EMRN
The European Metaverse Research Network

OCEAN
Open Council of Europe Academic Networks

Présidence de / Presidency of
MONACO

COUNCIL OF EUROPE
CONSEIL OF EUROPE

HUMAN RIGHTS IN IMMERSIVE REALITIES (XR): FREEDOM OF EXPRESSION, JUSTICE, CHILDREN'S RIGHTS AND WELL-BEING

PROCEEDINGS

18–19 May 2026, Strasbourg, France

Editors

Proceedings

Human Rights in Immersive Realities (XR): Freedom of Expression, Justice, Children's Rights and Well-Being

Council of Europe, Strasbourg, 18–19 May 2026

Conference Organisers

Freedom of Expression and CDMSI Division, Directorate General of Democracy and Human Dignity, Council of Europe
Mel Slater, EMRN and University of Barcelona

Technical Manager

Freedom of Expression and CDMSI Division, Directorate General of Democracy and Human Dignity, Council of Europe

Poster and Demos Organiser

Alex Fuentes Raventós, University of Barcelona

The papers in these proceedings have adopted the style of the journal Frontiers for convenience, though the conference has no relationship with the journal. However, it is expected that there will be a later Frontiers Research Topic based on this conference.

Preface

Immersive technologies such as virtual, augmented and mixed reality (XR) are increasingly shaping how individuals communicate, express themselves, interact with public authorities and participate in democratic life. While these technologies offer new opportunities for creativity, inclusion and participation, they also raise emerging challenges for human rights, democratic values and the rule of law.

Building on the 2023 Council of Europe- IEEE report on the Metaverse and the Feasibility Study on Freedom of Expression in Immersive Realities adopted by the Steering Committee on Media and Information Society (CDMSI), this conference is conceived as a concrete follow-up to one of the study's key conclusions: the need to continue and strengthen the Council of Europe's alerting and awareness-raising function vis-à-vis member States at a stage when immersive technologies are still evolving.

The conference also contributes to the consultation process towards a New Democratic Pact, by addressing how emerging technologies may affect democratic participation, trust in information spaces, inclusion, human dignity and the effective enjoyment of human rights.

Contents

Session 1 — Freedom of Expression in Immersive Realities

Wearable, Everyday AI+Augmented Reality: Perceptual Mediation, Harms, and Governance 9
Joseph O'Hagan, Thomas Goodge, Graham Wilson, Jan Gugenheimer, Orlaith Choo, Mark McGill

From the Metaverse to the Editable World: Rethinking Freedom of Expression in Phygital Reality 21
Tehilla Shwartz Altshuler

Freedom of Expression in Urban Spaces: A VR Platform for Inclusive Civic Participation 31
Krzysztof Nowakowski, Agnieszka Rumież, Piotr Skrzypczyński

From the Forest to the Metaverse: Immersive Experiences and Co-Creation with Indigenous Peoples 37
Florence Dravet

Exploring XR Policies in the Metaverse with an Intelligent Embodied Virtual Agent 41
Ke Li, Jeroen Prinsen, Julia Hertel, Frank Steinicke

Session 2 — Children's Rights, Safety and Protection of Vulnerable Users

Applying Child Rights to Product Design Through Co-Created Ethical Principles 49
Beckett LeClair, Nishita Yerraguntla, Salome Toliashvili

Embodied VR for Children: A Cross-disciplinary Agenda for Future Research 55
Xueni Pan, Oscar Sill, Marco Gillies, Dorothy Cowie

From Screen to Body: Rethinking Child Safety through the Concept of Embodied Harm in Immersive Environments 61
Burcin Genis-Ergün

Youth Co-Creation and Ethical Storytelling in Immersive 360° Video 71
Danai Mikelli

Session 3 — Justice, Law and Jurisdiction in Immersive Environments

Building the Lex Metaversi: A Regulated Self-Regulatory Approach to Dispute Resolution in Virtual Worlds 83
Aurelio López-Tarruella Martínez, Raül Ruiz Rodríguez

Virtual Reality to Quantify Rehabilitation of Violent Behaviour 91
Maria V. Sanchez-Vives, Tania Johnston, Nicolas Barnes

Virtual Reality and Rehabilitation: A Human Rights-Based Approach for the Treatment of Gender-Based Violence in the Penitentiary System of Catalonia 97
Nicolás Barnes Méndez

Session 4 — Health, Ethics and Well-Being in XR

Contents (continued)

Workers' Rights, Health and Safety, and Employer Responsibilities in Immersive Realities (XR) 105
Verity McIntosh

Supporting Recovery with Mental Health Teleclinics in the Metaverse 115
Yannick Prié, Ketty Steward

Empowering Oncology Patients Through XR Self-Observation and Mixed Reality Exercise 121
Alex Fuentes-Raventos, Caterina Calderon, Mel Slater

Immersive Realities (XR) Stress Management for Protection of the Human Right to Mental Health 123
Olga Maslova, Vasiliy Pyatin

Rebuilding Human Presence in Immersive Realities: Holoportation, Accessibility, and the Future of Human Connection 129
Stefano Beccaletto

Session 1

Freedom of Expression in Immersive Realities

Wearable, Everyday AI+Augmented Reality: Perceptual Mediation, Harms, and Governance

Joseph O'Hagan [1,*], **Thomas Goodge** [1], **Graham Wilson** [1], **Jan Gugenheimer** [2], **Orlaith Choo** [1], **Mark McGill** [1,*]

[1] *School of Computing Science, University of Glasgow, Scotland, UK*
[2] *TU Darmstadt, Germany*

Correspondence*:
Corresponding Authors
joseph.ohagan@glasgow.ac.uk & mark.mcgill@glasgow.ac.uk

1 INTRODUCTION

Smartphones are the predominant technology by which we communicate, engage with society and consume digital content. Approximately 88% of the EU population and 90% of children in Europe and Asia, have access to a smartphone, and these devices shape our perception, acting as an ever-present lens into digital and augmented worlds (1). But once a smartphone is in your pocket, the power this lens holds over the user is greatly diminished as they no longer perceive or interact with visual digital content but instead are limited to haptic or audio content. Furthermore, the sensing capability of the device (cameras and microphones) and any associated multimodal AI agents and models is severely restricted, e.g. allowing GPS tracking permissions means the device can locate where the user is but crucially is no longer able to perceive and capture the user's surrounding environment and reality. And though smartphone addiction is a concern (2, 3), this offers a natural limit and barrier for how this technology can mediate attention, activities, perception, and the extent to which it is aware of a user's surroundings at any given moment.

Wearable technologies, however, by design reduce and/or remove such limits, providing an everyday/all-day wearable and fashionable form factor (e.g. smart glasses (4, 5)) that circumvents the barriers imposed by the user's pocket, instead placing the device between the user's eyes/ears and their surrounding reality (6). This switch enables these devices to continually surveil the wearer's actions and environment (aurally and visually through cameras and microphone arrays), process and interpret that data (through multimodal local and cloud AI models), and act upon it in how they mediate our aural and visual perception of the real world (through visual and auditory augmented reality, heads-up smart glasses, projective displays and more), all in discreet form factors e.g. stylised as a pair of glasses (6). And such wearable technologies are seeing surging commercial interest with companies such as Meta (7) (RayBan Display), Google (8), and Samsung (9) (Android XR) releasing, or soon expected to release, AI-enabled smart glasses to consumers.

For users, this combination of continuous sensing and oversight, AI comprehension, display modalities, and form factor, promises a transformative experience and increase in technical capabilities. Such devices offer the prospect of AI-driven insights interleaved into user's everyday lives, e.g. augmenting memory, intelligence, cognition and perception, and personalization of how we perceive ourselves, others, and our surrounding world (10), empowering users, communities, business, governments, and others to alter, augment, diminish or otherwise mediate our perception of reality (11). In doing so, the expectation is the sheer utility, and eventual necessity, of this technology will force adoption, with users being no more able to opt out from using such devices in the future as they can opt out of owning smartphones today (12). But

alongside the envisioned benefits of these devices, there are clear, inherent potential harms which have yet to be addressed (4, 13). For example, if we consider those who control these devices, their underlying software/hardware platforms, and the predominant applications and ecosystems that emerge, there are clear motivations around: instrumenting individuals to better understand their likes, dislikes, behaviours and attitudes (14); aggregating this data as a form of *worldscraping*; and then ultimately leveraging this to inform or influence those very same behaviours and attitudes, both individually and collectively. Such a position of power could be utilised for persuasion and coercion, deceptive and manipulative design, information disorder, censorship, and more (15, 16). Moreover, the necessary technical capabilities to achieve this will already be present in these devices, i.e. the same technology needed to enable real-time text comprehension assistance could also be used for enacting information disorder and censorship (6, 15).

The question then is: what recommendations, regulations, laws, and rights should govern the development and use of a technology that largely does not yet exist but has such potential, given that failure to act before widespread adoption may render meaningful regulation and protections slow and difficult to enact (10). Our motivation for addressing this now is in examining how companies are pursuing this technology, e.g. with Meta, Google, Microsoft, Apple, Bytedance, and more, all investing significantly in AI, wearables and immersive technologies, anticipating their eventual convergence and the influence they will potentially exert on users, alongside the significant economic value and data collection systems to be obtained. And these companies pursue this from a regulatory perspective that is arguably not aligned with the principles of the EU. For example, companies at UnitedXR 2025 openly questioned the restrictions imposed by EU law around privacy and data processing whilst suggesting that existing measures for protecting bystanders from these devices remain sufficient (17), despite contrary scientific evidence (18) and public sentiment (19, 20).

In this agenda-setting paper, we offer insight into concerns around everyday AI-enhanced wearable augmented reality (AR) from the perspective of Human-Computer Interaction (HCI, a field of computing science that evaluates, interrogates, and anticipates the capabilities, usage, and human impact of technology), highlighting some key topics and provocative works both we, and others within HCI, have contributed around the scope of perceptual tensions and vulnerabilities/harms. In particular, we focus on less widely covered considerations around the concept of *perceptual mediation*, where the wearable device leverages visual and auditory displays and AI inputs to mediate the user's perception of reality. Informed by this, we then discuss our on-going work on what we termed *perceptual rights*, and the approach we are exploring to formalise our framework through cross/multi-disciplinary expert-led consensus around actionable recommendations, regulations and prospective rights that could inform governance of these technologies, and ensure we anticipate and mitigate the harms exposed by wearable AI/AR prior to their realisation.

2 PERCEPTUAL AND COGNITIVE VULNERABILITIES

Works in HCI have increasingly started to interrogate augmented cognition (21, 22, 23), perceptual vulnerabilities (24, 25), manipulation (26, 27, 28), inter-reality (29) and societal risks (30, 4, 31), ethics and responsible practice (32, 33, 34) and more, going beyond more grounded and near-term considerations around e.g. privacy risks posed by AR+AI (6). We draw upon these in discussing a handful of example topics as provocations to consider when addressing prospective regulation of this technology.

2.1 Social Identity and Appearance

For individuals, the capacity to augment how we, and others, are perceived (35, 36). could risk provoking a range of psychologically damaging reactions as users feel a pressure to conform their appearance to perpetuated ideals (37). And for malicious actors, this capability could enable new forms of abuse. Beyond

'identity hacks' such as identity theft (34), it is easy to envision a convergence of AR sensing and cheap/deep fake technology (38) to, for example, sexualise (39) or otherwise appropriate the identity of others for socially unacceptable reasons (e.g. blackface filters (40)). Lemley *et al.* considered the legality of this ability to augment our personal sensescape and the sensescapes of others, asking: "What if people use this... to make [you] appear ridiculous... without your knowledge or consent? Or what if they want to make you appear naked" (41). But as O'Hagan et al discussed, even when used with consent there exists the potential for sexualised use cases to facilitate unhealthy sexual relationships where partners become a "haptic proxy" and are reduced to placeholders for the fantasies re-enacted by the other person, or even be mislead as to how they are being augmented, e.g. believing they are being augmented in a way they have requested but in actuality are being portrayed as a child during sex (42).

2.2 Maladaptive Use of Augmented Cognition

The combination of AR and AI offers the prospect of real-time AI support around memory, intelligence, cognition, social interactions, and more. E.g., Meta's "Persistent Assistant" AI (43) autonomously infers about the user's goals and analyses the environment to provide AR-rendered guidance on objects of interest. Whilst in "Memoro" (44), participants interacted with an autonomous "query-less" AI assistant that provided auditory information about their current context, and preferred the passive query-less version over an approach requiring active user input. AI such as Visual Language Models (VLMs) could be used to understand our activities, environments, intent and needs and react to them autonomously - in effect providing new and greater levels of cognitive offloading than any previous technology. Given the increasing capability and automation of AI agents more than 50 authors from academia and industry came together to publish a position paper on "The Ethics of Advanced AI Assistants" (45). The paper covers a range of topics however, there is no mention of potential positive or negative effects (or ethical concerns) regarding human memory, cognition or everyday living.

In contrast, in the HCI domain, we have begun to consider what *responsible* cognitive tools and support - or cognitive integration (46) - might be. E.g., for memory augmentation, the possibility of AI providing perfect recall/support across semantic, procedural, episodic, spatial and prospective memory risks potential maladaptive use, such as over-reliance that degrades human memory and even biological brain matter (21), as well as overt manipulation and interference with how memory is encoded, experienced, and retrieved (22). Research is already showing how generative AI can imperceptibly alter a person's memory of events (47, 48) or the presentation of information online (16, 49, 50). And critically, the proper consolidation of memory requires time and repetition (51), and if technology has supportive or disruptive effects on consolidation it may occur over long time periods. Meaning by the time research has grappled with how to responsibly augment such capabilities, these harms may already have occurred to long-term users.

2.3 Pervasive Advertising and Personalization

Unlike digital or print advertising, AR ads are not confined to screens or physical objects, but can be augmented anywhere within a user's perceived environment (52). As such pervasive advertising's potential is far greater with AR as any "surface", including building facades, vehicles, people, food, and even the sky (53, 54, 55, 56), are available to be monetised. It is highly likely AR platforms will leverage this given advertising is central to their underlying platform economics (57). In practice, platforms may subsidise the cost of their technology (and associated inference costs) with AR adverts, operating under similar models as used in social media today, where applications attract interaction and engagement and then sell access to this audience to advertisers (58). This is already reflected in industry, with Google having obtained patents for "pay per gaze" and "pay per emotion" systems, adapting current "pay-per-click" models to AR (53).

HCI research has begun to explore the risks of pervasive advertising in AR. Mhaidli & Schaub identified pervasive advertising as one of five manipulative XR advertising techniques, describing a scenario where a company augments its brand onto other people's clothing leading passers-by to perceive widespread adoption of that brand (59). And while some have considered mitigation strategies, e.g. AR ad-blocker (60) or how to adapt existing regulations and consumer protection laws (53), the vast majority of work in this topic does not focus or address the potential risks and manipulations of pervasive advertising. Instead, research efforts are concentrated on understanding how AR can more deeply influence customer behaviour (61, 62), with de Ruyter *et al.* arguing there has been too much reliance on the novelty of AR advertisements and we have yet to harness the true potential for AR advertisement, i.e. one that closely aligns adverts to the user's context and spatial environment (63) by, for example, utilising real-time gaze tracking to learn about the significance of objects in the user's home (63). Such works point towards a potential convergence of contextualisation and personalisation in AR advertising, highlighting the risk of AR advertising not only being more pervasive in where it can be augmented but also more precise and targeted in its ability to influence users at the level of perception and decision making.

2.4 Deceptive and Manipulative Design

In being able to track and understand our pre-existing likes and attitudes (e.g. through biometric psychography (64)), actions (e.g. through body tracking, context awareness, etc (6)) and even intention to act (e.g. through EEG-based readiness potential (65)), AR headsets also offer the possibility for enhanced behavioural nudging (66, 67, 68, 69), deceptive design (28), manipulation of actions (70), and preference change (71, 72, 73). This directed behaviour change may be voluntary (enacted through persuasion, nudging, positive reinforcement) or involuntary (via coercion or imperceptible manipulation). E.g., Ruocco *et al.* examined how a supermarket shopper's attention could be redirected to manipulate purchasing decisions (27). Pandjaitan *et al.* through *AuctentionAR* suggested users' attention to real-world objects could become a quantifiable and manipulable property "to auction off" to the highest bidder (26). And Krauß *et al.* outlined how AR could be used to amplify existing deceptive designs, and enable new ones, leveraging perception, spatiality, physical/virtual barriers, and device sensing (28).

2.5 Information Disorder and Censorship

Significant also is the potential capacity to comprehend and then alter our perception of our environments and the media we consume that exist outside of the system in use. E.g., a real-world political advert such as a billboard or poster by one party could be rebutted, undermined, or obfuscated by another. LLMs have been shown to be capable of re-writing web content to, for instance, alter the perceived political bias of content (49), and it is reasonable to expect VLMs might enact similar alterations on physical and digital media perceived by AR users. Indeed, political parties might create their own metaversal media layers over reality for their supporters to experience (4), e.g. the *Trumpverse* (13). And, going beyond physical media, particular social groups and demographics could even be visually "othered" (74) based on personal characteristics to confirm and amplify bias, and political opponents could be visually and/or aurally "blocked" *Black Mirror* style (75) or otherwise censored. Going further, censorship and information disorder may even collide. E.g., Eghtebas *et al.* discussed using AR to 'beautify' the physical environment, and despite their positive intentions (e.g. to support digital regeneration), this could lead to the "removal of poverty from sight", a form of censorship of the real state of a given society or community (31).

2.6 Perceptual Illusions and Manipulations

Beyond the above examples, everyday AR will open the door to new perceptual attacks and targeted augmentation of *any perceivable visual or auditory element of reality*. E.g., *SwitchAR* demonstrated how passthrough AR devices could imperceptibly switch from real to digital twin environments, exposing the user to physical safety risks and harms if this approach were abused (25). Similarly, Tseng *et al.* leveraged Virtual-Physical Perceptual Manipulations (VPPMs) of physical movements to redirect VR users into taking harmful actions (e.g. walking into obstacles, hitting other users aligned with virtual elements) and we could imagine similar illusions or manipulations occurring given sufficiently capable devices, ultimately demonstrating the potential for direct mediation and transformation of human perception (24, 76). This raises fundamental challenges for user autonomy, informed consent, and safety, reinforcing the need for anticipatory safeguards before such capabilities become widely deployed.

3 TOWARDS PERCEPTUAL RIGHTS

Given the potential of the combined use of AR and AI to reshape perception, we think it is necessary to consider the concept of *perceptual rights* - as a parallel to similar efforts in other fields such as around neuro-rights (77), as a means of codifying and unifying the disparate recommendations and guidelines made around the metaverse and immersive technologies, and as a bridge to support policy makers, regulators and law makers to explore the applicability of existing rights, laws and regulations to the concerns raised. In prior work (4, 13, 33, 21), we have speculated on prospective rights to be considered around the *right to metaversal access*; *perceptual agency and autonomy* around control of the user's sensecape; *cognitive autonomy* tensioning the right to free-will and independence of thought against (non-)consensual manipulation through perceptual mediation; and *perceptual integrity*, tensioning the right to freedom of expression against the risks posed by augmentations embedded in the world and perceived en masse. But we are far from alone in making such recommendations as many efforts have been made to provide guidelines and recommendations around the governance of immersive technologies including:

Multilateral Bodies, Standards and Regulators Such as the WEF (78), ITU Focus Group on the Metaverse (79), Council of Europe & IEEE (80), the European Commission (81), IEEE Standards (82), the Metaverse Standards Forum (83) and more;

Industry Coalitions and NGOs Such as XRSI, Future of Privacy Forum (FPF), the EFF etc.;

Academia With papers typically either semi-systematically reviewing, or speculating on, prospective governance challenges and recommendations (84, 85, 86, 87, 88, 89) or reflecting more broadly on ethics, risks, and the applicability of existing laws (e.g. XR4Human[1]).

However, we see a number of issues with existing efforts thus far.

3.1 The Metaverse and the Combined Treatment of AR, MR, VR

Metaverse, Web 4.0, virtual/immersive worlds, digital twins, and more, are all terms commonly used to describe the scope and context within which extended reality (XR) is situated. These terms are varyingly used to refer to different subsets of the mixed reality/reality-virtuality continuum. This combined treatment of the range of possible XR experiences causes a lack of specificity regarding recommendations attempting to be generalisable across this range. Additionally, this encourages a fallback to considering virtual reality more often, given it is the more immediate, "tangible" consumer experience available currently (90).

[1] https://xr4human.eu/outputs/

3.2 Methodological, Evidentiary, and Disciplinary Grounding

Where recommendations are made, there is often only consensus among select contributors risking bias towards the views of them. Methodologically, the majority of the papers rely on semi-systematic reviewing of existing works, or employ speculative design approaches with limited participant pools. This produces a range of whitepapers and publications making varyingly different recommendations that lack integration and a methodological approach which calls into question the validity and comprehensiveness of the made recommendations. Moreover, the evidentiary basis for recommendations, and the positionality/disciplinary grounding that underpins them, is often unclear. This risks creating a fragmented and insufficiently robust foundation for governance, limiting the reliability of existing recommendations and highlighting the need for more systematic, transparent, and interdisciplinary approaches to informing policy.

3.3 Scope and Breadth of Risks

Recommendations are often tied to specific domains, applications, use cases, or user demographics. Security, privacy, interoperability are recurring and important themes, while child safety is of paramount importance, but there is a lack of breadth in understanding the wider societal impact of specific technological innovations. Consider criminal or malicious use; or how information disorder might be enacted in a world where wearable AI mediate our perception of the world. Absent a firm grounding and understanding of the breadth of vulnerabilities exposed, it becomes challenging to understand if recommendations are sufficient and comprehensive and have reached a point of *saturation* regarding covering prospective harms to individuals and society overall given mass adoption and subsequent exploitation.

3.4 The Need for Integration and Consensus

In our view, there is a need for more detailed discussion and integrative efforts that aim to:

(a) Anticipate and map the prospective risks, harms and vulnerabilities posed by the adoption of AR+AI;
(b) Systematically track the associated and emerging evidence around efficacy and exploitation;
(c) Map potential mitigations and protective measures;
(d) Achieve multi-disciplinary consensus on recommended mitigations (rights, laws, regulations and guidelines) and the strength of recommendation, agnostic of existing (inter)national policy.

In doing so, we would have a basis for transnational consideration of the fundamental principles that should govern this powerful interaction of technologies to assist and inform policy and law makers in-practice. Addressing perceptual harms in everyday AI-enhanced AR at a regulatory level requires us to first anticipate and map these prospective risks, based on cutting-edge scientific evidence of potential exploitation of this technology. This can only be done through multi-disciplinary discussion and collaboration in identifying protective measures and mitigations, in order to reach consensus on recommended rights-based, legal, regulatory, and guideline-driven responses, agnostic of existing national and international regulations and policy. For example, a Delphi protocol study (91, 92) to address points (*a* - mapping risks, *c* - mapping mitigations, *d* - consensus), looking specifically at everyday augmented reality, and arriving at expert consensus from an XR practitioner perspective across disciplines in the EMRN would provide the foundation to prevent these harms being realised. The intended output is a cohesive set of governance recommendations grounded in a replicable, open methodology to inform policy changes, e.g. initially by extending existing legislation to explicitly cover issues such as real-time perceptual mediation or mandate perceptibility of augmentations (e.g., subtle disclosure mechanisms), and longer term by formalising core perceptual rights as a natural extension of neuro-rights frameworks and integrate them into EU AI Act and Council of Europe initiatives.

4 ACKNOWLEDGEMENTS

This work was supported by UK Research and Innovation (UKRI) under the UK's Horizon Europe funding guarantee [EP/Z000068/1] (AUGSOC). This work was also supported by the AI Security Institute (AISI) under the WearAI project (UKRI856); and by the Research centre on Privacy, Harm Reduction and Adversarial Influence online [EP/V011189/1] (REPHRAIN) under the project Design Against Augmented Reality Crime (DAARC).

REFERENCES

1.GSMA Intelligence. The mobile economy europe 2025 (2025). Media release.

2.European Parliament. New eu rules needed to address digital addiction (2023). Press release 20231208IPR15767.

3.World Health Organization, Regional Office for Europe. Teens, screens and mental health (2024). Media release.

4.O'Hagan J, Gugenheimer J, Bonner J, Mathis F, McGill M. Augmenting People, Places & Media: The Societal Harms Posed by Everyday Augmented Reality, and the Case for Perceptual Human Rights. *Proceedings of the 22nd International Conference on Mobile and Ubiquitous Multimedia* (New York, NY, USA: Association for Computing Machinery) (2023), MUM '23, 225–235. doi:10.1145/3626705.3627782.

5.Al-Taie A, Matviienko A, O'Hagan J, Pollick F, Brewster SA. Around the World in 60 Cyclists: Evaluating Autonomous Vehicle-Cyclist Interfaces Across Cultures. *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems* (New York, NY, USA: Association for Computing Machinery) (2025), CHI '25. doi:10.1145/3706598.3713407.

6.O'Hagan J, Saeghe P, Gugenheimer J, Medeiros D, Marky K, Khamis M, et al. Privacy-Enhancing Technology and Everyday Augmented Reality: Understanding Bystanders' Varying Needs for Awareness and Consent. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* **6** (2023) 177:1–177:35. doi:10.1145/3569501.

7.Meta. en-USIntroducing Orion, Our First True Augmented Reality Glasses. *Meta Newsroom* (2024).

8.Google. enLearn more about Android XR (2025).

9.Samsung. Unlock the infinite possibilities of xr with galaxy ai (2025).

10.Steyvers M, Tejeda H, Kumar A, Belem C, Karny S, Hu X, et al. What large language models know and what people think they know. *Nature Machine Intelligence* **7** (2025) 221–231. doi:10.1038/s42256-024-00976-7.

11.Strecker J, Mayer S, Bektas K. Personalized reality: Challenges of responsible ubiquitous personalization. *Mensch und Computer 2024-Workshopband* (Gesellschaft für Informatik eV) (2024), 10–18420.

12.Hon A. Digital sight management, and the mystery of the missing amazon receipts (2020). Accessed: 2026-03-29.

13.O'Hagan J, Gugenheimer J, Mathis F, Bonner J, Jones R, McGill M. A Viewpoint on the Societal Impact of Everyday Augmented Reality, and the Need for Perceptual Human Rights. *IEEE Security and Privacy* (2024).

14.Zuboff S. Big other: Surveillance capitalism and the prospects of an information civilization. *Journal of Information Technology* **30** (2015) 75–89. doi:10.1057/jit.2015.5.

15.Wang X, Lee LH, Bermejo Fernandez C, Hui P. The Dark Side of Augmented Reality: Exploring Manipulative Designs in AR. *International Journal of Human–Computer Interaction* **40** (2024) 3449–3464. doi:10.1080/10447318.2023.2188799.

16.Krauß V, McGill M, Kosch T, Thiel YM, Schön D, Gugenheimer J. "create a fear of missing out" - chatgpt implements unsolicited deceptive designs in generated websites without warning. *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems* (New York, NY, USA: Association for Computing Machinery) (2025), CHI '25. doi:10.1145/3706598.3713083.

17.Takhshid Z. Wearable ai, bystander notice, and the question of privacy frictions. *BUL Rev.* **104** (2024) 1087.

18.Bhardwaj D, Ponticello A, Tomar S, Dabrowski A, Krombholz K. In focus, out of privacy: the wearer's perspective on the privacy dilemma of camera glasses. *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* (2024), 1–18.

19.BBC News. Smart glasses: 'I was secretly filmed and trolled article' (2026).

20.BBC News. Women filmed in secret for tiktok content – then harassed article (2026).

21.Wilson G, Goodge T, O'Hagan J, McGill M. Towards Responsible Memory Augmentation: Surveying the Intended Uses of Everyday Augmented Reality and Artificial Intelligence for Memory Processes. *Proceedings of The Augmented Humans International Conference* (Okinawa, Japan) (2026).

22.Bonnail E, Tseng WJ, Mcgill M, Lecolinet E, Huron S, Gugenheimer J. Memory Manipulations in Extended Reality. *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (New York, NY, USA: Association for Computing Machinery) (2023), CHI '23, 1–20. doi:10.1145/3544548.3580988.

23.Wilson G, Goodge T, O'Hagan J, McGill M. What Role (s) Will GenAI Play in Human Memory? Cognitive Theories for the Design of Responsible "Tools For Memory" (2026).

24.Mei Y, Wendt S, Müller F, Gugenheimer J. ShadAR: LLM-driven shader generation to transform visual perception in Augmented Reality. *2025 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)* (2025), 959–960. doi:10.1109/ISMAR-Adjunct68609.2025.00267.

25.Wombacher J, Li Z, Gugenheimer J. SwitchAR: Perceptual Manipulations in Augmented Reality. *Proceedings of the 38th Annual ACM Symposium on User Interface Software and Technology* (New York, NY, USA: Association for Computing Machinery) (2025), UIST '25, 1–17. doi:10.1145/3746059.3747595.

26.Pandjaitan A, Strecker J, Bektas K, Mayer S. AuctentionAR - Auctioning Off Visual Attention in Mixed Reality. *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems* (New York, NY, USA: Association for Computing Machinery) (2024), CHI EA '24, 1–6. doi:10.1145/3613905.3650941.

27.Ruocco M, Saeghe P, Kerber F, Gugenheimer J, McGill M, Khamis M. From Redirected Navigation to Forced Attention: Uncovering Manipulative and Deceptive Designs in Augmented Reality through Retail Shopping. *2024 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)* (2024), 720–729. doi:10.1109/ISMAR62088.2024.00087.

28.Krauß V, Saeghe P, Boden A, Khamis M, McGill M, Gugenheimer J, et al. What Makes XR Dark? Examining Emerging Dark Patterns in Augmented and Virtual Reality through Expert Co-Design. *To Appear in ACM ToCHI* (Association for Computing Machinery) (2024), ToCHI.

29.Baldry MK, Happa J, Steed A, Smith S, Glencross M. From Embodied Abuse to Mass Disruption: Generative, Inter-Reality Threats in Social, Mixed-Reality Platforms. *Digital Threats* **5** (2024) 38:1–38:36. doi:10.1145/3696015.

30.Katins C, Chen K, Lazik C, Knierim P, Marky K, Kosch T. Public Perceptions on Psychological and Social Impacts of Virtual and Augmented Reality. *Proceedings of the Mensch Und Computer 2025* (New York, NY, USA: Association for Computing Machinery) (2025), MuC '25, 163–180. doi:10.1145/3743049.3743079.

31.Eghtebas C, Klinker G, Boll S, Koelle M. Co-Speculating on Dark Scenarios and Unintended Consequences of a Ubiquitous(ly) Augmented Reality. *Proceedings of the 2023 ACM Designing Interactive Systems Conference* (New York, NY, USA: Association for Computing Machinery) (2023), DIS '23, 2392–2407. doi:10.1145/3563657.3596073.

32.Regenbrecht H, Knott A, Ferreira J, Pantidi N. To See and be Seen—Perceived Ethics and Acceptability of Pervasive Augmented Reality. *IEEE Access* **12** (2024) 32618–32636. doi:10.1109/ACCESS.2024.3366228.

33.McGill M, O'Hagan J, Goodge T, Wilson G, Khamis M, Krauß V, et al. Do We Need Responsible XR? Drawing on Responsible AI to Inform Ethical Research and Practice into XRAI / the Metaverse (2025). doi:10.48550/arXiv.2504.04440.

34.Slater M, Gonzalez-Liencres C, Haggard P, Vinkers C, Gregory-Clarke R, Jelley S, et al. The Ethics of Realism in Virtual and Augmented Reality. *Frontiers in Virtual Reality* **1** (2020) 1. doi:10.3389/frvir.2020.00001.

35.Bonner J, Mathis F, O'Hagan J, McGill M. When Filters Escape the Smartphone: Exploring Acceptance and Concerns Regarding Augmented Expression of Social Identity for Everyday AR. *ACM Symposium on Virtual Reality Software and Technology (VRST 2023)* (Christchurch, New Zealand) (2023).

36.Szittya DG, O'Hagan J, Pöhlmann KMT, Wilson G, McGill M. Exploring the Social Challenges and Opportunities of Group Use of Augmented Reality Through Speculative Design Workshops. *Proceedings of the ACM International Conference on Interactive Media Experiences (IMX '26)* (ACM) (2026), 1–11. doi:10.1145/3788851.3815029.

37.Barker J. Making-up on mobile: The pretty filters and ugly implications of Snapchat. *Fashion, Style & Popular Culture* **7** (2020) 207–221. doi:10/gh2sn2.

38.Chesney B, Citron D. Deep fakes: A looming challenge for privacy, democracy, and national security. *Calif. L. Rev.* **107** (2019) 1753.

39.Keats Citron D. Sexual privacy. *Yale LJ* **128** (2018) 1870.

40.Joshi S. It's 2021, but People Are Still Using a Racist Blackface Filter on Instagram. *Vice* (2021).

41.Lemley MA, Volokh E. Law, virtual reality, and augmented reality **166** (2017) 1051.

42.O'Hagan J, Williamson JR, McGill M, Khamis M. Safety, Power Imbalances, Ethics and Proxy Sex: Surveying In-The-Wild Interactions Between VR Users and Bystanders. *2021 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)* (2021), 211–220. doi:10.1109/ISMAR52148.2021.00036.

43.Cho H, Fashimpaur J, Sendhilnathan N, Browder J, Lindlbauer D, Jonker TR, et al. Persistent assistant: Seamless everyday ai interactions via intent grounding and multimodal feedback. *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems* (New York, NY, USA: Association for Computing Machinery) (2025), CHI '25. doi:10.1145/3706598.3714317.

