Tech Trends Position Paper

Immersive technologies

Australian Government | eSafetyCommissioner

eSafety.gov.au

Contents

Executive summary

Content warning: *This report discusses child sexual exploitation and abuse, sexual violence, terrorism and violent extremism, and other material some people may find distressing. Please take care when reading and consider whether this report is right for you at this time.*

Immersive technologies, such as augmented reality and virtual reality, enable users to experience and interact with digital content in three dimensions (3D). These technologies offer transformative benefits for sectors like health, sciences and education. However, they also have the capacity to amplify existing and emerging harms, including child sexual exploitation and abuse, terrorism and violent extremism, and tech-facilitated gender-based violence.

Before immersive technologies become even more ubiquitous, we have a unique opportunity to proactively implement safeguards that prevent abuse and misuse, while also promoting their benefits.

All players in the digital ecosystem must recognise this imperative and work together to make sure immersive technologies are designed with safety at their core. Tech companies, policymakers, regulators, academia, law enforcement, educators, practitioners, parents, and carers all have a critical role to play. Equally important is ensuring the voices of those with lived experience of harms in immersive environments, including those from negatively impacted and marginalised communities, are elevated at every stage of the design, development, and deployment process.

This position statement begins by defining key types of immersive technologies, before exploring the associated risks, harms, and opportunities. It then examines regulatory challenges and how immersive technologies fall within eSafety's remit. The final section outlines Safety by Design measures for industry to mitigate the safety risks of immersive technologies.

Background

Who we are

The eSafety Commissioner (eSafety) is Australia's independent regulator and educator for online safety.

Under the *Online Safety Act 2021* (Cth) (the Act), we coordinate government efforts to improve online safety for all Australians. We conduct research, provide education, and enforce regulatory schemes to combat online harm. We work with government agencies, businesses, and organisations – both nationally and internationally – to make the internet a safer place.

As an anticipatory regulator, a core part of our role is to stay abreast of technological developments and anticipate the emerging safety risks and opportunities they present. We also stay at the forefront of online safety issues through research, stakeholder engagement, and dialogue with experts across disciplines.

Our Tech Trends and Challenges program helps us understand existing and emerging online threats and opportunities. This knowledge shapes our regulatory guidance, resources, and advice to industry.

Our multidimensional regulatory remit gives us a range of powers to help protect Australians from harm and promote safer online experiences. These include:

- Our four complaints schemes, which apply to material which constitutes adult cyber abuse, child cyberbullying, image-based abuse, or illegal or restricted content that is class 1 and class 2 material.
- Industry codes and standards, which contain mandatory and enforceable compliance measures for key sections of online industry. There are currently six codes and two standards which contain measures to address the most seriously harmful online content, such as child sexual exploitation material and pro-terror material. A second phase of industry codes is in development, focusing on online pornography and other material inappropriate for children.
- The Basic Online Safety Expectations (the Expectations), which set out the Australian Government's expectations for certain services to take

reasonable steps to keep users safe. eSafety is also empowered to compel information from certain services about their compliance with the Expectations and to publish transparency reports.

- Implementation of the *Online Safety Amendment (Social Media Minimum Age) Act 2024* (Cth), which will require age-restricted social media platforms to take reasonable steps to prevent children under 16 from having accounts. Legislative rules may be made to specify that a service is or is not an age-restricted social media platform. The age restriction requirements will take effect on a date to be determined by the Minister, no later than 10 December 2025.

This position statement

As part of our commitment to remain current on and anticipate technological change, we review and update our published position statements where necessary.

In December 2020, we published a position statement on immersive technologies. Since then, immersive technologies have evolved and expanded in commercial availability, with major companies such as Apple[1] and Meta[2] launching flagship devices and platforms. While these developments present new opportunities across various sectors, they also introduce new safety risks.

As an online safety regulator, eSafety takes a harms-minimisation and risk-based regulatory approach. We also adopt a strengths-based and intersectional approach to understanding and responding to a person's online experience. This approach also guides our framing and discussion of emerging technologies. This updated position statement outlines current and emerging risks and harms, identifies opportunities, explores regulatory challenges, provides updated guidance on Safety by Design measures, and grounds this in what it means for the online safety of all users.

