a problem for anti-democratic agitation but also offers the opportunity to make democracy education better for more people. Racism, anti-Semitism, and other threats to democracy are present in all areas of life, and they are quickly conquering the new digital worlds. With every new space, democracy is also renegotiated, and political education is called upon to intervene in this negotiation.

Immersive environments (IE) offer potential to reach more people at a lower threshold and with a different degree of involvement. There is great interest from the entire field of political education to become effective in these new worlds. It may even be possible to preserve important testimonies and witnesses of history by recreating them in virtual realities. Therefore, it is primarily historical-political education, memorial sites, and museums that are launching projects despite the insufficiently explored possibilities of immersive environments.

Especially memorials feel the pressure to make places of remembrance alive, and of course, especially in the field of Shoah Education: the last eyewitnesses are leaving us. First-hand accounts will no longer be possible. Many actors in political education see Immersive User as a possibility to reconstruct the authenticity of a contemporary witness encounter. For example, there is the Anne Frank House in Amsterdam, whose digital VR exhibition allows a tour of the family's hiding place. There is also the app "Inside Auschwitz", developed by WDR, which combines the accounts of three eyewitnesses with a virtual visit to the memorial (Planet Schule, 2022). Additionally, there is the project "Dimensions in Testimony" by the USC Shoah Foundation (USC SF), which lets eyewitnesses speak to us as AI-based holograms (Dimensions in Testimony, n.d.). There are also the first projects in which students "travel back" in the metaverse to key moments in Black history (Education in the Metaverse, 2023). What all actors in the field have in common is that they perceive the new possibilities

of IE as very important, sometimes invest a lot in the field, have high expectations of the possibilities, but at the same time deal with them rather restrictively. For example, the VR Anne Frank House has no protagonists at all, but instead presents the hiding place furnished, in contrast to the physical exhibition. Other applications only allow certain characters, who usually have a bystander role. Behind this lies the still unanswered question of how to unite the possibilities of new technology with the principles of political education. It will depend on the quality of the answers whether IE will become suitable for widespread use in democracy education and history teaching.

No matter how it is approached, IE will come up against limits – technological ones (which I will address later), but often also educational-theoretical ones: Is "artificial authenticity" even possible? Does it make pedagogical sense? Or is an immediacy that cannot be reconstructed being fetishized here?

How Realistic Can Remembering Be?

The limits of immersion are often discussed in such applications, a debate that is older than IE. In memorial education, long before the Internet, there was a debate about how deep the experience of the visitors should go. The main issue was whether such an experience should be accompanied or unaccompanied. In German-speaking contexts, the so-called Beutelsbach Consensus was used as a guideline, which emphasizes the prohibition of overpowering, the requirement of controversy, and student orientation. It is difficult to apply Beutelsbach to IE, as such environments usually overwhelm from their very game logic. In the USA, however, the focus is often on "empathy education", the development of compassion and understanding, often with methods that are quite "overwhelming" (Besand et al., 2019). The problem of "artificial authenticity", which can lead to distancing or empathy, is obvious: history can potentially be distorted. A purely imagined image of the past emerges that is neither congruent with historical knowledge nor with the victims' world of experience.

Apart from the inherent pedagogical risks of the technology, the controversy also goes to the content. The biggest point of contention is the possibility for non-victims to play victims of racism or anti-Semitism in IE. What some may view as an opportunity for empathy and a change of perspective can be seen by others as insensitive, cultural appropriation,

false alignment of perpetrators and victims, violation of the prohibition of overwhelming and a potential trigger for (transgenerational) trauma. Complex feature compositions are truncated and cannot be explored in depth, especially when, for example, a typical VR session usually does not exceed one hour. The socio-political effect on the dominant society must also not be underestimated: Memory-political problematics and exoneration narratives can be reinforced if the descendants of the perpetrators can feel even more strongly as part of the persecuted through IE.

The balance of distance and empathy that political education has to maintain in IE could be understood as a kind of variant of the "uncanny valley" effect – at a certain level of reality of the simulation it becomes strange, "creepy". The mentioned example of the WDR app "Inside Auschwitz" illustrates how difficult this balancing act is. On the one hand, it deliberately tries to eliminate the distance: "The users should consciously not take a distanced position but explore the memorial site as if they were visiting it on foot." (Nägel & Stegmaier, 2019). On the other hand, the app's producers warn against too much closeness and even recommend not using VR glasses in a school context, as the lack of distance can disturb students.[1]

It appears that applications are better accepted when they use alienation effects, creating a mixed space that is not real, but not completely fictional either. It is not about identification, but rather sympathy. This seems to apply to the medIEm as a whole: while VR/AR are still niche products, MMORPGs[2], for example, are functioning excellently as a mass medIEm. Here, it often appears to be the artificiality of the environment that makes it possible to deal with it confidently. There is never any doubt that one is not in reality; an avatar that does not resemble the user enables greater identification than the clinical perfection of VR/AR, which often does not enable the user to experience their body at all, as when looking down with VR goggles, one does not see their body, but the floor.

It seems almost impossible to consider every potential user and every possible multiple concern. Can't we operate similarly to our everyday experience – with limited knowledge, but with room to learn? Our daily

1 "The stories told by the witnesses and the images from the camp can be challenging for sensitive students or young people with their own experience of war and flight. It can help to watch the videos with a little more distance (for example, over the shoulder of a classmate holding the tablet). In addition, it is advisable not to use headphones or VR." (Planet Schule, 2022).

2 Massively Multiplayer Online Role-Playing Games, online video game worlds with many thousands of users.

behaviour is also based on limited knowledge, yet we are getting very far here as well.

Educational science findings on the pedagogical impact of IE are still manageable: mixed results were found for "Anne Frank House VR" compared to desktop learning – while identification with those involved was stronger, at the same time less content was absorbed (Mulders, 2023). Learning dynamics are complex and require more intensive research.

Metaverse and Beyond

Certainly, the best-known IE application is Mark Zuckerberg's "Horizon Worlds", widely known simply as "Metaverse"[3] – known both in terms of ambition and for negative headlines. What is less important for our field is whether "the Metaverse," i.e., Horizon Worlds has been successful, but it is uncertain whether a similar platform for IE applications could be successful. Civic education must prepare for the possibility that such a platform could become a mass medIEm. Comparisons can be drawn with the failed "Second Life" platform.

The same questions that apply to social media also apply to the metaverse: Who leads the discourse? Who dominates? Do discriminated people feel safe in the metaverse? The higher degree of reality of IE puts the classic problems of social media into sharper focus. There is still little research on whether IE has an even greater potential for radicalization than social media already does. Heuristically, it is probably safe to assume a similar threat level as social media. Furthermore, adaptive and reactive content in IE will increasingly be produced by AIs, as human content development

3 The term "metaverse" in the narrower sense is theoretically meant to encompass all immersive environments and their interconnection – in this sense, "the metaverse" does not yet exist, but individual IEs Horizon Worlds, Minecraft, etc. do. ChatGPT offers the following definition of the metaverse (prompt: write a scientific definition of the metaverse): „The Metaverse is a comprehensive, large-scale virtual environment that integrates multiple interconnected digital spaces, facilitating seamless interactions, communication, and collaboration among users, who are represented by digital representations known as avatars. This virtual framework merges aspects of virtual reality, augmented reality, and interconnected networks to create an immersive and persistent ecosystem for users, with wide-ranging applications spanning entertainment, social interaction, education, commerce, and various other domains. By transcending physical limitations and geographical boundaries, the Metaverse fosters novel opportunities for creative expression, innovation, and interaction in the digital realm."

cannot keep up with the demand for these demanding environments – this is also the view of a contribution by the World Economic Forum, for example (AI Is Shaping the Metaverse, 2023). In this scenario, it is conceivable that AI language models could create personalized worlds that essentially affirm users' attitudes, reinforce them in problematic attitudes, force ideological cocooning effects, and drive them into radicalization tunnels. The already low inhibition threshold in the metaverse would do the rest.[4] These structural aspects of AI racism could have an exponential effect in immersive environments: Racist echo chambers in virtual worlds would inevitably leak outward, shaping society – and in turn serving as a source of AI content generators.

The likelihood of this scenario also depends on the people who develop the metaverse. They would need to be diversity- and discrimination-sensitive from recruitment through beta testing in all areas of work. Experience has shown that this cannot be achieved through decree, but through the active participation of the people involved. The question of who uses the metaverse and why has not yet been answered. What inhibitions might exist for certain parts of society to use the Metaverse? Why might this be the case? Do corporations like Meta actively counteract such tendencies? What is being done to provide representation and protection? How do Meta and Co. prevent right-wing infiltration, as can also be observed on other digital platforms that operate more in the niche?

Political Education in the Metaverse

Apart from these questions, this task arises in our field of action: How should civic education respond to these and similar questions? In our opinion, the question should not be whether political education should now "invest in the metaverse." Such a generalist approach is doomed to fail. However, it should be experimented with – the metaverse should be seen as software in which concrete, limited settings are tried out. Thus, the maxim should not be "We're putting our whole memorial into the Metaverse," but rather, "This specific exhibition can (also) be visited in the Metaverse."

4 cf. only Yinka Bokinni: A barrage of assault, racism and rape jokes: my nightmare trip into the metaverse. (25.04.22.) URL: https://www.theguardian.com/tv-and-radio/2022/apr/25/a-barrage-of-assault-racism-and-jokes-my-nightmare-trip-into-the-metaverse (12.09.23)

The potential of such steps can already be observed today, for example in Minecraft. For example, a library of black literature was saved not only to Meta's "Horizon Worlds" but also to a Minecraft server. One explanation for Minecraft's success could be found in the "uncanny valley," as it does not try to imitate reality. There is currently a project underway to transfer the Yad Vashem memorial to Minecraft (Yad Vashem Comes to Minecraft, 2023), as well as private projects that strive to reflect the Holocaust on the platform (Jcirque25, 2020). However, this should not hide the fact that Minecraft is also used by the opposite side: users often point out servers on which Nazis glorify Nazi atrocities.

Excursus: AI, Discrimination and Justice

The increasing importance of AI in various environments is platform-independent and requires its own analysis. The language used to discuss "artificial intelligence" promises a revolution in all aspects of life, from the world of work and macroeconomics to literature and art. At the same time, a promise of justice based on procedural rationality and the absolute neutrality of the algorithm is being made. Exclusions resulting from misanthropic attitudes can thus be avoided automatically – the assessment is made on the basis of an alleged objectivity that cannot be achieved by humans.

In fact, the publicly known language models, first and foremost ChatGPT, do not seem to be suitable for reproducing discriminatory language – ChatGPT in particular stands out here for its robust ethical framework. ChatGPT strongly opposes racism and discrimination. All attempts to use the software to produce racist texts are supposed to fail theoretically already in the approach (although this can be circumvented with the appropriate Prompts). It should not be forgotten that there are projects that specifically train language models to combat misanthropy, such as the international research project "Decoding Antisemitism"[5].

At the same time, there are increasing reports that structural racism is being reinforced by intelligent image recognition software. The algorithm bases its decisions on patterns it finds in society and evaluates non-white people differently than white people. This means that racism is avoided on a linguistic level, while it is further perpetuated, automated, and industrialized on an operational level, usually without transparency for those affect-

5 https://decoding-antisemitism.eu/ (12.09.23).

ed. If AI is used to make decisions about credit and housing allocation, job applications, and social forecasts in the future, the problems are foreseeable. Unlike human decisions, the apparent objectivity of the machine still makes it immune to criticism.

The framework conditions of the AI industry also need to be critically examined – for example, the underpaid Kenyan workers whose performance was essential for the success of ChatGPT (Perigo, 2023). On an economic level, AI reproduces capitalism instead of critically questioning it, and this is done along racial distinctions. Who owns the systems? Who is allowed to market their services, and who benefits from the increased productivity? Who owns AI art, whose creation is inspired by the names of famous artists? Who can work in this new AI world, whose job may be replaced by automation, and whose life may be managed by AI? All indications suggest that those affected by racism will not receive positive answers to these questions, and that the negative effects of the revolution will disproportionately affect them.

Furthermore, the implications for antisemitism should not be underestimated. Already, the portrayal of Jewish people is subject to a historical bias – image-creation programs draw primarily on historical representations and images of ultra-Orthodox groups. As a result, Jewish people, who today participate in modern life as a matter of course, are absent from the collective imagination.

Last but not least, Artificial Intelligence (AI) systems can already be used to generate an endless stream of fake news, deep fakes, and conspiracy theory content – with a plausibility and apparent authenticity that current conspiracy ideologues can only dream of. Likely, the effects of this will be felt particularly strongly by Jewish people.

All these questions arise with increased urgency when AI is not just a nice toy within Information Units (IEs) but is instead constitutive of them. How will educational IEs change when they no longer present content that is prefabricated by humans, but instead AI dynamically generates such content? What will happen when AI-based avatars target users and radicalize or intimidate them with personalized propaganda? Where is the control? And how can automated patterns of reproduction of anti-human attitudes be interrupted?

This requires asking questions that naturally point to the need for further exploration. Beyond the field of IE in civic education, namely: what conditions for success do we need for a discrimination-sensitive AI ethics? How can AI be taught to recognize power asymmetries? Here, limits become

apparent that are already inherent in the current way of working of AI – it is fundamentally reproducing.

A Counter-Model: IE as Lived Utopia

Perhaps it is not expedient to reconstruct stories of persecution and experiences of discrimination for political education in IE. Perhaps it would make sense to simulate worlds in IE in which a certain utopian hope has become real. Instead of putting learners into a pedagogical capsule with certain preconceived content and learning goals, they could independently explore worlds that have already been liberated. The learning effect would arise precisely from the absence of a context of violence that is now commonplace.

An example of this is the everyday simulation game "The Sims" (released in 2000) and its role in promoting homosexual/queer emancipation. In this PC game, queer relationships could be experienced for the first time in a mainstream title without it being a constant theme. Many queer people today report having experienced queer representation and empowerment for the first time simply because of the possibility that was not imposed or used as a teaching tool but was explored and motivated intrinsically.