44.Zulfikar WD, Chan S, Maes P. Memoro: Using large language models to realize a concise interface for real-time memory augmentation. *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* (New York, NY, USA: Association for Computing Machinery) (2024), CHI '24. doi:10.1145/3613904.3642450.

45.Gabriel I, Manzini A, Keeling G, Hendricks LA, Rieser V, Iqbal H, et al. The ethics of advanced ai assistants (2024).

46.Heersmink R, Sutton J. EnglishCognition and the web: extended, transactive, or scaffolded? *Erkenntnis* **85** (2020) 139–164. doi:10.1007/s10670-018-0022-8.

47.Pataranutaporn P, Archiwaranguprok C, Chan SWT, Loftus E, Maes P. Slip through the chat: Subtle injection of false information in llm chatbot conversations increases false memory formation. *Proceedings of the 30th International Conference on Intelligent User Interfaces* (New York, NY, USA: Association for Computing Machinery) (2025), IUI '25, 1297–1313. doi:10.1145/3708359.3712112.

48.Pataranutaporn P, Archiwaranguprok C, Chan SWT, Loftus E, Maes P. Synthetic human memories: Ai-edited images and videos can implant false memories and distort recollection. *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems* (New York, NY, USA: Association for Computing Machinery) (2025), CHI '25. doi:10.1145/3706598.3713697.

49.O'Hagan J, Wilson G, McGill M. WebPerceptor: An Open Source Chromium Plugin for Real-Time LLM-Based In-Line, In-Browser Re-Writing of Website Content. *Extended Abstracts of the 2026 CHI Conference on Human Factors in Computing Systems (CHI EA '26)* (Association for Computing Machinery) (2026). doi:10.1145/3772363.3798730.

50.O'Hagan J. Evaluating Personalised Generative User Interfaces on an AI Mediated Web (2026).

51.Shettleworth SJ. *Cognition, Evolution, and Behavior (2nd Ed)* (Oxford University Press) (2009).

52.Rauschnabel PA, Babin BJ, Tom Dieck MC, Krey N, Jung T. enWhat is augmented reality marketing? Its definition, complexity, and future. *Journal of Business Research* **142** (2022) 1140–1150. doi:10.1016/j.jbusres.2021.12.084.

53.Wassom BD. Chapter 4 - Advertising, Marketing, and eCommerce. Wassom BD, editor, *Augmented Reality Law, Privacy, and Ethics* (Boston: Syngress) (2015), 71–100. doi:10.1016/B978-0-12-800208-7.00004-1.

54.Reynolds M. Augmented reality graffiti will lead to advertising ambush wars. *New Scientist* (2017).

55.O'Hagan J, Wilson G, Goodge T, Drago D, McGill M. The BrutARlist: Exploring How and Why Individuals Alter a Known Brutalist Building with Augmented Reality. *Proceedings of the 2026 International Conference on Advanced Visual Interfaces* (New York, NY, USA: Association for Computing Machinery) (2026), AVI '26. doi:10.1145/3811427.3811475.

56.Joy A, O'Hagan J. Acceptance of an Augmented Society: Initial Explorations into the Acceptability of Augmenting Real World Locations (2025).

57.Heller B, Bar-Zeev A. enThe Problems with Immersive Advertising: In AR/VR, Nobody Knows You Are an Ad. *Journal of Online Trust and Safety* **1** (2021). doi:10.54501/jots.v1i1.21.

58.Gudowsky N, Kowalski J, Bork-Hüffer T. enAugmented futures? Scenarios and implications of augmented reality use in public spaces. *Futures* **151** (2023) 103193. doi:10.1016/j.futures.2023.103193.

59.Mhaidli AH, Schaub F. Identifying Manipulative Advertising Techniques in XR Through Scenario Construction. *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (New York, NY, USA: Association for Computing Machinery) (2021), CHI '21, 1–18. doi:10.1145/3411764.3445253.

60.Katins C, Strecker J, Hinrichs J, Knierim P, Pfleging B, Kosch T. Ad-Blocked Reality: Evaluating User Perceptions of Content Blocking Concepts Using Extended Reality. *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems* (New York, NY, USA: Association for Computing Machinery) (2025), CHI '25, 1–18. doi:10.1145/3706598.3713230.

61.Sung EC. The effects of augmented reality mobile app advertising: Viral marketing via shared social experience. *Journal of Business Research* **122** (2021) 75–87. doi:10.1016/j.jbusres.2020.08.034.

62.Scholz J, Duffy K. We ARe at home: How augmented reality reshapes mobile marketing and consumer-brand relationships. *Journal of Retailing and Consumer Services* **44** (2018) 11–23. doi:10.1016/j.jretconser.2018.05.004.

63.de Ruyter K, H Jonas, H Tim, C Matthew, K Debbie I, , et al. Seeing with the Customer's Eye: Exploring the Challenges and Opportunities of AR Advertising. *Journal of Advertising* **49** (2020) 109–124. doi:10.1080/00913367.2020.1740123. _eprint: https://doi.org/10.1080/00913367.2020.1740123.

64.Heller B. Watching Androids Dream of Electric Sheep: Immersive Technology, Biometric Psychography, and the Law. *Vanderbilt Journal of Entertainment & Technology Law* **23** (2020) 1.

65.Schurger A, Hu PB, Pak J, Roskies AL. What Is the Readiness Potential? *Trends in Cognitive Sciences* **25** (2021) 558–570. doi:10.1016/j.tics.2021.04.001.

66.Hummel D, Maedche A. How effective is nudging? A quantitative review on the effect sizes and limits of empirical nudging studies. *Journal of Behavioral and Experimental Economics* **80** (2019) 47–58. doi:10.1016/j.socec.2019.03.005.

67.Schmidt AT, Engelen B. The ethics of nudging: An overview. *Philosophy Compass* **15** (2020) e12658. doi:10.1111/phc3.12658.

68.Cho H, Edgar D, Lindlbauer D, O'Hagan J. Evaluating dynamic delivery of audio+visual message notifications in xr. *2025 IEEE Conference Virtual Reality and 3D User Interfaces (VR)* (2025), 277–287. doi:10.1109/VR59515.2025.00052.

69.Henderson C, O'Hagan J, Goodge T, McGill M. MovieMetaViewer: Extending TV Movie Browsing Interfaces to Visualise and Filter Based on Hidden or Unlinked Metadata using Augmented Reality Overlays. *Proceedings of the ACM International Conference on Interactive Media Experiences (IMX '26)* (ACM) (2026), 1–7. doi:10.1145/3788851.3815025.

70.Tseng WJ, Bonnail E, McGill M, Khamis M, Lecolinet E, Huron S, et al. The Dark Side of Perceptual Manipulations in Virtual Reality. *CHI Conference on Human Factors in Computing Systems* (New York, NY, USA: Association for Computing Machinery) (2022), CHI '22, 1–15. doi:10.1145/3491102.3517728.

71.[Dataset] Franklin M, Ashton H, Gorman R, Armstrong S. Recognising the importance of preference change: A call for a coordinated multidisciplinary research effort in the age of AI (2022). doi:10.48550/arXiv.2203.10525.

72.Drago D, O'Hagan J, Kawsar F, Brewster SA. Exploring avatar seat allocation strategies in social ar. *Proceedings of the Extended Abstracts of the CHI Conference on Human Factors in Computing Systems* (New York, NY, USA: Association for Computing Machinery) (2025), CHI EA '25. doi:10.1145/3706599.3719867.

73.O'Hagan J, Medeiros D, Wilson G, McDermid RJ, McGill M. Improving ergonomic viewing of spatial xr workspaces through 2d rotational assistance. *Proceedings of the Extended Abstracts of the CHI Conference on Human Factors in Computing Systems* (New York, NY, USA: Association for Computing Machinery) (2025), CHI EA '25. doi:10.1145/3706599.3719920.

74.Jensen SQ. Othering, identity formation and agency. *Qualitative Studies* **2** (2011) 63–78. doi:10.7146/qs.v2i2.5510.

75.Tibbetts C. White Christmas. *Black Mirror* (2014).

76.Eghtebas C, Kiss F, Koelle M, Woźniak P. Advantage and Misuse of Vision Augmentation – Exploring User Perceptions and Attitudes using a Zoom Prototype. *Augmented Humans Conference 2021* (New York, NY, USA: Association for Computing Machinery) (2021), AHs'21, 77–85. doi:10/gmk86w.

77.Yuste R, Genser J, Herrmann S. It's Time for Neuro-Rights. *Horizons: Journal of International Relations and Sustainable Development* (2021) 154–165.

78.World Economic Forum. Shared Commitments in a Blended Reality:Advancing Governance in the Future Internet. Tech. rep. (2024).
79.ITU Focus Group on metaverse (FG-MV). *ITU* (2024).
80.Conseil de l'Europe, IEEE-CPMT, editors. *The Metaverse and Its Impact on Human Rights, the Rule of Law, and Democracy* (Strasbourg: Council of Europe and IEEE) (2024).
81.Web 4.0 and virtual worlds (RP 2025) | Interoperable Europe Portal (2025).
82.IEEE Standards for the Metaverse. *IEEE Standards Association* (2023).
83.The Metaverse Standards Forum. *Metaverse Standards Forum* (2022).
84.Chen Z, Wu J, Gan W, Qi Z. Metaverse Security and Privacy: An Overview (2022). doi:10.48550/arXiv.2211.14948.
85.Tukur M, Schneider J, Househ M, Dokoro AH, Ismail UI, Dawaki M, et al. The Metaverse digital environments: A scoping review of the techniques, technologies, and applications. *Journal of King Saud University - Computer and Information Sciences* **36** (2024) 101967. doi:10.1016/j.jksuci.2024.101967.
86.Rosenblat MO. NYU Stern Center for Business & Human RightsReality Check: How to Protect Human Rights in the 3D Immersive Web (2025).
87.Tu X, de Castro e Silva B. Are We Ready for the Metaverse? Implications, Legal Landscape, and Recommendations for Responsible Development. *Digital Society* **4** (2025) 9. doi:10.1007/s44206-025-00163-0.
88.Hine E, Rezende IN, Roberts H, Wong D, Taddeo M, Floridi L. Safety and Privacy in Immersive Extended Reality: An Analysis and Policy Recommendations. *Digital Society* **3** (2024) 33. doi:10.1007/s44206-024-00114-1.
89.Yang L. Recommendations for metaverse governance based on technical standards. *Humanities and Social Sciences Communications* **10** (2023) 253. doi:10.1057/s41599-023-01750-7.
90.Pöhlmann KMT, Wilson G, Bajorunaite L, McGill M, O'Hagan J. "Everyone Knows You're Watching P*rn": Reflecting on First-Hand User Accounts on the Use and Public Perception of in-Transit Passenger XR. *2025 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)* (2025), 560–570. doi:10.1109/ISMAR67309.2025.00066.
91.Tetzlaff JM, Moher D, Chan AW. Developing a guideline for clinical trial protocol content: Delphi consensus survey. *Trials* **13** (2012) 176.
92.Slade SC, Dionne CE, Underwood M, Buchbinder R. Standardised method for reporting exercise programmes: protocol for a modified delphi study. *BMJ open* **4** (2014) e006682.

From the Metaverse to the Editable World: Rethinking Freedom of Expression in Phygital Reality

Tehilla Shwartz Altshuler[1,2*]

[1] Israel Democracy Institute, Jerusalem, Israel

[2] Jio Institute, Mumbai, India

***Correspondence:**
Corresponding Author
tehilla@idi.org.il

Keywords: metaverse, phygital reality, editable world, freedom of expression, perceptual mediation, smart glasses

Abstract

The legal and normative discourse surrounding immersive technologies has largely been shaped by the metaverse: a vision of persistent virtual environments in which individuals enter digitally constructed worlds. This literature has generated important discussions regarding freedom of expression, governance, moderation, and democratic participation in immersive digital spaces. Yet emerging phygital technologies, particularly AI-enabled smart glasses and perceptually integrated wearable systems, introduce a different challenge. Rather than relocating users into virtual environments, these technologies embed computational mediation directly into physical reality, transforming how individuals perceive, interpret, and navigate the world around them.

This paper argues that existing legal frameworks, including the extension of Article 10 protections into immersive digital spaces, do not adequately capture this transition. While current scholarship on smart glasses has focused primarily on privacy, surveillance, data governance, and consent, it has not systematically addressed the implications of computationally mediated perception for freedom of expression and democratic life.

To address this normative gap, the paper distinguishes between phygital reality as an emerging socio-technical condition and the editable world as its normative interpretation. In environments where physical spaces, human bodies, and everyday experience become subject to real-time annotation, filtering, and algorithmic personalization, expression extends beyond speech and symbolic communication into the organization of perception itself.

In response, the paper proposes an emerging framework of Reality Rights: a rights architecture spanning relational privacy, representational integrity, perceptual reliability, and shared reality. Together, these rights seek to protect not only expressive participation, but the perceptual and democratic conditions that make meaningful expression possible.

1 Introduction: the end of the metaverse as we know it

For nearly seventy years, the screen functioned as the boundary between the physical and the digital world. We watched, browsed, scrolled, and interacted, but always through an interface that separated us from reality. The metaverse vision sought to dissolve that boundary by inviting users to step inside the screen: immersive environments, virtual rooms, and persistent digital worlds in which social, economic, and political life could unfold.

Recent developments suggest a different trajectory. Meta's strategic shift away from its metaverse vision and toward AI-driven systems, alongside parallel developments in spatial computing and wearable devices, signals a shift from users entering virtual environments to digital infrastructures extending outward into the physical world (1,2,3).

This shift marks the transition from immersive virtual worlds to what may be termed "phygital reality": environments in which digital layers are embedded directly into the physical environment: streets, buildings, transportation hubs, and everyday spaces, accompanying users through daily life. Rather than replacing the physical, these technologies overlay, annotate, and reinterpret it.

2 Smart glasses as a new perceptual infrastructure

Smart glasses represent a contemporary instantiation of a broader trajectory toward wearable, embedded, and perceptually integrated computing. They intensify this trajectory by creating a particularly intimate mode of perceptual mediation (4).

This paper builds on an earlier conceptualization of such systems as operating through three interconnected layers of sensing and interpretation: outward-facing environmental capture and augmentation, networked contextualization, and inward-facing biometric sensing (5).

First, outward-facing systems align with human vision, capturing and processing the physical environment in real time. They recognize objects, read text, map spaces, and detect patterns beyond human thresholds. Yet they do not merely interpret the world - they augment it. Digital overlays can highlight, annotate, filter, or replace elements of the physical environment. For example, a storefront may appear brighter than its competitors, a sign may be instantly translated, or a person may be tagged with contextual information. Perception becomes a dynamic and modifiable process (6).

Second, these devices are connected to vast networked infrastructures. What is seen is immediately contextualized through access to the internet and AI models. A building becomes searchable, a sign becomes translatable, and a face becomes a node within a network of information. Perception merges with computation.

Third, smart glasses increasingly turn inward. Through biometric and physiological sensing, they detect cognitive load, emotional states, attention, and stress. In this mode, perception extends not only outward but inward, transforming the user's body into a continuous stream of data.

3 The metaverse precedent: extending rights into new environments

The legal and academic discourse surrounding the metaverse offers an important precedent. The literature on the metaverse consistently framed it as a qualitatively new socio-technical environment: immersive, persistent, and interconnected, requiring a reconsideration of existing legal frameworks

(7,8). Within this paradigm, key questions concerned how rights should be extended and adapted to virtual environments. Scholars and policymakers addressed issues such as embodied expression through avatars (9), behavioral moderation, and the governance of digital spaces. The central framework focused on balancing expression and safety, often through the extension of existing doctrines such as Article 10 of the European Convention on Human Rights (10,11). At the same time, this literature identified broader democratic risks: concentration of platform power, algorithmic curation, disinformation, and the fragmentation of public discourse (12,13,14). Importantly, however, these risks were conceptualized as occurring within discrete virtual environments, analogous, albeit more immersive, to social media platforms.

4 From virtual worlds to phygital reality: a normative gap

The shift to smart glasses requires a similar normative effort, but under fundamentally different conditions. Unlike the metaverse, which assumed a transition into virtual space, phygital technologies operate directly within the physical world. They do not replace reality; they reconfigure it.

Yet the emerging literature on smart glasses has not fully made this transition. A dominant strand focuses on privacy, data collection, and information governance (15). Smart glasses are treated as sensor-rich devices that enable continuous capture of audio, video, location, and biometric data, raising concerns about surveillance, profiling, and inference. The central question becomes the legitimacy of data flows—what is collected, who accesses it, and under what conditions (16).

More advanced accounts extend this analysis to the breakdown of consent. Concepts such as "ambient consent collapse" describe the erosion of traditional mechanisms of awareness and agreement in public space, driven by real-time identification, data enrichment, and asymmetries between observer and observed (17). Empirical studies further highlight concerns about bystander privacy, lack of control, and inadequate signaling (16).

Alongside this, a smaller body of philosophical literature recognizes that smart glasses mediate perception and reshape human-world relations (18). Drawing on phenomenology and post-phenomenology, it characterizes these technologies as producing a mediated version of reality and transforming the shared "life-world." (19).

However, these strands remain largely disconnected. The first focuses on surveillance and data; the second on experience and mediation. Neither systematically addresses how these transformations relate to freedom of expression or to the conditions under which meaning is constructed and shared (20,21).

The shift discussed here should not be understood as the emergence of entirely new technological forms. Augmented reality, mixed reality, digital twin models, and broader traditions of ambient, pervasive, and ubiquitous computing have long explored the integration of computational systems into physical environments and everyday life (22, 23). Earlier scholarship also examined how such technologies mediate perception, interaction, and the experience of space. The contribution of this paper lies elsewhere: not in claiming technological novelty, but in identifying a specific normative shift in which layered computational mediation becomes embedded in ordinary public life in ways that require a reconsideration of freedom of expression.

5 The editable world: a new normative category

This paper distinguishes between two related but analytically distinct concepts. Phygital reality describes the socio-technical condition in which digital layers are embedded into ordinary physical environments and encountered through perceptually integrated technologies. The editable world, by contrast, captures the normative implications of that condition: a way of understanding how computational mediation transforms the legal and democratic conditions of perception, representation, and expression.

A “phygital space” can thus be understood as a physical environment overlaid with digital information, visible only to those equipped with appropriate devices (5). The digital layer is superimposed onto the physical world, transforming streets, classrooms, and public institutions into sites of annotation, curation, and intervention. In such environments, perception itself becomes computational, and the distinction between seeing, sensing, and interpreting begins to collapse (4).

In phygital environments, physical environments become sites of annotation, curation, and potential manipulation. Three dimensions illustrate this transformation. First, objects become editable: anything within the field of vision can be modified, annotated, hidden, or replaced—including people, who may be tagged, classified, visually altered, or even erased from view through forms of “social blocking.”

Second, places become editable: streets, buildings, and public spaces can be reconfigured and reorganized based on individual profiles. A grey urban space may appear green and vibrant; a historical site may be reframed to align with a particular narrative; commercial environments may be selectively highlighted or obscured.

Third, the flow of reality itself becomes configurable: navigation cues, alerts, and recommendations dynamically shape attention, movement, and decision-making in real time. Together, these shifts mark a transition from representation to intervention: reality is actively constructed rather than passively perceived.

This paper conceptualizes this transformation through the notion of "the editable world": a condition in which physical environments, public spaces, and human bodies become subject to real-time digital annotation, filtering, and personalization. All of this unfolds within the physical world itself, where there remains a social and cognitive expectation of a shared, fact-based reality. Under conditions of phygital reality, the conceptual boundaries of freedom of expression must be reconsidered. Traditionally, freedom of expression has been understood as protecting speech, content, and symbolic acts—forms of communication that can be produced, transmitted, and received. However, in phygital environments, expression increasingly operates through the organization of perception itself.

The ability to highlight, suppress, reorder, or transform elements of reality becomes an expressive act. What is made visible or invisible, what is foregrounded or marginalized, and how individuals and spaces are represented are shaped by design decisions, algorithmic processes, and platform governance. Expression thus extends beyond the production of messages to include the construction of the perceptual conditions under which meaning is generated.

This shift also introduces a new asymmetry. Individuals encounter versions of reality tailored to their data, preferences, or inferred states. As a result, expression is differentiated not only at the level of content, but at the level of experience itself.

6 Reality rights: a new rights architecture for phygital environments

This transformation gives rise to a broader rights architecture for perception-mediated environments, grounded in—but extending beyond—the traditional framework of Article 10. While this paper is anchored in freedom of expression, the emerging normative framework cuts across adjacent human rights domains, including privacy, dignity, autonomy, democratic participation, and interpersonal integrity. This reflects the fact that phygital environments do not simply affect speech as a discrete act; they reshape the broader conditions under which expression, interpretation, visibility, interpersonal relations, and participation become possible. These emerging claims may be understood as a developing framework of Reality Rights: a bundle of rights concerned with the conditions under which reality itself is encountered, interpreted, negotiated, and shared.

6.1 The right to relational privacy

The first layer concerns interpersonal integrity: the normative conditions under which individuals become subjects of technologically enhanced observation and interpretation in public space. Traditional privacy frameworks were largely developed around vertical surveillance: observation by states, platforms, or institutional actors. Phygital environments introduce a distinct condition of interpersonal co-veillance (24) in which individuals may continuously observe, identify, classify, infer, record, and computationally augment one another through wearable systems.

The challenge here is not observation alone, but intrusiveness. Technologies capable of revealing hidden information, inferring identity or emotional state, attaching contextual data, or visually transforming another person create forms of interpersonal exposure far more invasive than ordinary public visibility.

This suggests the emergence of a right to relational privacy: a claim that individuals should not become passive objects of technologically enhanced interpersonal scrutiny without meaningful normative constraints.

This is not a claim that all observation in public space requires explicit consent, but rather that the asymmetries and intrusions created by perception-enhancing systems require rethinking privacy as a relational rather than purely institutional right.

6.2 Representational integrity

The second layer concerns individual dignity and the integrity of appearance. Phygital systems make it possible not merely to observe individuals, but to alter how they appear within others' mediated perceptual environments. A person may be visually modified, tagged, classified, filtered, sexualized, erased, or contextually reinterpreted in real time.

This gives rise to a right to representational integrity: a dignity-based claim against becoming a manipulable object within computationally mediated systems of perception. Unlike traditional privacy rights, which focus on secrecy or information control, representational integrity concerns the preservation of the person as a subject of appearance rather than an editable visual artifact.

6.3 The right to perceptual reliability

The third layer concerns cognitive and epistemic autonomy in relation to the physical environment itself. If representational integrity addresses the editing of persons, this layer concerns the editing of the world.

Phygital systems make it possible not only to augment perception, but to selectively transform the environments through which individuals orient themselves. Explanatory text in public spaces may be altered or translated selectively; contextual information may be hidden, amplified, or replaced; prices, directions, symbols, or historical markers may be computationally reinterpreted in real time. The issue is not merely personalization, but the instability of reality as an object of orientation.

In such environments, individuals increasingly rely on algorithmically mediated perception in navigating physical reality. If these perceptual conditions are selectively filtered, manipulated, or computationally distorted, autonomy itself becomes unstable. Meaningful decision-making presupposes sufficiently reliable conditions of orientation.

This suggests a right to perceptual reliability: the ability to understand when one's perceptual environment is being shaped, to retain meaningful awareness of such mediation, and to preserve cognitive agency in relation to reality-shaping systems. This extends existing commitments to autonomy into environments where perception itself becomes computationally structured.

6.4 The right to shared reality

The fourth layer concerns collective democratic integrity. Freedom of expression has traditionally protected the ability to speak, receive information, and participate in public discourse. Yet these protections presuppose a minimal shared communicative environment: a sufficiently common world in which individuals can encounter one another, contest facts, and deliberate over shared concerns. In Habermasian terms, democratic public discourse depends not only on formal rights of participation, but on the existence of a shared communicative space within which public reasoning can meaningfully occur.

Phygital environments destabilize this condition. Individuals may inhabit radically different perceptual versions of the same physical environment, such that physical co-presence no longer guarantees shared experience. Public space itself may become fragmented into personalized realities, undermining the common experiential reference points upon which democratic deliberation depends.

This suggests a right to shared reality: not a claim to uniform perception, but to preserving sufficiently common conditions of visibility, information, and appearance to sustain democratic participation. If freedom of expression protects communicative participation, this right protects the perceptual infrastructure that makes such participation possible.

7 Conclusion: from speech governance to perception governance

The transition from immersive virtual environments to phygital, perception-mediated realities marks a shift in the object of legal and normative concern. The central challenge is no longer confined to governing speech within platforms or virtual spaces but extends to the infrastructures that shape how reality itself is encountered, interpreted, and shared.

This paper has argued that the metaverse literature offers an important precedent, but not a sufficient framework for this transition. While earlier debates focused on expression within immersive digital environments, phygital reality raises a different challenge: the governance of perception in physical space, where digital mediation increasingly structures interpersonal relations, cognitive orientation, and democratic participation.

If phygital reality names the emerging socio-technical condition, the editable world captures its normative implications. In response, this paper proposes an emerging framework of Reality Rights: a rights architecture spanning relational privacy, representational integrity, perceptual reliability, and shared reality. Together, these rights shift the focus from expression as discrete communicative content toward the broader conditions that make meaningful expression possible.

This move has important governance implications. Future work should examine how these emerging claims translate into institutional design, platform governance, interoperability standards, public infrastructure, and democratic accountability. If freedom of expression has historically protected the communicative public sphere, the next challenge may be protecting the perceptual conditions upon which that sphere depends.

References

1. Murti L. Meta is shutting down VR social platform Horizon Worlds in further pivot away from the metaverse. *CNBC*. (2026, March 18). Available online at: https://www.cnbc.com/2026/03/18/meta-horizon-worlds-metaverse-vr.html

2. Roth E. Apple is reportedly planning to launch AI-powered glasses, a pendant, and AirPods. *The Verge*. (2026, February 17). Available online at: https://www.theverge.com/tech/880293/apple-ai-hardware-smart-glasses-pin-airpods

3. Snow J. Meta is killing off the metaverse. It lost $80 billion. *Quartz*. (2026, March 19). Available online at: https://qz.com/meta-metaverse-mark-zuckerberg-horizon-worlds-ai

4. Zuidhof N, Ben Allouch S, Peters O, Verbeek PP. Defining smart glasses: A rapid review of state-of-the-art perspectives and future challenges from a social sciences perspective. *Augmented Human Research*. (2021) 6:15. doi: 10.1007/s41133-021-00053-3

5. Shwartz Altshuler T, Aridor Hershkovitz R, Mikulinsky R, Müller B. Governing phygital spaces: Human rights by design meets speculative design. *Internet Policy Review*. (2025) 14(4). doi: 10.14763/2025.4.2048

6. Huang K, Isaac M, Hill K. Meta plans to add facial recognition technology to its smart glasses. *The New York Times*. (2026, February 13). Available online at: https://www.nytimes.com/2026/02/13/technology/meta-facial-recognition-smart-glasses.html

7. Council of Europe, IEEE Standards Association. *The metaverse and its impact on human rights, the rule of law and democracy*. DG1 (2023)06. Strasbourg: Council of Europe (2024).

8. Munawar M. Ethical considerations and privacy in metaverse. In: Gupta B, editor. *Unveiling Social Dynamics and Community Interaction in the Metaverse*. Hershey, PA: Information Science Reference (2025). p. 159–182. doi: 10.4018/979-8-3693-8628-6.ch008

9. Ebers M. Avatars and the protection of digital identities in the metaverse. *MetaverseUA Research Paper*. (2024) 3. Available online at: https://metaversechair.ua.es/working-papers/

10. Council of Europe (2025) Immersive realities: Benefits and challenges to freedom of expression. Feasibility study adopted by the Steering Committee on Media and Information Society (CDMSI).

11. Ex L, Hijstek B, van Huijstee M. Human rights risks from immersive technologies. *Business and Human Rights Journal*. (2025) 10(2–3):440–447. doi: 10.1017/bhj.2025.10020

12. Lucas JF, Galissaire J, Isaac H. Navigating democratic challenges in the age of metaverses. In: *Democratic Culture in the Metaverse: Interdisciplinary Perspectives on the Potentials and Perils*. (2025). p. 71–82. doi: 10.5771/9783748948117-71

13. Tu X, de Castro e Silva B. Are we ready for the metaverse? Implications, legal landscape, and recommendations for responsible development. *Digital Society*. (2025) 4:9. doi: 10.1007/s44206-025-00163-0

14. Vidaki AN. The metaverse as an AI-powered society of control: Ethical concerns on autonomy, self-determination and surveillance. In: Slater M, editor. *A Metaverse for the Good: Exploring the Intersection of AI and the Metaverse*. Proceedings of the 2nd International Congress. Zenodo (2025). doi: 10.5281/zenodo.15847422

15. Gallardo A, Choy C, Juneja J, Bozkir E, Cobb C, Bauer L, Cranor L. Speculative privacy concerns about AR glasses data collection. *Proceedings on Privacy Enhancing Technologies*. (2023) 2023(4):416–435. doi: 10.56553/popets-2023-0117

16. Chock TM, Kim SJ, Lee Y. Privacy concerns about AI smart glasses: A cross-cultural study of the U.S. and South Korea. (2026). Available online at: https://hdl.handle.net/10125/111729

17. Shukanayev D. Permissioned reality: Smart glasses, real-time identification, and the collapse of ambient consent. *SSRN Working Paper*. (2026). Available online at: https://ssrn.com/abstract=6100683

18. Iqbal MZ, Campbell AG. Adopting smart glasses responsibly: Potential benefits, ethical, and privacy concerns with Ray-Ban Stories. *AI Ethics*. (2023) 3:325–327. doi: 10.1007/s43681-022-00155-7

19. Tretter M, Hahn M, Dabrock P. Towards a smart glasses society? Ethical perspectives on extended realities and augmenting technologies. *Front Virtual Real*. (2024) 5:1404890. doi: 10.3389/frvir.2024.1404890

20. Council of Europe. *Immersive realities: Benefits and challenges to freedom of expression: Feasibility study adopted by the Steering Committee on Media and Information Society (CDMSI)*. Strasbourg: Council of Europe (2025).

21. Fornasier MdeO. Freedom of expression and the metaverse: On the importance of content creation for the emergence of a complex environment. *Revista de Investigações Constitucionais*. (2023). doi: 10.5380/rinc.v10i1.87584

22. Weiser M. Ubiquitous computing. *Computer*. (1993) 26(10):71–72. doi: 10.1109/2.237456

23. Benford S, Giannachi G. *Performing Mixed Reality*. Cambridge, MA: MIT Press (2011).

24. Mann S, Nolan J, Wellman B. Sousveillance: Inventing and using wearable computing devices for data collection in surveillance environments. *Surveillance & Society*. (2003) 1(3):331–355.

Freedom of Expression in Urban Spaces: A VR Platform for Inclusive Civic Participation

Krzysztof Nowakowski[1], **Agnieszka Rumież**[2], **Piotr Skrzypczyński** [1,*]

Poznań University of Technology, [1] *Institute of Robotics and Machine Intelligence,*
[2] *Institute of Architecture and Heritage Protection, Poznań, Poland*

Correspondence*:
piotr.skrzypczynski@put.poznan.pl

ABSTRACT

Immersive technologies offer new possibilities for civic participation, yet urban design processes remain difficult to access for non-expert users. This extended abstract presents EcoUrbanComposer VR, an immersive tool that enables users to explore and directly modify realistic urban environments through intuitive interaction. Participants can place, remove, and adjust spatial elements such as vegetation or street furniture, expressing their preferences through action rather than formal design language. This interaction model supports more inclusive participation and can be interpreted as a form of spatial expression. The system is part of a broader platform integrating data-driven modelling and web-based access for diverse stakeholders. The extended abstract outlines the interaction concept and system architecture, and presents a live demonstration scenario. The work contributes to discussions on XR and freedom of expression by providing a concrete, demonstrable use case.

Keywords: virtual reality, urban design, participatory planning, freedom of expression, civic engagement, human-computer interaction

1 INTRODUCTION AND MOTIVATION

Immersive technologies, including virtual, augmented, and mixed reality (XR), are increasingly influencing how people communicate, interact, and participate in society. In this paper, XR is primarily addressed through VR-based interaction. In urban planning, these technologies offer new opportunities for engaging citizens in the design and transformation of public spaces. In particular, virtual reality (VR) has emerged as a powerful medium for presenting complex spatial scenarios in an intuitive manner, improving comprehension and engagement among non-expert stakeholders (Wolf et al., 2020; Fegert et al., 2021). However, despite these advances, traditional participatory processes remain largely inaccessible to non-experts, relying on technical drawings, specialised terminology, and formal consultation procedures.

This gap has significant implications for the practical exercise of freedom of expression. As articulated in Article 10 of the European Convention on Human Rights, individuals have the right to hold opinions and to receive and impart information and ideas without interference by public authority. In the context of urban planning, however, the effective exercise of this right is often constrained by structural and procedural barriers. Many citizens lack the tools, language, or institutional access necessary to articulate their perspectives within established planning frameworks. As a result, participation is frequently limited to a relatively narrow group of stakeholders with professional expertise, reinforcing existing asymmetries in representation and influence. Prior research highlights this “participation paradox,” where early-stage engagement is crucial but difficult to achieve due to barriers in understanding and communication (Wolf et al., 2020).