This position statement reflects eSafety's position as of **May 2025**.

[1] Apple, *Introducing Apple Vision Pro: Apple's first spatial computer*, 5 June 2023, https://www.apple.com/au/newsroom/2023/06/introducing-apple-vision-pro/
[2] BBC, *Meta releases social VR space Horizon Worlds*, 10 December 2021, https://www.bbc.com/news/technology-59609996

What are immersive technologies?

Immersive technologies enable users to experience and interact with digital content in 3D, in ways that look, sound, and feel almost real. They include **augmented reality (AR), virtual reality (VR), and mixed reality (MR)** – technologies that may be collectively referred to as **extended reality (XR)**.[3] They can be combined with **haptic sensory technologies**.

Definitions of key technologies vary. The following definitions and examples are not exhaustive and may evolve as the technologies continue to develop.

Definitions

Augmented reality (AR) overlays a user's view of the physical world with digitally generated, real-time sound and vision through devices such as smartphones or AR glasses. Google Lens and Pokémon Go are examples of commonly used AR apps.

Virtual reality (VR) uses computer hardware and software to create an artificial environment that looks and sounds as if the user is really present in that environment. Devices such as VR headsets and handheld controllers with sensors track head and hand movements to support interaction. PlayStation VR2 is an example of a VR headset.

Mixed reality (MR) devices combine elements of AR and VR by blending digital content into the physical world, allowing users to interact with virtual elements as an extension of reality. Virtual objects or characters behave as if they are real, interacting with light, sound, and space. Meta Quest Pro is an example of a commercially available MR device.

Extended reality (XR) is an umbrella term for any technology, such as AR, VR, or MR, that digitally alters the physical environment.

Haptic sensory technologies (haptics) simulate sensory experiences such as touch. When combined with XR, haptics enhance user experience by providing sensory feedback and allowing users to 'feel' interactions in the virtual environment. Simple haptics can be built into phones, controllers, and

[3] B Marr, Forbes, *What is extended reality technology? A simple explanation for anyone*, 10 December 2021, https://www.forbes.com/sites/bernardmarr/2019/08/12/what-is-extended-reality-technology-a-simple-explanation-for-anyone/

wearable devices, while more advanced experiences are delivered through full-body haptic suits with multiple sensors. Commercial examples include HoloSuit, Teslasuit, and Plexus VR Glove.

Social virtual reality (social VR) refers to shared virtual spaces where users interact using VR technology, often replicating in-person social interactions. These platforms can offer social experiences, such as games, parties, and other activities. Examples include Horizon Worlds, VRChat, Oculus Rooms, and Rec Room.

The metaverse is a continually evolving concept with varying definitions. One common definition describes the metaverse as a virtual world where users interact in a 3D space that closely resembles reality. Others define it as a fully immersive version of the internet. Some definitions suggest the metaverse will only be fully realised when an unlimited number of users can move freely between virtual environments. This vision of the metaverse has not yet been achieved, and technical or commercial barriers may impede its realisation.[4]

Exploring benefits and harms

Immersive technologies enable 3D interactions and real-time experiences that transcend geographical barriers and the limitations of traditional 2D media. These capabilities offer significant benefits, such as facilitating new opportunities for people living with disability, supporting therapeutic applications, and enhancing educational experiences. We explore positive uses in a later section of this paper.

However, alongside such opportunities, immersive technologies also present new ways to perpetrate abuse. eSafety has long recognised that the online world offers a mix of both benefits and harms. Adequate safeguards are essential to prevent immersive technologies being weaponised and to mitigate potential harms.

