

# VIRRAC TOOLKIT REPORT

VIRTUAL REALITY RISKS AGAINST  
CHILDREN

A REPORT FOR PRACTITIONERS, POLICY MAKERS, LAW  
ENFORCEMENT & INDUSTRY



## LEADS

Prof. J Davidson  
Asc Prof. E Martellozzo

## TEAM

Dr. R Farr  
P Bradbury  
B Meggyesfalvi

**REVIEWED BY:**  
*ChildNet & KABUNI*

REPHRAIN  
Protecting citizens online



University of  
East London



THE INSTITUTE FOR CONNECTED COMMUNITIES



Middlesex  
University



# Executive Summary

## Research Aims of VIRRAC

The primary aim of VIRRAC was to raise awareness of child safety in the metaverse, and to better understand the harms that children and young people face in this immersive environment in order to explore practical solutions.

The four pillars of investigation within VIRRAC were as follows:

- To develop **an understanding of the existing challenges that are faced by tech companies**, using MV apps in their drive to ensure the greatest number of children, including neurodivergent children, can navigate platforms safely.
- To develop a greater understanding of the **knowledge gap and resource needs of professionals and practitioners** working with children at risk of abuse and exploitation(CASE) in the MV.
- To improve **safeguarding and police understanding of the MV** and facilitate their ability to report and investigate child sexual abuse offences that occur in the MV.
- To develop a **greater understanding, from the voices of children**, regarding their MVsafety and wellbeing support needs.

## Research Context

Young people typically spend long periods of time online every day, and the time spent on digital devices (including smart phones and video gaming) is largely perceived as too long to spend online by parents (PEW research centre, 2020). The number of children and young people known to use virtual reality headsets and engage in the metaverse is rapidly increasing across the world (IET, 2023). The metaverse brings unprecedented opportunities to learn, communicate and game for thousands of users worldwide, offering new forms of immersive communication and the opportunity to play, develop skills and to meet new friends from all over the world. These exciting opportunities also result in increased potential for users to experience exposure to risk and harm. This is especially the case for more vulnerable or impressionable audiences, such as children and young people. Metaverse platforms are already reporting instances of online grooming and cyberbullying (BBC, 2019) and we know that online sex offenders are already using 2D metaverse platforms as spaces to identify victims, and lure children into private arenas (Afzal et al, 2021). However, society is not yet equipped with the understanding, tools or the skills that are needed to try to mitigate these harms for

young metaverse users. For instance, educational awareness programmes across the UK do not typically address (or even acknowledge) harms in the metaverse. As the metaverse continues to evolve and to grow in access and popularity worldwide, it is expected that children will continue to encounter risk and harm. Academics, industry developers, police forces, policy makers and third sector organisations will need to adjust, collaborate and ultimately respond tactfully to ensure child safety in the metaverse is prioritised and accounted for.

## Methods

The VIRRAC team established a Stakeholder Advisory Board comprised of industry and practice experts, conducted an extensive review of the wider literature and undertook a series of focus group interviews with children and young people and with experts in the field.

## Summary of Findings

The findings from the focus groups with children and young people suggested that encountering harm, experiencing pressure to share personal information, and being exposed to unfavourable content was relatively common. The experts and young people who participated in the VIRRAC research studies unanimously recognised the need for more guidance, awareness, support and platform online safety features to keep young users safe from harm in the metaverse.

- **Psychological & physical risks** - The focus groups with children and young people indicated that despite their age, the participants are acutely aware of a range of potential psychological and physical risks that engaging in immersive technologies might lead to and make attempts to actively avoid these risks.
- **Positive aspects** - The focus groups with children and young people uncovered a unique, broad and interesting range of perspectives about what young people like about immersive technology, and what they perceive to be especially enticing about spending time in the Metaverse.
- **Experiencing harm** - The Expert Roundtable raised concerns about virtual assault incidents and the possible impact on childhood development in relation to cognitive effects of VR use, neuroplasticity, and even potential physical harm from prolonged VR headset use. Experts voiced concern about specific vulnerabilities among children with special needs and those on the neurodiverse spectrum, and requirements for closer caregiver attention.
- **Age verification** - Addressing age verification, parental involvement and understanding and safety protocols is crucial to safeguarding children in the Metaverse.
- **Further research** - The literature review evidenced the potentially

detrimental psychological and physical impacts that immersive technologies might have on frequent users, but further highlighted the need for additional empirical research, as well as the application of academic theory (for instance criminological, psychological and sociological perspectives) to aid current understandings of youth safety in the Metaverse, and possible avenues for the mitigation of harm.

- **Safety by design** -VIRRAC has further emphasised the current lack of investment and encouragement in safety by design, and the crucial need for Industry to develop evidence-based, user friendly and user-tested safety features for immersive technology.