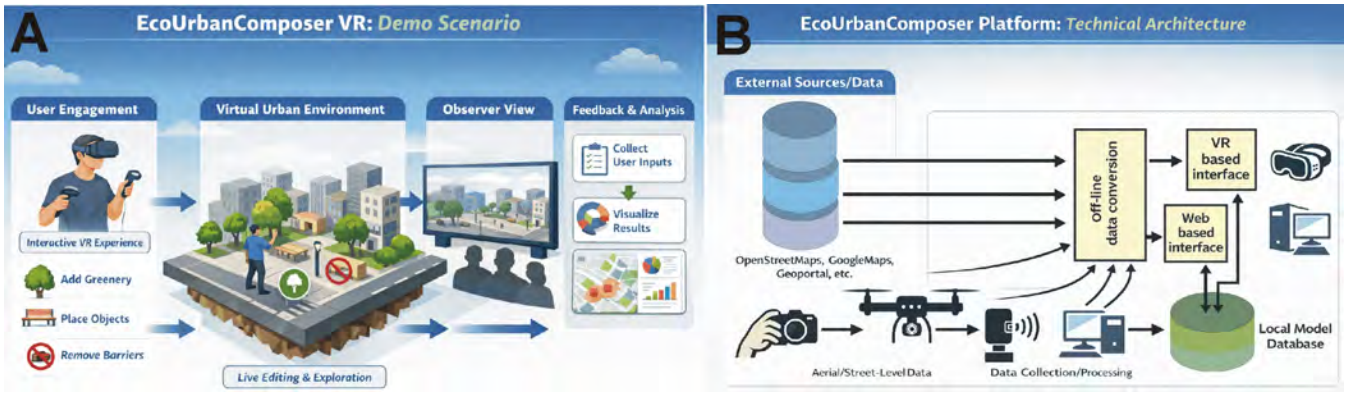

Figure 1. Conceptual overview of EcoUrbanComposer: VR for spatial expression in participatory urban design (A), and platform architecture integrating data, web, and VR interaction layers (B).

This work investigates how immersive environments can address these limitations by enabling alternative forms of expression. As illustrated in Fig. 1A, instead of relying on verbal or written communication, users can directly interact with spatial representations of their environment. Such interaction supports what may be understood as embodied or spatial expression, where placing, modifying, or removing elements within a virtual scene becomes a meaningful way of communicating preferences and concerns. Studies in XR-based participatory planning demonstrate that immersive interaction improves both user motivation and the quality of feedback, particularly when interfaces are designed to be accessible and low-threshold (Saßmannshausen et al., 2021; Holtmann et al., 2025).

At the same time, the use of XR in participatory contexts raises important considerations related to fairness, representation, and governance. The ways in which user contributions are captured, filtered, or prioritised may influence whose voices are effectively represented, while interface design choices can introduce unintended biases in how options are perceived or explored. Moreover, the moderation and interpretation of user-generated content in immersive environments require careful consideration to ensure proportionality, transparency, and respect for diverse perspectives. These concerns have been increasingly recognised in recent policy-oriented analyses of immersive environments, including the Council of Europe's work on freedom of expression in XR and broader metaverse governance frameworks (Council of Europe CDMSI, 2025; Council of Europe, 2024).

The presented approach is situated within the broader context of the Driving Urban Transitions (DUT) Partnership and the EDITUA project (EDITUA Consortium, 2025), which aim to develop innovative tools for ecologically oriented urban transformation. A central objective of this research is to create an integrated instrument that combines technological, procedural, and social innovations to support sustainable placemaking. The EDITUA framework focuses on transforming underutilised or vacant urban spaces through targeted, small-scale interventions, including approaches grounded in urban creativity and spatial activation (Pazder, 2018), inspired by the concept of urban acupuncture, where localised actions can generate wider systemic impact (Wesołowski, 2021).

Within this framework, the EcoUrbanComposer platform is conceived as a multi-level, data-driven environment that integrates XR technologies with analytical and decision-support components, supporting not only analysis but also the interpretation of user-generated spatial input. It enables stakeholders, including citizens, planners, and local authorities, to interact with digital representations of urban environments in an intuitive manner. By combining realistic spatial modelling (e.g., LiDAR-based reconstruction), interactive editing, and structured data export, the system supports both qualitative and quantitative

evaluation of proposed interventions. This aligns with recent research emphasising the importance of traceable, interactive, and data-integrated participatory tools in urban planning workflows (Luo et al., 2022; Daher et al., 2021).

EcoUrbanComposer VR represents the immersive interaction layer of this broader platform. It extends previous work on a VR Urban Scene Editor by explicitly addressing inclusivity, accessibility, and civic participation. The system enables users to explore realistic urban environments and to propose changes through direct manipulation of spatial elements, without requiring professional design skills. In this sense, it functions not only as a visualisation tool but as a medium for expressing spatial ideas and preferences.

By framing these capabilities in relation to freedom of expression in immersive urban environments, this work contributes to ongoing discussions on the role of XR technologies in democratic participation and the rule of law in digital contexts. It demonstrates how immersive systems can lower barriers to engagement, support more inclusive representation of perspectives, and create new channels for communication between citizens and decision-makers. At the same time, it underlines the importance of embedding such tools within transparent and accountable participatory processes, ensuring that digitally mediated expression can meaningfully inform real-world urban transformation. The work is further presented as an interactive demonstration to illustrate these concepts in practice.

2 SYSTEM ARCHITECTURE AND IMPLEMENTATION

EcoUrbanComposer VR constitutes the immersive interaction layer of a broader multi-component platform designed for participatory urban analysis and co-creation. The system architecture (Fig. 1B) integrates heterogeneous data sources, including public geospatial datasets (e.g. OpenStreetMap, municipal and/or national GIS data portals, like Geoportal in Poland) and site-specific data acquired through UAV-based photogrammetry and LiDAR scanning (Fig. 2A). These inputs are processed through an offline data conversion pipeline, producing optimised 3D models and structured datasets suitable for interactive use.

The overall architecture combines two complementary user interfaces: a VR-based environment for immersive interaction and a web-based interface supporting data exploration, contribution, and access for users who do not engage with XR technologies. Both interfaces are connected through a shared data layer, enabling consistent storage and retrieval of spatial configurations and user-generated input.

Within this architecture, EcoUrbanComposer VR is implemented in Unreal Engine 5 with OpenXR support, enabling deployment on modern VR headsets. The system allows users to explore high-fidelity representations of urban environments and to interact with them using direct manipulation techniques (Fig.

Figure 2. World model and navigation in EcoUrbanComposer VR: general view of a fragment of the Poznań University of Technology Warta Campus used in the demo (A), a 2D map of the entire scene to assist user navigation in the large environment (B), and a representation of teleportation used as the main navigation mode (C).

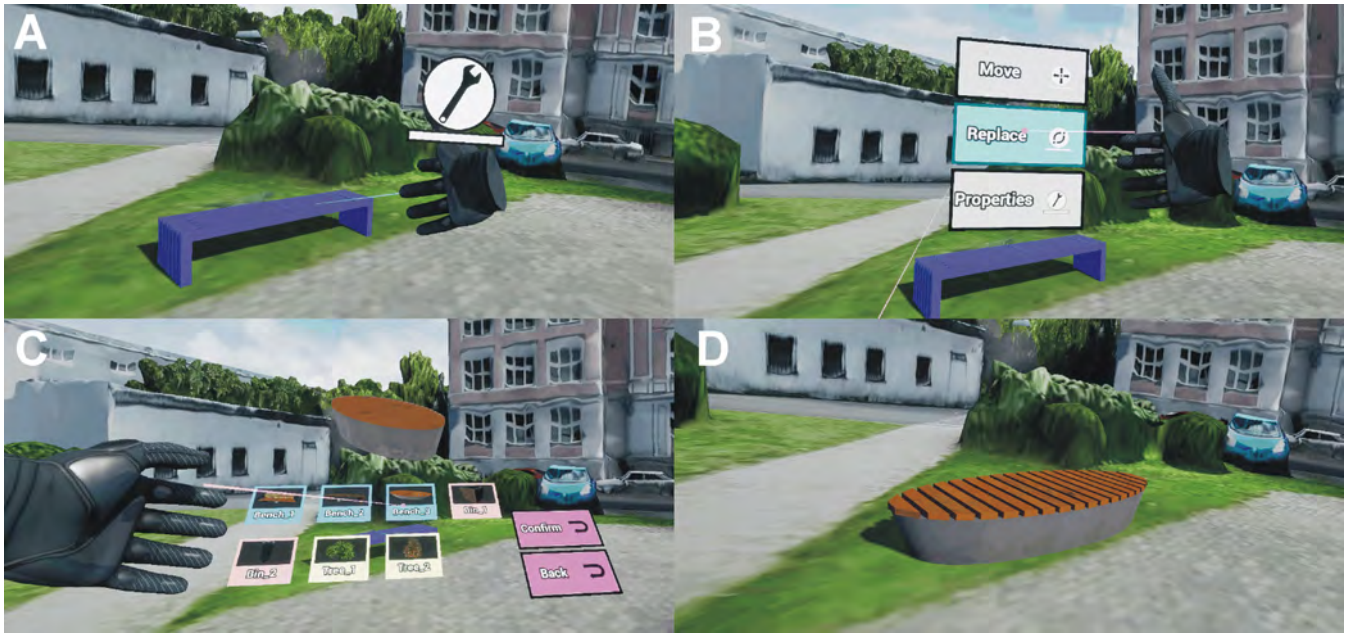

Figure 3. Scene editing example: the edit model tool (A), the main menu with the "Replace" option selected (B), the menu showing objects that can be placed in the scene (C), and a new bench object replacing the removed one (D).

2B). Navigation is based on teleportation, reducing motion discomfort and supporting accessibility for non-expert users (Fig. 2C). The interaction model is intentionally simplified and supports three primary operations (Fig. 3):

- Object placement: introducing elements such as vegetation or street furniture,
- Object removal: deleting existing or proposed features,
- Object modification: adjusting object properties and configurations.

These interactions enable users to construct and compare alternative spatial scenarios in an intuitive, experience-based manner. All actions performed in the VR environment are recorded in structured formats (e.g. JSON), forming a persistent representation of user input that can be analysed, shared, and integrated into broader planning and evaluation workflows.

At the platform level, this interaction layer is complemented by data-driven modelling and assessment components, including scenario evaluation and comparison of alternative interventions. In this way, EcoUrbanComposer supports not only visualisation and exploration, but also participatory co-creation and evidence-informed decision-making in urban design processes.

3 DEMONSTRATION SCENARIO

The proposed demonstration will present EcoUrbanComposer VR as an interactive platform for expressing spatial preferences in urban environments. Participants will be invited to explore a realistic virtual reconstruction of a selected urban site representing a real-world context (PUT Warta campus fragment, 120m×180m) and to modify it using simple interaction techniques. During the demonstration, users will:

- navigate the environment using immersive VR interaction,

- place and remove urban elements such as benches, solar panels and street furniture,
- experiment with alternative spatial configurations reflecting their preferences.

The session will illustrate how non-expert users can express ideas about public space without relying on technical or professional language. Audience members will be able to observe how different design choices emerge through interaction and how these choices can be recorded and analysed.

The demonstration will also highlight how such interactions can serve as a form of civic expression, making visible individual and collective perspectives on urban environments. In this way, the system provides a concrete example of how immersive technologies may support more inclusive participation in decision-making processes.

4 DISCUSSION AND CONCLUSIONS

This work presents EcoUrbanComposer VR as an immersive tool supporting inclusive civic participation in urban design through intuitive, spatial interaction. By enabling users to modify virtual environments directly, the system allows preferences and concerns to be expressed through action rather than formal descriptions, aligning with approaches that emphasise visual and imaginative modes of communication in complex societal contexts (Ravetz, 2023). Such interaction can be interpreted as a form of embodied expression, extending the ability to "impart information and ideas" in line with the principles underlying Article 10 of the European Convention on Human Rights.

By lowering the threshold for engagement, the system supports participation by individuals who are often underrepresented in conventional planning processes, including those without technical expertise or facing linguistic and accessibility barriers. The ability to explore and modify realistic urban environments also facilitates more informed and experience-based contributions to design discussions.

At the same time, the use of XR technologies introduces challenges related to representation, interpretation, and governance. The translation of user-generated input into planning decisions must remain transparent and proportionate to ensure trust and legitimacy, as emphasised in recent policy discussions on immersive environments (Council of Europe CDMSI, 2025; Council of Europe, 2024). Addressing these challenges requires embedding immersive tools within broader institutional frameworks that support accountability and meaningful consideration of citizen input.

EcoUrbanComposer VR contributes to ongoing discussions on human rights in immersive realities by providing a concrete, demonstrable example of how XR technologies can expand opportunities for expression and participation in urban contexts. The system will be presented in an interactive demonstration, illustrating how users can explore, modify, and communicate spatial ideas within a shared virtual environment.

Future work will focus on extending the platform towards multi-user collaboration, integration with decision-support and evaluation modules, and evaluation of how spatial interaction in XR can be incorporated into formal urban decision-making processes.

ACKNOWLEDGEMENT

This project has been funded by NCBR under the Driving Urban Transitions Partnership, co-funded by the European Commission, DUT/2023/101/EDITUA/2024. Work of P. Skrzypczynski was supported by the PUT internal grant no. 0214/SBAD/0252

REFERENCES

[webpage] Council of Europe (2024). Risks and opportunities of the Metaverse. https://assembly.coe.int/nw/xml/XRef/Xref-XML2HTML-en.asp?fileid=33712lang=en

[webpage] Council of Europe CDMSI (2025). Immersive realities — benefits and challenges to freedom of expression feasibility study. https://rm.coe.int/cdmsi-2025-10-draft-feasibility-study-benefits-and-challenges-to-foe-i/488029b847

Daher, E., Maktabifard, M., Kubicki, S., Decorme, R., Pak, B., and Desmaris, R. (2021). Tools for citizen engagement in urban planning. In *Holistic Approach for Decision Making Towards Designing Smart Cities*, eds. G. C. Lazaroiu, M. Roscia, and V. S. Dancu (Springer), vol. 18 of *Future City*. 115–145. doi:10.1007/978-3-030-85566-6_7

[webpage] EDITUA Consortium (2025). DUT EDITUA project. https://dutpartnership.eu/projects/editua

Fegert, J., Pfeiffer, J., Reitzer, P., Götz, T., Hariharan, A., Pfeiffer-Leßmann, N., et al. (2021). Ich sehe was, was du auch siehst. über die möglichkeiten von Augmented und Virtual Reality für die digitale Beteiligung von Bürgerinnen in der Bau- und Stadtplanung. *HMD (in German)* 58, 1180–1195. doi:10.1365/s40702-021-00772-6

Holtmann, C., Stepczynski, J., Wiesner, H., Jung, T., Seifi, M., Keppler, S., et al. (2025). Participatory development of AR and VR systems to improve engagement in urban planning processes. In *Proceedings of the 21st International Conference on Culture and Computer Science: From Humanism to Digital Humanities*. KUI '24

Luo, J., Liu, P., and Cao, L. (2022). Coupling a physical replica with a digital twin: A comparison of participatory decision-making methods in an urban park environment. *ISPRS International Journal of Geo-Information* 11, 452. doi:10.3390/ijgi11080452

Pazder, D. (2018). *Obszary kreatywności creative syntax jako czynnik ożywiania śródmieść (in Polish)* (Poznan University of Technology)

Ravetz, J. (2023). Visual imagineering: graphic mind-games for the emerging collective climatic intelligence. In *On the Interplay of Images, Imaginaries and Imagination in Science Communication* (Leo S. Olschki). 287–297

Saßmannshausen, S. M., Radtke, J., Bohn, N., Hussein, H., Randall, D., and Pipek, V. (2021). Citizen-centered design in urban planning: How augmented reality can be used in citizen participation processes. In *Designing Interactive Systems (DIS) 2021*. 250–265. doi:10.1145/3461778.3462130

Wesołowski, P. M. (2021). Urban acupuncture — ephemeral arrangements of space. *BUILDER* 284, 44–47. doi:10.5604/01.3001.0014.7421

Wolf, M., Söbke, H., and Wehking, F. (2020). Mixed reality media-enabled public participation in urban planning. In *Augmented Reality and Virtual Reality*, eds. J. Jung and M. Rauscher (Springer). 125–138. doi:10.1007/978-3-030-37869-1_11

From the Forest to the Metaverse: Immersive Experiences and Co-Creation with Indigenous Peoples

Florence Dravet[1,2]

[1]Lab Metaverse UCB, PPG Inovação em Comunicação e Economia Criativa, Universidade Católica de Brasilia, Brasilia, Federal District, Brazil

*** Correspondence:**
Florence Dravet
florence@p.ucb.br

Keywords: Virtual Reality[1], Indigenous people[2], Narrative[3], Immersion[4], Freedom of expression[5].

Abstract

This proposal is part of an academic project focused on pedagogical innovation in Brazilian higher education, aiming to present immersive technologies through dedicated demonstration sessions. The initiative is based on the 360° film *Kairau: The Seed of Chants*, produced in the Amazon rainforest, in the Mutum village (Acre), in collaboration with the Yawanawá people. Developed through a co-creation methodology with the Indigenous communities involved, the project expands its scope by including the design of a gamified metaverse, incorporating narratives and knowledge from three Indigenous groups: Pataxó, Yawanawá, and Tukano. Through these immersive and interactive experiences, the project seeks to bring students closer to Brazilian Indigenous realities, fostering engagement, cultural awareness, and new forms of transdisciplinary learning.

1 Introduction

The film *Kairau: The Seed of Chants* (Andreoni; Dravet, 2025) tells the story of a magical seed that gives rise to songs and enabled the Yawanawá community to develop a self-sustainable way of life based on the valorization of their culture.

So we have a traditional narrative, a profecy received by a yawanawá leader from a spiritual creature from the sea (although there is no sea in the middle of the forest, an image of the sea does exist). And based on this profecy all the village started to mobilize so that the village could grow again around the Yawanawá culture.

Through traditional songs, the community has engaged in festivals, album recordings, and eco-ethnotourism activities, strengthening both their cultural identity and economic autonomy.

The project was conceived through a co-creation approach, directly involving Indigenous peoples in the creative process, ensuring respect, representation, and cultural legitimacy (Duarte *et al.*, 2025; Moraes, 2010).

As an extension, the project also proposes the development of an immersive, gamified metaverse environment in which users can interact with cultural elements, narratives, and symbolic territories of the Pataxó, Yawanawá, and Tukano peoples.

In this context, fundamental questions emerge: Who controls forms of expression in these environments? How can distortions or misappropriations be prevented? What models of governance and curation are compatible with Indigenous epistemologies?

2 Objectives

Integrate immersive technologies into higher education as an innovative pedagogical tool; Promote understanding of Brazilian cultural diversity; Raise students' awareness of contemporary Indigenous realities; Encourage learning experiences based on immersion and emotional engagement; Foster transdisciplinary approaches in education; Value co-creation practices with traditional communities; Explore the potential of the metaverse as an educational and intercultural space.

3 Pedagogical Approach

The proposal is grounded in a transdisciplinary (Nicolescu, 1999; Dravet et al, 2023) approach that promotes:
A sense of presence and participation;
Experiential learning;
Development of empathy and social awareness;
Bridging academic knowledge and traditional knowledge systems;
Recognition of the importance of co-creation with Indigenous communities.

6 Target Audience

Higher education students;
Faculty and researchers;
Participants in academic events;
The broader university community.

7 Expected Outcomes

Increased student engagement through innovative methodologies;
Greater awareness of cultural diversity and Indigenous peoples;
Strengthening of transdisciplinary pedagogical practices;
Integration of immersive technologies into educational contexts;
Valuation of co-creation methodologies with traditional communities;
Exploration of virtual environments (metaverse) as learning spaces;
A meaningful and memorable educational experience.

8 Conclusions

The use of immersive technologies combined with the film *Kairau: The Seed of Chants*, alongside a co-creation methodology with Indigenous peoples and the development of an intercultural gamified metaverse, represents an innovative approach in higher education. By providing a sensory, participatory, and culturally situated experience (Haraway, 1988), the proposal contributes to the development of more aware, empathetic students who are better prepared to engage in a diverse and plural society.
The co-creation model presented is aligned with Indigenous epistemologies (Krenak, 2022) and prioritizes the decentralization of power, ensuring that communities themselves retain authority over their forms of expression and knowledge. In this context, meaning is neither universal nor isolated, but deeply rooted in relationships and the territories from which it emerges. Thus, rather than expanding unrestricted access, this model emphasizes responsibility in the use, circulation, and interpretation of knowledge, respecting cultural protocols and collectiveties.

9 Conflict of Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

10 Author Contributions

All the text and the demonstration proposal were written and conceived by Florence Dravet.

11 Funding

This project was funded by FAPDF and CNPq (Brazil)

12 Acknowledgments

I acknowledge the contributions of Rafael Andreoni for the film and Tata Portal for the metaverse. I also acknowledge the contributions of Rasu Yawanawá, Tanara Pataxó and Luvan Tukano in the whole co-creation process.

For some examples please click here.

For more examples of citing other documents and general questions regarding reference style, please refer to the Chicago Manual of Style.

13 References

Andreoni, R., Dravet, F. (2025). The seed of chants. 360° Film. Brasília/Rio de Janeiro. Cadenza Filme, FAPDF, Universidade Católica de Brasília.

Nicolescu, B. (1999). O manifesto da transdisciplinaridade. São Paulo: Triom.

Moraes, D. (2010). Metaprojeto: o design do design. São Paulo: Blücher.

Dravet, F., Pasquier, F., Collado, J., Castro, G. (2023) Transdisciplinarité et éducation du futur. Paris: L'Harmattan.

Duarte, C., Rocha, S., Bessa, L., Dravet, F. (2025). Do mito ao metaverso: uma metodologia transdisciplinar de Cocriação de narrativa. XIX Encontro de Pesquisadores em Comunicação e Cultura, realizado pelo Programa de Pós-Graduação em Comunicação e Cultura da Universidade de Sorocaba (Uniso).

Krenak, A. (2022). Futuro ancestral. São Paulo: Companhia das letras.

Haraway, D. (1988). Situated Knowledges: The Science Question in Feminism and the Privilege of Partial Perspective. Feminist Studies. 14(3). https://doi.org/10.2307/3178066

Exploring XR Policies in the Metaverse with an Intelligent Embodied Virtual Agent

Ke Li [1], Jeroen Prinsen [2], Julia Hertel [1], Frank Steinicke [1,*]

[1] *Human Computer Interaction Group, Hamburg University, Hamburg, Germany*
[2] *SOUND, Amsterdam, Netherlands*

Correspondence*:
Frank Steinicke, Bundesstraße 56b, 20146 Hamburg, Germany
frank.steinicke@uni-hamburg.de

ABSTRACT

This demo invites policy makers, legal scholars, and XR researchers to hold short, semi-structured conversations with a virtual agent that embodies a real person that has been grounded in a retrieval-augmented knowledge base built from the conference programme and from publicly available information about the embodied person. The agent runs on the open-source PRESENCE Intelligent Virtual Human SDK, powered by the Gemini Live realtime API. To minimise cybersickness, the agent is displayed on a large extended display rather than a head-mounted display. The conversation follows a five-node decision tree (one of four scripted opener / follow-up pairs covering freedom of expression, safety of vulnerable users, justice and law, and health and well-being) while permitting open-ended adaptive responses in between. Because the underlying Gemini model family supports multilingual dialogue in context, participants are explicitly encouraged to discuss the topics in their own native language.

Keywords: Intelligent Virtual Agents, XR Policy, AI Ethics, Large Language Model

1 INTRODUCTION

Immersive technologies are increasingly shaping how individuals communicate, express themselves and interact with others, raising emerging challenges for human rights, democratic values and the rule of law. The 2024 Council of Europe / IEEE report on the metaverse (Council of Europe and IEEE Standards Association, 2024) and the subsequent CDMSI Feasibility Study on Freedom of Expression in Immersive Realities (Council of Europe Steering Committee on Media and Information Society (CDMSI), 2025) identified an urgent need to strengthen anticipatory dialogue between policy makers, legal experts and the research community while these technologies are still evolving.

We argue that this dialogue benefits not only from policy reports and panel discussions, but also from *experiential* interventions that allow policy makers to encounter the issues in the medium where they unfold. To that end, we deploy an embodied IVA at the conference on *Human Rights in Immersive Realities*, jointly organised by the CDMSI and the European Metaverse Research Network (EMRN). The agent is designed to engage attendees which primarily consist of policy makers, XR researchers and lawyers and focused conversations about the four conference themes: **(A)** freedom of expression, **(B)** safety of vulnerable users, **(C)** justice, law and jurisdiction, and **(D)** health and well-being.

Figure 1. Screenshots of the demonstrated embodied virtual agent. **Left:** the agent invites the participant to choose one of the four conference themes. **Right:** the agent's universal closing question after the topic dialogue.

The purpose of the demonstration is twofold. First, it gives attendees a concrete, first-person encounter with an LLM-driven embodied agent which represents the kind of system whose human-rights implications the conference is convened to discuss. Second, it generates qualitative and quantitative research data on how domain experts engage with such systems when the conversation is grounded in their own professional concerns.

2 SYSTEM DESCRIPTION

The agent is built on the open-source PRESENCE Intelligent Virtual Human (IVH) SDK (Li et al., 2026; Mostajeran et al., 2025; Li et al., 2025), a Unity package developed at the Human-Computer Interaction group of the University of Hamburg for authoring embodied agents driven by large foundation models.

Embodiment. With his explicit consent, the agent visually embodies Prof. Dr. Frank Steinicke. The avatar's facial geometry was reconstructed via the Reallusion Character Creator 4 (CC4) headshot pipeline from public photographs, paired with a CC4 facial-animation rig driving FACS-based blendshapes for expressive behaviour. Body, gaze and micro-behaviours (blinking, idle motion) are handled by the SDK's behaviour subsystems Li et al. (2025).

Real-time Voice. The conversational backbone is Google's Gemini Live API over a persistent WebSocket connection, with native-audio input and output at 16 kHz / 24 kHz respectively. We deploy on the *Vertex AI* enterprise tier of the service: under the Google Cloud terms applicable to that tier, customer audio is used solely for inference and is not retained for training the foundation model. The SDK additionally enables server-side transcription for both the participant's input and the agent's reply, supporting a live UI log and the post-session transcript export.

Multilingual Interaction. The Gemini Live family inherits the multilingual in-context understanding of the underlying Gemini model (Gemini Team, Google, 2024), and the SDK exposes Gemini Live's automatic multilingual switching capability. Conference attendees are therefore explicitly encouraged to converse with the agent in their own native language including, but not limited to, English, German, and French.

Retrieval-augmented grounding. To improve factual accuracy on questions about the demo, the conference and the person being embodied, the agent integrates a Retrieval-Augmented Generation (RAG) pipeline (Lewis et al., 2020). Two curated Markdown corpora are pre-chunked, embedded via the `gemini-embedding-001` model, and stored as a local vector index baked into the build:

- the **conference operational corpus**, which includes programme, session agenda, consent terms, data-handling policy, project provenance;
- the **persona corpus**, which includes a hand-curated document derived from publicly available information about Prof. Steinicke (academic profile, public talks, publication highlights);
- a **self-description corpus** which includes a topic-headed description of the SDK so the agent can answer "how do you work?" questions accurately.

At runtime each user utterance is embedded and the top-k most similar chunks are prepended to the system prompt with a natural-speech instruction. The agent does not verbalise source markers, keeping the spoken dialogue fluent.

Display and input modality. The agent is rendered on a large extended display (rather than a head-mounted display) for two reasons: it permits walk-up engagement at a conference booth without device fitting, and it avoids the cybersickness commonly induced by short, ad-hoc HMD sessions in untrained users (LaViola Jr., 2000; Stanney et al., 2020). Participants speak to the agent via a microphone.

3 INTERACTION FLOW

Each session lasts between five and eight minutes and follows a five-node decision tree that combines scripted nodes (delivered verbatim for protocol consistency and research reliability) with LLM-guided adaptive nodes (which allow open-ended responses across multiple directions of inquiry):

N0 *Scripted introduction.* The agent discloses its synthetic nature, briefly acknowledges the cloud-based processing, and invites engagement.

N1 *LLM-guided topic detection.* The agent asks which of the four themes is closest to the participant's work and routes accordingly.

N2a–d *Scripted opener.* The agent delivers a provocative first-person statement on the selected branch (e.g. "*My creator, my server and you are each in a different country. Which law governs what I just said to you?*" for Branch C). The statements were curated from primary Council of Europe and academic source material.

N3a–d *LLM-guided follow-up.* The agent briefly acknowledges the participant's response and delivers a paired follow-up question, then engages adaptively. Per-branch adaptation logic preserves the dialogue's intent while honouring divergent participant views.

N4 *Scripted close.* The agent asks "*How did it feel to have this conversation with a virtual human?*" and closes with a fixed reflective line. The operator then conducts the post-interaction debrief.

4 ETHICAL PROCEDURES

The demonstration is deliberately positioned as a self-referential research object: asking participants to consent to data collection by a virtual agent is precisely the kind of situation the conference is convened to discuss. The consent and data-handling protocol therefore models the practice it advocates.

Before each session, a human operator briefs the participant and obtains written informed consent. The consent form covers (i) the synthetic nature of the agent; (ii) the streaming of voice input to Google's Gemini service on the Vertex AI enterprise tier and the corresponding guarantee that the audio is not retained for model training; (iii) the local recording and pseudonymised storage of the conversation transcript, the detected topic branch, the session duration, and the verbal response to the closing question; (iv) the use of a curated RAG knowledge base together with the explicit acknowledgment that the large language model's output cannot be fully controlled by the developers and should not be taken as a Council of Europe or EMRN position; and (v) participant rights, including the right to withdraw at any time, the right to request deletion of collected data, and the five-year maximum retention period.

Data is stored on PRESENCE project servers under GDPR and the University of Hamburg's data-protection policy. Transcripts are pseudonymised; no name, affiliation or biometric identifier is attached. No facial recognition, eye tracking, or voice prints are used for identification. Microphone audio is used to drive conversation only.

5 CONCLUSION

The demonstration aims to bring the human-rights questions raised by immersive AI agents into the room where they are being debated. By embodying a real researcher, grounding the agent in a curated RAG knowledge base, and balancing scripted protocol fidelity with open-ended LLM-driven exchange, we hope to produce conversations that are intellectually substantive, ethically responsible, and analytically useful.

ACKNOWLEDGMENTS

This work has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No. 101135025, PRESENCE project (`https://presence-xr.eu/`). Generative AI tools (Claude and Gemini) were utilized during the preparation of this manuscript to enhance language and readability. The authors maintain full responsibility for the original ideas and final content.

REFERENCES

Council of Europe and IEEE Standards Association (2024). *The Metaverse and its Impact on Human Rights, the Rule of Law and Democracy*. Joint report, Council of Europe

Council of Europe Steering Committee on Media and Information Society (CDMSI) (2025). *Feasibility Study on Benefits and Challenges to Freedom of Expression in Immersive Realities*. Feasibility study, Council of Europe. Adopted at the 27th CDMSI Plenary Meeting, 4 December 2025

[Dataset] Gemini Team, Google (2024). Gemini 1.5: Unlocking multimodal understanding across millions of tokens of context

LaViola Jr., J. J. (2000). A discussion of cybersickness in virtual environments. *ACM SIGCHI Bulletin* 32, 47–56. doi:10.1145/333329.333344

Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., et al. (2020). Retrieval-augmented generation for knowledge-intensive NLP tasks. In *Advances in Neural Information Processing Systems*. vol. 33, 9459–9474

Li, K., Mostajeran, F., Rings, S., Hertel, J., Schmidt, S., Arz, M., et al. (2026). Anthropomorphic AI: A toolkit for authoring and interacting with intelligent virtual agents for extended reality. *Frontiers in Virtual Reality* 7. doi:10.3389/frvir.2026.1794720

Li, K., Mostajeran, F., Rings, S., Kruse, L., Schmidt, S., Arz, M., et al. (2025). I hear, see, speak & do: Bringing multimodal information processing to intelligent virtual agents for natural human-AI communication. *2025 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)* , 1648–1649

Mostajeran, F., Li, K., Rings, S., Kruse, L., Wolf, E., Schmidt, S., et al. (2025). A toolkit for creating intelligent virtual humans in extended reality. *2025 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)* , 1648–1649

Stanney, K., Lawson, B. D., Rokers, B., Dennison, M., Fidopiastis, C., Stoffregen, T., et al. (2020). Identifying causes of and solutions for cybersickness in immersive technology: Reformulation of a research and development agenda. *International Journal of Human–Computer Interaction* 36, 1783–1803. doi:10.1080/10447318.2020.1828535

Session 2

Children’s Rights, Safety and Protection of Vulnerable Users

Applying Child Rights to Product Design Through Co-Created Ethical Principles

Beckett LeClair [1,*], **Nishita Yerraguntla** [1] **and Salome Toliashvili** [1]

[1] *5Rights Foundation, Brussels, Belgium*

Correspondence*:
Beckett LeClair
beckett@5rightsfoundation.com

ABSTRACT

This work outlines approaches to operationalise ethical principles in the development of Extended Reality (XR) applications impacting children, through the novel application of ethical principles (Ethics4Children) co-created with a global cohort of youth. This is supported through comparison to analysis of an existing consumer XR product. This represents the first application of the Ethics4Children principles to XR contexts, and explores the feasibility for developer organisations of applying principles grounded in children's rights and age-appropriate design in practice. This work also represents an early-stage attempt to uphold children's right to have a say in all matters impacting them, in a still relatively nascent technology space.