[4] eSafety, *The metaverse: A snapshot of experiences in virtual reality*, December 2023, page 7, https://www.esafety.gov.au/research/the-metaverse

Case study: Benefits and risks of interoperability in the metaverse

Interoperability is the ability for easy and frictionless user interaction and information exchange across different systems, platforms, and technologies. It is forecast to be fundamental to the operation of the metaverse, as opposed to having a separate cluster of multiple virtual worlds. If interoperability is achieved across major platforms, the metaverse could become one extended 3D environment. It may also enable a seamless digital identity experience across multiple platforms and may accelerate the growth of the metaverse economy.[5]

Interoperability has implications for online safety. A potential online safety benefit of interoperability is enabling safety settings to carry over from one platform or environment to the next. However, if service providers fail to implement appropriate safeguards, interoperability may also carry online safety risks. For example, it may enable children to move between platforms without restriction or detection, increasing their risk of exposure to age-inappropriate content and experiences. Safety must be a core consideration in the design and implementation of interoperable systems.

Interoperability may also have implications for data management. For example, a user may consent to sharing personal information with one platform but not with another.[6] Privacy and security are essential considerations alongside safety if immersive technologies move towards interoperability.

Online safety risks and harms

Both current and emerging immersive technologies present a range of risks and harms.

Some of these harms are beyond eSafety's remit. These include issues such as information integrity[7] and physical injury associated with the use of immersive

[5] World Economic Forum, *Interoperability in the metaverse*, 18 January 2023, https://www.weforum.org/publications/interoperability-in-the-metaverse/

[6] World Economic Forum, *Metaverse interoperability is essential. How will regulation play a part?*, 6 August 2024, https://www.weforum.org/stories/2024/08/metaverse-interoperability-regulation/

[7] M Del Castillo, Forbes, *Facebook's metaverse could be overrun by deep fakes and other misinformation if these non-profits don't succeed*, 29 August 2022, https://www.forbes.com/sites/michaeldelcastillo/2022/08/29/facebooks-metaverse-could-be-overrun-by-deep-fakes-and-other-misinformation-if-these-non-profits-dont-succeed/

devices. Some experts predict the metaverse could give rise to new forms of crime – sometimes referred to as metacrimes[8] – including property crimes (such as theft), cybercrimes (such as the use of ransomware), and financial crimes (such as money laundering or scams).[9]

This position statement focuses on online safety risks that fall within eSafety's remit under the Act, including child sexual exploitation and abuse (CSEA), terrorism and violent extremism, and cyberbullying. It also examines related risks, which are interconnected and can exacerbate online safety risks within eSafety's remit.

Understanding the risks

Immersive technologies can mimic interactions in the physical world, in real time. Like other online harms, the risks associated with immersive technologies also draw upon and extend existing social issues and conditions. This means they draw upon the same forms of discrimination, oppression, and inequality in society, including sexism, racism, ableism, homophobia, ageism, and transphobia.

It is therefore important to understand the relationship between harms that exist in society, which are then amplified, extended, or facilitated through immersive technologies.

As immersive technologies evolve, tech companies, policymakers, and regulators must proactively examine the evolution of risks within the context of new and emerging technologies. This underscores the need for collaboration across the digital ecosystem to better understand and respond to the emerging harms landscape in the context of evolving societal risks and harms.

Amplification of existing harms

Existing harms which are experienced both offline and in 2D online environments can be amplified in immersive environments. The features of immersive technologies – such as their cross-dimensional, borderless, and hyper-realistic nature – can lead to new manifestations of existing harms.

---

[8] The term 'metacrimes' is not universally defined, but broadly refers to criminal events in the metaverse. See: Y Zhou et al., *Metacrime and cybercrime: Exploring the convergence and divergence in digital criminality*, 9 August 2024, https://doi.org/10.1007/s11417-024-09436-y

[9] Interpol, *Metaverse: A law enforcement perspective*, January 2024, page 13, https://www.interpol.int/en/News-and-Events/News/2024/Grooming-radicalization-and-cyber-attacks-INTERPOL-warns-of-Metacrime

Harms that occur in immersive environments must not be minimised because they do not take place in the physical world. In fact, the heightened experiences of immersive environments can have real and significant impacts on users, with the impact of immersive experiences likened to being 'drawn on the brain in permanent ink'.[10]

The unique interactive features of immersive environments mean the impacts of interactions and experiences can be distinct from 2D online experiences and similar to physical-world experiences. For example, users may experience 'embodiment', where they feel that their avatar or virtual body is their physical body and, therefore, what has happened to their virtual self also happened to their physical self.[11]

In 2024, researchers at the Stanford Virtual Human Interaction Lab conducted a study on shared body sensations via haptic feedback in social VR. Shared body sensations occur when an individual experiences the sense of touch while seeing that touch directed to another person. Participants in the study (n=32) observed conversations between two virtual agents, sharing touch with one virtual agent. The findings indicated that shared body sensations can induce stronger body illusion and empathy,[12] which may point to the unique psychological impacts of interactions in immersive environments.