## Outputs

This Toolkit Report includes a literature review, findings and evidence-led guidance and recommendations for policy, industry and practice. This report provides a detailed account of the VIRRAC project, the findings from which form the basis of 6 evidence-led tools and a Toolkit that will inform policy and practice in a range of areas. The six tools have been developed specifically for use by educational institutions (tools 2 and 3), frontline practitioners (tool 2), industry (tool 4), parents and caregivers (tool 1), and children (video resource). The outputs provide a starting point and will inform the future development of a full and interactive 'Child Safety in the Metaverse' educational programme which the team aims to develop in full as a priority in their continued work in the area of child safety in the metaverse.

## Recommendations

VIRRAC has highlighted a lack of literature available specifically focusing upon child safety in the metaverse, but also a lack of dedicated resources and awareness raising initiatives for educational institutions, frontline practitioners, parents and carers, and, more strategically, for policy and practice.

### Key Recommendations and Considerations for Future Research, Practice, and Policy

These recommendations aim to guide future efforts towards creating a safer environment for users, particularly children and young people, in the evolving landscape of the metaverse.

**Incorporate Children's Voices:** Research is required to integrate the experiences, perspectives, and recommendations of children and young people concerning safety measures within the metaverse (MV).

**Investigate Harms First-hand:** Further investigation into the first-hand harms experienced by children and young people in the MV, incorporating their lived experiences, is essential.

**Explore AI Moderation:** Insights from discussions on AI moderation, as discussed at the Westminster e-forum conference (8<sup>th</sup> February 2024) should be leveraged to inform future initiatives.

**Review Safety Features:** Conducting a comprehensive review of existing safety features already implemented (safety by design) on platforms could provide valuable insights for industry best practices, particularly benefiting small and medium-sized enterprises (SMEs) and start-ups.

**Consider Legislative Scope and effectiveness:** Although in theory the OSA should cover VR harms and offences, there is some debate regarding the effectiveness of the legislation in addressing these issues.

**Integrate into Education:** There is a need to incorporating the MV into educational resources on online safety for schools.

**Integrate into Training:** Training programs for police forces and the Crown Prosecution Service should include modules addressing the unique challenges and safety considerations of the metaverse.

**Increase Investment:** Increased investment is needed to enhance detection and mitigation efforts for harms within the metaverse.

**Facilitate Cross-Sector Collaboration:** Prioritising and facilitating effective cross-sector collaboration across key domains such as education, health, criminal justice, academia, and industry is vital.

**Revision of Safety Efforts:** Safety by design and prevention efforts by platforms in VR and AR environments should be reviewed, exploring potential applications of AI moderation and the feasibility of implementing age verification for VR headsets.

**Collectively Develop Society's Understanding** of the technological components that comprise the metaverse.

**VR Headsets:** Ensuring that collected data is used responsibly, and at the minimum, especially when children are involved. Headsets collect a staggering amount of personal data such as motion or heartbeat data, that can provide unique insights into the users' physical and mental state. Ensuring that this data is used responsibly, and at the minimum, especially when children are involved, should be a priority.

### **Evidence-Led Tools for Safety in the Metaverse**

Based on the research conducted within the VIRRAC project, the following evidence-led tools for practical use have been produced as an introductory sample of what is to come in the full VIRRAC Toolkit that the team anticipates completing in the continuation of their work focusing on safeguarding children in the Metaverse. The VIRRAC tools include a series of posters sharing the key messages regarding online safety for parents, children (both primary and secondary), and teachers. There are many iterations of general online safety 'checklists' for children and adults to use help them to stay safe online, but very few tools are available that are Metaverse specific. The checklist tool (TOOL 1) is an evidence-led checklist for parents and caregivers to use to support them in keeping children safe in the Metaverse. As stated in the Executive Summary and Recommendations sections of this Report, a safety by design approach is vital to safeguarding children and young people in the Metaverse. The industry checklist tool (TOOL 4) has been developed by the VIRRAC team. It is a quick to use summary to aid platform and game developers in adopting a safety by design approach to new Metaverse features and future updates.

The VIRRAC team has also worked closely with a stakeholder board member, Nina Jane Patel (KABUNI) to produce a series of 5 short video reels that tackle emerging themes and topics of concern in child safety in the metaverse. These videos are unique as they have been created using the adapted experiences, perspectives and solutions that children and young people shared during the VIRRAC focus groups.

**The following tools have been developed by the VIRRAC team.**

- 1) TOOL 1: Parent and Caregiver Checklist
- 2) TOOL 2: Metaverse Safety Poster for Teenagers
- 3) TOOL 3: Metaverse Safety Poster for Children
- 4) TOOL 4: Metaverse Safety Industry: Safety by Design Checklist
- 5) TOOL 5: Metaverse Safety Poster for Schools
- 6) TOOL 6: VIRRAC Film Resource for 8-12 Year Olds

Adapted quotes from the research will be used within the video content. These videos were disseminated on Safer Internet Day 2024, and via the VIRRAC stakeholders, including Childnet and KABUNI.