Perhaps political education for IE should take a similar approach: provide a concrete utopia that can be explored, with possibilities users can experiment freely – and in the process reduce fears and prejudices. The virtual world should be a model for reality, not the other way around.

This approach would possibly have the advantage of avoiding the numerous pitfalls mentioned above that threaten political education in IE from the outset. The boundaries of the user experience would also have to be drawn less tightly; there would be less need for pedagogical guidelines, and less top-down control overall.

References

AI Is Shaping the Metaverse – But How? Industry Experts Explain. (2023, May 9). World Economic Forum. https://www.weforum.org/agenda/2023/05/generative-ai-and-how-can-it-shape-the-metaverse-industry-experts-explain/

Besand, A., Overwien, B., & Zorn, P. (Eds.). (2019). *Politische Bildung mit Gefühl*. Bundeszentrale für politische Bildung.

Dimensions in Testimony. (n.d.). Speak with survivors and other witnesses to the Holocaust and other genocides through their interactive biographies. USC Shoah Foundation. https://sfi.usc.edu/dit

Education in the Metaverse Takes Students Back in History. (2023, December 1). WBUR. https://www.wbur.org/hereandnow/2022/12/01/black-history-metaverse

Jcirque25. (2020). Using Minecraft as a medIEm to reflect on the Holocaust. *Reddit*. https://www.reddit.com/r/Minecraft/comments/htlst6/using_minecraft_as_a_medIEm_to_reflect_on_the

Mulders, M. (2023, March 6). Learning about victims of the Holocaust in virtual reality: The main, mediating, and moderating effects of technology, instructional method, flow, presence, and prior knowledge. *MDPI*. https://www.mdpi.com/2414-4088/7/3/28

Nägel, V. L., & Stegmaier, S. (2019, October 8). AR und VR in der historisch-politischen Bildung zum Nationalsozialismus und Holocaust – (Interaktives) Lernen oder emotionale Überwältigung? *Bundeszentrale für politische Bildung*. https://www.bpb.de/lernen/digitale-bildung/werkstatt/298168/ar-und-vr-in-der-historisch-politischen-bildung-zum-nationalsozialismus-und-holocaust-interaktives-lernen-oder-emotionale-ueberwaeltigung/

Planet Schule. (2022, December 10). Inside Auschwitz | Erzählungen Überlebender und 360 Grad-Videos. *Planet Schule*. https://www1.wdr.de/schule/digital/unterrichtsmaterial/dreisechzig-inside-auschwitz-100.html

Perrigo, B. (2023, January 18). OpenAI used Kenyan workers on less than $2 per hour to make ChatGPT less toxic. *TIME*. https://time.com/6247678/openai-chatgpt-kenya-workers/

Yad Vashem Comes to Minecraft for International Holocaust Day. (2023, January 26). i24NEWS. https://www.i24news.tv/en/news/israel/technology-science/1674714331-yad-vashem-comes-to-minecraft-for-international-holocaust-day

Police Handling of Hate Crime: A Pilot Project to Use VR Technology for Professional Development in Sensitizing Police Officers to the Experiences of Victims of Bias Crime in Hamburg

Eva Groß, Ulrike Zähringer & Anabel Taefi

Immersive technologies, such as virtual reality-based training, offer new opportunities for experience-based learning. Learners can immerse themselves in virtual or fictional worlds, making the experience feel real and allowing them to fully participate in it. These technologies also have great potential for basic and advanced police training in various subject areas, particularly in addressing bias crime, which is the focus here. The quality of police interactions with minority communities can greatly impact the level of trust these communities have in both the police and the state. Trust in institutions is an important measure of the quality of a democratic society, as it is intertwined with diversity and participation. This article demonstrates the potential benefits of immersive technologies in helping the police handle victims of bias-motivated crimes, which are often targeted at members of minority communities. Improving the professionalism of the police in this manner is a crucial step in building trust within minority communities and strengthening democratic resilience. The article then moves on to discuss a pilot project carried out at the Hochschule der Akademie der Polizei Hamburg, using virtual reality technology, and the initial evaluation results.

1 Bias Crime and How the Police Handle them in Germany[1]

Bias-based acts[2] and hate crimes are major issues when it comes to promoting diversity and inclusivity in many European countries. These acts specifically target individuals based on their social group membership

1 This section builds on an earlier publication by the first author (Groß & Häfele, 2021) and contains text fragments that have already been published there.

2 From a criminological perspective, the term bias crime is more appropriate than hate crime, especially as the acts are an expression of group-related devaluation and discrimination (group-focused misanthropy) or negative prejudices against social groups

and are based on protected characteristics, such as skin color, religious beliefs, or sexual orientation, under constitutional law (Groß & Häfele, 2021). The concept of bias crime (BC) is closely linked to the concept of group-focused enmity (Gruppenbezogene Menschenfeindlichkeit, GMF) (Heitmeyer, 2002), both of which revolve around the prejudiced belief in the inequality of different population groups (ideology of inequality) (Heitmeyer, 2002; Zick, Küpper & Heitmeyer, 2009; Zick, Wolf, Küpper, Davidov, Schmidt & Heitmeyer, 2008). This inequality can manifest in discriminatory attacks and hostile acts, such as racism, anti-Semitism, or transphobia. Therefore, the concept of bias crime theoretically covers the territory where prejudiced attitudes turn into concrete actions (Zick & Küpper, 2021). Although GMF and bias crime share a similar foundation, they are not identical. GMF focuses on attitudes, while bias crime deals with actions. In her study, Krieg (2022) was able to empirically prove the theoretical connection between GMF attitudes and the commission of discriminatory behavior, with a representative sample of students (N= 2824). What makes bias-motivated crime especially concerning is that, according to the logic of GMF, victims are not just targeted as individuals but also because of their group membership, either actual or perceived. These crimes not only harm the direct victim but also send a threatening message to the entire group to which the victim belongs (Groß & Häfele, 2021; Williams, 2021), making them damaging on a wider scale and threatening to democracy.

Since 2001, the police have officially recorded these crimes as "hate crimes" (Lang, 2014, p. 54). This includes crimes directed against a person or group of people due to their political stance, attitude and/or commitment, nationality, ethnicity, skin color, religious affiliation, ideology, or social status, physical and/or mental disability and/or impairment, gender/sexual identity, sexual orientation, or physical appearance (BKA, 2023). These acts can be directly committed against a person or group of persons, an institution, or an object/thing associated with one of the aforementioned social groups (actual or perceived affiliation), or against any target in connection with the perpetrator's biases (BKA, 2023b). In Germany, as well as other European countries, there was a significant increase in officially reported hate crimes between 2014 and 2018 (Riaz et al., 2021). For the

that are linked to social structures of power and oppression, see also Fuchs, 2021, p. 270.

year 2022, 11,520 offenses were recorded in Germany, representing a 10 % increase from the previous year (2021) (BMI & BKA, 2023, p. 10). The few available studies on the unreported cases of bias crime in Germany indicate a very high number of unreported cases, ranging from 50 % to 90 % (e.g. Church & Coester, 2021; Fröhlich, 2021[3]; Groß et al., 2019).

Just like the GMF concept, the BC concept and its reporting in Germany is continuously evolving to reflect social debates and developments. Since 2017, the new version of the police recording system includes "gender/sexual identity, sexual orientation" instead of the restrictive "sexual orientation" category, explicitly including trans* individuals in the police counts. In addition, the category of "race" was removed from the police recording system, and "physical and/or mental impairment" was added in 2017 (Groß & Häfele, 2021). Since 2017, law enforcement authorities have also been required to consider the perspective of the victim, among other aspects, when assessing the circumstances of the crime (Kleffner, 2018, p. 35).

Although the German police counting system is theoretically well-positioned, there are indications of quality gaps in the recording and perception of bias-motivated crime in police practice (e.g. Kleffner, 2018: 36; Habermann & Singelnstein, 2018; Groß & Häfele, 2021). There are significant differences in the statistics between those recorded by the police and those reported by independent counseling centers, suggesting an underestimation of bias crime based on police counts (e.g. Lang, 2015). For instance, specialized counseling centers registered approximately one third more acts of right-wing violence in 2017 than law enforcement agencies and offices for the protection of the constitution. When considering homicides linked to right-wing and racist violence as subcategories of the BC, the discrepancy between civil society and security authority counts is particularly significant, as noted in a comprehensive study of right-wing political homicides in Berlin from 1990 to 2008 (Feldmann et al., 2018). There could be various reasons for this high discrepancy. Errors in the initial recording by the police can lead to distortions. In addition, although some cases are reported to victim support services, they are not reported to the police. In a recent study[4], it was found that the most common reasons for not reporting bias-motivated crime were fear that it wouldn't make a difference, that the police wouldn't be able to solve the case, and concern about not being

3 https://stadt.muenchen.de/dam/jcr:c19e83da-eca8-48b0-920e-e6e37791d4e7/Kurzfassung_DRUCK_final.pdf.

4 Groß, E., Häfele, J. & Peter S. (i.E.). Kernbefundebericht zum Projekt HateTown.

taken seriously. These frequently cited reasons suggest a lack of trust in the police by those affected. Empirical evidence also supports this assumption, as victims of bias-motivated crime have been shown to have the least trust in the police compared to victims of non-bias-motivated crime and non-victims. This lack of trust is likely a key reason for the high estimated level of unreported crime in Europe.

Existing research in Germany confirms that many victims of bias crime feel that they are not taken seriously enough by the police or are treated inappropriately. This may be due to a lack of empathy, understanding, and proper training for police officers in dealing with bias crime. Unfortunately, this is not unique to Germany and is also seen in an international comparison. In 2018, the Greater Manchester Police (GMP) in England attempted to improve police officers' sensitivity and professionalism towards those affected by implementing virtual reality training. This is where the VR headset-based training, aptly named "Affinity", comes into play.

2 Research Objective and Methods of the Pilot Project: Sensitize Police Officers to the Experiences of Victims of Bias Crime

This virtual reality-based training, developed in cooperation with Mother Mountain Productions CIC (MMP) and Greater Manchester Police (GMP), allows officers to experience bias crime in a virtual setting by putting them in the role of victims in an immersive[5] environment. Real-life stories of individuals who have faced anti-Semitism, trans-hostility, and attacks as people with disabilities are recreated using 3D technology and experienced through virtual goggles. The goal is to help officers understand how their initial reactions and language can affect victims and recognize stereotypes and justifications used by perpetrators in these types of crimes. The purpose of the training is to establish a connection and empathy with victims, in order to build trust in the police and increase reporting rates. Data collected so far[6] has shown that VR training has a significant and long-term effectiveness in changing attitudes and behavior among police officers, and also reinforces these changes among those who have already been sensitized.

5 "Immersive" means "to dive into" or "to immerse oneself in something" and refers to the effect that occurs in the consumer when they "dive into" the virtual or fictional world and it feels as if the experience is real and they are actually part of it.

6 MMP & GMP, 2023.

As part of the "Immersive Democracy Project"[7], a quasi-experimental VR-based affinity training was conducted on August 21 and 22, 2023 at the Hochschule der Akademie der Polizei Hamburg in the Department of Sociology for future police officers.[8] The VR experience included original scenarios developed in collaboration with the GMP. The students were shown three English-language films where they experienced discrimination as victims of different marginalized groups, such as anti-Semitism, transphobia, and assault on a visually impaired person. The participants also experienced being questioned by the police after the incident from the perspective of the victim. The films were accompanied by re-enacted interviews with the actual victims describing their experiences during the assault and with the police.

The response patterns of the participating police students were measured through scenario-based questions designed to assess empathy before and after using the VR goggles as part of the Affinity project.[9] The results of the quasi-experimental[10] study on the changes in the participants' attitudes towards victims of bias-motivated crimes are presented in the following section.

7 The projekt „Die immersive Demokratie“ is part of the European Metaverse Research Network with members in the Netherlands, Poland, Spain, Italy, Sweden and France. The aim is to conduct research that demonstrates both the liberalizing and highly beneficial opportunities of the metaverse idea as well as the potential risks and challenges associated with it. Further information on the project can be found at https://www.metaverse-forschung.de/en/about-the-project/ [17.10.2023], the European Metaverse Research Network is presented at https://www.metaverse-research-network.info [17.10.2023].

8 This was followed by a one-day expert workshop with representatives from the police, academia, victim protection organizations, police students and teachers.

9 Specifically, the "Toronto Empathy Questionnaire (TEQ)", a measurement method developed by Spreng et al. (2009), was used in the scenario-related development of the question formulation. It is a one-dimensional, concise and valid instrument for assessing empathy. As empathy itself counts as a comparatively stable characteristic, related behavioral and scenario-related questions were formulated that are suitable for a before-and-after measurement.

10 In addition to the lack of randomization of the experimental group, no parallel control group could be investigated; the experimental element in the setting is therefore exclusively the treatment and the before-after measurement of the dependent variables.

3 Results

25 third-semester students from the police protection program, who were not randomly selected, took part in the pilot project. Out of them, 21 students completed the closed questionnaire before and after the VR application. The quantitative analysis is based on these 21 cases, and is not representative of the entire population of police students at the Hochschule der Akademie der Polizei Hamburg. Also, no control group was included in the study, and the lack of randomization only allows for a weak quasi-experimental design. Despite these methodological limitations, the comparison of the response patterns in the dependent variables before and after the VR application provides initial insights into the changes in the attitudes of the police students towards victims of bias-motivated crimes and the prevalence of such crimes.

Figure 1 shows the difference in the response patterns of the 21 students before and after the application to the statement: "Hate crime should be a priority in police work."