This section explores several harms that are likely to be heightened in immersive environments:

- Terrorism and violent extremism
- Child sexual exploitation and abuse
- Exposure to age-inappropriate content and experiences for children
- Sexual violence and gender-based violence
- Negative online social experiences

[10] B Heller, *Watching Androids Dream of Electric Sheep: Immersive technology, biometric psychography, and the law*, 2020, page 22 https://scholarship.law.vanderbilt.edu/jetlaw/vol23/iss1/1/

[11] E J Ramirez et al., *XR embodiment and the changing nature of sexual harassment*, 2 February 2023, https://doi.org/10.3390/soc13020036

[12] Y Tao, J Egelman, J N Bailenson, *I feel you: Impact of shared body sensations on social interactions in virtual reality,* 21 October 2024, https://vhil.stanford.edu/publications/social-interaction/i-feel-you-impact-shared-body-sensations-social-interactions

Terrorism and violent extremism

Perpetrators of terrorism and violent extremism adapt their methods and exploit technological advancements to enhance their coordination, propaganda, and indoctrination efforts.[13] Interpol anticipates perpetrators may misuse the metaverse to commit four categories of harm: radicalisation and indoctrination; financing terrorism; cyberterrorism through cyber-physical attacks; and coordination of physical-world terrorist attacks.[14]

Immersive technologies could be weaponised to perpetrate terrorism and violent extremism in various ways:

- The interactive and embodied nature of immersive environments may amplify the effect of echo chambers,[15] enabling terrorist and violent extremist groups to perpetuate their ideologies and leading to increased risks of recruitment and radicalisation.[16] Additionally, there have been instances of users recreating and role-playing terror attacks on immersive platforms which are popular amongst children and young people. This has raised concerns regarding the exposure of children and young people to extremist ideologies and potential radicalisation.[17]
- The reach of terrorist and violent extremist groups may expand, as immersive spaces may facilitate meaningful connections between bad actors, supporters, and potential recruits.[18]
- Perpetrators could access realistic virtual recreations of physical-world sites to plan, rehearse, or coordinate terror attacks.[19]

---

[13] OECD, *Transparency reporting on terrorist and violent extremist content online: Fourth edition*, 24 June 2024, https://www.oecd.org/en/publications/transparency-reporting-on-terrorist-and-violent-extremist-content-online_901cb8cf-en.html

[14] Interpol, *Metaverse: A law enforcement perspective*, page 15.

[15] Amplification of any content necessarily involves giving it preference over other content. By amplifying some content, recommender systems can end up deprioritising or excluding different viewpoints or valuable ideas contrary to a person's existing beliefs, contributing to what are commonly known as 'echo chambers'. eSafety's position statement on recommender systems and algorithms provides additional information, eSafety Commissioner, *Recommender systems and algorithms – position statement,* December 2022, https://www.esafety.gov.au/industry/tech-trends-and-challenges/recommender-systems-and-algorithms

[16] G Weimann and R Dimant, *The metaverse and terrorism: Threats and challenges*, Perspectives on Terrorism, June 2023, pages 95-96, https://pt.icct.nl/article/metaverse-and-terrorism-threats-and-challenges?

[17] Australian Federal Police, *Holiday season warning: Extremists infiltrating online and gaming platforms to recruit young Australians*, 3 December 2023, https://www.afp.gov.au/news-centre/media-release/holiday-season-warning-extremists-infiltrating-online-and-gaming

[18] Weimann, *The metaverse and terrorism*.

[19] Interpol, *Metaverse: A law enforcement perspective*, page 15.