Keywords: Extended Reality (XR), Child Rights, Ethics, Age-Appropriate Design, Participatory Design

1 INTRODUCTION

Extended and virtual reality products have seen peaks and troughs in popularity over recent years, with new devices and games still being developed and sold through major websites and physical retailers. Some actors within society have expressed concerns over potential unintended impacts on children and teenagers, while others have promoted the technologies as ways to provide more positive immersive experiences for them.

Rather than operating on adult assumptions of children's ethical expectations for these experiences, this work analyses a commercial headset product through the lens of ethical principles co-created with children and experts from around the world. Using principles made by children themselves facilitates the child's right to be heard in all decisions impacting them, an internationally-agreed right which can often go undervalued (particularly in technology contexts).

2 BACKGROUND

2.1 Extended Reality (XR)

Extended Reality (XR) can be considered as a collection of technologies which add a digital layer over physical environments. It is commonly taken to include Augmented Reality (AR), Virtual Reality (VR) and Mixed Reality (MR) within its remit.

The device focused on in this work is a consumer hardware device released in 2023, marketed explicitly as an XR product. It can be used as a standalone system or in conjunction with existing personal computer

(PC) platforms for VR gaming, teleconferencing, creative arts, wellness (e.g. guided meditation) and augmenting the workspace. This work will only focus the use cases most likely to be relevant to children.

2.2 Children's Rights

Children's rights are internationally agreed and defined by the United Nations Convention of the Rights of the Child (United Nations, 1989), which classes a child as anyone under the age of 18. Children have a variety of rights under the Convention including (but not limited to) privacy, safety, freedom from exploitation, a good standard of education, play and relaxation, identity, culture and the arts, and to have a say in all matters affecting them.

Furthermore, General Comment 25 (United Nations, 2021) outlines how the rights in the UNCRC apply equally to online spaces as they do offline ones. Indeed, as the line between online and offline becomes increasingly blurred (especially so in the context of XR), General Comment 25 becomes an increasingly key document for rights-respecting design, development, governance and policymaking.

Despite this, many digital products and services children use or are otherwise impacted by were not designed with their rights in mind. This has led to a variety of harms including but not limited to commercial exploitation (5Rights, 2023), online radicalisation (Resolver, 2026), and services being implicated in instances of suicide (France-Presse, 2026).

2.3 Age-Appropriate Design

Age-Appropriate Design can be defined as the creation of artefacts that take into account what is suitable for a user in light of the increased risks that may present by virtue of their age. When done properly, this requires an awareness that each child's context is different and there is no one uniform developmental pathway for all. In this sense, what is appropriate for one child of a given age may vary from what is appropriate for a different child of the same age. Age-Appropriate Design should not be taken to mean that there is a uniform experience for all users of a given age, rather, it demands an approach to design that is sensitive to factors of age and the vulnerabilities that commonly present with them. Age-Appropriate Design has been applied to digital services for children in technical standards such as IEEE 2089:2021 (IEEE, 2021), which itself is founded in an awareness of the UNCRC.

2.4 Ethics4Children

The 'Ethical Principles for Engineering Digital Services for Children' (shortened to Ethics4Children) project was run as an Industry Connections activity at IEEE SA over the course of 2025 (IEEE SA, 2025). The goal of the activity was to work with a cohort of youth from around the world to co-create a set of ethical principles embedding their expectations of conduct for organisations creating the digital services they engage with. This was further used to propose a short and memorable oath for technical experts, similar in vein to the modern Hippocratic Oath. The working group was multinational and multidisciplinary, including experts from industry, academia, policymaker and civil society backgrounds. Final reports from the Ethics4Children project are pending online publication in 2026.

3 METHOD

The first part of this work comprised a desk-based team exercise to brainstorm how the co-designed principles may apply to the context of an XR device. This produced an initial list of questions and features to explore. The second part of this work involved hands-on experimentation with the chosen hardware

device and analysis of available policy and terms documents, using the list as a starting point while not precluding the addition of new questions in light of ongoing findings. Finally, the team returned to the original principles and considered the findings through the lens of children's wants, needs, and internationally-agreed rights.

Focus was given to the device and operating system environment itself, along with the associated terms and policies, rather than the installation of additional third party software (as this would have vastly increased the scope and created opaque questions around chains of accountability for facets of experiences). However, the scope did include software pre-existing on the device 'out of the box'.

4 RESULTS

Revisiting the Ethics4Children principles through an analytical lens of VR and XR experiences resulted in a set of 67 questions broken down into logical categories. This list was then cross-checked back against the original principles to ensure comprehensiveness. The researcher that prepared the original set of questions was not given prior knowledge of the specific device being used in order to avoid inadvertent biasing of questions. During the course of the work, one additional question (K1) was added in response to experimentation with the device, giving a final total of 68. The set of questions is included as Supplementary Material to this work.

The list was used as a questionnaire to be filled out during device testing, with space for open-ended answers. The completed questionnaire is also included as Supplementary Material to this work.

5 DISCUSSION

5.1 Device case study

Analysis was able to reveal some positive information about how the device may respond to the needs and wants of children and teens. General safety features like boundary setting, re-centring, seated usage options, reduced-motion choices, and accessibility options facilitate positive experiences for youth and adults alike. The conscious design of calm 'home screen' environments, easy onboarding explanations with visual cues, and a lack of attention exploitation features are also worth noting as positives.

Furthermore, pre-packaged software contains opportunities for non-manipulative, unguided and open-ended creative play (Livingstone and Pothong, 2021) (one particular example being the art program Open Brush). There are no clearly exploitative reward loops here, and no forceful nudges to engage in riskier activities in the home environment. Third party data collection from the front cameras is also forbidden by the Software Development Kit (SDK), protecting privacy. It is possible for guardians of younger children to set up usage restrictions, screen time reminders, and real-time monitoring.

However, analysis also revealed a considerable number of concerns. While the device's default privacy posture is reasonable, support for privacy awareness is lacking. Despite plain-language settings, the privacy policy itself is extremely complex and hard to parse even for adult readers, let alone children. Furthermore, although users can opt out of advertising cookies, one of the cookies present is Meta's Pixel, which has been implicated in multiple serious privacy violations in the past (e.g. (Girish et al., 2025)). Users may not fully comprehend the implications of allowing such cookies unless they read the available policies in their entirety, which is unlikely (especially for younger audiences). There is also a lack of granularity over which particular data is collected for which purpose(s), and limited clarity on if/how minors' personal data is exempted from profiling, advertising and behavioural prediction.

Age safeguards seem to rely on a policy statement that the company does not 'knowingly' process the data of under-14s, but this is not validated in practice so there is no way of robustly assuring this statement. Therefore, this statement seems to contribute more to absolving accountability than it does ensuring age-appropriate experiences.

The storefront system for downloading new apps has underdeveloped safety features. Despite appearing to have some degree of curation (though the extent of this is unclear), we could not find any option to report rule-violating content in the store. Generic age ratings are provided for apps with no mandatory detail on how these were determined or what specific age-restricted experiences may be encountered in each app.

While the parental controls appear reasonable and not violating of the child's dignity, they require the setup of an additional mobile app on the guardian's device. This relies on guardians having a compatible device and the requisite level of digital literacy, as well as an awareness of if/when their child is playing in XR.

Finally, it is hard to determine whether the device was independently audited for child rights or age-appropriateness, or tested with a variety of children through meaningful and ethical consultation. This implicates the child's right to be heard, and the need to have the best interests of children be a key consideration in all matters affecting them. It is also unclear whether the design decisions taken were grounded in evidence and expert knowledge on key topics such as child development. Given the further lack of a publicly available child-specific risk registry, we may reasonably assume that child-specific risk assessment was not a priority. This is a key concern that could potentially point to missed risks and gaps in mitigation strategies, beyond what has been uncovered here.

5.2 Applying the principles

Analysis through the lens of the principles was able to uncover positive aspects which align with the expectations, needs and wants of a diverse group of youth. This has use as a way of detecting and highlighting good practice, which can both increase user trust in products and encourage better organisational conduct. The activity also uncovered considerations similar to what authors have experienced while conducting platform investigations based on the UK Age-Appropriate Design Code (Information Commissioner's Office, 2020) and the EU Digital Services Act (European Union, 2022), highlighting this approach as having potential additional use for compliance practice.

More urgently, the method also uncovered a large number of things which have seemingly not been considered by the organisation. These findings could be used to create a concrete action plan with clearly desired outcomes. Remedying the issues could increase end user trust and steer innovation in more beneficial ways that respect the wants and needs of young users. In conclusion, we found the E4C-based approach to yield findings which seem to have not yet been addressed by organisations through their standard design or risk management processes. While this is only a first-pass example application of the principles, the findings suggest value in continuing to refine an approach.

5.3 Recommendations

For technical staff creating XR experiences, using the principles should not replace other key practices like standardised risk management or Child Rights Impact Assessments, but they are a helpful complement to tie design features back to what real children expect in practice. We therefore encourage others to examine their own work (or the work of others) through the Ethics4Children outputs.

We also strongly urge developing organisations to pay greater attention to providing accessible privacy information, and to children's right to be heard in all matters impacting them. While the context of game design may seem more mundane compared to other contexts such as civil proceedings and national policymaking, social activities such as gaming still have profound impacts on children's lives, and it is clear that children have profound things to say. Inviting younger perspectives in meaningful ways can provide new avenues for rights-respecting innovation.

We encourage policymakers, certification providers, and collaborative industry bodies to place greater emphasis on age-appropriate design, including by incentivising best practice where applicable. International standards and best practice have already begun to emerge in this area, and knowledge sharing could uplift the XR gaming ecosystem as a whole. This would create better experiences for all consumers, but especially children.

6 LIMITATIONS

One clear limitation of this work is in the amount of information the team could access as third party researchers. Sometimes information was missing or hard to find. This in itself could be considered a finding of note, however, as it underscores the need for increased transparency from device developers.

A second limitation is scope. The scope had to be constrained for the sake of time, meaning only one type of device was tested. The results may not generalise to other hardware. This work should not be considered a comprehensive child rights audit of the device under consideration, as it was intended to instead explore novel applications of ethical principles designed primarily by young people. Additionally, in an ideal world auditing would have have been co-created with youth in order to embody the spirit of the principles we worked with. It is likely that youth participants would surface questions the researchers may not have considered. However, given the resource constraints it was infeasible to involve a cohort of external participants and distribute multiple devices.

This links to the third limitation, that being the availability of devices for testing. The cost of hardware meant it would have been prohibitively expensive to procure multiple headsets, so testing was limited to one example device. The analysis should not be taken to describe what 'will' happen whenever a child uses the device, rather they indicate the kinds of experiences that 'may' happen.

CONFLICT OF INTEREST STATEMENT

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest. The lead author played a role in the IEEE SA Ethics4Children working group but did not receive any payment from IEEE SA or other incentives to either partake in the group or submit this work.

AUTHOR CONTRIBUTIONS

The first author for this work performed the on-device testing and compiled the background information and limitations. The second author for this work analysed the implications of the results for the wider discussion. The third author worked with the Ethics4Children principles to create the comprehensive questionnaire used in testing. All authors played a part in the final write-up.

FUNDING

This project did not receive specific funding.

ACKNOWLEDGMENTS

The authors would like to acknowledge the contributions of the IEEE SA Ethics4Children working group, especially the youth that took part.

REFERENCES

5Rights (2023). Disrupted Childhood - The cost of persuasive design.

European Union (2022). Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a single market for digital services and amending Directive 2000/31/ec (Digital Services Act) (text with EEA relevance). *EUR-Lex.*

France-Presse, A. (2026). Google and AI startup to settle lawsuits alleging chatbots led to teen suicide. *The Guardian.*

Girish, A., Acar, G., Vallina-Rodriguez, N., Weerasekara, N., and Vlummens, T. (2025). Disclosure: Covert web-to-app tracking via localhost on android. *Local Mess.*

IEEE (2021). IEEE Standard for an Age Appropriate Digital Services Framework Based on the 5Rights Principles for Children. *IEEE Standards Association.*

IEEE SA (2025). Ethical Principles for Engineering Digital Services for Children. *Industry Connections.*

Information Commissioner's Office (2020). Age-appropriate design: A code of practice for online services. *ICO.*

Livingstone, S. and Pothong, K. (2021). Playful by design: free play in a digital world. *Digital Futures Commission.*

Resolver (2026). Critical Harm Intelligence Briefing - Weaponised Loneliness. *Kroll.*

United Nations (1989). Convention on the Rights of the Child. *Office of the High Commissioner for Human Rights.*

United Nations (2021). General comment No. 25 (2021) on children's rights in relation to the digital environment. *Office of the High Commissioner for Human Rights.*

Embodied VR for Children: A Cross-disciplinary Agenda for Future Research

Xueni Pan [1,*], **Oscar Sill** [2], **Marco Gillies** [1] **and Dorothy Cowie** [2]

[1] *See VR Lab, Goldsmiths, University of London, London, United Kingdom*
[2] *Department of Psychology, Durham University, Durham, United Kingdom*

Correspondence*:
Xueni Pan
x.pan@gold.ac.uk

Keywords: Children, Embodiment, Developmental Psychology, Metaverse, Virtual Reality

1 INTRODUCTION

A recent survey showed that 6% of households in the UK own a Virtual Reality device Ofcom (2024), while almost a third of the population had experienced VR by 2021 alone Bennett et al. (2021). Use of at-home VR is increasingly common, and many schools have adapted or are considering using VR headsets as part of their curriculum Department for Education (2023). It is therefore increasingly important to understand how virtual experiences work for children. With the aim of bringing together researchers from different disciplines interested in this topic, we organised a workshop as part of the IEEE VR 2026 Conference in Daegu, South Korea, under the theme: VR Embodiment for Children (VEC). The reason for choosing this conference is not only that it is the biggest annual research conference in this area, but also it is known to be well-attended by interdisciplinary researchers. Although it is still a mainly computer science conference, in recent years, it has started to attract researchers from cognitive sciences and humanities subjects such as art and design. In this paper, we highlight the motivation behind our workshop, describe the format used and papers presented, and provide an overview of the workshop discussion session.

2 MOTIVATION OF OUR WORKSHOP

Recent developments in Virtual Reality have enabled users to inhabit a variety of realistic human-like avatars as well as varied non-human bodies: someone struggling with obesity can experience how it might be to be “living” in the slimmer version of themselves Anastasiadou et al. (2024), individuals struggling with negative thoughts and self-doubts can take a different form of cognitive behavioural therapy by becoming their own conversational partner Slater et al. (2019), and, using a lip-tracker, those suffering from loneliness can be living in the body of a ray-fish, whose “smile” has the power to transform gloomy undersea world into a wonderland of friendly sea creatures every time the user smiles Li et al. (2025). The potential opportunities enabled by embodiment and VR could fundamentally improve many aspects of our lives. However, most VR applications focus only on adult participants. There is not enough understanding of how virtual reality embodiment works for younger users. Therefore, in our recent workshop we set out to collate the most recent research on how these embodied VR experiences work for younger users. In doing this we address several overarching challenges in the field.

One of the key challenges is to measure embodiment in children. Even for adults, there has been recent controversy on how to accurately capture a subjective sense of ownership over a virtual or fake body Lush

et al. (2020). To complement subjective ratings, researchers have recently used more objective measures such as proprioceptive drift Cowie et al. (2013), skin conductance Dewe et al. (2025), or movement kinematics Johannsen et al. (2025). How can we use these novel measures in child samples? Finally, different aspects of embodiment may follow different developmental trajectories Cowie et al. (2013): how can we understand their inter-relations or independence to paint a coherent picture of VR embodiment at a given age? We know that both multi-sensory information and top-down expectations play a crucial role in adults' ability to embody virtual avatars. For example, viewing the movement of a virtual body at the same time as feeling the movement of one's own limbs can elicit a strong sense of body ownership. While multi-sensory information is used in this way at all ages, children's lower sensitivity may mean that they require richer multi-sensory information than adults to establish ownership over virtual bodies Cowie et al. (2022). Likewise, a user may expect an avatar to look similar to their own in various ways. Specifically, children seem more accepting of virtual bodies which differ in size Keenaghan et al. (2020), texture Preston and Kirk (2022), or function Johannsen et al. (2025). They also integrate some of these factors less than adults do Dewe et al. (2026). In this workshop we examine crucial developmental differences in how children use multi-sensory information and top-down expectations to drive a sense of self within an embodied VR experience. We ask how these factors are balanced, and how this balance changes with age.

A second overarching challenge is that embodied VR experiences require users to spend time in a physically and psychologically unusual situation. The experience of inhabiting another character for a period of time could lead to after-effects in the user's behaviour or attitudes, known as the Proteus effect Yee and Bailenson (2007). Despite these concerns, VR remains a promising tool for both cognitive and motor skill learning, across all ages. What are the realistic risks to child users, and how can these may be mitigated to fulfill VR's potential as a tool for children's education, entertainment, and healthcare?

Key topics set out for discussion in our workshop therefore include:

- How can we best measure embodiment in children?
- How do sensory information and top-down expectations drive successful embodiment experiences in children?
- How can we promote physical and psychological safety in children's VR embodiment experiences?
- How do we best design VR embodiment experiences for children?

3 THE WORKSHOP

3.1 Format and talks

The workshop was organised on the afternoon of the 22nd of March, 2026, with a total of four hours allocated to us in a room filled with 30-40 attendees. A Miro virtual interactive whiteboard was set up in advance to facilitate our discussion, with the following child embodiment related themes pre-populated: (1) Today's talks; (2) Measurements (ideas, challenges); (3) Applications (possibilities, barriers); (4) Limitations or hurdles for our field; (5) Ethics and Safety; (6) Future Directions. For each topic, some initial questions were also embedded in the Miro board to facilitate the discussion (for instance, under "Today's talk" we asked, "Which findings were surprising to you?", and under "Measurements" we asked, "What kinds of data are most meaningful for our field?"). The Miro Board link was shared at the beginning of the workshop, and all attendees were encouraged to make contributions (see Figure 1 (A)).

The workshop had two invited keynote talks, from developmental psychology and communication studies respectively. They shared their work, which was in related areas, such as how children differ from

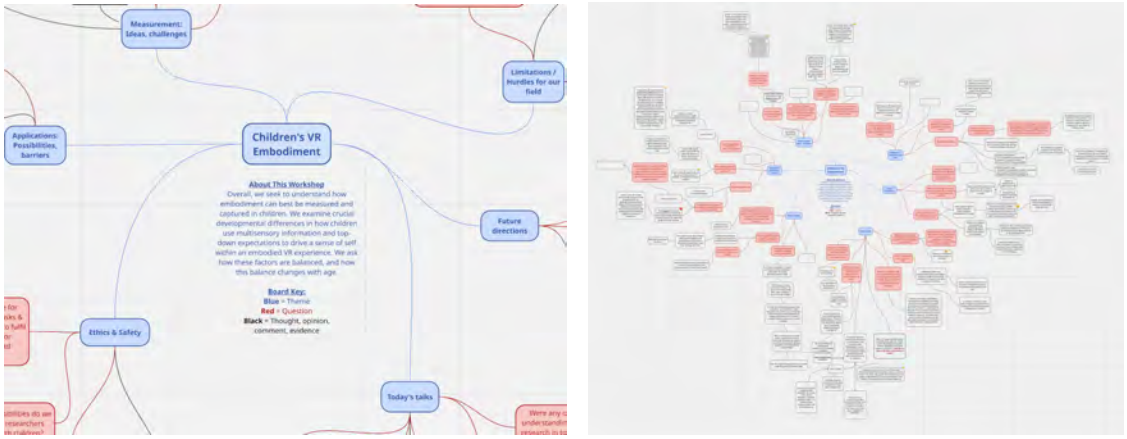

Figure 1. Left: Discussion Topics. Right: Overall view of the Miro board at the end of the workshop; initial Miro board contains only the blue and red parts.

adults in their capacity and cues for adapting a virtual body in VR Keenaghan et al. (2022), and insights from a large-scale community-facing project with over 1,000 school-aged children using XR Ball et al. (2021). Another six papers were presented in either short or long format (5 or 12 mins, plus time for questions), covering a variety of topics including: exploring the use of water-based interactive environment for adolescence's mental health, co-designing mental health interventions for adolescence, the use of Social VR for primary school children for art exhibitions of their own work, understanding racial prejudice in school-aged children, designing super-functional virtual bodies for embodiment research for children, and using VR to teach history to young children. For a full list of papers and authors names please see https://sites.google.com/view/xrkids.

The final 45 minutes was reserved as discussion time, which was led by one of the co-authors. We discussed all six topics listed above, with many attendees contributing towards the discussion. The Miro board was also further developed during this stage.

3.2 Summary of our workshop discussion

We had very active participation in the discussion, as reflected in our Miro board notes (Fig 1(B)). In the following, we attempt to give some example comments for each of the topics from our Miro board which will serve as our initial steps for a longer version research paper. It is by no means an exhaustive summary of comments, nor was it our objective to provide answers to the questions raised in our workshop. The current paper is written by the four workshop chairs, and we plan to invite all paper-authors and attendees who expressed their interests to be co-authors for our longer paper.

3.2.1 Inspiration from workshop talks

Several participants expressed that they were surprised by the result from one of the keynotes, where the speaker discussed how children were found to be better at adopting a bigger and growing body in VR than a smaller one – perhaps due to their brain being wired to deal with their continually growing body. The same pattern was not observed in adults. Some attendees commented that perhaps children would be equally likely to be interested in activities in VR that are impossible or even "dangerous" in real life.

One of the key discussion points was about the need to make VR experiences for children engaging – perhaps through methods of gamification and collaboration with artists and game designers. However,

others also pointed out the validity of using VR for research studies with children - how do we know if the results were not merely a novelty effect, and that findings from these engaging, immersive VR experiences are reliable and replicable?

3.2.2 Measurements

It was generally agreed that measurements in children can be very different from measurements in adults, and that self-report questionnaires need to be altered and checked for understanding for different age groups. Unfortunately, there is no simple recipe – pre-test training can be used, and we should always check children's understanding of the questions asked, but one needs to also be careful not to influence them in any way while doing so. Some attendees mentioned using sad and happy faces as answers – but one researcher cautioned this approach, noting that children are more likely to just choose the happy faces regardless of the questions being asked simply because "they like happy faces".

It was also agreed that behavioural data in the form of system logs (task completion time, accuracy, eye gaze, proxemics) is a good way to objectively obtain measurements; however, to meaningfully interpret those types of data, they need to be used in combination with qualitative methods.

3.2.3 Applications

The use of qualitative methods is again mentioned here – many expressed the need to have children being actively involved in the co-design of VR applications aimed at children. It was also highlighted that there is the need for applications to be more inclusive for neurodiverse children, such as those with ASD and ADHD – indeed some of these children might be the population to potentially benefit from VR applications the most, since they can potentially be engaging and interactive yet free of classroom distractions and social stresses.

In general, the attendees felt very positive and enthusiastic about future VR applications for children, and felt that VR could be a great way to encourage positive behavioural or attitudinal outcomes. However, concerns were raised in terms of how "positive" is defined and, importantly, by who. Although applications designed for children to foster pro-nature behaviours and to reduce bias are well-intended and harmless, it could be a slippery slope – how do we draw the line to prevent VR from becoming a mass manipulation tool for children?

3.2.4 Limitations and hurdles

One of the biggest challenges is the technical hurdle in terms of devices access – current HMDs are not built for children (IPD, neck strain), and hand/body tracking may fail for a child-size body. Although this is generally not a problem for children aged 13 and above, this population is more likely to be more affected by the constantly evolving gaming industry and therefore have higher expectations when asked to participate in VR studies, which are likely be developed in-house with lower graphics and animation fidelity.

Attendees also raised issues around cultural differences when developing applications for children, although it was also pointed out that VR could be an ideal tool for experiencing another culture without physical travelling.

3.2.5 Ethics and safety

Despite the common concerns about VR for motion sickness, lack of official recommendations for age-related guidance, and privacy/safeguarding issues, attendees felt VR could be a good way to amplify

or challenge existing inequalities. For instance, many commented about being anonymous in a social VR environment could be a great way for children to express themselves more freely, although this should be done with caution to prevent online bullying. We also noted that good-quality existing data suggests no significant impacts on children's visual systems Tychsen and Foeller (2020).

However, we all agreed for the need for more research into the biological and cognitive impacts of VR on children (e.g., what would be a harmful form of embodiment? see Bexson et al. (2024)). This would fill the gap in terms of regulatory needs for both VR hardware and applications.

3.2.6 Future directions

Several attendees expressed their strong interests in developing applications for adolescents – their bodies are almost adult-like, and therefore eliminating hardware concerns (e.g., HMD size, tracking), yet their developing brains are very different from adults'. There is the space for research and applications to help them cope with the fast development of the psychological self, and the need to develop better understanding of the cognitive impact from existing commercial VR applications.

Similarly, more research needs to be done and applications to be developed to serve neurodiverse children Tariq (2026) or those with special needs. As researchers, we should constantly question ourselves about the specific needs we are trying to address, but we must do so while reassured of the importance of our work as it is filling important gaps of societal needs. Findings from our work are of importance to both developmental psychology, VR technology, and human wellbeing in general.

4 CONCLUSIONS AND FUTURE WORK

Children embody virtual bodies differently — not less. The findings collected in this workshop make clear that they more readily inhabit growing bodies than shrinking ones, more easily accept avatars of unusual size or function, and may need richer multi-sensory cues to do so. "Embodied VR for children" is not a slower, smaller version of adult VR research — it is a distinct programme with its own measures, methods, and ethical questions. Our workshop converged on four such questions — measurement, drivers, safety, and design — and on two priority populations: adolescents, whose bodies fit current hardware but whose brains are still developing; and neurodiverse children, who may benefit most and should be designed with, not retrofitted around. The most pressing gap is regulatory: there is no agreed age guidance for children's VR use, and limited data to support such guidance. This paper is offered as the seed of a longer, multi-author synthesis. We invite workshop speakers, attendees, and the wider community to join us.

FUNDING

This work is supported by the ESRC grant ES/W003120/1.

REFERENCES

Anastasiadou, D., Herrero, P., Garcia-Royo, P., Vázquez-De Sebastián, J., Slater, M., Spanlang, B., et al. (2024). Assessing the clinical efficacy of a virtual reality tool for the treatment of obesity: Randomized controlled trial. *Journal of Medical Internet Research* 26, e51558

Ball, C., Novotny, E., Ahn, S. J., Hahn, L., Schmidt, M. D., Rathbun, S. L., et al. (2021). Scaling the virtual fitness buddy ecosystem as a school-based physical activity intervention for children. *IEEE Computer Graphics and Applications* 42, 105–115

Bennett, J., Dalton, P., Goriunova, O., Preece, C., Whittaker, L., Verhulst, I., et al. (2021). *Audience Insight Report—The Story of Immersive Users 2021*. Tech. rep., StoryFutures, Royal Holloway, University of London

Bexson, C., Oldham, G., and Wray, J. (2024). Safety of virtual reality use in children: A systematic review. *European Journal of Pediatrics* 183, 2071–2090

Cowie, D., Gottwald, J. M., Bird, L.-A., and Bremner, A. J. (2022). The role of hand size in body representation: A developmental investigation. *Scientific Reports* 12, 19281

Cowie, D., Makin, T. R., and Bremner, A. J. (2013). Children's responses to the rubber-hand illusion reveal dissociable pathways in body representation. *Psychological Science* 24, 762–769

Department for Education (2023). *Technology in Schools Survey Report: 2022 to 2023*. Tech. rep., Department for Education, UK, London, UK

Dewe, H., Brenton, H., Castelow, I., Grimshaw, E., Gillies, M., and Cowie, D. (2026). Children's embodiment of non-human virtual hand forms. *Developmental Science* 29, e70131

Dewe, H., Sill, O., Thurlbeck, S., Kentridge, R. W., and Cowie, D. (2025). The role of visuomotor synchrony on virtual full-body illusions in children and adults. *Journal of Neuropsychology* 19, 57–74

Johannsen, L., Pi, C.-Y., Thurlbeck, S., Gillies, M., Pan, X., and Cowie, D. (2025). Embodiment of a functionally altered virtual arm in adults and children. *bioRxiv* doi:10.1101/2025.01.03.631079. 2025.01.03.631079

Keenaghan, S., Bowles, L., Crawfurd, G., Thurlbeck, S., Kentridge, R. W., and Cowie, D. (2020). My body until proven otherwise: Exploring the time course of the full body illusion. *Consciousness and Cognition* 78, 102882

Keenaghan, S., Polaskova, M., Thurlbeck, S., Kentridge, R. W., and Cowie, D. (2022). Alice in Wonderland: The effects of body size and movement on children's size perception and body representation in virtual reality. *Journal of Experimental Child Psychology* 224, 105518

Li, C., Watson, Z., Pan, X., and Gillies, M. F. P. (2025). Forced smile creates positive mood: A qualitative analysis with a VR experience. In *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*. 1–8

Lush, P., Botan, V., Scott, R. B., Seth, A. K., Ward, J., and Dienes, Z. (2020). Trait phenomenological control predicts experience of mirror synaesthesia and the rubber hand illusion. *Nature Communications* 11, 4853

Ofcom (2024). *Technology Tracker 2024*. Tech. rep., Office of Communications (Ofcom), London, UK

Preston, C. and Kirk, E. (2022). Exploring the development of high-level contributions to body representation using the rubber hand illusion and the monkey hand illusion. *Journal of Experimental Child Psychology* 223, 105477

Slater, M., Neyret, S., Johnston, T., Iruretagoyena, G., Crespo, M. Á. d. l. C., Alabèrnia-Segura, M., et al. (2019). An experimental study of a virtual reality counselling paradigm using embodied self-dialogue. *Scientific Reports* 9, 10903

Tariq, M. U. (2026). Designing immersive learning experiences for neurodiverse learners: A practical framework for inclusive educational technology. In *Practical Perspectives on Inclusive and Immersive Learning Environments* (IGI Global Scientific Publishing). 59–86

Tychsen, L. and Foeller, P. (2020). Effects of immersive virtual reality headset viewing on young children: Visuomotor function, postural stability, and motion sickness. *American Journal of Ophthalmology* 209, 151–159

Yee, N. and Bailenson, J. (2007). The Proteus effect: The effect of transformed self-representation on behavior. *Human Communication Research* 33, 271–290

From Screen to Body: Rethinking Child Safety through the Concept of Embodied Harm in Immersive Environments

Burcin Genis-Ergün [1,2]

[1]Department of Psychology, Umeå University, Umeå, Sweden

[2]TAIGA – Center for Transdisciplinary AI, Department of Computing Science, Umeå University, Umeå, Sweden

*** Correspondence:**
Burcin Genis-Ergün
burcin.genis-ergun@umu.se

Keywords: extended reality (XR), embodied harm, online safety, digital bodily integrity, children's rights, safety-by-design, presence, haptic ethics

Abstract

Immersive technologies, including virtual and augmented reality are reshaping digital interaction by creating environments in which users do not merely observe content but experience it through embodied presence. This shift has significant implications for the protection of children and vulnerable users. Existing legal and regulatory frameworks, largely developed for text- and screen-based media, conceptualize harm primarily in terms of exposure to unlawful or harmful content. However, in immersive environments, harm is increasingly experiential, arising from spatial interaction, perceived proximity, and the simulation of physical presence (Slater 2009, p. 3552; Bailenson 2018, pp. 45-46).

This paper introduces the concept of "embodied harm" to capture the qualitative transformation of risk in immersive realities. It argues that forms of misconduct such as harassment, coercion, or grooming may be intensified in XR due to their affective and sensory immediacy (Madary and Metzinger 2016, pp. 5-7; Citron 2014, pp. 112-114), particularly for children, who may have a limited capacity to distinguish between virtual and physical experiences. As a result, existing safeguards, such as content moderation and notice-and-takedown mechanisms, are insufficient to address these risks.

Drawing on principles derived from the United Nations Convention on the Rights of the Child (United Nations 1989, Art. 3), particularly the best interests of the child, and the jurisprudence surrounding Article 10 of the European Convention on Human Rights (Council of Europe 2024, p.12), the paper argues for a shift from content-based regulation to experience-based protection. It proposes that immersive platforms should be subject to enhanced duties of care, including safety-by-design measures that account for embodied interaction. By reframing harm in experiential terms, this paper contributes to emerging debates on human rights in immersive realities and calls for the development of regulatory approaches that reflect the unique affordances and risks of XR technologies.

1 Introduction: The Collapse of Digital Distance

Immersive technologies, including virtual, augmented, and mixed reality (XR), are transforming the nature of digital interaction. Unlike traditional online environments, which are primarily screen-based and content-driven, XR environments create a sense of presence in which users experience interactions as if they were physically co-located with others (Slater 2009, p. 3551; Bailenson 2018, pp. 45-46). This shift from observation to participation has significant implications for how harm is understood and regulated, particularly in relation to children and vulnerable users.

Existing child safety and online protection frameworks are largely premised on exposure to harmful content, such as violent, sexual, or otherwise inappropriate material (UNICEF 2017, p. 34; Livingstone & Third 2017, pp. 12-15). However, immersive environments challenge this paradigm by introducing forms of interaction that are spatial, embodied, and affective. In these contexts, harm is not merely seen or read but experienced through simulated proximity, movement, and social presence (Madary & Metzinger 2016, pp. 3-5; Biocca 1997, pp. 10-12).

This paper argues that immersive technologies transform harm from a content-based phenomenon into an embodied experience, requiring a fundamental rethinking of child protection frameworks. It introduces the concept of "embodied harm" to capture this shift and examines its implications for law, platform governance, and the protection of children's rights (Council of Europe, 2024, p. 5).