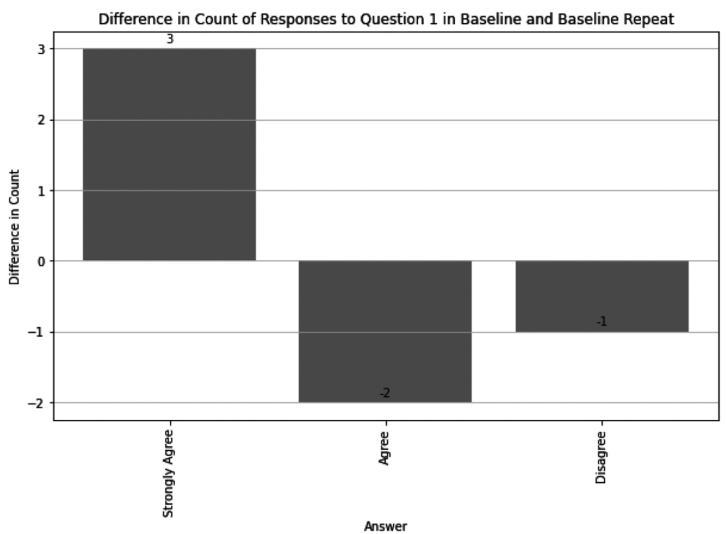

Figure 1: The difference in response patterns of the 21 students before and after the VR application to the statement: "Hate crime should be a priority in police work."

Among the 21 cases, a strong agreement on the importance of prioritizing hate crimes in police work was clearly seen, with three cases showing an increase in agreement, while the other two categories showing a decrease. This suggests that the students became more attuned to the issue of hate crimes after the VR application.

The statement, "Victims of hate crime should be more resilient and able to deal with the situation without reporting it to the police," lost overall agreement among students after the VR application. However, the rejection of this statement increased (see Figure 2), indicating a higher sensitivity towards victims of bias-motivated crime.

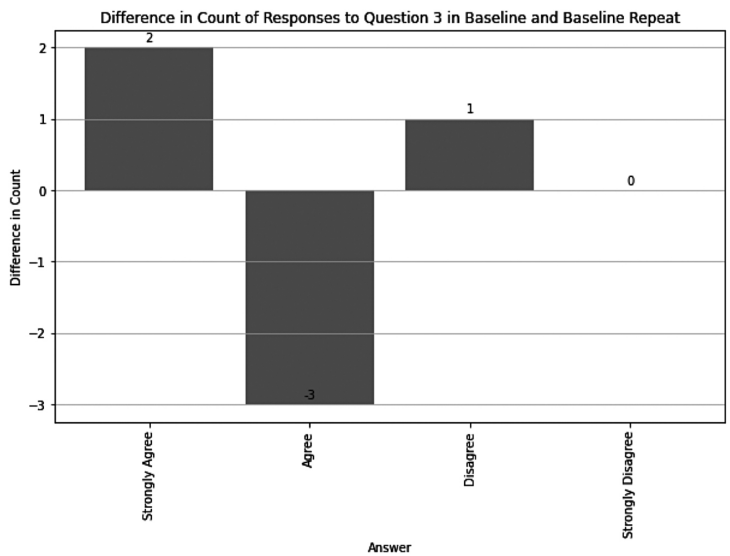

Figure 2: Difference in the response pattern of 21 students before and after the VR application to the following statement: "Victims of hate crime should be more resilient and able to deal with the situation without reporting it to the police."

The statement following the VR application gained approval: "I believe that my way of interacting with a victim of a bias crime can influence that person's ability to deal with what has happened." (see Figure 3).

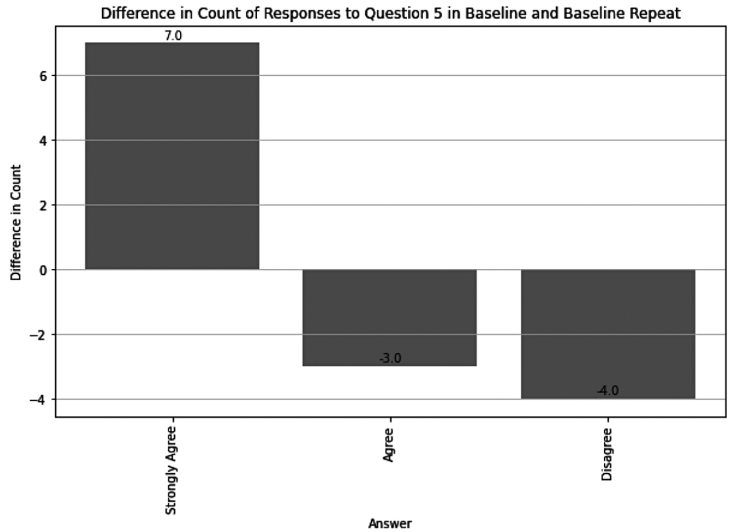

Figure 3: Difference in the response pattern of 21 students before and after the VR application to the following statement: "I believe that my way of interacting with a victim of a bias-motivated crime can influence that person's ability to deal with what has happened."

As intended, this result indicates an increased sensitivity towards one's own behavior as a police officer as a result of the VR application, and therefore, an increased reflection on one's own behavior in the investigated situations. It can be assumed that the VR application leads to an increased sense of self-efficacy as a police officer when dealing with the sensitivities of victims of bias-motivated crime.

The change in the response behavior of the 21 students to the following statement points in a very similar direction: "The way I deal with a victim has the potential to improve the victim's experience." (see Figure 4).

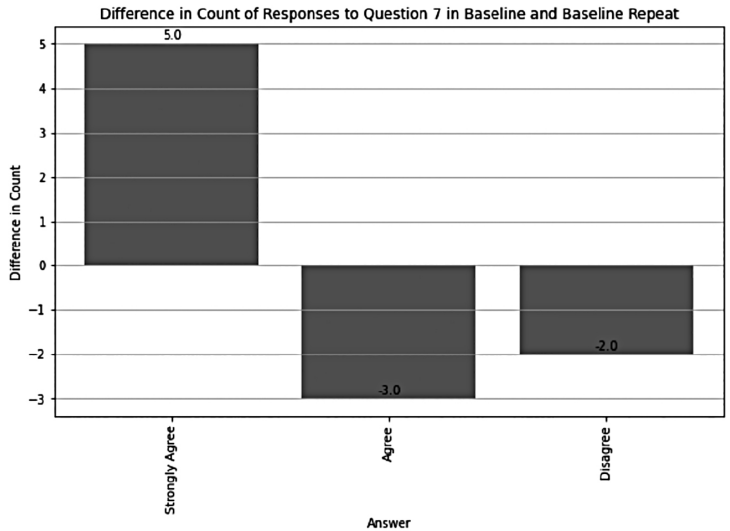

Figure 4: Difference in the response pattern of 21 students before and after the VR application to the following statement: "The way I deal with a victim has the potential to improve the victim's experience."

Overall, the collected quantitative data indicate an effect of the VR application in the intended direction for the Hamburg students, similar to what MMP and GMP found in their study of police officers from Manchester after 3 years (MMP & GMP, 2023).

In a qualitative design, the students were asked to reflect on their experience with the Affinity application and to share their thoughts and feelings. Many participants had a positive experience, as shown in the following examples:

- "It felt like I was there myself."
- "I was directly engaged."
- "I paid more attention compared to a movie, where I usually look away and get distracted."
- "This made it easier to understand how people feel."
- "This method is a good way to gather people's points of view, which is important when they come into contact with the police. I immediately felt a tension that I'm sure those affected also feel, as they know it could happen again."

4 Summary and Outlook

Hate crimes and bias-based crimes do not only affect individual victims, but also have a broader impact on corresponding social subgroups and the quality of democracy as a whole. Studies show that victims of hate crimes often do not report them to the police due to low trust. It must be in the interest of the police to undergo appropriate training and education to ensure that vulnerable victims, such as those affected by hate crimes, feel comfortable turning to the police for support. This could also help address the issue of unreported cases in this area.

To address this, the virtual reality-based training program "Affinity" was developed by MMP in collaboration with Greater Manchester Police. This training allows police officers, including trainees, to experience real-life stories of bias crime in an immersive environment. They can take on the perspective of the victims and witness the effects of verbal and non-verbal communication through the reactions of the interviewing officers. Additionally, they learn about stereotypical justifications associated with these acts, which helps them better identify and classify them. The training also aims to improve empathy among police officers. Initial data from England shows that this training has a high and lasting effectiveness on changes in attitude and behavior among police officers and reinforces the sensitivity of those already aware of the issues associated with bias crime.

Similarly, results from a pilot project at the Hochschule der Akademie der Polizei Hamburg with 25 police students also show the effectiveness of this training. The students strongly agreed that hate crime is a priority for police work, and their perceived resilience decreased while their awareness of their responsibility as police officers in dealing with hate crime increased. After the VR experience, they were also more likely to believe that their interactions with victims of bias crime can influence how they handle the crime and how they perceive and evaluate it in general. Thus, overall, a positive effect was also evident among police students in Hamburg – although the group of individuals tested was very small due to the pilot nature of the project, and thus the results cannot be generalized.

It should be noted that a simulated VR experience cannot truly replicate the perspective of those affected by bias crime and racism. However, other methods used in police training, such as role-playing or training with colleagues, have similar limitations, if not worse. The disconnection from the environment that occurs in the VR experience, due to the combination of 3D glasses and headphones, results in participants experiencing the simula-

tion more intensely compared to other training courses. This was reflected in the comments of the students immediately after the VR experience.

Overall, this suggests that VR training in immersive environments holds promise in sensitizing police officers to dealing with victims of hate crime. The increased sensitivity of police officers may also lead to greater trust in the police and a higher likelihood of crime reporting by victims. In Germany, where different training concepts are used for prospective police officers compared to England, further research is necessary to assess the effectiveness of VR-based training with larger groups, using control group designs and longitudinal approaches. Additionally, scenarios designed specifically for Germany, including language, clothing, and environment, should be developed to make the experience as realistic as possible.

References

BKA (2023, 9. Mai). *Politisch motivierte Kriminalität 2022 – Vorstellung der Fallzahlen in gemeinsamer Pressekonferenz*. Bundeskriminalamt. https://www.bka.de/SharedDocs/Kurzmeldungen/DE/Kurzmeldungen/230509_PMK_PK.html

BKA (2023b). *Politisch motivierte Kriminalität (PMK) – rechts – Phänomen – Definition, Beschreibung, Deliktsbereiche*. Bundeskriminalamt. https://www.bka.de/DE/UnsereAufgaben/Deliktsbereiche/PMK/PMKrechts/PMKrechts_node.html

Church, D., & Coester, M. (2021). *Opfer von Vorurteilskriminalität. Thematische Auswertung des Deutschen Viktimisierungssurvey 2017*. Forschungsbericht. Kriminalistisches Institut, Kriminalistisch-kriminologischen Forschung, BKA.

Feldmann, D., Kohlstruck, M., Laube, M., Schultz, G., & Tausendteufel, H. (2018). *Klassifikation politisch rechter Tötungsdelikte – Berlin 1990 bis 2008*. Universitätsverlag der TU Berlin.

Fröhlich, W. (2021). *Hasskriminalität in München. Vorurteilskriminalität und ihre individuellen und kollektiven Folgen*. Kurzfassung der Studie des Sozialwissenschaftlichen Instituts München. https://stadt.muenchen.de/dam/jcr:c19e83da-eca8-48b0-920e-e6e37791d4e7/Kurzfassung_DRUCK_final.pdf

Fuchs, W. (2021). Vorurteilskriminalität – Konzept, Auswirkungen auf Opfer, Rechtsgrundlagen und verbesserte statistische Erfassung. *Journal für Strafrecht,* 8(3), 279–294. https://doi.org/10.33196/jst202103027901

Groß, E., & Häfele. J. (2021). Vorurteilskriminalität. Konzept, Befunde und Probleme der polizeilichen Erfassung. *Forum Politische Bildung und Polizei,* 1/2021, 20–30.

Groß, E., Dreißigacker, A., & Riesner, L. (2019). Viktimisierung durch Hasskriminalität. Eine erste repräsentative Erfassung des Dunkelfeldes in Niedersachsen und in Schleswig-Holstein. *Wissen schafft Demokratie, 4*, 140–159. https://doi.org/10.19222/201804/13

Habermann, J., & Singelnstein, T. (2018). Praxis und Probleme bei der Erfassung politisch rechtsmotivierter Kriminalität durch die Polizei. *Wissen schafft Demokratie,* 4, 20–31. https://doi.org/10.19222/201804/02

Heitmeyer, W. (2002). Gruppenbezogene Menschenfeindlichkeit. Die theoretische Konzeption und erste empirische Ergebnisse. In Heitmeyer, W. (Eds.), Edition Suhrkamp: Vol. 2290. *Deutsche Zustände. Folge 1* (S. 301). Suhrkamp Verlag.

Kleffner, H. (2018). Die Reform der PMK-Definition und die anhaltenden Erfassungslücken zum Ausmaß rechter Gewalt. *Wissen schafft Demokratie,* 4, 30–37. https://doi.org/10.19222/201804/03

Krieg, Y. (2022). The Role of the Social Environment in the Relationship Between Group -Focused Enmity Towards Social Minorities and Politically Motivated Crime. *Kölner Zeitschrift für Soziologie und Sozialpsychologie,* 74, 65–94. https://doi.org/10.1007/s11577-022-00818-7

Lang, K. (2014). *Vorurteilskriminalität. Eine Untersuchung vorurteilsmotivierter Taten im Strafrecht und deren Verfolgung durch die Polizei, Staatsanwaltschaft und Gerichte*. Nomos-Verlag.

MMP (Mother Mountain Productions CIV) & GMP (Greater Manchester Police) (2023). *Affinity. VR Hate Crime Training Program 2020–2023* [Impact Review Präsentation]. Workshop in Hamburg 2023.

Riaz, S., Bischof, D., & Wagner, M. (2021). Out-group Threat and Xenophobic Hate Crimes: Evidence of Local Intergroup Conflict Dynamics between Immigrants and Natives. *OSF Preprints*. https://doi.org/10.31219/osf.io/2qusg

Spreng, R., McKinnon, M., Mar, R., & Levine, B. (2009). The Toronto Empathy Questionnaire: scale development and initial validation of a factor-analytic solution to multiple empathy measures. *Journal of Personality Assessment,* 91(1), 62–71. https://doi.org/10.1080/00223890802484381

Williams, M. (2021). *The Science of Hate: How prejudice becomes hate and what we can do to stop it.* Faber & Faber.