- Wearable XR devices, such as AR smart glasses, could be used to livestream terror attacks.[20] Terror attacks could also be broadcast in immersive environments, where bystanders could be unwittingly exposed to such attacks, the content of which are likely to be more violent and highly visceral, intensified by the immersive nature of their environment.

eSafety has powers under the Act relating to terrorist and violent extremist online content. We explore these regulatory powers in the later section, **'How do immersive technologies fit within eSafety's remit?'**.

Child sexual exploitation and abuse (CSEA)

Immersive technologies can create new avenues for abusers to have illegal and inappropriate contact with children. Though this section specifically discusses illegal harms towards children, given the fluid and individual nature of harms and the complexities of adolescent cognitive development, many of these harms also apply to young adults aged 18 and above.

Early iterations of XR products have emerged through the gaming market, including platforms popular with children. Abusers can weaponise these platforms and use the anonymity afforded by avatars to approach and groom children, who may believe they are interacting with a peer, rather than an adult. Immersive spaces allow anonymity to combine with expressive features such as gesture, movement, and body language, which may enhance the ability for abusers to build rapport with children.[21]

Children are particularly vulnerable in immersive spaces that are not age-appropriate or that enable adults to have contact with children. Once contact and trust have been established, the abuse may escalate. Abusers may engage in sexualised conversations, perpetrate sexually exploitative or abusive activities towards children, or force children to act out sexual movements.[22] Abusers may move their interactions from the immersive platform to somewhere they may be less likely to be detected, such as 2D messaging

[20] National Counterterrorism Innovation, Technology, and Education Center, *Meta's Glasses, ChatGPT, and the Rise of High-Tech Extremist Violence*, 8 January 2025, https://www.unomaha.edu/ncite/news/2025/01/hunter-meta-glasses.php

[21] C Allen and V McIntosh, Limina Immersive commissioned by NSPCC, *Child safeguarding and immersive technologies: An outline of the risks*, September 2023, page 22 https://learning.nspcc.org.uk/research-resources/2023/child-safeguarding-immersive-technologies

[22] M Guggisberg, The Conversation, *Virtual reality grooming is an increasing danger. How can parents keep children safe?*, 1 February 2024, https://theconversation.com/virtual-reality-grooming-is-an-increasing-danger-how-can-parents-keep-children-safe-221608

platforms that are encrypted,[23] where abusers may coerce children into creating CSEA material of themselves or coerce children into offline interactions.

There are emerging concerns that pre-existing harms to children may find new manifestations in immersive environments and risks to child safety will continue to grow as immersive technologies develop. There have already been cases of abusers using VR headsets to view and store CSEA material.[24] Some experts predict that the metaverse may be used to facilitate immersive streaming of CSEA material, including AI-generated CSEA material. This may take place in metaverse spaces where multiple abusers could congregate to view the material.[25] Additionally, there is anecdotal evidence that presents concerns about adults using avatars appearing as children to simulate sexual activity occurring in immersive environments,[26] which amounts to CSEA given the nature of the depiction.[27] Further research into the harms and effects of such emerging behaviours are required to consider their impacts, including their potential adverse influence on norms, in immersive environments.

Exposure to age-inappropriate content and experiences for children

Immersive technologies may amplify the risk and impact of children encountering age-inappropriate content, such as pornography or violent material. When interacting in a 3D, immersive environment, such material can be more visceral than when viewed on a 2D screen.

Further studies into the impact of immersive technologies on children and young people are required. However, existing research indicates that children

---

[23] End-to-end encryption can prevent or limit the detection of online grooming or CSEA by blocking technologies that can identify illegal material and activity. For more information, refer to eSafety's position statement on end-to-end encryption.

[24] A Crawford, BBC, *Child abuse material found on VR headsets, police data shows*, 23 February 2023, https://www.bbc.com/news/uk-64734308

[25] J Gomez-Quintero et al., *A scoping study of crime facilitated by the metaverse*, 12 February 2024, Pages 15, 18, https://doi.org/10.1016/j.futures.2024.103338.