2 From Content to Experience: Defining Embodied Harm

Traditional digital harms have largely been understood through the lens of content exposure. Existing legal and regulatory frameworks focus primarily on identifying, removing, or restricting access to harmful material, supported by mechanisms such as content moderation, reporting systems, community standards, and age-based access controls (Gillespie, 2018, pp. 25-30; Citron, 2014, pp. 88-90). These frameworks were developed for screen-based environments in which harm is typically mediated through text, images, audio, or video. While such approaches remain important, they are increasingly insufficient in immersive environments where interaction itself becomes the primary vector of harm.

XR technologies enable users to inhabit avatars, navigate persistent virtual spaces, and engage with others through gestures, voice, eye contact, movement, and spatial proximity (Slater, 2009, p. 3553; Bailenson, 2018, pp. 48-49). These affordances create a powerful sense of "presence" and social co-location, producing experiences that are psychologically immediate and emotionally consequential (Madary and Metzinger, 2016, pp. 5-6). Unlike conventional digital media, where users remain conscious of the distance between themselves and the screen, immersive systems blur the boundary between representation and lived experience. As a result, interactions in XR can provoke emotional, cognitive, and physiological responses comparable to those experienced in physical environments (Biocca, 1997, pp. 10-12).

The concept of embodied harm captures this qualitative shift. Embodied harm may be defined as any adverse psychological, physiological, emotional, or developmental effect arising from the spatial, sensory, affective, or interactional dynamics of immersive environments (Table 1). Crucially, embodied harm does not depend upon the existence of illegal or explicitly harmful content. Instead, it emerges from the experiential qualities of immersion itself: proximity, movement, gesture, voice, social presence, and sensory feedback. By focusing on experience rather than informational content, embodied harm highlights a critical gap in existing regulatory approaches, which are not designed to

address harms arising from interactional dynamics rather than discrete pieces of information (Gillespie, 2018, p. 30).

Table 1. Comparison between Traditional (online) Harm and Embodied Harm

Feature	**Traditional Online Harm**	**Embodied Harm**
Medium	Text/Image/Video	Spatial Interaction
User Role	Observer	Participant
Harm Mechanism	Content Exposure	Embodied Experience
Moderation Target	Content	Behavior and Interaction
Evidence	Screenshots	Spatial Telemetry
Child Risk	Exposure	Psychological Immersion

Examples of embodied harm include unwanted proximity, simulated physical contact, persistent following, coordinated social exclusion, verbal intimidation in immersive settings, and manipulative uses of avatar embodiment (UNICEF, 2017, p. 36; Livingstone and Third, 2017, pp. 15-16; UK Information Commissioner's Office, 2020, pp. 10-12). Although such conduct may not always involve unlawful speech or prohibited content, it can nevertheless produce fear, humiliation, distress, anxiety, or psychological trauma, particularly for children and vulnerable users whose emotional and cognitive capacities are still developing (UNICEF, 2017; Livingstone and Third, 2017).

3 Children and Amplified Vulnerability in XR

Children are uniquely vulnerable to embodied harm in immersive environments due to developmental, cognitive, psychological, and social factors. Unlike adults, children and adolescents are still developing executive functioning, emotional regulation, risk assessment, and stable identity formation. Their comparatively high levels of neuroplasticity and ongoing social development make immersive experiences particularly influential and potentially harmful (Steinberg, 2014, pp. 66-68). As XR technologies increasingly blur the boundaries between virtual and physical experience, these developmental vulnerabilities become significantly amplified.

Younger users may possess a more limited capacity to distinguish between virtual and physical experiences, particularly within highly immersive environments that generate strong sensations of presence and embodiment (Slater, 2009, p. 3552; Bailenson, 2018, pp. 45-47). The realism, immediacy, and sensory intensity of XR interactions can blur these boundaries further, intensifying emotional and physiological responses to virtual events (Madary and Metzinger, 2016, pp. 6-7). Consequently, negative experiences in XR may be processed not as distant or symbolic events, but as personally lived experiences, increasing their psychological impact and potential long-term consequences.

Immersive environments also amplify social dynamics in ways that are especially consequential for children and adolescents. The sense of co-presence, real-time interaction, and avatar embodiment can increase peer pressure, conformity, emotional dependency, and the desire for social validation. For

younger users whose identities and interpersonal relationships are still forming, these dynamics heighten susceptibility to manipulation, exclusion, coercion, and social anxiety (Livingstone and Third, 2017, pp. 17-20). Because XR interactions are spatial and embodied rather than merely textual, experiences of humiliation, ostracization, or intimidation may feel more immediate and emotionally overwhelming than comparable interactions on traditional social media platforms.

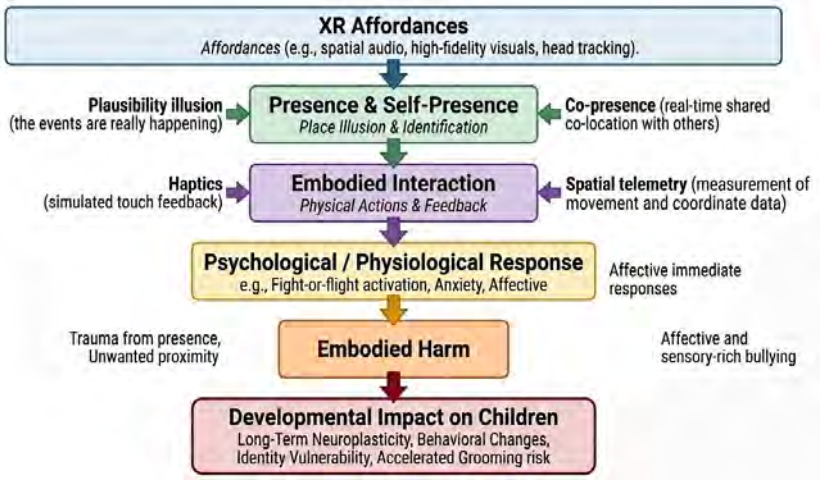

Figure 1. The XR Pathway: Causation and Developmental Vulnerabillity in Immersive Environments, Genis-Ergun, Burcin (2026).

The flowchart as Figure 1 above, depicting "The XR Harm Pathway" aims to visualize the precise casual chain that transforms basic technological affordances, such as co-presence and haptics, into profound and lasting developmental consequences, explicitly mapping the psychological mechanisms through which virtual embodiment manifests as the relational vulnerabilities and identity leakage detailed in the following sections.

3.1 The Proteus Effect and Self-Presence

The "Proteus Effect" describes how users' behaviour and self-perception can begin to align with the traits and expectations associated with their avatars (Yee and Bailenson, 2007). For children and adolescents, whose identities are still developing, repeated experiences of degradation, exclusion, or abuse through an avatar may shape real-world self-image and social understanding. This is intensified by "self-presence," the extent to which users experience their avatar as an extension of themselves (Aymerich-Franch, 2010). In immersive XR, harm directed at the avatar may therefore be experienced as harm directed at the person, with potential implications for body image, consent, gender roles, and interpersonal boundaries.

3.2 Compressed Trust and Spatial Grooming

Immersive XR environments can accelerate interpersonal trust and emotional intimacy through embodied interaction, including voice, gesture, eye contact, spatialized audio, and shared virtual activity (Citron, 2014, pp. 112-115). This "compressed trust" may cause children and adolescents to form emotional bonds unusually quickly, increasing vulnerability to manipulation and exploitation. XR may therefore enable new forms of grooming that bypass traditional timelines and evade moderation systems focused on text rather than behaviour. These risks engage the "best interests of the child" principle under the United Nations Convention on the Rights of the Child (United Nations,

1989, Art. 3), requiring attention not only to content exposure but also to children's psychological experiences in immersive environments (UNICEF, 2017, pp. 34-35).

4 Legal and Regulatory Gaps: The Content-Experience Divide

Current legal and regulatory frameworks are poorly equipped to address embodied harm in XR. Existing online safety models, including the Digital Services Act (2022) and the Online Safety Act 2023, were designed around content moderation focused on posts, images, videos, and text-based communication (Gillespie, 2018, pp. 31-32; Citron, 2014, pp. 113-114). By contrast, XR harms are often relational, spatial, embodied, and context-dependent, arising through avatar behavior, proximity, gesture, movement, and social dynamics rather than discrete content. Many forms of embodied harm, such as virtual stalking, cornering, or repeated invasions of personal space, may cause serious distress while leaving little conventional evidence (UK Information Commissioner's Office, 2020, pp. 11-12). Existing moderation systems and reporting mechanisms remain largely reactive and poorly suited to interpreting these context-rich interactions (Gillespie, 2018; Citron, 2014).

4.1 The "Expression-as-Conduct" Problem

XR blurs the distinction between expression and conduct. While Article 10 ECHR strongly protects freedom of expression, restrictions are permitted to protect the rights and safety of others (Council of Europe, 2024, p. 12). In immersive environments, harmful behavior may occur through avatar movement, gestures, spatial positioning, and simulated physical interaction rather than traditional speech or media. Violent or sexually threatening avatar behavior toward children therefore raises uncertainty over whether it constitutes protected expression or abusive conduct (Lemley and Volokh, 2018). Existing legal frameworks remain poorly equipped to address these embodied harms, creating a protection gap intensified by children's heightened sense of presence and embodiment (Madary and Metzinger, 2016, p. 7).

4.2 The Evidentiary Void

XR also creates major evidentiary challenges. Unlike conventional online harm documented through screenshots or messages, immersive harms are often spatial, embodied, and transient. Experiences such as threatening gestures, invasive proximity, or virtual cornering are difficult to capture through traditional evidence. Proper documentation may require extensive spatial, biometric, and contextual data, raising significant privacy and data protection concerns, especially for children. This creates an "evidentiary void" in which serious harms may be deeply experienced yet difficult to prove or reconstruct. Existing legal and enforcement systems, designed for more easily documented digital harms, remain poorly suited to these experiential and context-dependent interactions.

5 Experience-Based Protection: Rethinking Safety and Privacy in the Metaverse

Addressing embodied harm requires a fundamental shift from content-based regulation toward experience-based protection. Existing online safety frameworks were largely designed for environments in which harm could be identified through discrete pieces of content, such as text, images, or video. Immersive environments challenge this logic because the primary source of harm is often experiential, spatial, and interactional rather than informational. Consequently, effective regulation of XR environments must recognize that the design of immersive systems itself plays a central role in shaping user behavior, social interaction, and exposure to risk (Slater, 2009; Bailenson, 2018).

This transition requires a broader move toward "Safety-by-Design" approaches in XR governance. Rather than relying solely on reactive moderation and post hoc reporting systems, immersive platforms should incorporate preventative safeguards directly into the architecture of virtual environments. Such measures may include default personal-space boundaries, restrictions on certain forms of avatar interaction involving minors, age-sensitive interaction settings, and environment-specific safeguards designed to reduce the likelihood of harmful encounters before they occur (UK Information Commissioner's Office, 2020, pp. 10-12). In immersive environments, design choices are not neutral technical decisions; they actively structure the conditions under which embodied harm becomes possible. At the same time, experience-based protection requires a clearer articulation of platform responsibility and duty of care in relation to children. XR platforms increasingly function not merely as communication intermediaries, but as architects of embodied social environments. As such, they may bear obligations to assess, mitigate, and continuously monitor risks associated with immersive interaction, particularly where children and vulnerable users are involved (Gillespie, 2018, pp. 30-32; Citron, 2014, pp. 112-113). This may include obligations to prioritize user well-being, dignity, and developmental protection over engagement-driven design incentives that encourage prolonged immersion or intense emotional interaction.

The concept of embodied harm itself may also provide the foundation for developing new legal and regulatory categories more capable of addressing the realities of immersive technologies. Existing legal frameworks remain heavily oriented toward content moderation and informational harms, leaving substantial gaps where harmful experiences arise through avatar behavior, spatial interaction, or affective manipulation. By acknowledging the experiential dimension of harm, policymakers and regulators can begin to formulate responses that are proportionate to the distinctive risks posed by immersive environments (Council of Europe, 2024, p. 5).

5.1 The "Personal Bubble" and Virtual Bodily Integrity

One proposed safeguard is mandatory non-bypassable "personal bubbles" for minors, such as default spatial boundaries preventing avatars from entering a protected radius around a child's avatar. These measures could reduce harm such as virtual groping, intimidation, or coercive proximity by protecting virtual bodily integrity. However, overly restrictive boundaries may also reduce meaningful social interaction and participation in XR environments.

5.2 The Privacy-Safety Paradox: Telemetry and Surveillance

Preventing embodied harm may require platforms to monitor highly sensitive XR data, including movement, eye tracking, gestures, voice, and biometric responses. While such telemetry may help detect grooming or proxemic violations, it also risks creating pervasive behavioral surveillance, especially for children. The United Nations Convention on the Rights of the Child and General Comment No. 25 stress that children retain rights to privacy, dignity, and protection in digital environments (United Nations, 1989, Art. 3; UN Committee on the Rights of the Child, General Comment No. 25). One possible solution is privacy-preserving moderation, such as on-device AI systems that detect harmful behavior locally without transmitting raw biometric data.

5.3 Discussion

The conceptual shift from content-based harm toward embodied experience raises several unresolved legal, ethical, and policy questions. In the context of immersive environments, and particularly regarding the protection of children and vulnerable users, several issues merit further consideration:

- Should XR platforms be legally required to implement default spatial and interactional safeguards for minors as a form of Safety-by-Design (UK Information Commissioner's Office, 2020, pp. 10-12)?
- To what extent can existing human rights frameworks, particularly those developed around content-based harms, adequately address embodied and experiential forms of harm in immersive environments (UNICEF, 2017, pp. 34-36; Council of Europe, 2024, p. 5)?
- How should the balance between freedom of expression under Article 10 of the European Convention on Human Rights and the protection of children be recalibrated when "expression" takes the form of immersive, embodied conduct rather than conventional speech?
- Should certain forms of avatar behavior or spatial interaction be legally conceptualized as conduct rather than expression within immersive environments?
- How can regulators and platforms reconcile the need for behavioral monitoring to prevent embodied harm with children's rights to privacy, dignity, and freedom from excessive surveillance?
- What evidentiary standards should apply in cases involving immersive harms that are spatial, affective, and difficult to document through traditional means?
- To what extent should XR platforms bear a proactive duty of care for the psychological and developmental consequences of immersive interaction design?

Taken together, these questions illustrate that immersive technologies are not merely extending existing digital environments but transforming the underlying nature of online interaction itself. The emergence of embodied harm challenges foundational assumptions within current legal, regulatory, and platform governance models, particularly those premised on the moderation of content rather than the regulation of experience. As XR technologies continue to evolve and become increasingly integrated into children's social, educational, and recreational lives, the urgency of developing coherent experience-based protection frameworks will only intensify. The following conclusion argues that safeguarding children in immersive environments requires not simply adapting existing online safety models, but fundamentally rethinking how law, governance, and digital rights operate in an era of embodied virtual interaction.

6 Conclusion

Immersive technologies challenge fundamental assumptions about how harm occurs in digital environments. XR systems are not merely "new screens" or more advanced communication platforms; they create new forms of embodied and spatial experience that blur the boundary between the virtual and the physical. Through presence, avatar embodiment, and real-time social interaction, immersive environments transform digital engagement from observation into participation, producing experiences that can be psychologically, emotionally, and physiologically consequential (Slater, 2009; Madary and Metzinger, 2016, pp. 5-6). This transformation creates profound implications for children and vulnerable users. Existing legal and regulatory frameworks remain largely grounded in a content-centric paradigm focused on identifying and moderating harmful information. Yet many of the most significant risks in XR environments do not arise from unlawful content alone, but from the experiential dynamics of immersive interaction itself: proximity, gesture, social presence, embodiment, and affective realism. The concept of embodied harm provides a necessary conceptual framework through which to understand these emerging risks and to expose the limitations of current approaches to online safety.

Recognizing embodied harm requires a corresponding shift toward experience-based models of regulation, governance, and platform design grounded in the protection of children's rights, dignity, and well-being (UNICEF, 2017, pp. 34-36; Livingstone and Third, 2017, pp. 15-17). If harm in immersive environments is experienced through the body, then regulatory responses cannot remain confined to content moderation alone. Instead, they must engage with the structures, interactions, spatial architectures, and technological affordances that shape how users experience digital worlds (Bailenson, 2018, pp. 46-48; UK Information Commissioner's Office, 2020, pp. 10-12). The emergence of immersive technologies therefore demands more than incremental adaptation of existing online safety frameworks. It requires the development of a new architecture of digital rights capable of recognizing virtual bodily integrity, embodied vulnerability, and the experiential realities of immersive interaction. Concepts such as safety-by-design, spatial protections for minors, privacy-preserving moderation systems, and platform duties of care represent important initial steps toward this transition.

Ultimately, safeguarding children in immersive environments requires ensuring that the "best interests of the child" are embedded not only within legal doctrine and regulatory policy, but within the code, physics engines, interaction systems, and design logic of immersive platforms themselves. Failure to recognize the digital body as a site of potential trauma risks leaving an entire generation vulnerable to forms of harm that existing laws were never designed to perceive. As immersive technologies continue to evolve, the central challenge for law and governance will be to ensure that human dignity, autonomy, and child protection remain foundational principles within the emerging architecture of virtual experience.

7 Conflict of Interest

The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

8 Author Contributions

The author solely contributed to the conception, design, analysis, writing and revision of this work and approved the final manuscript.

9 Funding

This research received no external funding.

10 Acknowledgments

The author would like to acknowledge the support and resources in Umeå University.

References

Aymerich-Franch, Laura. "Presence and Emotions in Playing with Avatars." *PsychNology Journal* 8, no. 1 (2010): 117-40.

Bailenson, Jeremy N. *Experience on Demand: What Virtual Reality Is, How It Works, and What It Can Do.* New York: W. W. Norton, 2018.

Biocca, Frank. “The Cyborg’s Dilemma: Progressive Embodiment in Virtual Environments.” *Journal of Computer-Mediated Communication* 3, no. 2 (1997): 12-26.

Citron, Danielle Keats. *Hate Crimes in Cyberspace.* Cambridge, MA: Harvard University Press, 2014.

Council of Europe. *Human Rights in the Digital Environment: Immersive Realities and Extended Reality Technologies.* Strasbourg: Council of Europe, 2024.

European Union. *Digital Services Act (Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market for Digital Services and Amending Directive 2000/31/EC).* Official Journal of the European Union, 2022.

Freeman, Guo, et al. “Safety and Ethical Considerations for Social Virtual Reality Systems.” In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*, 1-8. New York: Association for Computing Machinery, 2020.

Gillespie, Tarleton. *Custodians of the Internet: Platforms, Content Moderation, and the Hidden Decisions That Shape Social Media.* New Haven: Yale University Press, 2018.

Hall, Edward T. *The Hidden Dimension.* New York: Doubleday, 1966.

Lemley, Mark A., and Eugene Volokh. “Law, Virtual Reality, and Augmented Reality.” *University of Pennsylvania Law Review* 166 (2018): 1051–1138.

Livingstone, Sonia, and Amanda Third. “Children and Young People’s Rights in the Digital Age.” *New Media & Society* 19, no. 5 (2017): 657–70.

Madary, Michael, and Thomas K. Metzinger. “Real Virtuality: A Code of Ethical Conduct: Recommendations for Good Scientific Practice and the Consumers of VR-Technology.” *Frontiers in Robotics and AI* 3, no. 3 (2016): 1-23.

O’Brolchain, Fiachra, et al. “The Convergence of Virtual Reality and Social Networks: Threats to Privacy and Autonomy.” *Science and Engineering Ethics* 22, no. 1 (2016): 1-29.

Slater, Mel. “Place Illusion and Plausibility Can Lead to Realistic Behaviour in Immersive Virtual Environments.” *Philosophical Transactions of the Royal Society B* 364, no. 1535 (2009): 3549-57.

Steinberg, Laurence. *Age of Opportunity: Lessons from the New Science of Adolescence.* Boston: Houghton Mifflin Harcourt, 2014.

UK Information Commissioner’s Office. *Age-Appropriate Design Code: A Code of Practice for Online Services.* London: Information Commissioner’s Office, 2020.

UK *Online Safety Act 2023.* London: The Stationery Office, 2023.

UNICEF. *The State of the World’s Children 2017: Children in a Digital World.* New York: UNICEF, 2017.

United Nations. *Convention on the Rights of the Child.* New York: United Nations, 1989.

UN Committee on the Rights of the Child. *General Comment No. 25 on Children's Rights in Relation to the Digital Environment*. Geneva: United Nations, 2021.

Yee, Nick, and Jeremy Bailenson. "The Proteus Effect: The Effect of Transformed Self-Representation on Behavior." *Human Communication Research* 33, no. 3 (2007): 271-90.

Youth Co-Creation and Ethical Storytelling in Immersive 360° video

Dr Danai Mikelli

School of Arts, Faculty of Arts, Humanities and Social Sciences, Oxford Brookes, UK

*** Correspondence:**

dmikelli@brookes.ac.uk

Keywords: immersive storytelling, XR ethics, youth participation, critical media literacy, co-creation, 360° video, immersive media pedagogy

Abstract

This paper examines ethical practice in XR through a practice-based research project exploring youth co-creation in immersive storytelling using 360° video. While young people are increasingly exposed to immersive platforms, ethical literacy has not developed at the same pace as technical access. This paper addresses this gap by bringing together young participants aged 14–20 from a community arts initiative, theatre practitioners, and digital media researchers to collaboratively produce an immersive video exploring social issues relevant to participants' lived experiences. Situated within the framework of Critical Media Literacy, the study positions immersive production as a site of inquiry in which questions of authorship, visibility, and responsibility are actively addressed. Through a series of facilitated workshops, including ideation, technical experimentation with 360° cameras, production, and reflective discussion, participants engage with immersive media not only as users but as co-authors of immersive narratives. The research further develops the concept of *reflexive immersion* as a mode of engagement that extends beyond affective response towards critical meaning-making. By embedding ethics as a design principle, the project highlights how participatory XR practices can support youth voice while foregrounding ethical approaches to co-creation. In doing so, it contributes to discussions on ethical design and governance in immersive environments.

1 Introduction

Immersive technologies are transforming the ways stories are created and experienced, introducing ethical, representational, and accountability considerations in digital environments. In this context, a key challenge emerges: how can ethical literacy be embedded into immersive media creation from the outset? While young people are increasingly exposed to immersive platforms through gaming, social media, and emerging creative industries, the development of ethical literacy has not kept pace with expanding technical access. Despite ongoing efforts to establish ethical frameworks for virtual reality (VR) (Madary and Metzinger, 2016; Kourtesis, 2024), significant gaps remain, particularly in relation to young people, as empirical studies involving children and immersive technologies are often difficult to conduct (McIntosh and Allen, 2022).

This study responds to this gap through a practice-based research project that brings together youth participants, theatre practitioners, and digital media students and researchers. Situated at the intersection of digital media practice, youth arts, and Critical Media Literacy, the project explores what immersive storytelling makes possible not only creatively, but also ethically and pedagogically.

By positioning young people as co-creators rather than users, it examines how immersive production processes can support reflective engagement with questions of voice, authorship, representation, and responsibility.

The research is guided by two questions: what ethical challenges and creative opportunities emerge when working collaboratively with young people in immersive practice contexts, and how immersive storytelling can support critical media literacy and reflective practice. To address these, the project pursues three interrelated aims: first, to explore youth-led immersive media production as a form of creative and ethical expression; second, to develop and test inclusive, practice-based methods for co-creation, in which shared decision-making over narrative direction, visibility, and representation is central; and third, to inform best practice in the ethical use of immersive technologies in community and educational settings, with particular attention to issues of consent, power, and participation.

The paper also engages with current policy debates, aligning with calls to strengthen awareness of the implications of immersive technologies for freedom of expression and fundamental rights (Council of Europe Steering Committee on Media and Information Society, 2025), and with recommendations for literacy-focused approaches that support ethical engagement (McIntosh and Allen, 2024). In doing so, it invites discussion on how participatory immersive media practices might inform emerging governance frameworks addressing children's rights and ethical design in immersive environments. As the co-creation workshops are currently underway, the paper presents the research design, conceptual framework and initial findings from the first workshops.

2 Immersive Storytelling and Youth Participation

This project draws on the tradition of Critical Media Literacy (CML), an approach to media education that extends beyond understanding how media operates to enabling learners to become active producers of media for civic participation (Kellner and Share, 2007). CML seeks to equip young people with the skills to interpret the multiple meanings of media texts and to develop critical responses to socially constructed forms of communication (Kellner and Share, 2007). Influenced by Paulo Freire's problem-posing model of education, it emphasises participation and dialogue, positioning learners as active agents capable of shaping media and society.

Building on this foundation, the present study introduces a methodological extension of CML through the use of immersive technologies, drawing on earlier work by the author (Mikelli and Dawkins, 2020). The VR Kaleidoscope project (2018), a community-based initiative involving young people from disadvantaged backgrounds, explored the co-creation of 360° films centred on themes of space and place. Findings from this study indicate that immersive media can support more inclusive and participatory forms of media production, particularly for marginalised groups. Two forms of empathy were identified: production-level empathy, emerging through collaborative storytelling and enabling mutual perspective-taking, and experiential empathy, arising from participants' first encounters with immersive environments. Importantly, this work introduced the concept of *intellectual immersion*, shifting the focus from affective engagement alone to processes of meaning-making. By foregrounding space and place as dynamic constructs, immersive storytelling enabled participants to articulate intersectional identities that challenge reductive media representations. The concept foregrounds the cognitive labour involved in engaging with situated experiences, perspectives, and social realities through immersive forms. *Intellectual immersion* therefore shifts emphasis from emotional absorption alone towards interpretive engagement with how environments, narratives, and perspectives are constructed and understood.

This project further develops CML by positioning immersive production itself as a site of inquiry, where questions of ethics, visibility, and audience positioning are actively addressed. While existing research has explored the role of VR in enhancing multiliteracy (Lazou et al., 2025) and the affective impact of immersive experiences on adolescents (Trudeau, 2023), the implications of youth co-creation in immersive contexts remain under-researched. Similarly, studies of autobiographical immersive storytelling (Bahn et al., 2023) demonstrate the expressive potential of these media, but offer limited insight into the ethical frameworks and safeguards required when young people are involved in production processes.

These gaps are particularly significant given emerging shifts in media literacy for immersive experiences. As outlined in *Understanding Virtual Reality: Challenging Perspectives for Media Literacy and Education* (Jones et al., 2022), new competencies are required, including understanding presence and immersion, recognising emotional influence, analysing embodied narratives, and critically interrogating design and interaction structures. These developments signal a broader pedagogical shift from analysing media texts to reflecting on immersive experiences, and from passive critique to active, participatory engagement.

While a full analysis of immersion and presence is beyond the scope of this paper, these concepts are central to understanding the ethical and pedagogical implications of immersive media. Immersion can be understood as the extent to which a virtual environment surrounds the user and supports the illusion of reality, while presence refers to the user's subjective experience of "being there" within that environment (Slater, 2003; Jensen and Konradsen, 2018). These qualities have particular significance in youth contexts, as immersive experiences have the potential to amplify emotional responses and intensify engagement, while raising new questions about responsibility, representation, and agency. It is within this shifting landscape that the present study examines how young people respond to these dynamics through collaborative immersive storytelling.

3 Context and Methodology

This project adopts a practice-based research design that combines co-creation with reflective inquiry. Co-creation provides a participatory framework in which young participants actively shape the themes, narratives, and representational choices of the final immersive work. As an alternative to single-author models of media production, co-creation encompasses a range of collaborative methods, iterative processes, and feedback systems through which projects emerge from within communities, rather than being produced for or about them (Cizek and Uricchio, 2022). Within this framework, participants are positioned not as subjects but as co-authors, contributing to both the creative and decision-making processes. Practice-based research underpins this approach by treating creative production as a mode of knowledge generation. New insights are developed through the process of making, alongside the documentation and critical reflection on that process. This dual emphasis on practice and reflection enables the project to examine not only the outcomes of immersive storytelling, but also the conditions under which it is produced.

The project is conducted in collaboration with the Black Excellence Project, a youth arts and empowerment initiative based in Oxford, UK, led by Euton Daley, a theatre practitioner, producer and director and the theatre company Kuumba Nia Arts. Working with five young people aged approximately 14–20, the research is embedded within an existing creative community context where expression and identity development are already central. The co-creation workshops take place at the The Story Museum in Oxford as part of the wider Noughts, Crosses and New Voices programme, a collaborative initiative that invites young people to engage with immersive storytelling through

creative responses to the novel Noughts & Crosses. Written by Malorie Blackman, the novel presents an alternative society structured around racial segregation and systemic inequality, reversing historical power relations to foreground issues of privilege, discrimination, and resistance. Its exploration of identity, belonging, and social justice provides a rich narrative framework for critical engagement, encouraging participants to reflect on contemporary social issues through re-imagined story worlds.

Prior to this project, participants had already engaged individually with the novel through creative writing and music/audio production, developing personal responses to its themes. The immersive workshops therefore mark a shift from individual to collaborative practice, as participants work together for the first time to co-create a shared immersive narrative. This transition is significant in methodological terms, as it enables the study to examine how collective authorship, negotiation, and shared decision-making emerge within immersive storytelling contexts.

The programme foregrounds co-design, cultural representation, and storytelling rooted in African, Caribbean, and Global Majority heritage, providing a context in which participants develop creative and analytical skills while engaging with issues of identity and visibility. Kuumba Nia Arts, a Black-led touring theatre company specialising in performance informed by African theatrical traditions, contributes to the project through its emphasis on creativity, purpose, and culturally grounded storytelling practices. These partnerships situate the research within a community-led framework, ensuring that the immersive workshops build on established relationships and practices rather than introducing an isolated intervention. The current concept explored by the group, re-imagining a different world, aligns closely with the affordances of 360° storytelling, which supports exploratory, non-linear, and spatial forms of narrative.

The co-creation process unfolds across six structured workshops: introduction and framing, idea development, technical exploration of 360° video, production, post-production, and reflection. Throughout these stages, participants engage with both the technical and conceptual dimensions of immersive storytelling, with particular attention to questions of representation, audience positioning, and ethical decision-making. The project also involves an undergraduate student technical mentor studying at Oxford Brookes University. The mentor contributes to the production and post-production of the co-created video, including the editing of 360° video materials. Their involvement reflects the collaborative and pedagogical ethos of the project, extending co-creative practice beyond participant workshops to include peer-supported knowledge exchange and emerging technical expertise in immersive media production.

Data collection comprises three primary sources: the co-created 360° video, workshop documentation, and semi-structured participant interviews conducted following the completion of the project. These interviews explore participants' experiences of authorship, visibility, and collaboration within the immersive production process. The analysis is informed by a reflexive methodological approach (Alvesson and Sköldberg, 2009), which foregrounds both interpretation and critical reflection. Interpretation involves recognising the theoretical assumptions and pre-understandings that shape the research process, while reflection entails a critical examination of how these perspectives influence the analysis of empirical material. This approach is particularly suited to practice-based research, where the researcher is actively involved in the production process and must account for their role in shaping both the data and its interpretation.

4 Ethics as Design in Immersive Co-creation

The study explores what constitutes an ethical question in the context of co-creating immersive media with young people, embedding ethics as a design principle in the creative process. In traditional media production, ethical considerations often centre on the relationships between producers, subjects, and audiences. However, immersive technologies introduce additional layers of complexity, raising new concerns around accessibility, data practices, privacy, and the intensified experiential impact of media (Rose, 2023).

Within this project, ethical practice is operationalised through a set of considerations, including consent and assent procedures, participant agency over representation, anonymity, and the management of immersive exposure. For participants under the age of 16, written parental or guardian consent was obtained alongside individual participant assent.

Participant agency is further supported through flexible approaches to representation. Young people are given the option to engage in either fictional or documentary storytelling and are not required to appear on camera. This enables participants to make informed decisions about visibility and self-representation within the immersive environment. In addition, the project recognises that immersive media may not be equally accessible or comfortable for all participants. To address this, non-headset engagement options are provided, including the ability to cast 360° content via screen-based projection. This approach not only mitigates potential discomfort associated with VR use, but also supports more collaborative modes of engagement, allowing participants to collectively reflect on and discuss immersive content.

By embedding these considerations within the design of the workshops, the project positions ethics as an integral component of immersive co-creation. This approach acknowledges that immersive storytelling intensifies both the creative possibilities and the responsibilities of media production, requiring attention to issues of consent and participant well-being throughout the research process.

5 Emerging Insights

The early stages of this research highlight the complexities of conducting practice-based academic research through community partnerships, particularly when working with young people in ethically sensitive contexts. A key insight relates to the procedural demands of institutional ethics approval, which required multiple revisions and extended timelines. Aligning these formal academic processes, such as ethics clearance and funding cycles, with the realities of community-based practice presents ongoing challenges, particularly where partner organisations operate under capacity constraints.

This tension became evident when an initial partner organisation withdrew from the project prior to data collection, citing the need to prioritise the wellbeing of their young participants, many of whom were managing multiple responsibilities. While this created practical challenges for the research, including the need to establish a new partnership, it also emphasised an important ethical principle: that community partners must retain the autonomy to disengage from research processes where participation may place pressure on participants. Respecting this decision was therefore integral to the ethical integrity of the project, reinforcing the need for flexible research design in community-based immersive practice.

These experiences point to the importance of understanding ethics not only as a set of protocols governing participant interaction, but as a broader relational practice that shapes how research is initiated, conducted, and sustained. They also highlight the need for methodological approaches that can accommodate uncertainty, adaptation, and the evolving conditions of collaborative work.