Zick, A., & Küpper, B. (2021). *Die geforderte Mitte. Rechtsextreme und demokratiegefährdende Einstellungen in Deutschland 2020/2*. Dietz.

Zick, A., Küpper, B., & Heitmeyer, W. (2009). Prejudices and group-focused enmity – a socio-functional perspective. In A. Pelinka, A., K. Bischof, K., & K. Stögner, K. (Eds.), *Handbook of Prejudice* (S. 273–303). Cambria Press.

Zick, A., Wolf, C., Küpper, B., Davidov, E., Schmidt, P., & Heitmeyer, W. (2008). The Syndrome of Group -Focused Enmity: The Interrelation of Prejudices Tested with Multiple Cross-Sectional and Panel Data. *Journal of Social Issues,* 64(2), 363–383. https://doi.org/10.1111/j.1540-4560.2008.00566.x

Real Participation in Virtual Environments: Navigating Public Participation in the Metaverse

Jonas Fegert

The Evolution of the Metaverse: Opportunities Beyond the Hype?

In 2021, Meta Platforms unveiled its ambitious vision for the metaverse, sparking a polarized debate about the future of digital environments (Dolata & Schwabe, 2023). Alongside this vision, the company announced its rebranding from Facebook to Meta, a move that underscored its commitment to virtual environments. This rebranding signaled to stakeholders that Meta's founder was serious about reshaping the digital landscape.

The timing of this shift was notable. At the time, Facebook was grappling with criticism surrounding its platforms, particularly Instagram, impact on mental health and societal polarization. During these challenges, Meta's rebranding seemed to serve as a strategic move to redirect public and industry attention. The announcement ignited a wave of excitement, with journalists and technology enthusiasts engaging in extensive debates about the potential of immersive systems.

Four years later, the reality appears more complex. Meta has already let go of some of the developers originally hired to bring this vision of the metaverse to life (Heath, 2024). While this might cast some doubt on the company's long-term commitment, it's undeniable that immersive environments are here to stay. In particular, the commercial sector has embraced the metaverse as a new frontier for advertising and selling products.

Yet, beyond its commercial potential, the metaverse offers intriguing possibilities for non-commercial applications. How might this digital sphere be harnessed for societal benefit? Could virtual environments facilitate meaningful public participation, allowing users to contribute as citizens and shape their physical surroundings through virtual engagement?

This article explores the current state of public participation in the metaverse and envisions how these virtual spaces could serve as platforms for civic engagement and societal contributions.

Meta did not coin the term "metaverse," nor did it invent immersive environments. The term was first introduced by author Neal Stephenson

in his 1992 science fiction novel *Snow Crash* (Stephenson, 1992), where the vision of an enhanced virtual world gradually becomes reality. The metaverse, as conceived today, represents a hypothetical, immersive, and interactive virtual space, envisioned as the next generation of the Internet (Dwivedi et al., 2022; Xi et al., 2023). Multiple virtual worlds exist today, operated by various platforms under the collective label of the metaverse. These environments consist of virtual spaces, objects, and entities, accessible through a wide range of devices. Thereby, the metaverse is more than a single technology; it represents a vision of a seamless, interconnected digital world where the boundaries between reality and virtuality blur, enabling new forms of social interaction, commerce, and entertainment (Dwivedi et al., 2022). The term "XR" (extended reality) encompasses all immersive technologies that aim to expand human perception of reality, including augmented reality (AR) and virtual reality (VR), and everything in between (Xi et al., 2023). As Dwivedi et al. (2022) note, immersion in the metaverse can be achieved through either AR or VR, though simultaneous use of both remains a challenge. However, rapid advancements in AR and VR hardware suggest that mixed-reality formats may become more prevalent.

Innovations like Apple's Vision Pro, which seamlessly blends the physical and digital worlds, exemplify how hardware is advancing. Although Meta succeeded in bringing the term "metaverse" into mainstream discourse, the development of XR technologies and platform infrastructure has been more challenging than anticipated. The initial wave of hype has diminished (Robinson, 2023), but the long-term potential of the metaverse remains. Nevertheless, the road to widespread adoption will be more complex than many expected.

This period of hardware and software development presents an opportunity to rethink how the metaverse can and should be used. The debates surrounding the metaverse today are reminiscent of discussions about early online social networks like Facebook and Twitter, now X, in the 2010s. Social media initially sparked great expectations, particularly regarding its potential for deliberative discourse, as seen during the Arab Spring (Wolfsfeld et al., 2013). Over time, however, online social networks have also revealed significant challenges, including the spread of hate speech, misinformation and disinformation. In the platform economy, mechanisms like network effects and microtargeting put advertisers – not users – as the primary customers at the center (Weinhardt et al., 2024). Data-driven

platforms offer advertisers new opportunities for personalized sales, and this has, to a large extent, also driven expectations around the metaverse.

Much of the research into immersive systems has focused on commercial applications. However, to avoid repeating the mistakes made with online social networks, it is essential to consider how the metaverse could be designed for civil society and political use. Early discussions should explore how democratic entities – governments, cities, and local communities – might engage in these new virtual spaces. One promising application is the use of immersive systems for public participation, such as urban planning. The demand for digital participation has grown in recent years, partly due to the COVID-19 pandemic (United Nations, 2020). Therefore, the potential for combining these participation tools with immersive systems in a metaverse context will be explored in the following discussion.

The Current State of Immersive Participation

Back in the 1990s, Lombard and Ditton (1997) found that telepresence and immersion could significantly enhance users' engagement and sense of involvement. This finding has profound implications for using XR technologies in public participation. A technology capable of generating excitement and captivating users in video game contexts seems highly relevant to public participation.

(Digital) Participation

In *Reconstructing Democracy* (2020), Taylor, Nanz, and Taylor emphasize the importance of local participation, giving citizens the opportunity to voice their concerns, even in representative democracies. Particularly in times of major transformation, such as in the energy or mobility sectors, the authors underscore the necessity of including citizens in decision-making processes. The village of Langenegg in Austria provides a positive example of this. Faced with population decline, the local authority engaged citizens in envisioning the region's future. This long-term consultation process not only made the village more attractive to residents but also had a positive impact on population growth, contrary to initial forecasts (Statistik Austria, 2021; Taylor et al., 2020).

In contrast, public participation is often discussed in the context of problematic cases, where participation is initiated to resolve conflicts stemming

from democratic deficits. A prime example is the controversy surrounding the Stuttgart 21 infrastructure project, which involved rebuilding Stuttgart's main railway station. After the partial demolition of the station began in 2010, protests emerged, and the conflict was only resolved through a public participation process and referendum. This case highlights the importance of early public involvement in construction planning, as the lack of transparency had led to widespread confusion and frustration (Brettschneider, 2013), failure to communicate the project's details, such as construction plans, to local citizens damaged trust in politicians and local authorities (Thaa, 2013).

Despite the lessons learned from Stuttgart 21, public participation processes are still frequently criticized for being costly, time-consuming, and complex. However, the knowledge gap between project initiators and affected citizens necessitates more accessible approaches. Since the 2010s, platforms like Consul, Liquid Democracy, LiquidFeedback, CitizenLab (Go Vocal), and Zebralog have emerged to facilitate digital citizen participation (Fegert, 2022), also known as e-participation (Macintosh, 2004). These platforms offer modular systems – some open-source – to streamline participation and voting procedures through digital platforms. Their relatively low cost and simplicity in implementation make them attractive compared to traditional outreach methods (Spirakis et al., 2010).

These digital tools remain niche, but the Pirate Party's use of LiquidFeedback for internal organization and the COVID-19 pandemic, which forced political parties in Germany and other countries to adopt digital platforms for party member participation, helped bring e-participation into the mainstream. However, studies by Rottinghaus and Escher (2020) and Novo Vázquez and Vicente (2019) show that prior political interest and personal relevance remain the key motivators for digital public participation. Additionally, gender-specific differences in e-participation have been observed, with male citizens being more likely to engage (Kim & Lee, 2019). Usability issues are cited as barriers to broader adoption, with e-participation platforms lacking user-friendliness (Fegert et al., 2021).

Given these challenges, there is a need to explore new approaches to digital participation that motivate users and make participation more accessible. The think tank Democracy Technologies predicts that the digital public participation market will grow from €100 million in 2022 to €300 million by 2027 (Democracy Technologies, 2023). As cities and local authorities begin to embrace online participation processes, the challenge lies in designing technologies that align with user preferences.

Digital Participation with Immersive Systems

To effectively enable public participation in local decision-making, it is essential to create intuitive visualizations and user-friendly feedback mechanisms. Immersive systems have the potential to greatly enhance participation platforms. Although no current e-participation platforms fully integrate immersive systems, their potential has been explored extensively, especially in industrial and commercial contexts. Research from these fields offers valuable insights that can be applied to public participation. For instance, Suh and Lee (2005) demonstrated that VR enhances users' knowledge about products, a finding that can be adapted to public participation contexts.

Funded by the Federal Ministry of Research, Technology and Space (BMFTR), the FZI Research Center for Information Technology has been developing and testing immersive participation applications through two projects: *Take Part* (2018–2021) and *VIRTUS* (2021–2024). These projects aim to design and evaluate the potential of immersive systems in public participation while engaging citizens and urban planners in the development process. Early prototypes were refined through various studies, with the goal of providing e-participation platform operators, cities, and local authorities with insights into immersive participation's potential.

In a qualitative interview study conducted in 2018 (n=27), we found significant interest among stakeholders in using immersive systems for public participation (Fegert et al., 2020). The majority of participants saw digital technologies as a valuable complement to traditional public participation, with two-thirds expressing a preference for 3D visualizations over traditional architectural plans. Our research demonstrated that immersive technologies, such as VR and AR, could help reduce the knowledge gap between citizens and experts in public construction projects.

A field study (n=339) from 2019 further supported these findings, showing that immersive systems, particularly VR, significantly enhanced participants' spatial understanding of building sites and urban planning. While both AR and VR improved engagement, VR outperformed AR in helping users visualize and comprehend spatial relationships (Fegert, 2022).

The studies also identified design principles for immersive participation platforms, which include: (1) **Accessibility**: ensuring the platform is easy to navigate and compatible with various devices; (2) **Information Quality**: leveraging appropriate visualization formats tailored to the strengths of different technologies; (3) **Motivation**: engaging users through incentives like badges or gameful design elements; (4) **Transparency**: communicating the

participation process clearly, also in immersive environments; and (5) **Data Protection and Sovereignty**: ensuring user anonymity and safeguarding sensitive data, such as eye-tracking data.

Figure 1: Public participation with immersive systems at an urban planning project in the city of Karlsruhe; 2021 within the research project Take Part

Following Meta Platforms' 2021 announcement, public interest in immersive systems surged, prompting a reassessment of their relevance to public participation within the VIRTUS research project. In early 2022, a qualitative interview study (n=14) explored stakeholders' knowledge and concerns regarding the metaverse. The study revealed that most participants had limited awareness of the metaverse, with those familiar mostly learning about it through media reports on Facebook's rebranding. While respondents saw potential in its use for e-commerce and gaming, they were skeptical about its application for civic participation, expressing concerns about alienation from reality and trust in platform operators. There were also fears of manipulation, a digital divide, and exclusion of individuals with visual impairments. However, some saw opportunities for engaging younger generations and modeling cities through digital twins. Overall, respondents reacted cautiously to the idea of the metaverse as a participation platform, with concerns primarily focused on the trustworthiness of the platform operator.

In contrast to these positive findings, a 2023 study found that the design elements of virtual participation platforms significantly influence the literacies users acquire during their engagement. The study argues that a

more nuanced evaluation of XR tools is necessary to ensure meaningful participation in urban planning processes (Stein, 2024).

In addition to government-funded research projects, student initiatives have creatively explored simpler implementations of immersive participation tools. For example, in 2020, Paulina Porten developed *Augmented Participation*, a tool that combines voice messages with immersive presentations, demonstrating innovative possibilities for enhancing citizen engagement.

Figure 2: Augmented Participation application (Paulina Porten, 2020)

Looking ahead, research on the use of immersive systems in the metaverse for public participation offers exciting possibilities for the future of digital engagement, particularly in urban planning. However, the current state seems far from being market-ready and thus not easily integrable into metaverse environments. As a result, the implementation of immersive participation in the metaverse remains a concept that is still far from widespread realization – lagging even further behind the mainstream adoption of the metaverse itself.

Conclusion: Opportunities and Challenges of Public Participation in the Metaverse

The current state of research indicates that while immersive technologies hold promise for digital public participation, they have yet to be integrated into existing software systems. Their use can greatly motivate public en-

gagement and enhance spatial visualization during participation processes. However, since such processes are typically initiated at the municipal or city level – where the digitization of administrative units, particularly in Germany, is lagging – it is unlikely that the metaverse will play a significant role in public participation within the next five years. While immersive technologies may be utilized in certain areas, e-participation platforms are unlikely to be early immersive systems due to their limited financial and staffing resources, particularly in development.

In contrast, small yet wealthy countries like Saudi Arabia are already focusing on digital twins, and it is more likely that these nations will explore digital twins in the metaverse as prestige projects, experimenting with participation initiatives. However, local contexts suggest that these participation efforts would not occur in democratic environments. It remains to be seen whether real citizen participation will be desired in such projects, or if simulated behavior, powered by generative AI – such as generative agents and memory streams (Park et al., 2023) – will increasingly be used as a substitute. This possibility should be closely monitored to ensure that democratic participation remains a priority if these generative agents become integrated into digital twins or the metaverse.

A more pertinent question is: What are the actual valuable applications for citizen participation in the metaverse? The most feasible applications may be for smaller, non-representative processes. For example, involving diaspora communities in the planning of memorial sites, regardless of their current location, could be a meaningful use case. Another promising scenario is participative meetings within citizens' councils, which have grown in importance in recent years, including at the federal level. The personal interactions central to these councils could translate well to the metaverse, where real-time interactions occur through avatars.