[26] M Chawki, S Basu and KS Choi, *Redefining boundaries in the metaverse: Navigating the challenges of virtual harm and user safety*, May 2024, page 6, https://doi.org/10.3390/laws13030033; C Reeves, *Fantasy depictions of child sexual abuse: The problem of ageplay in Second Life*, 2012, https://eprints.hud.ac.uk/id/eprint/13133/, C Reeves, *The virtual simulation of child sexual abuse: online gameworld users' views, understanding and responses to sexual ageplay*, January 2018, https://link.springer.com/article/10.1007/s10676-018-9449-5.

[27] The behaviour of consenting adults engaging with each other in sexual activity involving a roleplay or power exchange dynamic, where one or more participants assume the persona of a different age group, is sometimes referred to as 'sexual ageplay'. When carried out in a virtual environment, including immersive environments, this is sometimes referred to as 'virtual sexual ageplay'. However, where virtual sexual ageplay uses avatars who either are or appear as children, this constitutes CSEA, as it is depicting sexual abuse against a child.

may experience VR similar to how they experience physical reality.[28] The illusion of VR may be more effective on young children and those who are still developing the critical reasoning skills needed to distinguish between events which happen in the virtual world and events in the physical world.[29]

There are already reports of children being exposed to age-inappropriate experiences on immersive platforms, underscoring the urgent need for platforms to implement better safeguards.

In 2023, the Center for Countering Digital Hate (CCDH) conducted a study assessing children's exposure to child-inappropriate content in Meta's Horizon Worlds.[30] At the time of the CCDH's study, Horizon Worlds had a minimum age of 18 years.[31]

Researchers visited 100 virtual worlds for five minutes each. They identified minors present in 66 of these worlds. The CCDH recorded 19 incidents of abuse directed towards minors, including sexually explicit insults and harassment based on perceptions of gender, sexuality, and race. Researchers also identified minors in Mature Worlds – worlds which can contain sexual content, gambling, or intense violence[32] – including in worlds containing virtual strip clubs.

[28] J O Bailey and J N Bailenson, *Cognitive Development in Digital Contexts*, 2017, page 186, https://doi.org/10.1016/B978-0-12-809481-5.00009-2

[29] B Kenwright, *Virtual reality: ethical challenges and dangers [opinion]*, IEEE Technology and Society Magazine, 2018, https://ieeexplore.ieee.org/document/8558774

[30] Center for Countering Digital Hate, *Horizon Worlds Exposed*, 8 March 2023, https://counterhate.com/research/horizon-worlds-exposed/

[31] In 2023, Meta expanded Horizon Worlds to 13–17-year-old users, commencing in the US and Canada. In November 2024, Meta lowered the minimum age to 10–12-year-old users with parental consent and using a child account – which prevents children from voice chatting with other users – in the US and Canada. An investigation from July 2024 to April 2025 by children's organisation, Fairplay, found voices of children under 13 years in multiple games and experiences in Horizon Worlds, indicating such children were using non-child accounts which lacked privacy and safety settings. On 10 April 2025, Fairplay filed a request for the Federal Trade Commission to investigate Meta, based on allegations that Meta has knowingly let children under 13 years to improperly access Horizon Worlds and collected their personal data without parental notice and consent. See more: https://fairplayforkids.org/complaint-against-meta-filed-at-ftc-alleging-widespread-horizon-worlds-child-privacy-violations/.

[32] Meta Horizon, *Mature and prohibited worlds policy*, 2024, https://developers.meta.com/horizon-worlds/learn/documentation/save-optimize-and-publish/restrictions-to-worlds-in-horizon

Sexual violence and gender-based violence

Sextortion and image-based abuse

Immersive technologies present enhanced risks for sexual harms which already occur on traditional 2D platforms, such as sexual extortion.[33] For example, a user may engage in an intimate interaction within an immersive space which they believe is private, only for the interaction to be recorded without their consent and threatened to be shared unless the user complies with the demands of the other party.

Avatar-based interactions have also led to new ways to perpetrate sexual harms. eSafety's metaverse research found that 6% of metaverse users who were surveyed (n=259) reported someone had created a sexually explicit avatar of them to interact with, without their consent.[34] This may expand traditional perceptions of image-based abuse and raises important questions for legal definitions and regulation.