A key early insight relates to the complexity of facilitating co-creation in practice. While co-creation is often positioned as an inherently collaborative and organic process, the initial workshops revealed that collective authorship does not emerge automatically. During the story development workshop, participants brought forward distinct and competing ideas. One participant proposed developing a 360° narrative based on a poem he had written in response to the novel, reflecting on an experience of being followed by a security guard in a shop, while a younger participant expressed interest in creating a story centred on ballet. These divergent creative directions highlighted the challenges of aligning individual perspectives within a shared narrative framework.

Through a series of structured exercises and facilitated discussion, participants eventually converged on a collaborative storyline centred on a conflict between a brother and sister from the Crosses, addressing themes of identity and relationships drawn from Blackman's novel. However, this process required ongoing prompting and mediation, with facilitators repeatedly encouraging participants to engage in shared decision-making. Collaboration, in this context, did not emerge organically but was actively supported and negotiated throughout the workshop process.

This highlights an important methodological insight: co-creation in immersive storytelling contexts requires intentional scaffolding, particularly when participants are transitioning from individual creative expression to collective authorship. It also raises questions about how power, voice, and compromise are navigated within participatory creative processes, and how these dynamics shape both the narrative outcome and participants' engagement with ethical and representational concerns.

Finally, in the same workshop, participants began to conceptualise audience and purpose within immersive storytelling and were invited to consider what message they wished to communicate. One participant reflected that the story could encourage viewers to recognise parallels with their own lives, suggesting that audiences might "see something that can happen in real life" and reflect on similar experiences. This response indicates an emerging awareness of storytelling not only as a form of expression, but as a means of fostering connection and reflection.

This moment is significant in demonstrating a shift from personal or individual expression towards a more relational understanding of storytelling. Participants begin to consider how immersive narratives might be received and interpreted, and how they can prompt reflection in others. In this sense, storytelling becomes a space for anticipating audience experience, rather than solely articulating one's own perspective.

6 Discussion

Overall, the project contributes to efforts to align immersive innovation with the protection of fundamental rights, including freedom of expression and children's rights in evolving XR environments. It offers insights into how ethical XR practices might be designed and governed, demonstrating how ethical considerations can be embedded within creative processes rather than applied retrospectively. The findings suggest that participatory immersive storytelling can function as a model for ethical XR environments, where issues of consent, representation, and agency are actively negotiated through practice.

At the same time, the study highlights that co-creation is not inherently collaborative or democratic. Early workshops revealed that participants initially brought forward divergent and individually grounded ideas, requiring sustained facilitation to move towards a shared narrative. Collaboration did not emerge organically but was actively scaffolded through structured exercises and ongoing mediation. This insight challenges assumptions within policy and design discourses that participation

can be seamlessly achieved, instead emphasising the need for intentional frameworks that support negotiation, balance voices, and enable equitable contribution. In this sense, ethical participation must be understood not only as access, but as a process that requires support, structure, and care.

A second insight relates to participants' developing awareness of audience and purpose. As participants reflected on the potential impact of their story, they began to articulate how immersive narratives might prompt audiences to recognise parallels with their own lived experiences. This shift from individual expression to relational storytelling highlights how immersive media can support reflective engagement, encouraging participants to consider how meaning is constructed and received. Such moments indicate that immersive storytelling can extend beyond creative production to foster critical awareness of representation and interpretation.

These insights have direct implications for freedom of expression. Immersive storytelling expands the scope of participation by enabling young people not only to access content but to shape how stories are experienced. However, meaningful participation depends on the conditions under which co-creation takes place. While current policy discussions on children's rights in immersive environments tend to prioritise protection from harm, this study places emphasis on young people's rights to ethical participation and expression. It also highlights an ongoing tension between rapid technological innovation and the slower development of ethical and regulatory frameworks, pointing to the need for further empirical research grounded in practice.

Finally, this research extends the concept of *intellectual immersion* towards *reflexive immersion* by foregrounding reflective and analytical modes of engagement. Through co-creative practice, participants are not only immersed in narrative environments but are encouraged to question how these environments are constructed, whose perspectives are represented, and how audiences may interpret them. They also reflect on authorship, representation, audience positioning, and the ethical implications of shaping immersive experiences collaboratively. This shift from experiential to reflective engagement has important implications for both pedagogy and policy. It suggests that immersive technologies should not be understood solely as tools for emotional impact, but as environments that can foster critical awareness and ethical reflection. In this sense, *reflexive immersion* can be seen as a key component of ethical XR design, aligning with calls for literacy-focused approaches that enable users to engage with immersive media in informed and responsible ways.

7 Conclusion

This paper explores how immersive storytelling can function as a site for developing ethical literacy through youth co-creation. By situating the project within Critical Media Literacy and adopting a practice-based, participatory methodology, the study demonstrates how ethical considerations can be embedded within the design of immersive media production. The study highlights the value of positioning young people as co-creators, enabling not only creative expression but also critical engagement with the social and ethical dimensions of immersive storytelling. The concept of *reflexive immersion* further extends current understandings of immersive experience, foregrounding the role of reflection and meaning-making alongside affective engagement. While the study is ongoing, it points to the potential of participatory XR practices to inform more inclusive and rights-based approaches to immersive media. As immersive technologies continue to evolve, integrating ethical design, critical literacy, and youth voice will be essential in aligning innovation with the protection of fundamental rights.

8 Conflict of Interest

The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

9 Author Contributions

Dr Danai Mikelli conceived and designed the study, is responsible for data collection and analysis, and wrote the manuscript.

10 Funding

The study was funded by the Faculty of Arts, Humanities and Social Sciences, Oxford Brookes.

11 References

Alvesson, M. and Skoldberg, K. (2009). *Reflexive methodology: new vistas for qualitative research*. 2nd edn. London: Sage.

Bahng, S., McArthur, V. and Kelly, R.M. (2023). Designing immersive stories with novice VR creators: a study of autobiographical VR storytelling during the COVID-19 pandemic. *Front. Virtual Real*. 4:1174701. doi:10.3389/frvir.2023.1174701

Cizek, K. and Uricchio, W. (2022). *Collective wisdom: co-creating media for equity and justice*. Cambridge, MA: The MIT Press.

Council of Europe Steering Committee on Media and Information Society (CDMSI) (2025). *Feasibility Study on Benefits and Challenges to Freedom of Expression in Immersive Realities* (adopted 3–5 December 2025). Council of Europe. Available at: https://www.coe.int/en/web/freedom-expression/-/council-of-europe-steering-committee-for-media-and-information-society-adopts-the-feasibility-study-on-freedom-of-expression-in-immersive-realities (Accessed 30 Mar. 2026).

Jensen, L. and Konradsen, F. (2018). A review of the use of virtual reality head-mounted displays in education and training. *Education and Information Technologies* 23(4), 1515–1529. doi:10.1007/s10639-017-9676-0

Jones, S., Dawkins, S. and McDougall, J. (2022). *Understanding virtual reality: challenging perspectives for media literacy and education*. 1st ed. New York: Routledge. doi:10.4324/9780367337032

Kellner, D. and Share, J. (2007). Critical media literacy, democracy, and the reconstruction of education. In: Macedo, D. and Steinberg, S.R., eds. *Media Literacy: A Reader*. New York: Peter Lang Publishing, pp.3–23.

Kourtesis, P. (2024). A Comprehensive Review of Multimodal XR Applications, Risks, and Ethical Challenges in the Metaverse. *Multimodal Technologies and Interaction*, 8(11), 98. doi:10.3390/mti8110098

Lazou, C., Tsinakos, A. and Kazanidis, I. (2025). Enhancing multiliteracy through virtual reality: an inclusive approach to critical immersive-triggered literacy. In: *International Conference on Immersive Learning*, pp.198–215.

Madary, M. and Metzinger, T.K. (2016). Real Virtuality: a code of ethical conduct. Recommendations for good scientific practice and the consumers of VR-technology. *Front. Robot. AI* 3:3. doi:10.3389/frobt.2016.00003

McIntosh, V. and Allen, C. (2024). What do policymakers need to know about harassment in the metaverse? *Front. Virtual Real.* 5:1443384. doi:10.3389/frvir.2024.1443384

Mikelli, D. and Dawkins, S. (2020). VR Kaleidoscope: reconfiguring space and place through community-based media literacy interventions. *Media Practice and Education* 21(1), 54–67. doi:10.1080/25741136.2019.1681223

Rose, M. (2023). The ethics of VR documentary. Available at: http://opendoclab.mit.edu/presents/mandy-rose-ethics-vr-documentary/ (Accessed 30 March 2026).

Slater, M. (2003). A note on presence terminology. *Presence Connect* 3(3), 1–5.

Trudeau, A., Xie, Y., Ketsman, O. and Demir, F. (2023). “Breaking the fourth wall”: the effects of cinematic virtual reality film-viewing on adolescent students’ empathic responses. *Computers & Education: X Reality 2*, 100009. doi:10.1016/j.cexr.2023.100009

Session 3

Justice, Law and Jurisdiction in Immersive Environments

Building the *Lex Metaversi*: A Regulated Self-Regulatory Approach to Dispute Resolution in Virtual Worlds

Aurelio López-Tarruella Martínez[1], Raúl Ruiz Rodríguez[2*]

[1] Professor of Private International Law, University of Alicante, Spain

[2] Assistant Professor of Private International Law, University of Alicante, Spain

*** Correspondence:**
Corresponding Author
r.ruiz@ua.es

Keywords: virtual worlds, *lex metaversi*, online dispute resolution, Digital Services Act, platform governance, fundamental rights

Abstract

This paper explores the challenges of dispute resolution in digital and metaverse environments, highlighting the limitations of traditional legal frameworks in addressing borderless digital relationships. It proposes a shift from territorial to platform-based regulation, focusing on the European Union's approach to resolving disputes within digital platforms. While the framework offers benefits like speed and enforceability, the paper also discusses concerns regarding its scope, coherence, and due process. It advocates for a Regulated Self-Regulatory Approach to ensure fair and effective dispute resolution while protecting fundamental rights.

1 The Digital Paradox of Virtual Worlds: Law Without Territory

1. Resolving disputes in the Metaverse is extraordinarily complex. Virtual worlds operate without physical borders, fixed locations, or ties to any particular country. When a conflict arises (over a digital asset, a smart contract, or the activities of a decentralised organisation) there is no obvious court to turn to, no self-evident national law to apply, and no straightforward way to identify who has authority to decide anything. This structural difficulty is not a technical glitch that better legislation can easily fix; it reflects a fundamental mismatch between the way digital environments work and the way traditional legal systems are built.

National courts and conflict-of-laws rules, the tools States use to decide which country's law governs a cross-border dispute, were designed around territory. They ask questions like: *Where was the contract performed? Where did the harm occur? Where is the defendant domiciled?* These questions presuppose that human activity happens in physical places that can be pinned to a map and assigned to a sovereign. In the metaverse, that presupposition breaks down entirely. A smart contract executes automatically across thousands of computers located in dozens of countries simultaneously. A cryptocurrency theft occurs on a blockchain whose records are replicated across a global network, with no single point of origin. A Decentralised Autonomous Organisation has no registered office, no directors, and no physical presence anywhere. Applying territorially-structured legal rules to these situations does not produce wrong answers; it produces no meaningful answers at all. This is what we

call the *digital paradox:* the attempt to regulate inherently borderless situations through legal tools whose entire logic depends on connecting a dispute to a place.

The problem is not merely technical. Even in cases where existing rules can be stretched to yield a formally applicable law, the justification for privileging one State's legal order over another's evaporates when the relationship was conducted entirely online and no State has any materially stronger connection to it than any other. What is needed, therefore, is not a refinement of existing connecting factors but an entirely different regulatory approach, one that does not depend on territory at all.

2. It is precisely here that alternative dispute resolution mechanisms, and arbitration in particular, become not merely convenient but structurally necessary. Arbitration is already well established in international trade and commerce for familiar reasons: neutrality, specialisation, flexibility, speed, confidentiality, and cost-efficiency. Two features make it especially suited to digital environments. First, arbitral tribunals are not bound to apply the domestic law of any State; they can ground their decisions in trade usages, sectoral custom, and the logic of the technology itself. Second, arbitral awards are enforceable in over 170 countries under the 1958 New York Convention[1], a global enforcement infrastructure that no system of national court judgments can match. These advantages do not merely recommend arbitration for metaverse disputes; given the structural inadequacy of State-based jurisdiction described above, they make it the only realistic option.

2 The Platform as a Forum: ADR in the DSA Framework

3. If territorial categories can no longer adequately capture disputes arising in digital environments, the question becomes: what can? The answer proposed here is the platform itself. Disputes in the digital sphere do not arise ‘on the internet’ in the abstract. Rather, they emerge within a specific platform architecture, governed by particular rules, technical affordances, and enforcement mechanisms. It is therefore the platform, as the environment with which the dispute is most closely connected, that becomes the most plausible site for the organisation of dispute resolution through ADR.

This shift in perspective is evident, to some extent, in various instruments of the EU *acquis communautaire*. The most significant of these is the Digital Services Act (DSA)[2]. Reference can also be made to Regulation 2019/1150 on Platform-to-Business Relations[3], although, as discussed below, this instrument is now unlikely to remain in force. Other relevant examples include Art. 17.9 of Directive 2019/790 on Copyright in the Digital Single Market,[4] Art. 28b of Directive 2018/1808 on Audiovisual Media Services,[5] and Directive 2013/11 on Alternative Dispute Resolution and

2

[1] Convention on the Recognition and Enforcement of Foreign Arbitral Awards (adopted 10 June 1958, entered into force 7 June 1959) 330 UNTS 3

[2] Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market for Digital Services (Digital Services Act) [2022] OJ L 277/1 ('DSA').

[3] Regulation (EU) 2019/1150 of the European Parliament and of the Council of 20 June 2019 on promoting fairness and transparency for business users of online intermediation services [2019] OJ L 186/57.

[4] Directive (EU) 2019/790 of the European Parliament and of the Council of 17 April 2019 on copyright and related rights in the Digital Single Market [2019] OJ L 130/92.

[5] Directive (EU) 2018/1808 of the European Parliament and of the Council of 14 November 2018 amending Council Directive 89/552/EEC and Directive 2010/13/EU on audiovisual media services [2018] OJ L 303/69.

Consumer Protection.[6] Collectively, these instruments highlight the European legislator's increasing reliance on alternative dispute resolution mechanisms in the digital environment. More broadly, they reflect the gradual shift towards the privatisation of dispute resolution in the digital sphere, with such mechanisms being viewed as a fair and expedient means of resolving individual disputes while mitigating the challenges posed by the rapid and transnational nature of the internet for national courts.

4. The DSA offers the clearest expression of this regulatory orientation. It establishes a multi-tiered system of alternative dispute resolution.

At the first tier, Art. 14 requires online platforms to set out in their terms and conditions the procedures, measures, and tools used for content moderation. In doing so, it implicitly acknowledges that moderation may be carried out by algorithmic means, provided that users are informed accordingly. Such moderation may take place either motu proprio or in response to a notice submitted by a third party. In this regard, Arts. 16 and 17 create notice-and-action mechanisms allowing any person to flag illegal content; decisions must be accompanied by a reasoned statement and may be taken by automated means subject to transparency obligations.

At the second tier, Art. 20 creates an internal complaint-handling system for affected parties: appeals must be lodged within six months, free of charge, and the decision must involve 'adequately qualified staff' rather than purely automated means[7]. The Meta Oversight Board is a prominent illustration of such a mechanism[8].

At the third tier, Art. 21 allows users to challenge internal decisions before a certified out-of-court dispute settlement ('ODS') body of their choice, which must resolve the matter within ninety days (extended to one hundred and eighty in complex cases)[9]. In the first half of 2025 alone, ODS bodies reviewed over 1,800 disputes, overturning platforms' decisions in 52% of closed cases[10].

5. The advantages of this architecture over State-court litigation are considerable. Access to justice is simplified; proceedings are fast; the 'law' applied is the platform's own terms of use,[11] so no choice-of-law analysis is required; and enforcement is immediate, since sanctions (content removal, account suspension) are applied by the platform itself, eliminating recognition and enforcement issues entirely. Looking ahead, the accumulation of reasoned decisions by internal systems and ODS bodies could generate a body of interpretive practice, a nascent *lex metaversi*, capable of guiding

[6] Directive 2013/11/EU of the European Parliament and of the Council of 21 May 2013 on alternative dispute resolution for consumer disputes and amending Regulation (EC) No 2006/2004 and Directive 2009/22/EC

[7] Since the DSA entered into application, 30% of the 165 million content moderation decisions appealed through internal mechanisms have been reversed [European Commission, Two years of Digital Services Act allows 50 million content moderation decisions by platforms to be reversed (Digital Strategy, 17 February 2026) https://digital-strategy.ec.europa.eu/en/news/two-years-digital-services-act-allows-50-million-content-moderation-decisions-platforms-be-reversed accessed 27 March 2026].

[8] Meta Oversight Board, *About the Oversight Board* https://www.oversightboard.com/ accessed 27 March 2026.

[9] For further information on their functioning and a list of certified bodies, see: European Commission, *Out-of-Court Dispute Settlement under the Digital Services Act* (Digital Strategy, 28 May 2024) https://digital-strategy.ec.europa.eu/en/policies/dsa-out-court-dispute-settlement accessed 27 March 2026.

[10] European Commission (n 7).

[11] Notably, in the first half of 2025, platforms made 99% of content moderation decisions based on their own internal rules (such as community standards or terms of service), rather than making decisions to remove content that was considered illegal under EU or national law [ibid.].

future decisions and ultimately training the algorithms that process first-instance moderation requests.

3 Limitations and Structural Challenges

6. The DSA's ADR framework, for all its promise, suffers from significant limitations. Four merit particular attention:

3.1 Coherence and the risk of fragmented jurisprudence

7. A spontaneously emerging *lex metaversi* presupposes decisional coherence: like cases must be decided alike. Internal complaint systems, operating within a single platform, can in principle ensure this. The ODS tier is more problematic. Users may bring cases before any of the currently seven certified ODS bodies operating across the EU, with the consequence that divergent bodies may reach inconsistent outcomes on materially identical questions. The bi-annual reporting obligation imposed on Digital Services Coordinators (Art. 21.4) provides a potential harmonisation instrument, and the European Board for Digital Services could contribute through its mandate to 'contribute to the consistent application' of the Regulation (Art. 6.2.a). Yet, whether these soft mechanisms will prove sufficient to generate genuine uniformity remains an open question for the time being.

3.2 The privatisation of justice and due process concerns

8. The platform-centred model represents a form of privatisation of adjudication that raises serious due process concerns.

To begin with, where a dispute resolution mechanism is designed, administered, and ultimately controlled by the platform itself, there is an obvious risk of structural bias in favour of the platform's own interests. That risk is exacerbated by the marked asymmetry between platforms and users, since the imbalance in economic power, technical knowledge, and procedural control may significantly undermine the user's right to a fair process. Where moderation decisions are taken, or even assisted, by automated systems, the danger of algorithmic bias introduces a further layer of concern.

9. Emerging blockchain-based dispute resolution mechanisms (such as Kleros, Aragon and Metacourt) present concerns of a different order. In these systems, decisions are reached by a vote of token-holding jurors who are rewarded for voting with the majority. This architecture systematically incentivises predicting the modal outcome rather than reaching the legally correct one; the most popular outcome is not necessarily the most just, which raises serious questions about legal justification. Equally, there is no guarantee that jurors will possess subject-matter expertise, nor are arbitrators able to ground their decisions in judicial precedent. The model also raises concerns about effective protection. The parties' ability to challenge evidence submitted by the opposing side is significantly curtailed, and, more fundamentally, the system risks a denial of justice in low-value or unattractive disputes where too few jurors volunteer to participate. A further concern goes to transparency: both parties and arbitrators may act anonymously, which not only undermines the integrity of the process but may also preclude recognition of awards under the New York Convention.

10. Underlying all of these mechanisms is a deeper tension. Content moderation decisions often involve balancing competing fundamental rights, many of which are enshrined in the Charter of

Fundamental Rights of the European Union (CFREU)[12], such as: the right to freedom of expression (Art. 11) versus the right to privacy (Art. 7), intellectual property (Art. 17), child protection (Art. 24), non-discrimination (Art. 21), and the freedom to conduct a business (Art. 16). As digital platforms have increasingly become quasi-public spaces for social interaction and public discourse, the governance of speech on these platforms has taken on a constitutional dimension, one that purely contractual dispute resolution mechanisms are ill-equipped to address.

3.3 The inadequacy of existing moderation tools for immersive environments

11. A final significant challenge in metaverse dispute resolution stems from the limitations of traditional content moderation systems (a problem not exclusive to ODR mechanisms, but one that these mechanisms cannot avoid). These systems, designed primarily for static and text-based interactions, are ill-equipped to handle the dynamic, real-time, and immersive nature of metaverse environments. Harmful behaviours in the metaverse, such as harassment through gestures, voice chat, or manipulation in spatial settings, often evade detection by conventional moderation tools, leaving no persistent, reviewable record for the notice-and-action mechanisms under Arts. 16–17 DSA or the internal complaint-handling system under Art. 20 to act upon. When the harmful act is a gestural intrusion or a voice interaction in a shared virtual space, there may be nothing to flag, nothing to review, and no basis for issuing a reasoned decision under Art. 17.1 DSA.

This gap not only impedes effective enforcement but also raises fundamental questions about the adequacy of existing legal frameworks, which assume clear, persistent, and attributable content, assumptions that are structurally undermined by the nature of metaverse environments. Addressing this challenge will ultimately require purpose-built tools (including AI-driven detection, spatial moderation capabilities, and adapted evidentiary standards) designed from the outset for the realities of immersive environments rather than retrofitted from frameworks conceived for a different digital era. Yet the adoption of such tools raises its own questions, concerning their democratic legitimacy, their compatibility with fundamental rights, and their integration within existing legal frameworks that were not conceived with them in mind.

4 Towards Regulated Self-Regulation: Safeguards for a Legitimate *Lex Metaversi*

12. The difficulties canvassed above do not invalidate the platform-centred ADR model. They do, however, demonstrate that self-regulation without State oversight is insufficient. What is required is a Regulated Self-Regulation (SRS) framework in which public authorities set minimum standards of due process and substantive rights protection that platform-operated ADR systems must respect.

On the *procedural* side, the DSA already provides important guarantees. To begin with, must be informed of available avenues of redress and have effective access to them (Art. 16.1, Art. 20.3, Art. 21.1, second subparagraph). Decisions must be made in a timely, diligent, objective, and non-arbitrary manner (Art. 14.4, Art. 16.2, Art. 20.4). Where a restriction is imposed, Art. 17.1 requires a clear and specific statement of reasons, including the decision, the facts, the use of automated means, and the legal or contractual basis for the measure. There is no such requirement for decisions rejecting a notice (Art. 16.5). Users must also be informed of available remedies (Art. 17, Art. 20.5), and the final decision cannot be made solely by automated means (Art. 20.6). The system remains voluntary, allowing users to bring proceedings before a judicial body (Art. 21.1, third subparagraph), and parties must cooperate in good faith (Art. 21.2). The DSA also guarantees the quality of out-of-

[12] Charter of Fundamental Rights of the European Union (2000) OJ C 364/1.

court dispute settlement bodies, which must be certified by the Digital Services Coordinator (Art. 21.3). Certification depends on conditions ensuring impartiality, expertise, accessibility, and cost-efficiency, with biennial reports on their functioning. These requirements are without prejudice to those in Directive 2013/11.

On the *substantive* side, Art. 14.4 DSA represents a significant innovation: it imposes on private platforms an express obligation to give 'due regard' to the fundamental rights of service recipients when enforcing their terms. This is, to the current authors' knowledge, the first time a Union legislative instrument has expressly required private actors to respect fundamental rights *inter partes*, giving legislative expression to what the CJEU had previously recognised as horizontal direct effect[13]. Whether this general provision suffices to guarantee rights-compliant moderation in the metaverse remains doubtful. Additionally, AI content moderation systems were not classified as high-risk under the AI Act[14], limiting independent algorithmic scrutiny, a legislative choice whose wisdom is contestable[15]. This becomes all the more so given the role AI real-time tools may play in adapting content moderation to the ever-changing nature of metaverse interactions, from voice-based exchanges to gestural and spatial forms of harm.

13. As per the role of ODS, scholarship is divided on whether ODR bodies should serve as guardians of fundamental rights compliance. Several tensions and ambiguities in the DSA framework underlie this debate.

On one hand, the annual reports of ODR bodies and the biennial reports from coordinators could help assess platform moderation practices and suggest improvements. However, Art. 21.4 does not clearly require platforms to be informed of these reports or to act upon them, leaving their practical impact uncertain. Some scholars go further, arguing that ODR mechanisms should actively monitor compliance with Art. 14.4 DSA, ensuring platforms act diligently, proportionately, and with due regard for fundamental rights. Yet Art. 21 does not assign this responsibility to ODR bodies, leaving their role fundamentally unclear.

A related concern is the non-binding nature of ODR decisions. While this may appear troublesome at first glance, repeated non-compliance could expose platforms to reputational damage, and such patterns may, in turn, qualify as a systemic risk under Art. 34 DSA. Other scholars, however, caution against assigning fundamental rights oversight to ODR bodies altogether, on the grounds that they lack the authority, experience, and mandate for such a role. Instead, they suggest that compliance with Art. 14.4 should be assessed through the risk evaluation framework established under Art. 34.

6

[13] The horizontal effect of fundamental rights in EU law is recognised in cases such as Egenberger [*Case C-414/16 Egenberger v Evangelisches Werk für Diakonie und Entwicklung eV* (Court of Justice of the European Union, 17 April 2018) ECLI:EU:C:2018:257] and Viking Line [*Case C-438/05 Viking Line ABP v The International Transport Workers' Federation* (Court of Justice of the European Union, 11 December 2007) ECLI:EU:C:2007:772].

[14] 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) [2024] OJ L, 2024/1689 (AI Act)

[15] AI content moderation systems were included in the high-risk category in early legislative drafts but were ultimately removed from Annex III.

Regardless of this ongoing debate, it is worth emphasising that ODR proceedings are voluntary (Art. 21.2). Users retain the option to assert their rights before the Digital Services Coordinator in their place of residence or establishment (Art. 53), or before the judicial courts, ensuring that access to redress is not contingent on the resolution of these institutional questions.

5 Conclusion

14. This paper has argued that the structural inadequacy of PIL's territorial categories for digital and metaverse disputes is not a contingent deficiency capable of technical repair. It reflects a paradigmatic incompatibility between a regulation method premised on State-territorial connection and relationships that are constitutively deterritorialised. The response is to locate the regulatory forum at the platform, the environment with which the dispute is most closely connected, and to build ADR mechanisms within that platform.

The DSA's multi-tiered ADR framework represents a substantial step in this direction, offering meaningful advantages in terms of access, speed, and practical enforceability. The accumulation of reasoned decisions at both internal and ODS levels holds the promise of a spontaneously generated *lex metaverse*, this is, a body of soft law capable of guiding platform governance in the absence of applicable State law. Yet the system has important structural limitations: it remains largely confined to tortious disputes; it does not address contractual claims or provide for compensatory remedies; its coherence across multiple ODS bodies is not guaranteed; and the privatisation of adjudication it entails poses real risks to due process and fundamental rights, particularly where moderation decisions affect the exercise of free expression at scale.

The conclusion is not, by any means, that self-regulation should be abandoned, but rather, that it must operate within a RSR framework of minimum procedural and substantive standards, monitored by independent public authorities. The DSA provides the foundations; the task for legislators and courts alike is to strengthen them, while allowing users to retain at every stage, the right to submit their claims to a national court or to a competent Digital Services Coordinator. The platform may be the natural forum for digital disputes. It cannot, however, be the final one.

6 References

Husovec, Martin**,** *Principles of the Digital Services Act* (Oxford University Press 2024).

Ruschemeier, Hannah and Quintais, João Pedro**,** 'Fundamental Rights in Out-of-Court Dispute Settlement under the Digital Services Act' (10 December 2025) SSRN https://ssrn.com/abstract=5900422.

López Rodríguez, Ana Mercedes**,** 'Inmuebles virtuales y perspectivas innovadoras de resolución de conflictos en el metaverso' (2024) 16(1) *Cuadernos de Derecho Transnacional* 319.

— 'Towards a Lex Metaversi? The Law Applicable to Virtual RealEstate in the Metaverse' (2024) 7(2) *Interactive Entertainment Law Review* 110 https://doi.org/10.4337/ielr.2024.02.05

— *Resolución de Conflictos en el Metaverso* (Tecnos, 2025)

Fernández Masía, Enrique**,** 'Medios alternativos de solución de controversias entre las plataformas en línea y los destinatarios del servicio en el Reglamento europeo de servicios digitales',

in Castello Pastor, José Javier (ed), *Análisis del Reglamento UE de servicios digitales y su interrelación con otras normas de la Unión Europea* (Aranzadi 2024) 137.

Council of Europe and IEEE Standards Association, *The Metaverse and its Impact on Human Rights, the Rule of Law and Democracy* (2024) https://www.coe.int/en/web/freedom-expression/-/full-version-of-the-report-on-the-metaverse-and-its-impact-on-human-rights-rule-of-law-and-democracy-available-for-download accessed 27 March 2026.

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Virtual reality to quantify rehabilitation of violent behaviour

Maria V. Sanchez-Vives[1,2], Tania Johnston[1], Nicolas Barnes[3]

[1]Institute of Biomedical Research August Pi i Sunyer (IDIBAPS), Barcelona, Spain

[2]ICREA, Barcelona, Spain

[3]General Directorate of Prison Affairs, Justice Department, Government of Catalonia, Barcelona, Catalonia, Spain

*** Correspondence:**
Maria V. Sanchez-Vives
msanche3@recerca.clinic.cat

Keywords: rehabilitation programs, VR, justice, prisons, intimate partner violence, embodiment, perspective-taking, empathy

Abstract

Assessing the success of rehabilitation programs for violent offenders is often hindered by the social desirability bias inherent in self-reports and the ethical impossibility of observing behavior in high-stakes, real-world situations. This paper proposes immersive virtual reality (VR) as a standardized tool to quantify the degree of rehabilitation achieved, regardless of whether the intervention was VR-based or a traditional program. While VR has been successfully implemented for the rehabilitation of aggressors both in the community and within prison settings—leveraging unique possibilities for embodiment, perspective-taking, and the experience of aggression from the victim's point of view—its potential for quantification remains less explored. Research has already demonstrated that these immersive interventions significantly impact emotion recognition and brain activation patterns associated with empathy circuits. In this presentation, we discuss the use of VR as an environment to evaluate individuals who have completed a rehabilitation program. By placing them in standardized scenarios as either an observer or an embodied participant, their responses and behaviors can be analyzed to reveal objective pre-to-post rehabilitation changes. This approach could provide a scientifically rigorous framework for determining the ecological validity of therapeutic outcomes following rehabilitation of violent behaviour.

1 Introduction

Assessing the success of rehabilitation programs for violent offenders is not trivial, and tests and questionnaires are affected by the social desirability bias inherent in self-reports and the ethical impossibility of observing behavior in real-world situations. Currently, determining whether an offender has successfully integrated the cognitive and emotional tools provided by a rehabilitation program relies heavily on post-hoc metrics like recidivism or subjective scales. Recidivism though reflects a failure of the system to detect and respond to risk, and its seriousness increases with the level of violence involved. There is thus a critical need for standardized, objective tools that can quantify behavioral and neurocognitive changes in a safe and ecologically valid environment (Torao-Angosto & Sanchez-Vives, 2024).

Immersive virtual reality (VR) has emerged as a transformative technology in this field. The core of this potential lies in the phenomena of embodiment and perspective-taking. By inducing a "body ownership illusion," VR allows researchers to place an individual in a situation not only as an observer but also as a participant—or even as a victim—of an aggressive event (Seinfeld et al., 2018; Gonzalez-Liencres et al., 2020). Previous research has already established the efficacy of VR in the rehabilitation of intimate partner violence (IPV) perpetrators, demonstrating that embodying a victim can lead to significant improvements in emotion recognition (Seinfeld et al., 2018; 2023; Gonzalez-Liencres et al., 2020) and can modulate brain activation patterns within empathy circuits, such as the Default Mode Network (Seinfeld et al., 2021; deBorst et al., 2020).

While the therapeutic potential of VR is well-documented (Johnston et al., 2023), its use as a tool to quantify the degree of rehabilitation remains less explored. We propose that VR can function as a "behavioral stress test" to evaluate progress regardless of the intervention modality (Barnes et al., 2022). For instance, by placing an individual in standardized virtual scenarios—such as witnessing an escalating conflict or experiencing a domestic dispute from a child's perspective (Seinfeld et al., 2023)—clinicians can analyze real-time interventions, behavioral responses, and physiological markers. These objective metrics may deliver a more authentic reflection of an individual's neurocognitive state and their readiness for social reintegration, providing a scientifically rigorous framework for determining the validity of therapeutic outcomes in violent behavior.

1.1 The Neurobiology of Embodiment: Bypassing Cognitive Filters

Immersive Virtual Reality (IVR) offers a unique solution by leveraging the neurobiological mechanism of **embodiment**. Through the "body ownership illusion," the brain treats the virtual body as its own, shifting the user's perspective at a pre-reflective level (Slater & Sanchez-Vives, 2016). Unlike traditional role-playing, IVR bypasses the conscious "filters" of the offender. When a perpetrator embodies a victim, the experience triggers empathy circuits and modulates the Default Mode Network, as evidenced by changes in neural responses to emotional facial expressions (Seinfeld et al., 2021). By targeting these automatic responses, VR provides an implicit learning and window into the individual's emotional processing—such as the correction of hostile attribution biases—that is often access to transform through conscious effort (Seinfeld et al., 2018; Johnston et al., 2023).