Despite this potential, several challenges remain in enabling citizen participation in the metaverse. These include high hardware costs, hardware performance limitations (such as short battery life and sensitivity to light), and the need for interconnectivity and interoperability of platforms and hardware standards. These practical issues present obstacles to designing inclusive participation processes in the metaverse. As noted earlier, skepticism toward platform operators also hinders adoption. Furthermore, the tension between the desire for personal interaction through avatars and the need for anonymity in participation processes remains unresolved.

Nonetheless, it is crucial to explore how the metaverse could complement existing democratic procedures and analogue participation formats.

The potential to clearly visualize complex subjects and facilitate gatherings across various locations offers a unique opportunity. This presents a challenge for e-participation platform operators, metaverse developers, and researchers: they must consider how to design metaverse spaces that enable democratic participation. The mistakes made with online social networks – where content moderation and the prevalence of hate speech depend largely on the platforms' willingness to act – could be reimagined for the metaverse, provided there is sufficient interest from public entities in creating alternative participatory virtual environments. Ideally, the metaverse will become a space for democratic experimentation and inclusion, rather than a fertile ground for exclusion and hate.

References

Brettschneider, F. (2013). Großprojekte zwischen Protest und Akzeptanz: Legitimation durch Kommunikation. In F. Brettschneider & W. Schuster (Eds.), *Stuttgart 21: Ein Großprojekt zwischen Protest und Akzeptanz* (pp. 319–328). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978 -3–658–01380–6_12

Democracy Technologies. (2023). *Report on democracy technologies in Europe.* https://democracy-technologies.org/report-2023/

Dolata, M., & Schwabe, G. (2023). What is the metaverse and who seeks to define it? Mapping the site of social construction. *Journal of Information Technology*, 38(5), 978–996. https://doi.org/10.1177/02683962231159927

Dwivedi, Y. K., Hughes, L., Baabdullah, A. M., Ribeiro-Navarrete, S., Giannakis, M., Al-Debei, M. M., Dennehy, D., Metri, B., Buhalis, D., Cheung, C. M. K., Conboy, K., Doyle, R., Dubey, R., Dutot, V., Felix, R., Goyal, D. P., Gustafsson, A., Hinsch, C., Jebabli, I., ... Wamba, 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. https://doi.org/10.1016/j.ijinfomgt.2022.102542

Fegert, J. (2022). *Digital Citizen Participation: Involving Citizens Through Immersive Systems in Urban Planning*. https://doi.org/10.5445/IR/1000147640

Fegert, J., Pfeiffer, J., Peukert, C., Golubyeva, A., & Weinhardt, C. (2020, December 14). *Combining e-participation with augmented and virtual reality: Insights from a design science research project.* Proceedings of the International Conference on Information Systems (ICIS). https://aisel.aisnet.org/icis2020/digitization_in_cities/digitization_in_cities/4

Fegert, J., Stein, C., Peukert, C., & Weinhardt, C. (2021). Using e-participation for mission statement development: Promises and challenges. In P. Kommers (Ed.), *Proceedings of the 19th International Conference "e-Society"* (p. 233). https://publikationen.bibliothek.kit.edu/1000130538

Heath, A. (2024, June 18). *Meta forms new wearables group and lays off some employees.* The Verge. https://www.theverge.com/2024/6/18/24181420/meta-wearables-reality-labs-layoffs

Kim, S., & Lee, J. (2019). Gender and e-participation in local governance: Citizen e-participation values and social ties. *International Journal of Public Administration,* 42(13), 1073–1083. https://doi.org/10.1080/01900692.2019.1575669

Lombard, M., & Ditton, T. (1997). At the heart of it all: The concept of presence. *Journal of Computer-Mediated Communication,* 3(JCMC321). https://doi.org/10.1111/j.1083-6101.1997.tb00072.x

Macintosh, A. (2004). *Characterizing E-Participation in Policy-Making.* Proceedings of the Proceedings of the 37th Annual Hawaii International Conference on System Sciences (HICSS'04) – Track 5 – Volume 5, 50117.1-. http://dl.acm.org/citation.cfm?id=962753.962993

Novo Vázquez, A., & Vicente, M. R. (2019). Exploring the determinants of e-participation in smart cities. In M. P. Rodríguez Bolívar & L. Alcaide Muñoz (Eds.), *E-participation in smart cities: Technologies and models of governance for citizen engagement* (pp. 157–178). Springer International Publishing. https://doi.org/10.1007/978-3-319-89474-4_8

Park, J. S., O'Brien, J. C., Cai, C. J., Morris, M. R., Liang, P., & Bernstein, M. S. (2023). *Generative agents: Interactive simulacra of human behavior.* arXiv Preprint. https://arxiv.org/abs/2304.03442

Porten, P. (2020, March). *Paulina Porten—Augmented participation.* https://paulinaporten.com/augmented-participation-1

Rauschnabel, P. A., Felix, R., Hinsch, C., Shahab, H., & Alt, F. (2022). What is XR? Towards a framework for augmented and virtual reality. *Computers in Human Behavior,* 133, 107289. https://doi.org/10.1016/j.chb.2022.107289

Robinson, K. (2023, November 7). *Meta exec: "Metaverse hype is dead."* Fortune. https://fortune.com/2023/07/11/meta-vp-vishal-shah-happy-metaverse-hype-is-dead/

Rottinghaus, B., & Escher, T. (2020). Mechanisms for inclusion and exclusion through digital political participation: Evidence from a comparative study of online consultations in three German cities. *Zeitschrift für Politikwissenschaft,* 30(2), 261–298. https://doi.org/10.1007/s41358-020-00222-7

Spirakis, G., Spiraki, C., & Nikolopoulos, K. (2010). The impact of electronic government on democracy: E-democracy through e-participation. *Electronic Government*, 7(1), 75–88. https://doi.org/10.1504/EG.2010.029892

Statistik Austria. (2021). *Ein Blick auf die Gemeinde Langenegg.* https://www.statistik.at/blickgem/G0201/g80223.pdf

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

Stein, C., & Fegert, J. (2024). *Bridging realities: Exploring enablement factors for XR participatory urban planning.* Proceedings of the European Conference on Information Systems (ECIS).

Suh, K.-S., & Lee, Y. E. (2005). The effects of virtual reality on consumer learning: An empirical investigation. *MIS Quarterly,* 29(4), 673–697. https://doi.org/10.2307/25148705

Taylor, C., Nanz, P., & Taylor, M. B. (2020). *Reconstructing democracy: How citizens are building from the ground up*. Harvard University Press.

Thaa, W. (2013). „Stuttgart 21" – Krise oder Repolitisierung der repräsentativen Demokratie? *Politische Vierteljahresschrift*, 54(1), 1–20. JSTOR.

United Nations. (2020, April 15). *Embracing digital government during the pandemic and beyond*. Economic Analysis & Policy Division. https://www.un.org/development/desa/dpad/publication/un-desa-policy-brief-61-covid-19-embracing-digital-government-during-the-pandemic-and-beyond/

Weinhardt, C., Fegert, J., Hinz, O., & van der Aalst, W. M. P. (2024). Digital democracy: A wake-up call. *Business & Information Systems Engineering*, 66(2), 127–134. https://doi.org/10.1007/s12599-024-00862-x

Wolfsfeld, G., Segev, E., & Sheafer, T. (2013). Social media and the Arab Spring: Politics comes first. *The International Journal of Press/Politics,* 18(2), 115–137. https://doi.org/10.1177/1940161212471716

Xi, N., Chen, J., Gama, F., Riar, M., & Hamari, J. (2023). The challenges of entering the metaverse: An experiment on the effect of extended reality on workload. *Information Systems Frontiers*, 25(2), 659–680. https://doi.org/10.1007/s10796-022-10244-x

EU Platform Regulation and its Implications for the Metaverse: An Analysis of DSA, DMA, and Related Legal Acts

Martin Müller & Matthias C. Kettemann

1 Introduction[1]

In recent years, the European Union (EU) has significantly changed its attitude towards digital tools and technologies, platforms, services and markets. After years of adopting a more liberal stance, it has now strengthened its regulatory measures to meet the challenges of digitization. A decisive moment was the adoption of the General Data Protection Regulation (GDPR)[2] in 2016, which marked the beginning of a series of new regulations that have already been adopted as part of the Commission's strategy "A Europe Fit for the Digital Age" (European Commission, n.d). Against this background, the question arises as to whether these regulatory approaches already have implications for metaverses. The following article will first summarize the individual legal acts and then show the connections to the metaverse.

2 Legal Acts of European Platform Regulation

A total of five legal acts can be assigned to platform regulation of the EU (Steinrötter et al., 2025), but due to reasons of space only the Digital

1 This chapter that has been published in a longer version as Müller & Kettemann, '§ 7 Plattformregulierung' in Steege/Chibanguza (eds.), Metaverse. Rechtshandbuch, pp. 135–147 and has been translated and significantly updated for this volume. All online sources were retrieved on 31 July 2025 last.

2 Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons regarding the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation), OJ L 119, 1.

Services Act (DSA)[3] and the Digital Markets Act (DMA)[4] will be examined in detail. The other three acts, namely the Data Governance Act (DGA)[5], Data Act[6] and the Artificial Intelligence Act (AI Act[7]) will not be described in detail but their applicability to the Metaverse will briefly be analyzed in the second part of this article.

2.1 Regulating the Social Power of Platforms: Digital Services Act (DSA)

The DSA is an EU regulation that came into force in November 2022. The DSA and the DMA, which were co-negotiated at the same time, aim to create a safer digital space in which the fundamental rights of users of digital services are protected and to create a level playing field to promote innovation, growth and competitiveness in the European single market.

The DSA partially replaces the e-Commerce Directive[8] and further develops it through new due diligence obligations and the establishment of more uniform supervision throughout the Union, but retains the traditionally innovation-friendly liability regime (and the basic exemption from liability for third-party content if actual knowledge is lacking). The aim of the e-Commerce Directive was to create a legal framework that facilitates the free movement of intermediary services within the EU in order to promote innovation and e-commerce. The DSA now calls on digital platforms in

3 Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on an internal market for digital services and amending Directive 2000/31/EC (Digital Services Act), OJ L 277, 1.

4 Regulation (EU) 2022/1925 of the European Parliament and of the Council of 14 September 2022 on contestable and fair markets in the digital sector and amending Directives (EU) 2019/1937 and (EU) 2020/1828 (Digital Markets Act), OJ L 265, 1.

5 Regulation (EU) 2022/868 of the European Parliament and of the Council of 30 May 2022 on European data governance and amending Regulation (EU) 2018/1724 (Data Governance Act), OJ L 152, 1.

6 Regulation (EU) 2023/2854 of the European Parliament and of the Council of 13 December 2023 on harmonised rules on fair access to and use of data and amending Regulation (EU) 2017/2394 and Directive (EU) 2020/1828 (Data Act).

7 Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act).

8 Directive 2000/31/EC of the European Parliament and of the Council of 8 June 2000 on certain legal aspects of information society services, in particular electronic commerce, in the Internal Market ('Directive on electronic commerce'). OJ L 178, 1.

particular as responsible actors in the fight against specific illegal content and formulates important procedural, transparency and compliance obligations (Gerdemann & Spindler, 2023, p. 3).

2.2 Regulating the Economic Power of Platforms: Digital Markets Act (DMA)

The DMA seeks to limit the economic power of "Big Tech" platforms in digital markets. Although the competition law provisions at EU law (namely Articles 101 and 102 of the Treaty on the Functioning of the European Union, TFEU) and member state level apply to digital platforms also, they have been described as "*too little, too late*" in recent years. The justification for this is particularly evident in the long duration of proceedings by the European Commission against platforms that have not been able to improve competition in the market (Podszun et al., 2021, p. 60f.). The DMA moves from the so-called *ex-post* approach (i.e. that fines or other measures can only be imposed after a violation has been determined by the authorities) of classic competition law to an *ex-ante* approach and contains a total of 21 different due diligence obligations that are considered harmful to competition in digital markets. These due diligence obligations do not apply to all online platforms, but only to those designated as gatekeepers by the European Commission. For these general requirements, specific thresholds follow in Article 3 (2) DMA, all of which are fulfilled by the well-known Big Tech companies.

3 Applicability of the Legal Acts to Existing Metaverses

3.1 Applicability of Platform Regulation Acts for Metaverses

a) Application of the DSA

In the case of an offer for users in the Union, metaverses fall within the scope of the DSA due to the marketplace principle, even if they do not have an establishment in the Union. From a factual point of view, metaverses are to be regarded as hosting services in the tripartite division of intermediary services: In order to represent the virtual world and interact with it, the operators must store information of the users on their behalf, so that the

requirements of Art. 3 lit. g no. iii DSA for hosting services are met. Not directly relevant, but as an indication for the classification under hosting services, is also the fact that in most cases the design of the metaverses is technically and with regard to the applicable rules in the hands of their operators.

A special case is represented by *completely decentralised* metaverses. There is a lack *of a* hosting service that operates the metaverse, but there are *several* operators. Because the DSA continues the concept of the e-Commerce Directive, it will have to be assumed that the DSA cannot be applied here. This is also supported by the fact that there are no provisions for separate responsibilities, as is the case in Art. 26 GDPR, for example.[9] At the moment, however, no market-ready metaverse can be seen that does not allow attribution to a hosting service, so these considerations do not yet have any practical relevance.

At the heart of the DSA are due diligence obligations standardised in Chapter III, which initially affect all intermediary services and then individual regimes for all hosting services and then the even more specialised online platforms and very large online platforms and online search engines. Only if these obligations are complied with do the exclusions of liability now standardised in detail by Art. 4–6 DSA apply (previously Art. 12–15 e-Commerce Directive implemented in member states' kaw). In this respect, Art. 1 para. 2 lit. a DSA correctly speaks of a "conditional exemption from liability".