Access the VIRRAC video resource for 8–12-year-olds here:

<https://www.youtube.com/@UniversityofEastLondonvideos/videos>

### **Summary and Conclusion**

Young people typically spend long periods of time online every day, and, today, children as young as 2 years old are gaining their first digital devices (PEW research centre, 2020). The number of children and young people known to use virtual reality headsets and engage in the Metaverse is rapidly increasing (Han et al, 2022). The Metaverse brings unprecedented opportunities to learn, communicate and game for thousands of users worldwide. As the Metaverse continues to evolve and to grow in access and popularity worldwide, it is expected that children will continue to encounter risk and harm. Academics, industry developers, police

forces, policy makers and third 3rd sector organisations will need to adjust, collaborate and ultimately respond tactfully to ensure child safety in the Metaverse is both understood and prioritised.

Technological advancements play a pivotal role in shaping the current state of the Metaverse and its future trajectory. Platforms like Roblox, Minecraft, and Fortnite have

amassed billions of users worldwide, many of them children, providing a glimpse into the existing landscape of virtual worlds. However, the form that the Metaverse will take in future remains uncertain, with potential scenarios ranging from specialised platforms to a dominant, open, and interoperable Metaverse (UNICEF, 2023). Speculation includes theories about the continuous development of brain-computer-interface and Neuralink that can develop into a form that gives an experience that is difficult to distinguish from reality in the Metaverse (Park & Kim, 2022). Ensuring responsible development and ethical governance are crucial to address privacy, security, and social concerns that accompany the ever-evolving technological landscape. 'Recognising the need for continued openness and standards for the Metaverse, in 2021, IEEE, the world's largest technical professional organisation dedicated to advancing technology, committed to help set standards to define the operating principles of virtual worlds' (Faraboschi et al, 2022).

The Metaverse brings unprecedented opportunities for thousands of users worldwide. For instance, the Metaverse offers new forms of communication, multi-disciplinary learning, an array of new and exciting experiences, human development, and the chance for people to meet in virtual reality regardless of geography, from all ends of the globe who could never have met physically. These exciting opportunities also bring forth the increased potential for users to experience exposure to risk and harm. This is especially the case for more vulnerable or impressionable audiences, such as children and young people. As the Metaverse continues to evolve and to grow in access and popularity worldwide, it is expected that academics and policy makers will continue to learn and adjust alongside. In addition, while the emergence of the Metaverse represents opportunities of study in multiple fields of research, little has been done so far to evaluate the varying psychological impacts on the mental health of children and young people as it relates to the different levels of immersion represented by the various types of technologies in the Metaverse.

The Online Safety Act (OSA) encompasses various measures aimed at safeguarding children in the online sphere. While the government has affirmed that the Metaverse falls under the purview of the OSA, little attention has been paid at this stage to addressing the unique safety challenges posed by virtual reality (VR) and generative AI within the regulatory framework.

The legislation mandates platforms to conduct routine risk assessments, yet it remains unclear how these risks will be effectively mitigated within the Metaverse. For instance, moderating real time interactions in virtual reality presents challenges, particularly given the current limitations in automated detection of harmful behaviour in real-time online settings. Moreover, the capacity for human moderation is constrained (Spence et al, 2023), suggesting that AI moderation may become necessary in the future. Mitigating risk will be difficult as few

applications have been developed with child safety as a primary consideration. Furthermore, the true extent of the challenge remains unknown due to a lack of research in this area.

VIRRAC was a successful 1-year pilot project that began to address some of these gaps in current understanding. However, it is acknowledged that this project was small scale, and further investment and exploration in this area is vital for the future of online harms. A key aim of VIRRAC was to raise awareness of child safety in the Metaverse, and to better understand the harms that children and young people face, but, further, to explore potential practical solutions. VIRRAC includes the perspectives of stakeholder experts, children, and young people themselves, and industry, by applying critical thinking of the safety features currently available to Metaverse users, and championing elements of 'safety by design' (Pothong, K. and Livingstone, 2021). The VIRRAC team established a rich stakeholder advisory board made up of industry and practice experts, conducted an extensive review of the wider literature landscape, and designed and coordinated 3 separate qualitative research studies with professional stakeholders as well as young people. Utilising the team's existing expertise, as well as the new knowledge gained within VIRRAC, the team has developed an evidence-led and accessible video resource for children, as well as 5 dedicated Metaverse safety Tools. These tactile tools are an inceptive sample of the anticipated full interactive Toolkit the team aims to produce in the near future.

While the OSA serves as a mechanism to hold industry accountable, there exists a pressing need for tech giants to bear responsibility, prioritising an ethos of safety by design. This collective focus is crucial to ensuring children's safety in the Metaverse, equipping them with clear and practical safety features while promoting proactive incident reporting.

VIRRAC has shed light on the lack of relevant literature, applied theory, and empirical research dedicated to child safety in the Metaverse. It has also underscored the lack of dedicated resources and awareness initiatives targeting educational institutions, frontline practitioners, parents, guardians, and policymakers. This gap poses a significant challenge for practitioners in education, health, and social care settings, who must strive to remain abreast of children's evolving experiences through ongoing training.