1.2 VR as a Behavioral Assay

We propose that VR can function as a standardized "behavioral assay" to quantify rehabilitation progress through different approaches. Two of them are suggested here: 1) The Observer Perspective: Quantifying Bystander Intervention In this paradigm, the individual is placed in a virtual environment as a witness to a conflict (e.g., an escalating argument or a domestic dispute). This allows for the quantification of violence acceptance thresholds. By measuring if, when, and how an individual decides to intervene to de-escalate a situation, researchers can assess the internalization of prosocial norms. Studies on bystander behavior in VR show that individuals' actions in these environments are highly predictive of their underlying attitudes toward aggression, providing a metric of "prosocial readiness" (Tuente et al., 2019). 2) The Embodied Perspective: By placing the rehabilitated individual in the body of a victim or a child experiencing abuse (Seinfeld et al., 2023) sensitivity, tolerance and acceptability of violent behaviour can be detected. Metrics such as the accuracy and speed of identifying fear in the virtual antagonist, combined with physiological data may serve as objective indicators of empathy-related growth (Gonzalez-Liencres et al., 2020). Measurements pre and post rehabilitation trainings should demonstrate a heightened sensitivity to the victim's distress signals and a reduced "hostile attribution bias" (Tuente et al., 2020). Other situations that may trigger violent responses could be investigated as a standarized measure.

1.3 Ecological Validity vs. Ethical Safety: "High-Stakes" Testing Without Risk

One of the greatest challenges in forensic psychology is the "risk-safety" paradox: we cannot know if an offender will remain non-violent in a high-pressure situation until they are actually in one. VR resolves this by providing ecological validity—the psychological "presence" of being in a real situation—without the ethical or physical risk of actual harm (Blascovich et al., 2002). This allows for the simulation of "high-stakes" triggers (e.g., a heated provocation or an emotional conflict) to see if the individual utilizes the coping mechanisms learned during rehabilitation under realistic levels of arousal (Barnes et al., 2022).

1.4 From Research to Prison Systems: Practical Considerations

Implementing VR as a quantification tool in prison systems requires careful integration within well defined ethical frameworks in which participants retain voice, consent, and meaningful control over immersive experience and data. In this way, it can provide a standardized clinical assessment that supplements existing reports. Practical deployment (as discussed in Barnes et al., 2022) must consider: 1) **Standardization:** Using identical virtual scenarios for all evaluations to ensure forensic fairness. 2) **Multimodal Data:** Combining behavioral choices in VR with physiological markers. 3) **Ethical Frameworks:** Establishing clear guidelines on how VR behavioral data influences clinical decisions and parole evaluations (Torao-Angosto & Sanchez-Vives, 2024).

1.5 Conclusion: Towards a Quantifiable Model of Change

The transition from viewing VR solely as a medium for intervention to using it as a tool for quantification represents a paradigm shift in forensic rehabilitation. By moving from reports and questionnarires to behavioural measures in immersive environments, one can obtain a more accurate, neurobiologically grounded profile of an individual's progress. This model of change may offers a rigorous path forward for enhancing public safety and ensuring that rehabilitation efforts have truly transformed the individual's behavioral response to violence and it deserves further investigation. It is also important to define the ethical and legal framework of this approach, and to ensure that its use is bounded by dignity, transparency and accountability.

2 References

Barnes, N., Sanchez-Vives, M. V., & Johnston, T. (2022). On the practical use of immersive virtual reality for rehabilitation of intimate partner violence perpetrators in prison. *Frontiers in Psychology*, 13, 787483.

Blascovich, J., et al. (2002). Immersive virtual environment technology as a methodological tool for social psychology. Psychological Inquiry, 13(2), 103-124.

de Borst, A. W., Sanchez-Vives, M. V., Slater, M., & de Gelder, B. (2020). First-person virtual embodiment modulates the cortical network that encodes the bodily self and its surrounding space during the experience of domestic violence. Eneuro, 7(3).

Fisher, R. J. (1993). Social desirability bias and the validity of indirect questioning. Journal of Consumer Research, 20(2), 303-315.

Gonzalez-Liencres, C., Zapata, L. E., Iruretagoyena, G., Seinfeld, S., Perez-Mendez, L., Arroyo-Palacios, J., ... & Sanchez-Vives, M. V. (2020). Being the victim of intimate partner violence in virtual reality: first-versus third-person perspective. *Frontiers in Psychology*, 11, 507601.

Johnston, T., Seinfeld, S., Gonzalez-Liencres, C., Barnes, N., Slater, M., & Sanchez-Vives, M. V. (2023). Virtual reality for the rehabilitation and prevention of intimate partner violence–From brain to behavior: A narrative review. *Frontiers in Psychology*, 13, 788608.

Klein Tuente, S., et al. (2020). The effect of Virtual Reality Aggression Prevention Training (VRAPT) on various aspects of aggression. Journal of Clinical Medicine, 9(10), 3334.

Sanchez-Vives, M. V., Johnston, T., & Barnes, N. (2026). Virtual reality to quantify rehabilitation of violent behaviour. [Draft Manuscript].

Seinfeld, S., Arroyo-Palacios, J., Iruretagoyena, G., Hortensius, R., Zapata, L. E., Borland, D., ... & Sanchez-Vives, M. V. (2018). Offenders become the victim in virtual reality: impact of changing perspective in domestic violence. *Scientific Reports*, 8(1), 2692.

Seinfeld, S., Hortensius, R., Arroyo-Palacios, J., Iruretagoyena, G., Zapata, L. E., de Gelder, B., ... & Sanchez-Vives, M. V. (2023). Domestic violence from a child perspective: impact of an immersive virtual reality experience on men with a history of intimate partner violent behavior. *Journal of Interpersonal Violence*, 38(3-4), 2654-2682.

Seinfeld, S., Zhan, M., Poyo-Solanas, M., Barsuola, G., Vaessen, M., Slater, M., Sanchez-Vives, M. V., & de Gelder, B. (2021). Being the victim of virtual abuse changes default mode network responses to emotional expressions. *Cortex*, 135, 268-284.

Slater, M., & Sanchez-Vives, M. V. (2016). Enhancing our lives with immersive virtual reality. Frontiers in Robotics and AI, 3, 74.

Torao-Angosto, M., & Sanchez-Vives, M. V. (2024). Realidad virtual en salud mental y rehabilitación de conductas violentas. *Fonseca, Journal of Communication*, 28, 10-46.

Tuente, S., et al. (2019). Virtual Reality Aggression Prevention Training for forensic psychiatric patients (VRAPT). Trials, 20(1), 1-13.

van Gelder, J. L. (2020). Digital technology and the future of criminological research. Journal of Quantitative Criminology, 36(4), 651-665.

3 Conflict of Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

4 Author Contributions

All authors contributed to the experimental studies and the writing of the paper.

5 Funding

META-TOO project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No. 101160266.

Virtual Reality and Rehabilitation: A Human Rights-Based Approach for the Treatment of Gender-Based Violence in the Penitentiary System of Catalonia

Nicolás Barnes Méndez[1,2]

[1]General Directorate of Prison Affairs, Justice Department, Government of Catalonia, Barcelona, Catalonia, Spain

[2]Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), Barcelona, Catalonia, Spain

*** Correspondence:**
Nicolás Barnes Méndez
nbarnes@gencat.cat

Keywords: Virtual Reality, Gender-Based Violence, Rehabilitation, Reintegration, Human Rights, Mandela Rules, Embodiment, Criminal Justice

Abstract

This technical study analyzes the implementation of Virtual Reality (VR) as a neuropsychological tool in the rehabilitation of gender-based violence offenders in the penitentiary system of Catalonia. The research focuses on the embodiment technique, evaluating how visuomotor synchrony and sensorimotor contingencies facilitate the reduction of cognitive distortion and increase affective resonance toward victims. The intervention is structured through an "integral and multimodal itinerary," ensuring that the use of Extended Reality (XR) is not an isolated activity but a technical, pedagogical, and therapeutic process strictly supervised by professionals. The primary objective is to strengthen inmates' social competencies and mental health, facilitating their reintegration into an increasingly digital society within an ethical framework that protects fundamental rights. The main findings indicate a statistically significant improvement in empathic capacity and high clinical adherence, validating the effectiveness of immersive technologies for social reintegration. This work demonstrates that technological innovation, under a rigorous ethical framework, is essential for upholding the rule of law and democratic values in criminal justice, consistent with the Council of Europe's guidelines.

1 Introduction

The integration of immersive technologies (XR) into the justice system represents a paradigm shift in penitentiary treatment. Within the framework of the Strasbourg conference and following the recent "Feasibility Study on Freedom of Expression in Immersive Realities" by the CDMSI (Council of Europe/CDMSI, 2024), it is essential to analyze how these tools affect the protection of fundamental rights. The core challenge lies in developing rehabilitation methods for gender-based violence offenders that overcome the limitations of conventional therapy, particularly regarding cognitive resistance and lack of empathy.

The aim of this document is to present the success of the model applied in Catalonia as a benchmark for technological innovation. This project not only seeks therapeutic effectiveness but is grounded in

respect for human dignity, aligning with the Council of Europe's mandate on the use of technology to strengthen democratic values (Council of Europe/EMRN, 2026). The adoption of VR in Catalan prisons is presented here as an advanced tool that meets the ethical and legal demands of a modern justice system (Barnes et al., 2022; Barnes et al., 2024).

2 Regulatory Framework: The Mandela Rules and Human Rights

The implementation of VR in environments involving deprivation of liberty must be analyzed through the lens of the United Nations Standard Minimum Rules for the Treatment of Prisoners (United Nations, 2015).

Rule 1 (Inherent dignity): The VR intervention in Catalonia is designed under a strictly pedagogical and non-punitive principle. By avoiding re-victimization and ensuring that inmates are treated with inherent dignity, the program ensures that the technology becomes a means of human development, eliminating discriminatory bias in access to advanced treatment.

Rule 4 (Rehabilitation and social protection): The ultimate purpose of a sentence is the protection of society through reduced recidivism. VR justifies its application as a superior reintegration tool; by improving affective resonance and prosocial competencies in offenders, it fulfills the mandate of preparing inmates for a crime-free life, thus protecting potential victims and society at large. This approach directly responds to the need to ensure that immersive realities protect well-being and integrity within the justice system, as promoted by the Council of Europe.

3 Methodology and Project Context in Catalonia

The intervention program developed in Catalonia's penitentiary centers is based on experimental forensic psychology and the use of virtual embodiment.

3.1 The "Integral and Multimodal Itinerary"

The concept of the "integral and multimodal itinerary" represents the procedural and technical foundation of the intervention in Catalonia (Generalitat de Catalunya, 2021). This approach establishes that VR is a complementary tool that must be integrated into a broader treatment plan, combining advanced devices with traditional cognitive-behavioral therapy (CBT).

3.2 Technical Requirements and Device Security

Given the nature of the penitentiary environment, the methodology primarily employs standalone VR devices configured in offline mode. This ensures IT security and prevents unauthorized communications while allowing the necessary mobility for therapeutic or skills-training sessions. In specific cases (experimental projects), biometric indicators such as heart rate variability (HRV) are monitored to objectively evaluate inmates' stress responses to virtual stimuli and adjust session intensity in future redesigns of virtual scenarios.

3.3 Phases of the Itinerary

The intervention follows a rigorous pedagogical and clinical structure:

Criminological Suitability Assessment: Detailed analysis of the inmate's profile, mental health history, and specific rehabilitation needs. This determines whether VR is the optimal tool to address deficits such as low frustration tolerance or lack of empathy.

Preparation and Psychoeducation: Preliminary sessions explaining the technology's operation and establishing therapeutic objectives, reducing anxiety about device use.

Controlled Immersive Exposure: Execution of virtual sessions under the direct supervision of a professional (psychologist or social educator). Human presence is essential to guide the experience and regulate emotional impact. Using high-fidelity devices, inmates experience an aggression scenario from the victim's perspective. The success of this phase relies on sensorimotor contingencies and visuomotor synchrony, generating a body-ownership illusion that enables a first-person experience of vulnerability (Seinfeld et al., 2018).

Transfer and Clinical Debriefing: After the immersion, a processing session analyzes thoughts and behaviors that emerged in the virtual environment. The goal is to anchor digital learning and transform it into competencies applicable in the physical reality of prison and future life in freedom. This model ensures that technology is not an escape mechanism but a tool for prosocial confrontation and assisted learning.

4 Results: Success of the Immersive Intervention

Findings from the inmate population in Catalonia confirm the viability of XR in the criminological field (Barnes et al., 2022; Barnes et al., 2024). The success of the intervention is built on three pillars:

Effectiveness in empathy and perspective-taking: Technology enables deep affective resonance. By experiencing empathy bidirectionally through the victim's body, inmates show significant reductions in harm minimization and improved identification of early signs of violence.

Safety and control of the therapeutic environment: VR offers an ecologically valid yet safe environment where physiological and emotional responses to tension can be monitored without real-world risk, allowing precise modulation of therapeutic exposure.

Acceptance among inmates and technical staff: Clinical adherence surpasses that of traditional control groups. Staff value the transparency of the process and the tool's ability to break offenders' verbal defense barriers.

5 Global Perspective: Innovation in Justice and International Frameworks

The Catalan project aligns with international recommendations on the use of advanced technologies in criminal justice. A critical aspect is compliance with transparency, ensuring virtual scenarios are free of biases that could hinder rehabilitation. Likewise, the Catalan model applies the Human-in-the-loop principle. Unlike automated systems, VR in Catalonia remains subordinate to the therapist's judgment, who acts as ethical and professional guarantor (Johnston et al., 2023). This ethical implementation shows that technological innovation can and must serve justice without compromising direct human oversight.

6 Discussion and Conclusions

The integration of VR raises challenges within the framework of health, ethics, and well-being in virtual environments. However, Catalonia's experience shows that when technology is used for clear rehabilitative purposes and under human-rights-based standards, social benefits in terms of reduced recidivism and victim protection are undeniable.

The success in Catalonia confirms that VR is a viable tool for fulfilling the rehabilitative objectives of the Mandela Rules (United Nations, 2015), providing an operational and technological methodology that ensures the effective application of reintegration theory. Virtual embodiment technology is effective in addressing the cognitive roots of gender-based violence, enabling an intervention that spoken language alone cannot achieve (Seinfeld et al., 2018; Barnes et al., 2022). The adoption of these standards is recommended for other European penitentiary systems, ensuring that innovation is always supervised by human rights bodies to safeguard dignity and legal security.

7 Conflict of Interest

The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

8 Author Contributions

NB: Conceptualization, methodology, investigation, writing – original draft, writing – review and editing.

9 Funding

The author declares that no financial support was received for the research, authorship, and/or publication of this article.

10 Data Availability Statement

The data analyzed in this study are available upon reasonable request to the corresponding author, subject to applicable data protection regulations.

11 References

Barnes, N., Sanchez-Vives, M. V., and Johnston, T. (2022). On the practical use of immersive virtual reality for rehabilitation of intimate partner violence perpetrators in prison. Front. Psychol. 13:787483. doi: 10.3389/fpsyg.2022.787483

Barnes, N., Torao-Angosto, M., Slater, M., and Sanchez-Vives, M. V. (2024). Virtual reality for mental health and in the rehabilitation of violent behaviours. Fonseca, Journal of Communication, 28(1), 10–46. doi: 10.48047/fjc.28.01.04

Council of Europe/CDMSI. (2024). Feasibility Study on Freedom of Expression in Immersive Realities. Strasbourg: Council of Europe.

Council of Europe/EMRN. (2026). Call for Papers: Human Rights in Immersive Realities (XR): Freedom of Expression, Justice, Children's Rights and Well-Being. Council of Europe.

Generalitat de Catalunya, Department of Justice. (2021). The Multimodal Itinerary for the Treatment of Domestic Violence Offenders: Implementation Report. Barcelona: Government of Catalonia.

Johnston, T., Seinfeld, S., Gonzalez-Liencres, C., Barnes, N., Slater, M., and Sanchez-Vives, M. V. (2023). Virtual reality for the rehabilitation and prevention of intimate partner violence – From brain to behavior: A narrative review. Front. Psychol. 13:788608. doi: 10.3389/fpsyg.2022.788608

Seinfeld, S., Arroyo-Palacios, J., Iruretagoyena, G., Hortensius, R., Zapata, L. E., Borland, D., de Gelder, B., Slater, M., and Sanchez-Vives, M. V. (2018). Offenders become the victim in virtual reality: impact of changing perspective in domestic violence. Sci. Rep. 8:2692. doi: 10.1038/s41598-018-19987-7

United Nations. (2015). The United Nations Standard Minimum Rules for the Treatment of Prisoners (the Nelson Mandela Rules). New York: United Nations.

Session 4

Health, Ethics and Well-Being in XR

Workers' rights, health and safety, and employer responsibilities in immersive realities (XR)

Verity McIntosh[1]

[1]Digital Cultures Research Centre, University of the West of England, Bristol, United Kingdom

*** Correspondence:**
Verity McIntosh
verity.mcintosh@uwe.ac.uk

Keywords: work, immersive, virtual worlds, workers' rights, health and safety, teleoperation

Abstract

As immersive technologies move slowly into the mainstream, many employers are starting to incorporate tools such as virtual, augmented and mixed reality into their workflows. One of the more appealing 'features' of these technologies is the ability to simulate worlds, tasks and scenarios in virtual environments that would not be considered practical, or even ethical in the physical environment. Scenarios such as disaster recovery, crisis management, civil unrest and combat can be manifested in ways that are immersive, embodied, and psychologically engaging, whilst seemingly posing little physical risk to the user. This paper problematizes the prioritization of physicality in health and safety protections for modern workers, and foregrounds the need to revisit and improve current workplace safety provisions for workers who are expected to undertake duties in immersive and augmented environments.

A novel 'three planes of reality' framework is introduced in support of these aims, illustrating the layered dynamics of a worker's experience in immersive environments, and alluding to the distinct affordances and implications of working across 'corporeal', 'simulated' and 'actuated' realities. This is a preliminary framework designed to assist workers in advocating for their rights, employers in assessing and mitigating risks, and policymakers in considering the legal and regulatory instruments required to protect the rights of citizens as the modern workplace increasingly incorporates multi-layered hybridity.

1 Immersive Technology in the Workplace

In recent years, immersive technologies such as virtual, augmented and mixed reality headsets and smart glasses have begun to find their niches in sectors as broad as architecture, the arts, engineering, manufacturing, training and healthcare.

For many tech companies, the enterprise market has become central to technology development cycles and deployment. When it launched in 2023, Apple's mixed reality headset, the Apple Vision Pro was described in terms of productivity, offering an "infinite canvas for apps" and portraying images of workers layering virtual screens and 3D digital assets onto their regular field of view (Apple, 2023). In 2021 Microsoft fully pivoted their flagship mixed reality Hololens headset offer away from consumer, and towards enterprise markets, notably securing a contract with the US military worth a reported $22b (Nellis & Dave, 2021). And, in a 2025 economic impact report

commissioned by Meta, a series of bold, if not well evidenced claims are made regarding the financial and operational benefits of using their enterprise technology, including an apparent ‘average return on investment of 219%’ (Cavallaro, 2025).

For businesses, a key selling points of these technologies is the ability to simulate spaces, scenarios and interactions that would be considered ‘unsafe’ were they to be enacted in physical reality. However, the sales pitches often fail to engage with the decades-long research into the impact of the sense of presence that can be achieved using these same tools. Much less do they mention the incumbent risks of placing people in psychologically convincing environments that can trigger responses similar to those experienced in analogous physical scenarios. (Lee, 2004; Madary & Metzinger, 2016; Ratan, 2012; Slater, 2009; Yee & Bailenson, 2007). A few examples of the disparity in contemporary discourse:

“VR gives us the best of all worlds: the ability to create a believable, engaging and repeatable emergency but without the associated danger of the real thing” (Dalton, 2021)

and

“By moving would-be real-life training environments in potentially hazardous or consequential situations to virtual environments on Meta Quest, learners at the composite can practice and train more frequently in safe VR environments that might not be safe or feasible in person” (Cavallaro, 2025)

In sharp contract to:

“virtual reality pieces that hack the user’s brain into feeling present in a virtual space and embodied in a virtual body require a higher level of ethical interrogation” (Sinclair, 2020)

and

“Torture in a virtual environment is still torture. The fact that one’s suffering occurs while one is immersed in a virtual environment does not mitigate the suffering itself” (Madary and Metzinger, 2016)

The notion of immersive environments as ‘safe’ places to simulate unsafe practices is far from resolved, and yet applications in industry are proliferating. This presents challenges for employers and employees seeking to understand what responsible, or even legally compliant practice might be when immersive technology is integrated into the working environment.

A modest but quickly advancing segment of the immersive tech enterprise marketplace concerns the use of virtual simulations that are ‘digitally twinned’ with physical environments. Actions taken by a worker in a virtual or mixed reality simulation can have directly correlated consequences in the physical realm e.g. by a VR user embodying a robot or remotely piloting a drone.

In a recent study involving the author of this paper, researchers collaborated with robot teleoperation teams at Sellafield, a highly contaminated nuclear facility undergoing decommissioning and hazard clearance. Workers used virtual reality to embody and remotely operate industrial robots in contaminated spaces, allowing them to apply their spatial awareness and expertise to the task whilst protected from bodily risks such as radiation poisoning and falling debris. The study introduced a number of designed sounds into the virtual reality simulation, sonifying non-visual data such as

radiation, temperature and gas to help users in VR to make informed decisions whilst conducting their tasks. The findings of the study and subsequent citations focused on productivity measures such as task performance, accuracy and speed. However, results also indicated a negative correlation between task efficiency and user affinity, with several workers expressing antipathy towards arguably the most 'effective' sound state. Despite performing better, some participants spoke about this sound set as "disorientating" or "confusing" and feeling "overwhelmed". Notably all participants responded negatively to the idea that this could become part of their everyday working lives (Simmons et al., 2023, 2024). This was a very small sample and a very specific use case, however it raises the question of what should be done with immersive tech when there is disconnect between what is best for business, and what is best for workers.

2 Health and Safety and Workers rights

Although the convergence of technologies such as virtual reality, AI and teleoperation may be relatively novel, the balancing of corporate interests against workers' wellbeing is far from new. Around the world, health and safety legislation requires employers to assess workplace risks, and to safeguard workers against harms to health, irrespective of the impact on 'the bottom line'.

In the UK, the Health and Safety Executive (HSE) requires that employers undertake risk assessments and mitigate potential harms. Existing guidance does not make mention of working with virtual, augmented or mixed reality, with the closest related guidance relating to the use of display screen equipment (DSE). This includes some transferrable advice about eye strain and regular breaks but lacks insight into the specificity of immersive technologies. A report published in 2020 by the UK government's Department for Business, Energy & Industrial Strategy offers some useful perspectives on 'The safety of domestic virtual reality systems' (Department of Business, 2020), but focusses almost entirely on hardware, and the physical symptoms that can be associated with headsets e.g. nausea. There is no reference to the impact of the experience or activity that a user undertakes whist using a headset.

The European Agency for Safety and Health at Work has specific provisions now in place for workers using 'Smart Digital Systems'. This guidance targets sensor-based tech rather than simulative and/or tele-operative environments, however there is some helpful overlap as EU-OSHA recommends that future users are involved at the earliest stages of any technology design process.

In 2024, the US government added virtual, augmented and mixed reality headsets to its list of Nationally Recognized Testing Laboratory (NRTL) equipment. This new category (UL 8400) brings such devices under standard health and safety recommendations for workers and employees. Although XR is specifically provisioned, the focus is on the mechanics of the headset e.g. preparation of a safe area prior to use, maintenance and storage of batteries etc. No guidance yet exists to support employers in risk assessing simulated environments and/or digitally twinned physical environments. Nor is there offered any kind of support or protections for workers in terms of their rights and wellbeing in such heavily mediated circumstances.

Legislation can struggle to keep pace with new technologies, and it remains to be seen whether instruments now in place will prove adequate to the emerging use cases created when immersive technologies enter into the workplace, or whether new legislation is needed. Within this uncertainty, it remains challenging for employees asked to work in immersive realities to advocate for their rights, and challenging also for employers seeking to make informed decisions about how to configure healthy and productive working environments.

3 Framework: Three Planes of Reality

As exemplified in the Sellafield study, workers may be increasingly expected to use tools such as virtual and mixed reality, projecting their phenomenal self-model (PSM) into a place where they are not physically present (Madary & Metzinger, 2016), and take action in a simulation that corresponds to action in a remote physical location. If their experience were direct e.g. they were on site in a hazmat suit their rights would be clear. Employers would know what to put on the risk assessment and employees would be within their rights to demand their rights e.g. personal protective equipment, regular breaks, or to refuse the work entirely. Teleoperators in immersive environments have no such presumed rights. Their body is maintained at a safe distance, and their rights default to that of any other desk-based worker who might be conducting their activities via a keyboard and mouse.

To offer something constructive, it may be helpful to consider XR workers' safety in terms of their presence across three distinct, overlapping places, or planes of reality.

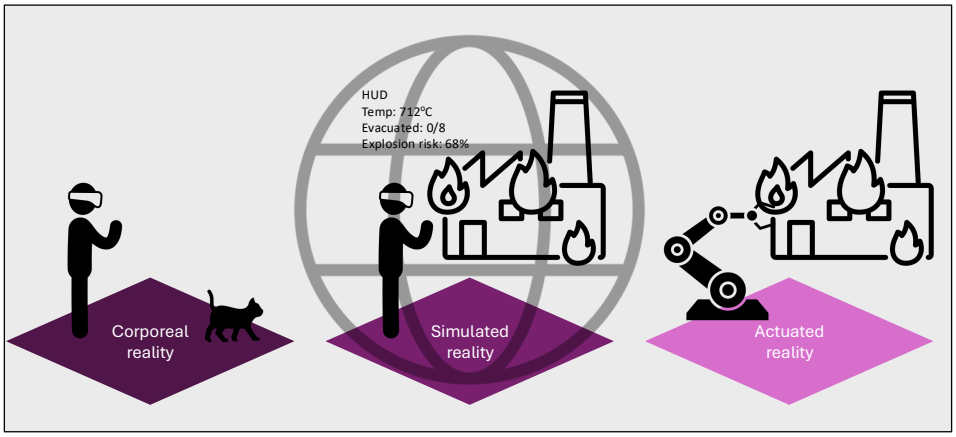

Fig 1. Three Planes of Reality

1. Corporeal Reality – a place where the worker is physically located.
2. Simulated Reality - a virtual environment in which the worker has a strong sense of 'being there'.
3. Actuated Reality - a digitally twinned physical environment. Action taken by the worker in the simulated reality has a correlated outcome in this physical space.

When conducting workplace safety reviews or risk assessments, each plane could be appraised in terms of its individual affordances and incumbent risks. Taking the nuclear decommissioning case study as a first example, a risk assessment might include versions of the following, making use of the three planes of reality model as a scaffold:

Corporeal reality

- Amount of space allocated to worker correlated to motion required in simulation
- Removal or making safe of physical obstacles e.g. furniture, pillars, trip hazards
- Controlled access and egress to space with only relevant personnel permitted in space
- Prior consent regarding communication including location and nature of touch if a colleague needs to attract the attention of the worker
- Prior consent regarding performance monitoring and recording of worker in headset
- Appropriate fire safety and evacuation protocols e.g ensuring alarms are audible
- Clear and appropriate hygiene practices, particularly for shared use equipment

- The key concepts of the task introduced prior to entering headset, previewing visual and audio landscape and offering guidance/strategies in support of psychological safety
- Protected 'decompression' time following simulation work to enable reflection and re-grounding before workers are expected to move on to other tasks
- Regular mental health and wellbeing support to support cognitive processing, and to monitor for signs of workplace trauma, or symptoms of derealization and depersonalization disorders

Simulated reality

- Assessment of the content, context, intensity and potential psychological impact of the simulation scenario at the earliest phases of design and development – co-creating with intended users wherever possible, and seeking expert guidance as appropriate
- Prior informed consent obtained from worker following clear communication of the content and context of the expected working conditions. Consent re-sought for each instance
- Worker able to withdraw consent at any time without the expectation of negative professional consequences
- Initial use of a 'sandbox' environment (not connected to actuated reality plane) in which worker can acclimatize to simulation, familiarizing self with the scenario and task
- Regular interval in-world alerts e.g. every 15-20mins, similar to in-car driver warnings recommending or mandating that workers take a break before resuming task
- Maximum time in simulation per day set and enforced
- Kill-switch facility in-world, allowing user to de-couple simulation from actuated reality at any time, giving agency, managing risk and enabling impromptu and regular breaks

Actuated reality

- Testing of correlation (e.g. accuracy and latency) between simulated and actuated reality planes prior to beginning task
- Prior agreement of risk and liability if action taken by the worker in simulated reality results in unexpected and/or unwanted outcomes in actuated reality plane
- Prior agreement of risk and liability should technical faults or failures result in unexpected and/or unwanted outcomes in actuated reality plane
- Supervision of actuation space by at least one other worker wherever possible
- 'Ship-to-shore' comms link in headset to enable live communication between workers in and out of simulation wherever possible
- Facility to de-couple physical action in actuation reality plane from simulated activity, correlated to in-world 'kill-switch' that can be used by worker at any time

This approach must also come with the understanding that, whist all three planes are distinct and confer their own risks, they are also interlinked and overlapping. The person doing the work is the connective tissue, and they are simultaneously present across all three types of spaces.

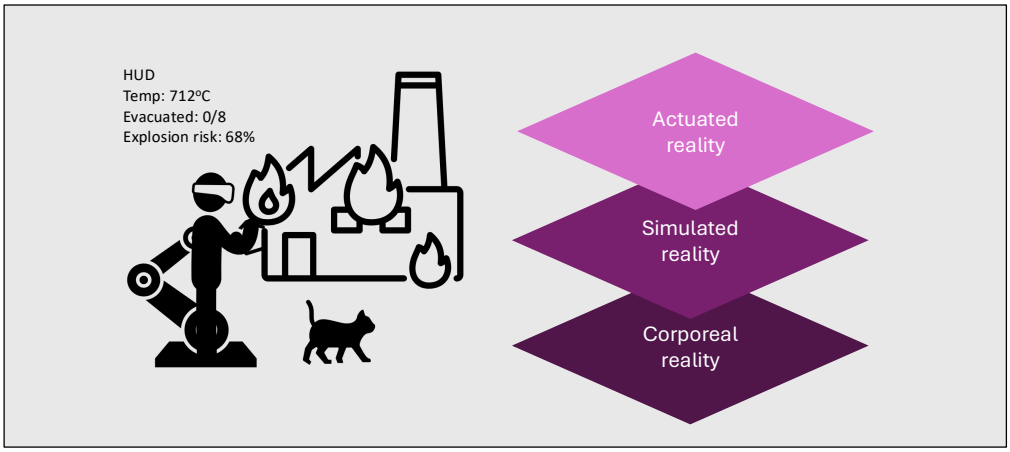

Fig 2. Three Planes of Reality as overlapping, simultaneous experiential realities

The act of spreading oneself thinly across multiple realities can require high levels of cognitive, emotional and phenomenological exertion (Thomas & Glowacki, 2018). The effort of sustaining dual, or even triadic consciousness across these attentional planes is significant, especially when the layered scenario a worker is attending to is inherently high risk, high stakes and or high stress (McIntosh, 2022). Some researchers have raised concerns regarding the risk of depersonalization and derealization disorders if such dynamics are not properly understood and managed (Peckmann et al., 2022).

For employers, creating a strategy for immersive tech use in the workplace could incorporate consideration of all applicable planes of reality, and the cumulative load involved for the individual in navigating each. Mitigation strategies might include co-design with workers as recommended by EU-OSHA. Regulatory bodies could intervene with practical safeguards such as the need for immersive literacy training to support informed risk assessment, a cap on total amount of time per day that workers can spend in simulative realities, definition and classification of intensity of experience taking simulated risk and threat into account, mandating mental health support for immersive workers, and offering protections against punitive measure if workers request adaptations or to refuse simulative work altogether. As technology horizons advance at pace, there is a leadership opportunity for government agencies in re-centring workers' rights, challenging the dominance of bodily safety in simulation and teleoperation work in immersive realities, and beginning this crucial work now, before workflows that compromise worker safety and wellbeing become entrenched.

4 Further research

The three planes of reality framework above is offered as a starting point. A relatively pragmatic response to a dilemma identified from limited prior research, and an extrapolation to socio-technical horizons. Further research would be welcomed to better understand the matrix of contextually dependent, intersectional and subjective experiences of workers across an array of industrial settings, and from myriad cultural contexts.

Further research would be welcomed to examine the impact and implications of multiple employees participating from separate corporeal reality locations, with the experience of being simultaneously present in a common simulated reality. The three planes model primarily involves a single worker, utilising XR tools that may be digitally twinned to a remote location. This paper generally therefore does not speak to the dynamics of multi-person experience in virtual space. Nor does it directly

address the use of devices such ‘smart glasses’ in a shared physical workplace or public space where the view of corporeal reality is preserved and overlayed by the simulation layer.

Another under-explored dynamic involves multiple employees simultaneously experiencing distinct simulated realities, each twinned to a common actuated reality. For practical or personal reasons, each worker might operate within a virtual environment tailored the specifics of their task, or based on their individual preference, comfort or accessibility requirements, resulting in an asynchrony of experience that is not yet fully understood in terms of its impact on worker experience and employer responsibilities. What is gained and lost when we no longer have a common point of reference to describe our experience of the working environment?

Broader and more detailed research along these lines could inform the development of specialist guidance, the empowerment of workers and employers, and the evolution of best practice in the fast-developing confluence of corporeal, simulated and actuated realities in the workplace.