Since metaverses fall under the definition of hosting services, the question arises as to whether they also meet the requirements for online platforms as special hosting services. According to the legal definition of Art. 3 lit. i DSA, the decisive factor here is that, in addition to the storage of users' information, "public dissemination" takes place, which does not constitute a secondary function. Based on the reshaped understanding of "communication to the public" within the meaning of Section 15 (2) UrhG, Article 3 (k) of the DSA is based on the fact that a "provision of information to a potentially unlimited number of third parties" takes place; unlike the UrhG, it is about information that is provided on behalf of users. In metaverses, this data can be found, for example in the generally accessible design of the avatar by individual users or in the transfer of virtual items

9 With reference to metaverses Kaulartz/Schmid/Müller-Eising, RDi 2022, 521 para. 39.

via NFTs. This means that **metaverses are online platforms within the meaning of the DSA.**[10]

First of all, metaverse operators must comply with the provisions of Art. 11–15 DSA, which apply to all intermediation services. Here, further obligations are now being introduced compared to the e-Commerce Directive: For example, contact points for authorities, the commission and users must be kept available (Art. 11, 12 DSA) In the absence of establishment in the European Union, a legal representative in the European Union must be appointed in accordance with Article 13 of the DSA. Finally, the general terms and conditions of all intermediary services must meet certain requirements (Art. 14 DSA). Finally, the transparency obligation for content moderation under Art. 15 DSA should be mentioned. According to the legal definition of Art. 3 lit. t DSA, this refers to how intermediary services deal with illegal content. As hosting services, providers of metaverses must also comply with the obligations of Art. 16–18 DSA and contain the provisions on the precise design of notification and remediation procedures for illegal content and the obligation for providers to report serious crimes to law enforcement or judicial authorities if they become aware of them.

In their function as online platforms, the provisions of Art. 19–28 DSA apply. For smaller providers of metaverses, reference should be made at this point to the provisions of Art. 19 para. 1 subpara. 1 DSA within the meaning of Recommendation 2003/361/EC, which limits the regime of obligations for online platforms to Art. 24 para. 3 DSA (Kraul, 2023); however, with the exception that these are not very large online platforms (= an average of 45 million active users in the Union and corresponding designation by the Commission, Art. 33 paras. 1, 4 DSA). Articles 19–28 of the DSA contain a variety of different obligations, such as the establishment of an internal complaint management system (Art. 20 DSA), other transparency obligations, including with regard to advertising and the use of recommendation systems (Art. 24, 26 et seq. DSA) or the online protection of minors (Art. 28 DSA). If metaverses are platforms that allow the conclusion of distance contracts, in other words online trading platforms, the provisions on the traceability of traders (Art. 30 DSA) or on the information of consumers under Art. 32 DSA must be followed.[11] Here, too, exceptions apply to micro and small enterprises, Art. 29 para. 1 DSA. Since the distinction between

10 Using the term, Kaulartz et al., RDi 2022, 521 para. 56, but without apparent further classification under the requirements of the DSA.

11 Kraul, Das neue Recht der digitalen Dienste, § 4 marginal no. 178.

online platforms and very large online platforms is based solely on the number of active users, it may also be possible in the future to apply the obligations of very large online platforms regulated in Art. 33–43 DSA.

b) Application of the DMA

For the application of the DMA to metaverses, there is no designation of metaverses as central platform services – regardless of the variable prerequisites of influence on the internal market and the consolidated and lasting position with a view to the future. Part of the ex-ante approach now being pursued is the possibility for the Commission to designate new services by means of a market investigation under Art. 19 DMA and thus avoid (quasi-)monopolistic digital markets for future platform services. Since this is an essential part of the DMA, an adjustment of the list of central platform services that can be justified on the basis of Art. 290 para. 1 subpara. 2 TFEU cannot be carried out by the Commission itself by means of a delegated act (Art. 288 TFEU), but by means of an adaptation of the DMA via the ordinary legislative process (Schmidt & Hübener*, 2023)*. The **DMA is therefore not currently applicable to metaverses**, but with the corresponding advancement of technology, the adaptation of the list of central platform services appears to be a suitable instrument to prevent market failure.

c) Applicability of the Other Provisions of Platform Regulation

The other provisions of platform regulation also do not directly address metaverses, but are applicable to them in view of their technology-open foundations. Among the two data-related legal acts, the planned provisions of the Data Act may be particularly relevant for operators of metaverses: Since access to the metaverse will already be partially, but even more so in the future, linked to the use of AR/VR glasses, the regulation on the "horizontal right" to hand over user data to data holders in the case of IoT devices will be relevant. It also seems possible that the interoperability regulations can ensure a certain standardization of data formats in metaverses. The provisions on the interoperability of data only apply to data intermediation services (Art. 26 paras. 3, 4 in conjunction with Art. 29 Data Act) and operators of data rooms pursuant to Art. 28 et seq. Data Act. However, if these services have the success predicted by the Commission (European

Commission, n.d),[12] large parts of the digital economy, such as operators of metaverses, will use data switching services and data spaces in the near future and will therefore inevitably have to follow the standardisation of these services.

In the area of the AI Regulation, individual metaverses will not be able to do without AI (Paal 2022, p.194), even if metaverses are once again not directly accessible to regulation. Based on the named risk classes, operators of metaverses will design the use of their algorithms in a legally compliant manner.

3.2 Platform Regulation and "Private Orders" of the Metaverse

The statutory requirements already mentioned to general terms and conditions or the "private regulations" of the operators of metaverses are changing due to new provisions in the legal acts of platform regulation.

First and foremost is Art. 14 DSA, which deals with the general terms and conditions of platforms. In addition to the minimum requirements for the content of the GTC described in Art. 14 para. 1 DSA, Art. 14 para. 4 DSA now requires that the interests of users must be taken into account in the moderation of content and in complaints handled by platforms. The users of fundamental rights, such as the fundamental right to freedom of expression, are explicitly mentioned here. Contrary to the previous case law of the Federal Court of Justice, this is a direct, "horizontal" fundamental rights binding of the platforms, regardless of their size (Spindler, 2021, p. 545 (551); Quintais et al., 2022.)

A further influence on private orders can be shown on the basis of the FRAND conditions, which can be found in Art. 6 para. 6, 12 DMA, Art. 6 para. 2 DGA and Art. 8 para. 1 Data Act. The overarching goal of these provisions, which are similar in content, can be seen in the fact that they shape access to "data treasures" in the same way and do not make them dependent on the position of power of the data holders – currently initially platforms, and in the future possibly also operators of metaverses. In the broadest sense, these requirements can be understood as safeguarding the fundamental rights of users, primarily their freedom of contract, which

12 European Commission, European Data Strategy, https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/europe-fit-digital-age/european-data-strategy_de (last accessed July 31, 2025).

is protected in terms of fundamental and human rights, for example by the general freedom of action under Article 2.1 of the German Basic Law (BVerfGE 95, p. 267 (303f)), the freedom of occupation or the freedom to conduct a business (Article 12 of the German Basic Law, Article 15 of the Charter of Fundamental Rights) or guarantees of property in Article 14 of the German Basic Law or Article 1 of the 1st Additional Protocol to the ECHR[13] is protected. In the context of practical concordance, this justifies the encroachment on the fundamental rights of the service holders.

4 Conclusion

In addition to the "normative power" of platforms, the Union legal acts on platform regulation also regulate (Mendelsohn, 2021, p. 857 (857f); Kettemann, 2020) the factors that justify the success of platforms with data and its algorithmic evaluation. With regard to metaverses, it can be seen that the legal acts are open to future digital developments and are already determining the design. Specifically, in the near future, the operators of metaverses will have to comply with the extensive due diligence obligations for online platforms within the meaning of the DSA and – provided that the concept of the "horizontal law" of the Data Act is not comprehensively changed – ensure that the data obtained can be made available to users. Not relevant at the moment, but possible with appropriate economic success, is also an application of the DMA to ensure sufficient competition between the various operators of metaverses. This is made possible by the technology-open design of the legal acts and the possibility of accelerated adaptation to technological developments.

References

European Commission. (n.d.). *A Europe fit for the digital age.* https://ec.europa.eu/info/strategy/priorities-2019-2024/europe-fit-digital-age_de (last accessed July 31, 2025).

Gerdemann, S., & Spindler, G. (2023). *Das Gesetz über digitale Dienste (Digital Services Act) (Teil 1). Gewerblicher Rechtsschutz und Urheberrecht* (GRUR), 2023(3), 3.

13 Human rights case law has so far been issued in the context of regulations of social tenancy law, most recently Pařízek v. Czech Republic, No. 76286/14, judgment of 12.01.2023, § 53f.

Kaulartz, M., Schmid, F., & Müller-Eising, M. (2022). *Das Metaverse – eine rechtliche Einführung. Recht & Digitalisierung* (RDi), 521.

Kettemann, M. C. (2020). *The Normative Order of the Internet: A theory of rule and regulation online*. Oxford University Press. (Oxford Scholarship Online; Print ISBN 978-0-19-886599-5; First published July 27, 2020) publikationen.ub.uni-frankfurt.de+9lehmanns.de+9researchgate.net+9rsw.beck.de+1beck-online.beck.de+1

Kraul, S. (Hrsg.). (2023). *Das neue Recht der digitalen Dienste* (1st ed.).

Mendelsohn, J. (2021). *Die „normative Macht" der Plattformen – Gegenstand der zukünftigen Digitalregulierung? MMR – Multimedia und Recht*, 857.

Paal, B. (2022). *Metaversum und Recht*. In S. Gräfe (Hrsg.), *Telemedicus Tagungsband zur Sommerkonferenz 2022 – Das Recht der Informationsgesellschaft* (pp. 185–195).

Podszun, R., Bongartz, R., & Langenstein, G. (2021). *The Digital Markets Act: Moving from competition law to regulation of large gatekeepers. European Competition and Market Law Review* (EuCML), 60.

Quintais, J. P., Appelman, P., & Fahy, E. (2022, December 14). *Using terms and conditions to apply fundamental rights to content moderation*. SSRN. https://ssrn.com/abstract=4286147

Schmidt, M., & Hübener, S. (Hrsg.). (2023). *Der neue Recht der digitalen Märkte* (1st ed.).

Spindler, Gerald (2021): Anmerkung. In: GRUR 2021, S. 545–551. Most recently BVerfGE 95, 267 (303f).

Steinrötter, J., Heinze, R., & Denga, S. (2025). *EU Platform Law*.

Regulation of the Metaverse

Caroline Böck & Matthias C. Kettemann[1]

The metaverse – a term that is associated with many associations in both positive and negative ways, although the *one* metaverse does not yet exist. Its exact design has therefore not yet been conclusively clarified. Mark Zuckerberg, founder and CEO of Meta, describes his vision of the metaverse as follows: "We will be able to do almost anything in the metaverse that is within our imagination: meetings [acquaintances] and family, working, learning, playing, shopping, creating content – and completely new things that we [...] not even imagine." (Facebook Connect, 2021) The aim is thus to create a link between the digital and real worlds. It is therefore clear that the metaverse is intended to represent a new type of internet (Europol, 2022) that will differ significantly from current internet use.

Such technological developments can have an impact on numerous areas of life, as the existing Internet has shown with the rise of platforms. A similar potential for change is attributed to the metaverse: Influences on the economy, especially trade, the real estate market and in the area of financial investments are conceivable, but also within social structures, primarily the world of work and leisure activities (Höfler & Krolle, 2023). These can have positive but also negative effects, such as social isolation, loss of reality or simulated abuse. In economic terms, significant effects are already evident. Within the virtual, blockchain-based platform "Decentralland", land sales are made that amount to the equivalent of several million euros (futurezone, 2021). In addition, according to empirical research by McKinsey & Company, approximately EUR 110 billion has already been invested in the development of the metaverse in 2022 (McKinsey & Company, 2022). The market volume of the metaverse is expected to be around EUR 507 billion by 2030 (statista, n.d). Therefore, from a legal point of view, the fundamental question arises as to whether and how the law should or can react to these technical, psychological, social and economic effects. Are

1 This paper represents a shorter translation and an update of the article Böck & Kettemann, Regulierung des Metaverse in Steege & Chibanguza, Metaverse Rechtshandbuch.

existing regulatory concepts sufficient or is a new strategy necessary? The article will provide an overview of these.

I. Definition in the Legal Sense

In order to present the legal implications of the metaverse, it is first necessary to define the term in more detail. The metaverse concept originates from science fiction literature: In his 1992 novel "Snow Crash", author *Neal Stephenson* created a world called the "Metaverse" for his characters, which was supposed to be a safe haven for them (Kaulartz et al., 2022, p. 521f.). The understanding of the term in the novel is thus very different from the technical metaverse that is to be created by Mark Zuckerberg and other digital companies, such as Animoca Brands from Hong Kong, as they dream of linking the real and digital worlds. From a technical point of view, the graphic online role-playing game "Habitat" by Lucasfilm Games from 1985 is considered the first metaverse (Bendel, 2021).

In the absence of an actual metaverse and the different ideas of the developers, no uniform definition can be found. The EU Commission and the EU Parliament have also developed their own definition: The metaverse is defined as "immersive and constant virtual 3D world where people interact through an avatar to enjoy entertainment, make purchases and carry out transactions with crypto-assets, or work without leaving their seat" (European Commission & Analysis and Research Team, Metaverse 2022, p. 3).

For further narrowing down, it makes sense to distinguish between the exact characteristics of the software, the hardware and the "content" (Park & Kim, 2022). In terms of hardware, the use, availability and development of virtual reality (VR), mixed reality and augmented reality (AR) devices will play an essential role (Kaulartz et al., 2022, p. 521f.), as these (glasses) provide the connection to the virtual world. The exact technical implementation of the software can be narrowed down in two directions: On the one hand, work is being done on platform solutions. Other companies, such as Animoca Brands, rely on solutions that offer an open, interoperable and decentralized infrastructure (Bendel, 2021). Within the third category of "content", it will depend on how the metaverse will be designed. Fantasy worlds as well as worlds that represent a reflection of the "real" world are conceivable here (Park & Kim, 2022).

II. Regulation of the Metaverse

Following the establishment of a definition, the question of which rules the metaverse follows from a legal point of view can be considered. It becomes clear that there is no such thing as a metaverse as such. Despite Facebook's rebranding as "Meta Platforms," which appears to suggest an ambition to create a unified metaverse, the actuality is divergent. As already mentioned, various companies are developing numerous metaverses that have individual similar features (Ball, 2020) but are nevertheless substantially different from each other. Therefore, the question now arises as to which rules apply to metaverses and whether government regulation exists for them. It is evident that the rules as well as state regulation depend on how the metaverses are specifically designed and which life circumstances are concretely realized (Kaulartz et al., 2022; Müller, 2022, 281).