5 References

Cavallaro, R. (2025). *The Total Economic Impact™ Of Meta Quest*. https://forwork.meta.com/ch/en/vr-insights/forrester-tei-study/

Dalton, J. (2021). *Reality Check: How Immersive Technologies can Transform your Business* (1st ed.). Kogan Page.

Department of Business, E. and I. S. (2020). *The safety of domestic virtual reality systems - A literature review*. https://assets.publishing.service.gov.uk/media/5f763502d3bf7f7c2bcf9eb9/safety-domestic-vr-systems.pdf

Introducing Apple Vision Pro: Apple's first spatial computer. (2023, June 5). Apple. https://www.apple.com/uk/newsroom/2023/06/introducing-apple-vision-pro/

Lee, K. M. (2004). Presence, Explicated. *Communication Theory*, *1*(1), 27–50. https://onlinelibrary-wiley-com.ezproxy.uwe.ac.uk/doi/epdf/10.1111/j.1468-2885.2004.tb00302.x

Madary, M., & Metzinger, T. K. (2016). Recommendations for Good Scientific Practice and the Consumers of VR-Technology. *Frontiers in Robotics and AI*, *3*. https://doi.org/10.3389/frobt.2016.00003

McIntosh, V. (2022). Dialing up the danger: Virtual reality for the simulation of risk. *Frontiers in Virtual Reality*, *3*. https://doi.org/10.3389/frvir.2022.909984

Nellis, S., & Dave, P. (2021). Microsoft wins $21.9 billion contract with U.S. Army to supply augmented reality headsets. *Reuters*. https://www.reuters.com/article/technology/microsoft-wins-219-billion-contract-with-us-army-to-supply-augmented-reality-idUSKBN2BN36B/

Peckmann, C., Kannen, K., Pensel, M. C., Lux, S., Philipsen, A., & Braun, N. (2022). Virtual reality induces symptoms of depersonalization and derealization: A longitudinal randomised control trial. *Computers in Human Behavior*, *131*. https://doi.org/10.1016/j.chb.2022.107233

Ratan, R. (2012). Self-presence, explicated: Body, emotion, and identity extension into the virtual self. In *Handbook of Research on Technoself: Identity in a Technological Society* (pp. 321–335). IGI Global. https://doi.org/10.4018/978-1-4666-2211-1.ch018

Simmons, J., Bown, A., Bremner, P., & Mitchell, T. J. (2023). Hear Here: Sonification as a Design Strategy for Robot Teleoperation Using Virtual Reality. *Virtual, Augmented, and Mixed-Reality for Human-Robot Interactions at HRI*, *2022-March*. https://doi.org/10.1145/nnnnnnn.nnnnnnn

Simmons, J., Bremner, P., Mitchell, T. J., Bown, A., & McIntosh, V. (2024). The ballad of the bots: sonification using cognitive metaphor to support immersed teleoperation of robot teams. *Frontiers in Virtual Reality*, *5*. https://doi.org/10.3389/frvir.2024.1404865

Sinclair K., Clark J. (2020). Making new reality: A toolkit for inclusive media futures [online]. Available from: https://makinganewreality.org/(Accessed March 14, 2022)Slater, M. (2009).

Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. *Philosophical Transactions. Biological Sciences*, *364*(1535), 3549–3557. https://doi.org/10.1098/rstb.2009.0138

Thomas, L. M., & Glowacki, D. R. (2018). Seeing and feeling in VR: bodily perception in the gaps between layered realities. *International Journal of Performance Arts and Digital Media*, *14*(2), 145–168. https://doi.org/10.1080/14794713.2018.1499387

Yee, N., & Bailenson, J. (2007). The Proteus Effect: The Effect of Transformed Self-Representation on Behavior. *Human Communication Research*, *33*(3). https://doi.org/10.1111/j.1468-2958.2007.00299.x

Supporting recovery with mental health teleclinics in the Metaverse

Yannick Prié[1], Ketty Steward[2,3]

[1]Nantes Université, École Centrale Nantes, CNRS, LS2N, UMR 6004, F-44000 Nantes, France

[2]Centre Ressource en Psychiatrie de la Personne Âgée, CHU de Nantes, France

[3]Laboratoire Psychopathologie et Processus de Changement, Université Paris 8, Saint Denis, France

*** Correspondence:**
Yannick Prié
yannick.prie@univ-nantes.fr

Keywords: Teleclinics, Metaverse, CHIME, Recovery, Elderly

Abstract: *We analyze how mental health teleclinics in the Metaverse may support patient recovery focusing on the connectedness, identity, and empowerment dimensions of the CHIME framework. We then discuss how healthcare may be impacted by such teleclinics, and some associated limits, risks, as well as open questions.*

1 Introduction

We have recently proposed (Prié et al, 2025; Slater et al., 2025) that immersive teleclinics could exist in the Metaverse as sets of care environments made available in social virtual reality (VR) by institutions such as hospitals. In those clinics, patients could carry out testing or training activities, have meetings with psychologists, meet peers in supervised discussion groups or during collective activities, etc. Following our first experiments on building and evaluation of a teleclinic for neuro-psychological assessment (Prié et al., 2025), and as part of the IPC4MH project[1], we have recently created a video illustrating our general vision of mental health teleclinics in the Metaverse for senior patients[2].

In this video we narrate the story of a depressed 75 years old widower who falls and is being offered a program in an immersive teleclinics in adjunction to conventional care. Some of the activities he has access to include: a Memory stroll (visit of a city and clinical interview) with a psychologist; Immersive Psychomotor Remotivation with the psychomotor therapist and her avatar, so as to gain sufficient confidence to walk again; supervised groups with peers and a nurse; consultations with a psychiatrist; as well as neuropsychological evaluations.

The story of Philippe, who participates in these various activities from home, exemplifies how such a technology could enhance current practice in psychiatry for the elderly. We propose to use the care program described in this video as a support to raise different issues around such teleclinics. We will do so first by studying how healthcare pathways in the Metaverse can participate in patient recovery,

[1] See https://ipc4mh.net/

[2] See https://www.youtube.com/watch?v=roRyGCQK_pw, we advocate the reader to look at the video before getting back to reading this paper.

before discussing how the evolution of healthcare can be foreseen with such teleclinics, as well as some risks, limits, and open questions.

2 Consequences on patients: using the CHIME framework

VR care can be seen as a contribution to the recovery process of the patient. The CHIME framework is used to identify the dimensions of recovery which appear in this experience. CHIME acronym stands for Connectedness, Hope and optimism about the future, Identity, Meaning of life, and Empowerment (Leamy et al., 2011). To us, teleclinics in the Metaverse have a direct effect on connectedness, identity, and the empowerment of (elderly) patients.

2.1 Connectedness

The connectedness dimension of recovering is based on peer support and support groups, the rebuilding of (or the building of new) relationships, and becoming part of a community. This is much needed for patients who face isolation problems such as the elderly population, in which isolation is a major problem highly associated with depression (Van As et al. 2022). Group activities, especially are known as good therapeutics, providing both group support and connection to others (Ba, 1991 ; Husaini et al., 2004).

Social VR care is obviously a means to such a process, for patients can have access to group activities from home, in safe environments, with the benefit of extended anonymity thanks to the avatars they embody (Baccon et al. 2019). Be it discussion groups monitored by professionals, or activity-based group interventions such as art therapy, these collective moments certainly participate in connectedness. Moreover, as we illustrated in our video, the relations that are built in the course of group activities within teleclinics might extend in reality. Without going that far, patients may well meet again in dedicated unsupervised spaces offered by an immersive teleclinic, or on external social VR platforms, within virtual environments and the associated communities[3].

What is key to connectedness in the Metaverse is the embodiment of avatars by users, and the types of relations one can build from interacting with other's avatars within 3D environments designed to conduct joint activity.

2.2 Empowerment

The empowerment dimension of recovering is mostly about regaining personal responsibility, control over life, e.g. by focusing upon one's strengths. Promoting empowerment of patients during recovery consists of helping them get awareness of their preserved capacities and strengths, not only to what is lost, more or less permanently.

As when Philippe visited with a psychologist a town where he spent his holidays years ago, VR social care offers the elderly the possibility to access places and activities that are hardly available to them in real life. Deciding to go to a specific place or to carry out a specific activity helps patients get back the power to decide for themselves.

Also, for some parts of the program, mostly when alone, the patient has a bit of latitude to choose when to practice.

[3] E.g. the Open Mind community on VRChat.

A data-based teleclinic can help practitioners get a better grasp on the strengths and weaknesses of the patient, allowing them to adapt the program to those.

2.3 Identity

The identity dimension of recovery is mostly concerned with the process of rebuilding and redefining a positive sense of identity. It often implies a change in one's identity: in most recovery narratives, people report the conviction of becoming a new person, different from the one they were before being ill (Ridge & Ziebland, 2006).

The fact that new collective practices can occur through social VR activity (in the case of Philippe, constructing a giant structure) may support changes in identity by facilitating the creation of new relationships. As illustrated by Philippe choosing an avatar reminding him of his amateur sailing years, creating one's own avatar can also be an opportunity to redefine oneself and determine which facets of their personality one wants to show and enhance.

Lastly, for elderly people, the use of top-tier technology may also be a driver of identity change.

3 Evolution of healthcare

Besides the benefits of social VR teleclinics for patient recovery, such new ways of providing integrated care may lead to changes in the healthcare system. Let us discuss four of them.

First, with regards to practitioners, some new professions could emerge with the need to handle the material, do maintenance or rental besides the actual care. The existing roles in the care system are also necessarily redefined in a virtual teleclinic. A nurse, a psychologist, a psychiatrist, have to take on new roles and identity, as they interact with the patients and VR environment and may have the feeling of working in another profession. In fact, it might be necessary to complete the training of caregivers with the expected new skills.

Second, the proximity of healthcare VR environments and classical ones, such as games, in the patient's immersive system may have an impact on healthcare practices, by facilitating the integration of non-healthcare apps into patients' programs, e.g. as extensions of sessions. Also the availability of non health-related social VR environments where both patients and practitioners can easily meet may extend the therapeutic possibilities, outside of the clinic.

Third, teleclinics in the Metaverse may well redefine what the notion of clinical structure means. Indeed, a teleclinic is a set of coherent care environments dedicated to specific activities, but also has an identity of its own, offered by an institution. This is different from what can be proposed with visio-conference consultations, because 3D environments and embodiment provide the impression of accessing places people can come to, that they can explore, where they can meet with one another and with practitioners, even if they live in different geographical positions. Within language and timezone constraints, immersive teleclinic exist as non-geographically located places. This may redefine the way care is accessed, as patients may now have more possibilities to choose the institution in which they will be treated.

Furthermore, these new healthcare practices will require the collection and management of new types of sensitive data integrated into patients' medical records, for example, behavioral data. This evolution makes it essential to establish an ethical and legal framework that protects individuals.

4 Discussion

Let us underline some limits and risks related to teleclinics in the Metaverse. An obvious limit is related to the capacity of patients to use immersive technologies such as HMDs and maintain them in working order. This raises the question of autonomy and independence, all the more so when it comes to older people. How autonomous do they have to be in order to be able to take advantage of this kind of technology that is supposed to increase their autonomy?

A risk is that some patients may have difficulties with such teleclinics, because they suffer from cybersickness, or are prone to forgetting they are in VR and may hurt themselves, e.g; by colliding with physical obstacles.

There are also a lot of open questions that may or may not be categorized as risks or limits, because they deserve further inquiry. For instance, there could be a negative effect related to the lack of temporal transition between healthcare and daily-life environments, because one may argue that the mandatory travel from home to a medical institution, and the associated wait are necessary to recovery. This however remains to be demonstrated.

Another open question is related to the possibility of having adequate fully distant care for several months, when the general opinion of psychologists might be that physical social relations with the patients should be maintained, e.g; by designing programs that mix real and virtual activities. A related question deals with the possibility that social relations in VR may be detrimental to physical social relations.

Also, the effects of the embodiment of avatars in psychotherapy, for instance of allowing the patients to choose how they should be represented, or even to changing their avatar in the course of a session with a psychologist, are still to be explored.

Lastly, the use of AI within such teleclinics, and its effect on care and patient recovery remains to be investigated. For example, what does it mean for a patient to have a VR exercise supervised by an AI avatar presented as the representative of their practitioner? How could this affect the therapeutic alliance between them?

5 Conclusion

We have discussed the potential of teleclinics in the Metaverse for psychotherapeutic care for elderly patients, showing that it may support connectedness, identity, and empowerment of the patients during their recovery process. We also highlighted a few evolutions such teleclinics may drive for mental healthcare, as well as some limits, risks, and open questions.

6 References

Ba, G. (1991). Group Therapy with Elderly Patients. *Psychotherapy and Psychosomatics*, 56(3), 157-161. https://doi.org/10.1159/000288549

Baccon L. A., Chiarovano E., MacDougall H. G. (2019). Virtual reality for teletherapy: avatars may combine the benefits of face-to-face communication with the anonymity of online text-based communication. Cyberpsychology, Behav. Soc. Netw.22, 158–165. 10.1089/cyber.2018.0247

Husaini, B. A., Cummings, S., Kilbourne, B., Roback, H., Sherkat, D., Levine, R., & Cain, V. A. (2004). Group Therapy for Depressed Elderly Women. *International Journal of Group Psychotherapy,* 54(3), 295-319. doi:10.1016/j.archger.2018.06.001

Leamy M, Bird V, Le Boutillier C, Williams, J., Slade, M. (2011) Conceptual framework for personal recovery in mental health: systematic review and narrative synthesis. Br J Psychiatry;199:445–52. doi:10.1192/bjp.bp.110.083733

Prié Y., Vigier T., Acier D., Bonneville H., Georges M., Thévenet S., Bulteau S. (2025) Towards Teleconsultations in the Metaverse: Design and Preliminary Study of an Immersive Teleclinic for Neuro-psychological Assessment, *Frontiers in Virtual Reality - Virtual Reality in Medicine*, Volume 6. doi:10.3389/frvir.2025.1545569

Ridge, D., & Ziebland, S. (2006). The Old Me Could Never Have Done That": How People Give Meaning to Recovery Following Depression. *Qualitative Health Research*, 16(8), 1038-1053. doi:10.1177/1049732306292132

Slater M, Di Dalmazi M, Friedman D, Galissaire J, Isaac H, Kobusinska A, Lopez-Tarruella-Martinez A, Lamberti L, Lucas J-F, Niamut OA, Pan X, Prié Y, Quent M, Ruiz Rodríguez R, Sanchez-Vives MV, Skrzypczyński P, Steed A and Wannerberg P (2025) A multidisciplinary approach to understanding the technology, opportunities, challenges, and implications for society of the metaverse. Front. Virtual Real. 6:1566419. doi: 10.3389/frvir.2025.1566419

Van As BAL, Imbimbo E, Franceschi A, Menesini E, Nocentini A. The longitudinal association between loneliness and depressive symptoms in the elderly: a systematic review. *International Psychogeriatrics.* 2022;34(7):657-669. doi:10.1017/S1041610221000399

Empowering Oncology Patients Through XR Self-Observation and Mixed Reality Exercise

Alex Fuentes-Raventos[1,2*], Caterina Calderon[1], Mel Slater[1]

[1]Department of Clinical Psychology and Psychobiology, Faculty of Psychology, University of Barcelona, Barcelona, Spain.
[2]Bridging Research in AI and Neuroscience, Computer Vision Center, Spain.

***Correspondence:**
Alex Fuentes-Raventos
alexfuentes@ub.edu

Keywords: virtual reality, supportive oncology, self-avatar, digital self-representation, exercise, mixed reality, XR demonstration

Abstract

Cancer patients are encouraged to exercise during treatment, yet severe fatigue, "kinesiophobia," and limited access to programs hinder adherence. This "Exercise Paradox" creates a cycle where fear of exertion exacerbates deconditioning.

This XR demonstration presents NEOBIC-VR, a virtual reality exercise platform for supportive oncology deployed on the Meta Quest 3. The system combines avatar-based self-observation, adaptive pacing, and mixed reality (MR) safety features. By inducing strong Body Ownership over a photorealistic virtual self-replica, the application allows patients to observe their digital self perform tasks prior to execution. This acts as a mental simulation to help users bypass psychological barriers associated with movement.

The demonstration showcases a structured home-based session:

1. Embodiment Warm-Up (5 min): Users synchronize movements with their avatar in a virtual mirror to solidify Body Ownership.
2. VR Upper-Body Phase (15-20 min): Set in immersive environments like a Roman Pantheon or beach , users engage in adaptively paced boxing and dodging tasks.
3. MR Lower-Body Phase (10 min): To ensure physical safety and reduce cybersickness during lower-body exertion (e.g., squats), the headset transitions to MR passthrough mode, blending a virtual trainer with the user's real physical space.
4. Cooldown (5 min): A guided breathing exercise concludes the session.

By presenting exercise through digital self-replication and transitioning between VR and MR, this demonstration explores how XR can foster confidence and enhance engagement in oncology rehabilitation.

2 Conflict of Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

3 Author Contributions

Alex Fuentes-Raventós conceived the submission and drafted the abstract.
Caterina Calderón contributed to the clinical framing and revised the text.
Mel Slater contributed to the conceptual framing of self-observation and XR and revised the text.
All authors approved the final version.

4 Funding

This research was funded by the Spanish Association Against Cancer (AECC), Strategic Projects Program 2024 (Project ID: PRYES246682CALD).

5 Acknowledgments

To the patients and researchers of the NEOTEAM team.

5 Data Availability Statement

No datasets were generated or analyzed for this study.

Immersive Realities (XR) stress management for protection the human right to mental health

Olga Maslova[1*], Vasiliy Pyatin[2]

[1]Research Unit Sociology of Rehabilitation and Participation, Department of Rehabilitation Sciences, TU Dortmund University, Dortmund, Germany

[2]Neurointerfaces and Neurotechnologies Laboratory, Neurosciences Research Institute, Samara State Medical University, Samara, Russia

*** Correspondence:**
Olga Maslova
olga.maslova@tu-dortmund.de

Keywords: human right, mental health, XR, metaverse, stress management

Abstract

The article argues that socioeconomic status constitutes a significant determinant of mental health issues. It demonstrates that human rights violations affect the physical, psychological and emotional well-being of service and nonservice recipients. In addition, the link between circadian rhythm disruptions and adverse mental health outcomes is highlighted. Within this framework, Immersive Realities (XR) and metaverse-based platforms are presented as a promising tool for person-centered stress management. Mainly, these technologies have enormous potential for realization of the human rights to mental health and contribute to the implementation a rights-based conception that accounts for the political and economic conditions that produce stress and disability. Accordingly, this study aims to highlight the innovative potential of XR and metaverse in advancing and protecting every individual's right to mental health.

1 Introduction

The Global Burden of Disease study, based on data from over 200 countries, including economically developed nations, highlights the extremely high prevalence of mental disorders worldwide (Fan et al., 2025). Mental health is a key factor in assessing individual and global well-being. The World Health Organization (WHO) has included mental health in the Sustainable Development Goals, highlighting that around 25% of the world's population faces these issues at some point in their lives (World Health Organization, 2022). Mental disorders currently constitute the second largest group of conditions leading to long-term incapacity for work (Orr et al., 2023). The European Commission's Comprehensive approach to mental health recognizes the need for a holistic response based on the guiding principles of effective prevention, access to high quality mental healthcare, and 'mental health in all policies' as stress-related aspects which are considered when developing policy areas related to education, employment and digital policy (European Commission, 2023). Under these circumstances, the realization of the human right to mental health has become a global imperative, which can be addressed at the individual level through XR technologies (Immersive Realities – Benefits and challenges to freedom of expression, 2025), as well as at a more innovative level—in the metaverse environment (Maslova et al., 2025).

2 Global and individual causes of mental health disorders

Socioeconomic status remains a key determinant of mental health outcomes. Given the diversity of cultures and population groups, including immigrants and minorities, challenges arise in the process of assimilation, which is a source of disproportionate socioeconomic inequality. At the same time, discrimination is intertwined both structurally and interpersonally, which increases the mental health risk of adverse consequences (Wang, Narcisse, 2025). In the traditional mental health care sector, significant accessibility limits persist despite the rising prevalence of mental diseases. These gaps are driven by a shortage of service providers, insurance limitations, high personal costs, language and cultural barriers, as well as limited digital literacy and the stigma surrounding mental health.

In the paradigm of mental disorders, stress is a key concept in the field of mental health. Over the past few decades, emotional stress has become a widespread problem, with significant variations across countries and regions, and its daily occurrence leads to anxiety disorders, depression, burnout and social isolation (da Silva et al., 2021). Subjective socio-economic factors such as financial situation and dissatisfaction with the education system have a significant impact on emotional stress (Lado et al., 2023). Work-related emotional stress negatively affects labour productivity, and in its chronic form, causes to reduced work efficiency and significant economic consequences for both individuals and economics (Ganster and Rosen, 2013). Notably, in 2026, social media are being used by almost 70% of the world's population (Datareportal, 2025). Study has shown that excessive use of social media is a risk factor associated with significant mental health issues including depression among working adults (Elangovan et al., 2026), the prevalence and incidence of which increase with age and have become a serious public health issue (Ponte et al., 2014).

Mental health impairments caused by socioeconomic status are further exacerbated by human rights violations, which affect the physical, psychological and emotional well-being of both service recipients and non-recipients with mental disorders (Mfoafo-M'Carthy and Huls, 2014). Finally, a wide range of adverse mental health consequences are caused by "circadian health" disruptions which affect all the Earth's population (Coelho et al., 2026). So, individuals with an evening chronotype are at a significantly increased risk of depression through multifactorial mechanisms, including circadian rhythm disruption, increased social isolation, and impaired stress response regulation (Dong et al., 2026).

It can be concluded that mental health disorders are multifactorial in nature, including a factor of violations of human rights for mental health support.

3 XR and metaverse as an innovative tool for human right to mental health

The classical theory of stress reduction posits that natural stimuli have a restorative effect in reducing psychological stress. In the XR era, the focus has shifted to the role of virtual stimuli in stress management. While evidence of the effectiveness of XR interventions has not yet reached maturity, their potential is clear as an innovative tool for human rights in mental health.

The literature highlights cutting-edge virtual reality (VR) and metaverse technologies as the tools for overcoming traditional barriers in the treatment of mental disorders, including anxiety, depression, post-traumatic stress disorder, phobias, and public speaking anxiety. It further demonstrates remote "Hospital to Home" VR therapy models designed for individuals who don't seek help due to the stigma surrounding mental health, and "just-in-time adaptive interventions", and new possibilities for personalized mental health management powered by AI-driven VR (Rajpal and Stevens, 2025). The advantages of VR therapy stem from its controlled and customizable environment for addressing mental health issues (Banakou et al., 2024).

By the end of 2026, stress management tools in XR will utilize generative artificial intelligence (GenAI) and large language models (LLMs) to create personalized interactive environments that adapt mindfulness and relaxation scenarios to individual needs, thereby enhancing emotional

resilience (Kong et al., 2026). The metaverse enables a personalized approach to mental health services through social interaction between users via 3D avatars. Using AI and neurotechnologies such as electroencephalography or functional near-infrared spectroscopy, one can dynamically control mental states in real time through interaction with avatars, creating continuous feedback with virtual agents in the metaverse (Maslova et al., 2025).

Consequently, XR and metaverse enable the individuals to receive personalized assistance in any type, time, or space, thereby representing new paradigms in the protection of the human right to mental health.

4 Discussion

The right to health is a fundamental human right that encompasses physical, mental and social well-being, and is essential to the realization of other fundamental human rights (World Health Organization, 2026). Notably, but the human rights violations are often the cause of mental ill-health (Cosgrove L., 2025). In an era marked by the global spread of XR as a human-centered digital technology for mental health management, the issue of the legal aspects of global human XR intervention has become not only particularly relevant but also solvable. Under these conditions, stress management requires human rights-based approaches, such as access to services, workplace safety, and social protection system.

The potential of XR-based therapy (Omisore et al., 2024) and metaverse (Maslova et al., 2025) provides a subject-centered solution to mental health challenges. Overall, XR is at a stage where evidence of the interventions' effectiveness in the context of mental health support has not yet reached maturity. This actualizes a human rights-based conception that accounts for the political and economic conditions that produce stress and disability. This applies equally to the experience of leveraging the technology of metaverse and circadian rhythms. It is evident that these represent new and promising path towards realizing the human right to mental health for the 25% of the world's population who face these challenges in their daily lives (World Health Organization, 2022). The legal aspects of addressing mental health issues are of critical importance to the well-being of a significant portion of the population in the EU countries, who face socioeconomic challenges, human rights violations, and circadian rhythm troubles. Protection the human right to mental health is a global imperative, which in the digital era can be addressed through the XR and metaverse.

5 Conclusion

Mental health disorders, particularly stress, are extremely common worldwide with causes that are both individual and global. This underscores the importance of addressing this issue within the context of broader social inequality and perceived shortcomings in public services aimed to alleviate emotional stress. In this context, XR demonstrate exceptional adaptability in providing psychological support to persons with mental health disorders. Leveraging the potential of XR and metaverse to promote human right to mental health is an innovative solution that supports this right through stress management facilitated by XR, AI, apps and metaverse. Thus, the integrative model highlights a convergence of such concepts as human rights norms in Europe, stress management, immersive technologies, ethical governance, with particular emphasis being put on preventive and personalized approach facilitated by XR and metaverse to ensure equitable and effective outcomes in mental health.

6 Conflict of Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

7 Author Contributions

OM: Writing – original draft, Writing - review & editing, Conceptualization, Methodology; VP: Writing – original draft, Writing - review & editing, Conceptualization, Analysis.

8 Funding

The authors declare that no financial support was received for the research, authorship, and publication of this article.

9 References

Banakou, D., Johnston, T., Beacco, A., Senel, G., Slater, M. (2024). Desensitizing anxiety through imperceptible change: Feasibility study on a paradigm for single-session exposure therapy for fear of public speaking. JMIR Form Res. 8:e52212. https://doi: 10.2196/52212.

Coelho, J., Micoulaud-Franchi, J.-A., Geoffroy, P.-A., Philip, P., Taillard, J., Bourgin, P. (2026). Defining circadian health: Conceptual framework and development of a circadian health index. Sleep Medicine Reviews. 86:102244. https://doi.org/10.1016/j.smrv.2026.102244

Cosgrove L. (2025). Addressing the Global Mental Health Crisis: How a Human Rights Approach Can Help End the Search for Pharmaceutical Magic Bullets. Health Hum Rights. 27(2):179-188.

daSilva A.W., Huckins J.F., Wang W., Wang R., Campbell A.T., Meyer M.L. (2021). Daily perceived stress predicts less next day social interaction: evidence from a naturalistic mobile sensing study. Emotion, 21(8):1760–70. https://doi.org/10.1037/emo0000994

Datareportal. (2025). Global social media statistics https://datareportal.com/social-media-users [Accessed March 24, 2026].

Dong, X., Liu, J., Wang, X., Liang, Z. (2026). Chronotype and depression: A global systematic review and meta-analysis. Psychiatry Research. 358:116945. https://doi.org/10.1016/j.psychres.2026.116945

Elangovan, P.D., Ishak, A.S., Rose, N.N. (2026). Social media leads to stress and depression for working adults. International Journal of Current Science Research and Review. 9(2). https://doi.org/10.47191/ijcsrr/V9-i2-39

European Commission. (2023). EU comprehensive approach to mental health. https://health.ec.europa.eu/non-communicable-diseases/mental-health_en[Accessed April 29, 2026].

Fan Y, Fan A, Yang Z, Fan D. (2025). Global burden of mental disorders in 204 countries and territories, 1990–2021: results from the global burden of disease study 2021. BMC Psychiatry. 25(1):486. https://doi.org/10.1186/s12888-025-06932-y

Ganster, D.C., Rosen, C.C. (2013). Work stress and employee health. J Manag. 39(5):1085–122. https://doi.org/10.1177/0149206313475815

Immersive Realities – benefits and challenges to freedom of expression. (2025). Strasbourg: Council of Europe. https://rm.coe.int/cdmsi-2025-10-draft-feasibility-study-benefits-and-challenges-to-foe-i/488029b847[Accessed April 29, 2026].

Kong, D., Feick, M., Liu, S., Maedche, A. (2026). CoEmpaTeam: Enhancing cognitive empathy using LLM-based avatars and dynamic role play in Virtual Reality. arXiv:2603.16614v1. https://doi.org/10.48550/arXiv.2603.16614

Lado M, Alonso P, Cuadrado D, Otero I, Martínez A. (2023). Economic stress, employee commitment, and subjective well-being. Rev Psicol Trab Organ. 39(1):7–12. https://doi.org/10.5093/jwop2023a2

Maslova O., Shusharina N., Pyatin V. (2025). The neurosociological paradigm of the Metaverse. Frontiers in Psychology. 15:1371876. https://doi.org/10.3389/fpsyg.2024.1371876

Mfoafo-M'Carthy, M., Huls, S. (2014). Human Rights violations and mental illness. SAGE Open. https://10.1177/2158244014526209

Omisore OM, Odenigbo I, Orji J, Beltran AH, Meier S, Baghaei N, Orji R. (2024). Extended Reality for mental health evaluation: Scoping Review. JMIR Serious Games. 12:e38413. https://doi.org/10.2196/38413.

Orr E, Arbel T, Levy M, Sela Y, Weissberger O, Liran O, Lewis J. (2023). Virtual reality in the management of stress and anxiety disorders: a retrospective analysis of 61 people treated in the metaverse. Heliyon. 9(7):e17870. https://doi.org/10.1016/j.heliyon.2023.e17870

Ponte, C., Almeida, V. Fernandes L. (2014). Suicidal ideation, depression and quality of life in the elderly: study in a gerontopsychiatric consultation. Spanish J. Psychol. 17:E14. https://doi.org/10.1017/sjp.2014.15

Rajpal, S., Stevens, J.P. (2025). Exploring the role of Virtual Reality and Metaverse in treating mental health: operating models and challenges in adoption. Journal of Computer and Communications. 13(10). https://doi.org/10.4236/jcc.2025.1310001

Wang, M.L., Narcisse, M.R. (2025). Discrimination, depression, and anxiety among US adults. JAMA Netw Open. 8(3):e252404. https://doi:10.1001/jamanetworkopen.2025.2404

World Health Organization. World mental health report. Transforming mental health for all. (2022). Geneva: World Health Organization. https://www.who.int/publications/i/item/9789240049338 [Accessed March 24, 2026].

World Health Organization. (2026). Constitution. https://www.who.int/about/governance/constitution [Accessed March 24, 2026].

Rebuilding Human Presence in Immersive Realities: Holoportation, Accessibility, and the Future of Human Connection

Stefano Beccaletto [1,*]

[1] *i2CAT Research Centre, Barcelona, Spain*

Correspondence*:
Stefano Beccaletto
stefano.beccaletto@i2cat.net

ABSTRACT

Immersive technologies are advancing rapidly, yet one essential dimension remains unresolved: genuine human presence. Current digital communication systems still struggle to reproduce the emotional, spatial, and social qualities that make interaction meaningful.

This demonstration presents a vision developed within the European research initiative PRESENCE XR, exploring how volumetric capture, AI-driven representation, next-generation networking, and multisensory interaction can enable realistic remote co-presence in extended reality.

The demo showcases a live holoportation system capable of capturing and transmitting people as volumetric holographic representations in real time, allowing users in different locations to interact as if physically together. Beyond entertainment or simulation, the project frames immersive communication as future social infrastructure: one that could reshape access to education, work, culture, healthcare, justice, and civic participation.

The societal implications are especially significant in a world marked by conflict, displacement, inequality, and restricted mobility. For people affected by war, political instability, or border restrictions, immersive communication technologies can become lifelines for education, collaboration, cultural exchange, emotional support, and human connection when physical movement is impossible or unsafe.

The same transformative potential applies to accessibility and inclusion. For individuals experiencing disability, chronic illness, mobility limitations, social isolation, or geographic exclusion, immersive co-presence technologies may redefine participation in society by enabling meaningful remote presence in classrooms, workplaces, cultural spaces, public institutions, and social life.

At the same time, this technological transition raises broader human rights questions aligned with the themes of this conference:

- Is meaningful internet access a fundamental human right?
- Should access to advanced immersive communication technologies also be considered essential for reducing exclusion and inequality?

- Can extended reality strengthen freedom of expression, intercultural understanding, and democratic participation?
- How can immersive environments remain safe, ethical, inclusive, and accessible for children and vulnerable communities?

The demonstration invites participants not only to experience a technological prototype, but also to reflect on the future architecture of the immersive internet. If previous generations of the internet connected information and devices, the next evolution may connect human presence itself.

In this vision, holoportation becomes more than a technical achievement. It becomes a tool for empathy, accessibility, inclusion, and global coexistence: a way to reduce distance rather than deepen inequality, and to preserve one of the most fundamental human experiences, the ability to meaningfully connect with one another across distance, borders, and circumstance.

Keywords: extended reality, holoportation, immersive communication, human presence, accessibility, digital inclusion, freedom of expression, human rights

CONFLICT OF INTEREST STATEMENT

The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

AUTHOR CONTRIBUTIONS

SB conceived the demonstration concept, prepared the abstract, and contributed to the development and presentation of the work.

FUNDING

This work was supported by the PRESENCE XR project. Grant agreement ID: 101135025. DOI: 10.3030/101135025

ACKNOWLEDGMENTS

The author acknowledges the contributions of i2CAT Research Centre and the PRESENCE XR consortium.

DATA AVAILABILITY STATEMENT

No datasets were generated or analyzed for this study.