It should be pointed out here that a separation of the term "rules" (norms) must be made from the term "regulation". In legal terminology, the term regulation means legal acts of state influence or lawmaking. Rules do not necessarily arise from state actors, but rather arise from social rules, such as moral and religious commands and prohibitions, conventions or customs or customs (Rüthers et al., 2022; Vesting, 2015). The necessity of distinguishing between terms arises as they have different legal effects.[2] This is important for the present case, as the metaverse is dependent on rules as well as regulation.

1. What Rules Apply?

a) Regulatory Approaches

As of May 2025, neither the national nor the Union legislators have developed a regulatory concept for the emerging metaverses. A concrete state, supranational or international one is therefore missing. Nevertheless, the individual characteristics of the metaverse are subject to existing national as well as european regulations, at least to some extent.

In relation to the design of the hardware, specifically the VR glasses or analogous connection objects, digital companies or manufacturers are bound by the prevailing regulations concerning product safety. The

2 Vesting, Rechtstheorie, § 2 marginal no. 33.

present-day placement of products on the market is primarily shaped by Union law, meaning that manufacturers – at least within the EU – are subject to specific regulations (Weber, 2022). To illustrate this point, one may consider the existence of general and specific legal acts, a phenomenon that is especially evident within the domain of electrical appliances (Langner et al., 2023). To the extent that the connection devices can still be classified as electrical devices, these regulations apply to this feature of the metaverse. In the meantime, broad catalogues of obligations for operators of online marketplaces have also been established (Kapoor & Klindt, 2022), so that these product liability regulations could also apply to operators in the metaverse. However, this depends on the software on which the metaverse is based, i.e. the question of whether platform solutions will prevail in the competition. This will already ensure that the hardware does not pose any significant risks to end users.

In addition, the characteristic of "content", i.e. the concrete design and use of the metaverse, is also bound to certain norms. These include, in particular, criminal law provisions. Testing the beta version of the Meta Horizons metaverse has revealed the first cases of potentially criminal acts (Der Standard, 2021). From a German point of view, the basic prerequisite for the prosecution of potentially criminal acts in the metaverse is therefore the applicability of the German Criminal Code. Due to the territorial principle, sovereign punitive power is limited to crimes on one's own territory (Mills, 2006; Schmalenbach & Bast, 2017). When determining the place of criminal offences on the internet, it is recognised that at least those offences are subject to German criminal law that were committed against or by a German citizen (Schönke & Schröder, 2019). In principle, this can be transferred to the metaverse if there will be the possibility of assigning the avatar of the metaverse to a specific real person (Kaulartz et al., 2022). From a German point of view, the prosecution of criminal offences is only possible to a very limited extent and excludes some cases (Beukelmann, 2012).

Finally, the question arises as to the extent to which the Digital Services Act (DSA) and the Digital Markets Act (DMA) already enable the regulation of the metaverse. DSA and DMA are two pieces of legislation adopted by the EU as part of its digital strategy (EU Commission, n.d). The DSA focuses on the fundamental rights of users by preventing much-discussed phenomena such as "hate speech" and "fake news". The addressees of the provision, namely the large digital companies, are subject to duties of care

and other duties of action aimed at content control within the platforms (Kettemann et al., 2021, p. 138). These include deletion obligations under Art. 9 (1) DSA if a state authority orders actions against illegal or unlawful content. These are to be complied with by threatening the addressees with fines, some of which are large (Kühling, 2021, p. 461). In addition, the DMA standardises competition law regulations for so-called gatekeepers who offer central platform services and hold a dominant position in the market (Gielen & Uphues, 2021, p. 627). According to Art. 2 No. 2 DSA, gatekeepers include online search engines such as Google and social media, such as Facebook and Instagram.[3] According to Art. 3 para. 1 DMA, the central characteristics of the gatekeeper are a significant influence on the internal market, the provision of a central platform service that serves as an important gateway to end users for business users, and a consolidated and lasting position of its business activity exists or it is foreseeable that it will achieve one.

The metaverses could fall under the scope of these two legislative acts and thus be classified as part of platform regulation (Kalbhenn, 2021; Paal, 2022 p. 194). However, this depends largely on the design of the software. The scope of application is only opened up to providers of metaverses that are also to be classified as intermediary services within the meaning of the DSA or as gatekeepers within the meaning of the DMA (Kaulartz et al., 2022, p. 521 ff.). This could be relevant to the emerging metaverses. The extent to which metaverse providers will prevail in the market with platform-based solutions remains to be ascertained. From the perspective of the EU legislator, it would be prudent to analyse the further development process and consider the extent to which the target group within existing legal acts can be adapted to encompass other types of metaverses, if necessary. This appears to be a desirable course of action, as it would enable the effective protection of users' rights against unlawful and illegal content, while also demonstrating the limits of competition law.

b) Rule-Making by Digital Companies

It turns out that the existing regulations only marginally regulate the essential features of the metaverse, namely the hardware, software and "content"

3 The EU Commission classified six gatekeeps in September 2023 - Alphabet, Amazon, Apple, ByteDance, Meta, Microsoft, cf. Art. 3 para. 6 DMA.

and are strongly dependent on which metaverses will prevail on the market and how they are specifically designed. Instead, the key players in shaping the metaverse are the digital companies previously mentioned, given their ability to determine the precise evolution of their metaverse without strong legal regulation.

This phenomenon has already been demonstrated with the platform economy (Mast, 2023, p. 287f.). Platform companies offer services in almost every area of life and play a key role in this – also because of their technical sovereignty (Spindler & Seidel 2022, p. 2733f.). Over time, so-called private orders have developed through the digital companies (and due to a certain legal and social pressure), which specify rules of conduct on the platforms. The term is the opposite of "public order" and describes independent sets of rules on subsystems and orders of a society that are independent of the state.

These private orders on digital platforms, such as social media, are nowadays subject to technical settings but also to the guidelines that the digital company itself has developed (Quintais et al., 2023). These rules are often called "community guidelines" and describe the user relationships to each other, but also the relationship between the user and the platform. Overall, these rules are structured and very extensive. They don't just set up a few superficial rules.

The question arises as to how these private regulations can be legally classified and what validity they have vis-à-vis the users. The principle of the rule of law posits that private entities, by virtue of their own authority, cannot legislate in a binding manner without the explicit consent of the individuals concerned (Rennert, 2009, p. 976 (p. 982ff.)). In this respect, the platform guidelines established cannot constitute a law in the classic sense. Rather, they are to be assigned to civil law and have a concrete effect on the basis of a legal relationship under private law. The admissibility of such regulations can in turn be derived from the Basic Law, specifically the fundamental rights and orders within the legal requirements (Teubner, 2012; Mast et al., 2024). This also applies to platform companies.

It follows from what has just been said that users must agree to the respective regulations in order for them to have legal effect vis-à-vis them in the first place. Nowadays, the use of a platform is only possible after prior agreement to the terms of use, which contain the Community Guidelines (so-called opt-in procedure) (Bräutigam & Rücker, 2017). This typically results in a so-called platform usage contract (Friehe, 2020, p. 1697; OLG München 2018, p. 3115f.; BGH 2021, p. 953 (957f.)). According to German

law, the terms of use themselves are regularly included in the contract as general terms and conditions in accordance with Section 305 (1) of the German Civil Code (BGB) and do not constitute the main contract (Spindler, 2019, p. 238 (240); Friehe, 2020, p. 1697; OLG München, 2018, 3115 (3116)). The individual clauses must stand up to judicial review by the general terms and conditions on their own.

These types of terms of service can also play a significant role within the metaverse, as long as they are used for the metaverse (Heet & Kalbhenn, 2021). This appears to be a probable outcome in view of a planned "mass business" with a large number of users, especially since the virtual worlds are also being developed by digital companies. Something else may result from the fact that other metaverse companies do not want to offer a platform solution, but want to create a decentralized virtual world. However, it is questionable how access to these metaverses will be structured differently and to what extent the operating companies will not still have a key position, as is the case in the context of operating a platform. If this exists, there is a likelihood that some legal responsibility will be attributed to the companies. In addition, it is also in the interest of these companies to generate a positive external perception through set standards.

2. Preliminary Conclusion

The findings show that developing companies play a central role in shaping the metaverse and most likely they are the same companies that developed the (social) platforms. If one looks at the existing legislation, especially at the level of EU law with regard to platforms with the DSA and the DMA, as well as the AI Act (Paal 2022, p. 194ff.), it becomes clear that these can at least partially apply to metaverses, but do not fully regulate them. Rather, they should be seen as a starting point that should be used to help determine the regulation of the metaverse as it emerges (Heet & Kalbhenn, 2021). While digital companies are instrumental in "regulating" the metaverse, lawmakers can exert indirect influence through legal acts and judicial control, as described earlier.

III. Regulation in the Metaverse

It is not only the question of whether and what regulation underlies the development of metaverses. Rather, from a legal point of view, the question

arises as to which rules and regulatory concepts apply *within* the metaverse(s). In this context, the question of the validity of existing legal areas in the metaverse is also raised. In the event of the implementation of the aforementioned platform solutions, the terms of use of the platform operators will be applicable, provided that they are incorporated correctly and demonstrate resilience to content control. In each individual case, it will be necessary to clarify the relationship between these terms and the legal requirements from various areas of law, including but not limited to civil law, regulatory principles, criminal law, domiciliary rights, the protection of minors and data protection. It is important to note that there are several additional points of contact. For reasons of space, an example will be discussed below to illustrate how diverse the questions are that arise within the metaverse.

For instance, data protection issues arise within the metaverse. Given the opportunities that the metaverses will offer, it is evident that even enabling access to the metaverse requires significant data provision. But the use of the metaverse will also produce a significant amount of sensitive data, such as movement data, physical and psychological reaction data, as well as visual and biometric data, collected by the VR or AG glasses (Paal, 2022, p. 191; Bender-Paukens & Werry, 2023, p. 127 (128)). The data collected by the hardware is so precise that it can permanently record and store breath, pulse and changes in the eye (Bender-Paukens & Werry, 2023, p. 127f.).

In principle, it can be assumed that the GDPR is applicable within the metaverses. However, individual transmission difficulties are evident in the further application of the GDPR. This can be seen, for example, in the fundamental question of responsibility under Art. 4 (7) GDPR, which is of particular importance in the assignment of the further catalogue of obligations of the GDPR. Due to the large number of metaverses that will emerge as well as the possibility of decentralized infrastructures and interoperability, it is not possible to clearly determine a controller within the meaning of the GDPR. The GDPR does not currently allow for a clear interpretation, as it is not tailored to the possible specifics of the metaverse.

IV. Need for Regulation

Within the development of the metaverse, it can be seen that metaverses that are comparable to the existing social platforms are being sought by the majority. Nevertheless, some developers are trying to offer metaverses

that work in a decentralized way. Regardless of which metaverses will prevail and how exactly interoperability between different metaverses will be designed, the operators behind them will play a major role in the design and thus the coexistence in the metaverses. The regulation or rules of the metaverses are therefore particularly dependent on private actors. In principle, state regulation plays a subordinate role in the structuring of private law relationships that is relevant here (Krönke, 2020, p. 195), especially since, as described above, the private actors themselves are obliged to set rules which gives them a strong role of power.

The need for regulation is therefore evident. The timing is just as favorable. Unlike platform regulation by the EU, which only took place after the platform operators had already achieved a dominant position in the market and the platforms were already used by billions of people, timely regulation, ready, could have a lasting impact on the development of metaverses. This allows legislative limits and values to be implemented more effectively (Kaulartz et al., 2022, p. 521 (531); Paal, 2022, p. 191). Law can thus be given the necessary ordering and shaping character, which is particularly necessary in situations of power asymmetry.

V. Summary and a Look into the Future

The metaverse(s) are still in the making. The exact extent and the main players cannot yet be predicted with certainty. Even a uniform definition cannot yet be determined from a legal point of view. For further narrowing down, it makes sense to divide the development and potential regulation into three areas: the software, the hardware and the content. Although there will be differences between the individual metaverses, all types will be due to these three areas. Such a division is therefore of particular importance for a potential regulation of the metaverse in order to create specific norms. Although the main players cannot be specifically named, a development as well as the later operation of the metaverses by the large digital companies seems likely. As part of the above-mentioned digital strategy, the EU wants to put a stop to the negative phenomena associated with social platforms. The same is desirable for the metaverse. From the point of view of the EU legislator, but also of the national legislators, it is important to clarify whether and to what extent the existing provisions of the legal acts can be applied to the current development of metaverses and at what point a modification is necessary.

Against the background of the explanations made here, central challenges of regulating the metaverse as a virtual, immersive and interactive spatial concept created by the fusion of the physical and digital worlds are:

- Data protection and privacy: The metaverse collects and processes large amounts of personal data of its users. It is important to ensure adequate data protection and develop privacy mechanisms to prevent the misuse of data.
- Security: The metaverse opens up new opportunities for cybercrime, fraud, and identity theft. It is crucial to implement security measures to protect users and their digital identities.
- Content regulation: The metaverse will include a variety of content created by users, but under the terms of the platform owners. It is a challenge to develop adequate regulatory mechanisms to identify and combat illegal or harmful content without restricting freedom of information and expression. It will be interesting to see to what extent the experience of algorithmic *content moderation at scale* can be made fruitful by the more classic digital communication services.
- Interoperability and open standards: The metaverse will consist of different platforms and applications. It is important to promote open standards to enable interoperability between platforms and prevent monopoly formation. The EU has already taken action in other areas, such as the interoperability of messenger services.
- Economic aspects: The metaverse will create new economic models and business opportunities. It is necessary to create appropriate framework conditions to ensure fair competition, consumer protection and the protection of intellectual property.
- Digital moats: The metaverse requires a reliable internet connection and access to tech devices. A key challenge is to ensure that the benefits of the metaverse are accessible to all populations and that existing digital divides are not deepened.

References

Ball, M. (2020, January 13). *The Metaverse: What It Is, Where to Find It, and Who Will Build It*.https://web.archive.org/web/20230418133149/https://www.matthewball.vc/all/themetaverse (last accessed July 31, 2025).

Bender-Paukens, J., & Werry, M. (2023). *Datenschutz im Metaverse*. ZD, 127.

Beukelmann, L. (2012). *Surfen ohne strafrechtliche Grenzen*. NJW, 2617.

Bundesgerichtshof [BGH]. (2021, July 29). *Decision in case III ZR 179/20*. ZUM 2021, 953, pp. 957–958. https://dejure.org/dienste/vernetzung/rechtsprechung?Aktenzeichen=III+ZR+179%2F20 (last accessed July 31, 2025).

Bräutigam, T., & Rücker, A. (Hrsg.). (2017). *E-Commerce*.

Dauses, M., & Ludwigs, S. (Hrsg.). (2022). *Handbuch des EU-Wirtschaftsrechts* (57th ed.). (Cited as: Dauses/Ludwigs EU-WirtschaftsR-HdB/Editor).

Der Standard. (2021, January 22). *First case of sexual harassment in Facebook's metaverse*. Der Standard. https://www.derstandard.at/story/2000132107900/erster-fall-von-sexueller-belaestigung-in-facebooks-metaverse (last accessed July 31, 2025).

European Commission & Analysis and Research Team. (2022, March 9). *Metaverse – Virtual World, Real Challenges* (p. 3). Publications Office of the European Union. https://data.europa.eu/doi/10.2860/432862 (last accessed July 31, 2025).

Europol. (2022, October 21). *Policing in the metaverse: What law enforcement needs to know* (p. 5).https://www.europol.europa.eu/cms/sites/default/files/documents/Policing%20in%20the%20metaverse%20-%20what%20law%20enforcement%20needs%20to%20know.pdf (last accessed July 31, 2025).

Futurezone. (2021, November 25). *Company buys virtual land in Decentraland for 2.5 million dollars* [Web article]. https://futurezone.at/digital-life/grundstueck-land-decentraland-25-millionen-dollar-metaversum/401818735 (last accessed July 31, 2025).

Friehe, T. (2020). *Löschen und Sperren in sozialen Netzwerken*. NJW, 1697.

Bendel, O. (2021). *Metaverse*. In *Gabler Wirtschaftslexikon*. Springer Gabler https://wirtschaftslexikon.gabler.de/definition/metaverse-123520 (last accessed July 31, 2025).

Gielen, S., & Uphues, T. (2021). *Digital Markets Act und Digital Services Act*. EuZW, 627.

Heet, A., & Kalbhenn, J. (2021). *MMR-aktuell*, 442089.

Höfler, A., & Krolle, S. (2023, February 23). *What is behind the metaverse hype*. Handelsblatt. https://www.handelsblatt.com/technik/metaverse-was-hinter-dem-metaverse-hype-steckt/28073180.html (last accessed July 31, 2025).

Kalbhenn, A. (2021). *Metaversum und Plattformregulierung*. MMR-Aktuell, 442089.

Kapoor, D., & Klindt, T. (2022). *Aktuelle Entwicklungen im Produktsicherheits- und Produkthaftungsrecht*. NJW, 1573.

Kaulartz, M., Schmid, F., & Müller-Eising, M. (2022). *Das Metaverse – Eine juristische Einführung*. RDi, 521.

Kettemann, M. C., Schulz, W., & Fertmann, M. (2021). *Anspruch und Wirklichkeit von Plattformregulierung*. ZRP, 138.

Kühling, J. (2021). *„Fake News" und „Hate Speech" – Die Verantwortlichkeit von Medienintermediären zwischen neuem NetzDG, MStV und dem Digital Services Act*. ZUM, 461.

Krönke, C. (2020). *Öffentliches Digitalwirtschaftsrecht: Grundlagen – Herausforderungen und Konzepte – Perspektiven* (pp. 195 et seq.). Mohr Siebeck. Jus Publicum, Band 295.

Langner, D., Klindt, T., Schucht, C., Dauses, M. A., & Ludwigs, R. (2023). Technical safety regulations and standards. In M. A. Dauses & R. Ludwigs (Eds.), *Handbuch des EU-Wirtschaftsrechts* (58th ed., chap. C.VI.1, para. 1). C.H. Beck.

Meta. (2021, October 28). *Quotes from Facebook Connect 2021* [Web post]. https://about.fb.com/de/news/2021/10/zitate-von-der-facebook-connect-2021 (last accessed July 31, 2025).

Mast, E. (2023). *AGB-Recht als Regulierungsrecht*. JZ, 287.

Mast, E., Kettemann, M. C., & Schulz, W. (2024). *Private Ordering von Medienorganisationen und Plattformbetreibern*. In Puppis, M., Mansell, R., & Van den Bulck, H. (Hrsg.), *Handbook of Media and Communication Governance*.

McKinsey & Company. (2022). *Value creation in the metaverse*. https://www.mckinsey.com/capabilities/growth-marketing-and-sales/our-insights/value-creation-in-the-metaverse (last accessed July 31, 2025).

Mills, A. (2006). *The private history of international law*. International and Comparative Law Quarterly, 55, 1.

Müller, S. (2022). *Rechtsprechung zur Vermietung von virtuellem Land und Implikationen für das Metaverse am Beispiel von Decentraland*. UR, 281.

Oberlandesgericht München [OLG München]. (2018). NJW, 3115, 3116.

Paal, B. (2022). *Metaversum und Recht*. In Gräfe, S. & Telemedicus e.V. (Hrsg.), *Tagungsband der Sommerkonferenz 2022*, Telemedicus – Recht der Informationsgesellschaft (pp. 186–195).

Park, Y., & Kim, S. (2022). *A Metaverse: Taxonomy, components, applications, and open challenges*. IEEE Access, 10, 4209.

Quintais, J. P., de Gregorio, G., & Magalhães, J. (2023). *How platforms govern users' copyright-protected content: Exploring the power of private ordering and its implications*. Computer Law & Security Review, 48, 105792.

Rennert, T. (2009). *Beleihung zur Rechtsetzung*. JZ, 64, 976.

Richter, H., & Gömann, K. (2023). *Private Durchsetzung des DMA – Ein Ausblick am Beispiel von Amazons*. NZKart, 208.

Rüthers, B., Fischer, C.-W., & Birk, R. (2022). *Rechtstheorie* (12th ed.).

Schmalenbach, K. (2017). *Völker- und unionsrechtliche Impulse für die Entterritorialisierung des Rechts*. VVDStRL, 76, 245.

Schönke, H., & Schröder, A. (Hrsg.). (2019). *Strafgesetzbuch* (30th ed.).

Spindler, G. (2019). *Löschen und Sperren von Inhalten auf Grundlage von Teilnahmebedingungen sozialer Netzwerke*. CR, 238.

Spindler, G., & Seidel, A. (2022). *Die Regulierung von Online-Plattformen*. NJW, 2730.

Statista. (n.d.). *Metaverse – weltweit*. https://www-statista-com.translate.goog/?_x_tr_sl=en&_x_tr_tl=de&_x_tr_hl=de&_x_tr_pto=rq&_x_tr_hist=true (last accessed July 31, 2025).

Steege, K., & Chibanguza, J. (2023). *Metaverse Rechtshandbuch*.

Teubner, G. (2012). *Verfassungsfragmente*.
Vesting, T. (2015). *Rechtstheorie* (2nd ed.).
Weber, R. (Hrsg.). (2022). *Juristisches Wörterbuch* (24th ed.).

List of Authors

Lenn Blaschke is a media artist and research associate at the Institute for Democratic Culture at Magdeburg-Stendal University of Applied Sciences. His work investigates how simultaneity and instantaneity across digital and physical worlds challenge contemporary norms and values. Using VR and installations, he explores extended realities as spaces of political and cultural reflection.

Caroline Böck is a PhD researcher at the Institute for Theory and Future of Law at the University of Innsbruck. Her research focuses on several topics in European and international digital law, such as the Metaverse and platform regulation.

Dr. Julia Ebner is a Calleva Researcher at the University of Oxford and leads the Violent Extremism Lab at Oxford's Centre the Study of Social Cohesion. She also serves as Co-Executive Director at the Institute for Strategic Dialogue. Her research focuses on radicalisation, violent extremism and threats to democracy.

Dr. Jonas Fegert is Head of the House of Participation at FZI Research Center for Information Technology. At his research group at Karlsruhe Institute of Technology, he investigates digital democracy, focusing on information manipulation, societal polarization, and public participation.

Jessica Galissaire is Senior Policy Researcher in the "Digital Public Sphere" programme at interface (formerly Stiftung Neue Verantwortung). During her time at French think tank Renaissance Numérique, she conducted research on virtual worlds, especially their governance, regulation and impact on fundamental rights and freedoms.

Dr. Eva Groß is Professor of Criminology and Sociology and Co-Chair of the Hamburg Institute for Interdisciplinary Research on Crime and Security (HIKS). She leads research projects in the fields of discrimination, racism and police. Her main research areas include bias-motivated crime, victimization, radicalization, empirical police research, perceptions of crime and antisemitism.

Henri Isaac holds a PhD in management sciences, and is an Associate Professor, Paris-Dauphine University, PSL Research University. He directs the Master Management Telecoms & Media. He was President of the Think Tank Renaissance Numérique" (2016–2020) and a member of National Digital Council (2018–2020). He was a member of Economic Observatory of Online Platforms (2021–2023).

Sandra Kero is a research assistant and PhD candidate at the Centre for Media, Communication and Information Research (ZeMKI). Her research focuses on long-term transformation processes in the public sphere, especially affective and activist communication and communication within fringe communities. Her PhD project examines competing futures of knowledge and contemporary metapolitical struggles.

Dr. Matthias C. Kettemann is Professor of Innovation, Theory and Philosophy of Law and head of the Department of Theory and Future of Law at the University of Innsbruck. He leads research groups at the Humboldt Institute for Internet and Society in Berlin and the Leibniz Institute for Media Research (Hans-Bredow-Institut) in Hamburg, and advises governments, the private sector, and foundations on the development and application of rules in digital constellations.

Jean-François Lucas, PhD in sociology, is the Managing Director of Renaissance Numérique. He previously worked as a consulting director, supporting local authorities and private-sector stakeholders in the development of inclusive and sustainable "smart" territories. He has also spent several years designing and managing innovative research projects at the intersection of the humanities, social sciences, engineering, and the arts.

Dr. Octavia Madeira researched Virtual Reality applications in clinical psychology before expanding her research focus on technology's role in radicalization and extremism at the Karlsruhe Institute of Technology (KIT). There, she also led a project on technology assessment and gained practical experience in forensic psychology alongside her academic work. Currently, she is training as a psychotherapist.

Martin Müller is PhD researcher (pre-doc) at the University of Innsbruck, Austria. His research focuses on International and European Law, particularly in matters of digitality and cybersecurity.

Dr. Georg Plattner is a secondary school teacher of politics and English and holds a PhD in political science. Before entering the teaching profession, he conducted research at the Karlsruhe Institute of Technology (KIT) on the nexus of technology and radicalization. He has published on far-right parties and tech-related extremism.

Mick Prinz is a social scientist and works at the Amadeu Antonio Foundation. He heads the "Good Gaming Support" project, which deals with strategies against the extreme right in gaming. His research focuses on right-wing extremist propaganda games, modifications and other anti-democratic and democratic agitations in video game spaces.

Dr. Matthias Quent is Professor of Sociology and Chair of the Institute for Democratic Culture at Magdeburg-Stendal University of Applied Sciences. He leads several research projects on education, ethics, and political culture in the context of the metaverse and extended reality. He also serves as Managing Director of Soziovista GmbH, a company specialising in analysis and consulting (www.soziovista.de).

Dr. Josephine B. Schmitt is a scientific coordinator and postdoctoral scientist at Center for Advanced Internet Studies (CAIS). Her expertise lies at the intersection of media psychology, communication sciences, media education, and science and technology studies. Her research focuses among others on the content and dissemination of hate speech and political education on the internet.

Dr. Deborah Schnabel is a psychologist and has been the director of the Anne Frank Educational Center in Frankfurt/Main since 2022.

Dr. Danielle Shanley is an Assistant Professor in the Philosophy Department at Maastricht University and a senior researcher at the Brightlands Institute for Smart Societies (BISS) in Heerlen. Dani works on interdisciplinary projects, exploring the ethics and politics of technoscience. Her research focuses on the question of responsibility in technological development and the ethical and societal implications of technological change.

Dr. Anabel Taefi is Professor of Criminology at the Police Academy of Lower Saxony, and a lecturer at the University of Hamburg. Her current research focusses on the dynamics of violent interactions and on various fields of police research (e.g. racism and discrimination, organizational sociology).

Esen K. Tütüncü is a PhD researcher at the University of Barcelona and an artist exploring human-AI collaboration, virtual embodiment, and immersive storytelling. Her work blends cognitive science, XR, and creative tech to design new forms of interaction, presence, and agency across digital and physical worlds.

Arne Vogelgesang is a performing artist focusing on politics of digital representation. With the label internil, and under his own name, his work has been combining documentary material, new media, fiction and immersive theatrical settings since 2005. He is currently pursuing an artistic research doctorate on virtual embodiment and desire at the University of Music and Performing Arts Vienna.

Sara Lisa Vogl is a Virtual Reality designer, immersive media researcher and creative technologist. As co-founder of *XR Future* and *Mindprise*, she explores how immersive technologies can expand human experience and cognition. Her work focuses on virtual embodiment, altered states of consciousness and the design of immersive, multisensory experiences. With a background in design and anthropology, she bridges scientific inquiry and creative practice to investigate the transformative potential of extended realities.

Dr. Ulrike Zähringer is a Professor of Criminology and Criminal Law at the Hamburg Police University of Applied Sciences and is founding member of the Hamburg Institute for Interdisciplinary Crime and Security Research. Her research focuses on group-focused enmity in the police force as well as on violence against children.