Tech-facilitated gender-based violence (TFGBV)

Gender-based violence (GBV) is any form of physical or non-physical violence or abuse against a person or group of people because of biased or harmful beliefs about gender. GBV mainly impacts women, girls and LGBTIQ+ people. However, gender stereotypes can also harm and affect men and boys.

TFGBV is where GBV is facilitated by technologies. It includes technology-facilitated behaviours such as sexual harassment, misogynistic comments, homophobic language, stalking, coercive control, image-based abuse, and threats of sexual violence. These behaviours are already prevalent online[35] and are now occurring in immersive spaces.

The hyperrealism and real-time experiences afforded by immersive technologies allow perpetrators to move beyond the limitations of 2D platforms and 'act out' threats of sexual violence and harassment. A survey of more than 600 users of popular VR devices found 49% of women respondents had experienced sexual harassment in VR environments. Reported behaviours

[33] R Sparrow and L Karas, *Teledildonics and rape by deception*, 2020, https://doi.org/10.1080/17579961.2020.1727097
[34] eSafety, *The metaverse: A snapshot of experiences in virtual reality*, page 17.
[35] A Powell, A Flynn, S Hindes, ANROWS, *Technology-facilitated abuse: National survey of Australian daults' experiences*, July 2022, https://anrows-2019.s3.ap-southeast-2.amazonaws.com/wp-content/uploads/2022/07/27172214/4AP.3-Flynn-TFa3-Survey-of-VS.pdf

included groping, stalking, catcalling, being shown explicit images, and receiving sexually explicit comments.[36]

These experiences compound the already disproportionate levels and impacts of other forms of violence experienced by women and LGBTIQ+ people. They also reinforce the harmful gender attitudes and social norms that drive all forms of GBV, by enabling and trivialising the use of violence.

Case study: Embodied harassment

'Embodied harassment'[37] is an emerging term that describes unwanted behaviours facilitated by avatars in immersive spaces. It can include actions such as unwanted touching or groping of another person's avatar, performing unwanted sexual gestures towards another person's avatar, or invading another avatar's personal space.

Haptic technologies, such as teledildonics, can amplify embodied harassment. Teledildonics are designed to stimulate sexual excitement through the remote control of sex toys. While they can be used consensually, where they are used without consent, victim-survivors may both perceive and feel unwanted touch. There is an additional risk of teledildonics being used alongside XR devices to perpetrate non-consensual sexual activity. For example, this could occur if a victim-survivor believes their intimate partner is controlling the device, when in fact it is being controlled or hacked by someone else who they have not provided consent to.[38]

There have already been multiple reports of embodied harassment in immersive spaces. For example:

- **October 2016:** Journalist Jordan Belamire published a post, 'My First Virtual Reality Groping' detailing her experience being groped in the VR game, QuiVr. She stated, 'the virtual groping feels just as real… it felt real, violating'.[39]

[36] J Outlaw, *Virtual harassment: The social experience of 600+ regular virtual reality (VR) users*, 4 April 2018, https://virtualrealitypop.com/virtual-harassment-the-social-experience-of-600-regular-virtual-reality-vr-users-23b1b4ef884e
[37] J E Gray, M Carter and B Egliston, *Governing Social Virtual Reality*, 2024, page 24.
[38] R Sparrow, *Teledildonics and rape by deception*.
[39] J Belamire, *My first virtual reality groping*, 21 October 2016, https://medium.com/athena-talks/my-first-virtual-reality-sexual-assault-2330410b62ee

- **November 2021**: a woman beta tester for Meta's Horizon Worlds was groped. She described her experience, '… being in VR adds another layer that makes the event more intense.'[40]
- **December 2021:** Nina Jane Patel, co-founder of XR education company Kabuni, was sexually harassed within a minute of joining Meta's Horizon Worlds. Her avatar was virtually gang raped by three avatars who appeared to be men, who yelled derogatory remarks at her as she attempted to get away. She stated that her 'physiological and psychological response was as though it happened in reality'.[41]
- **January 2024:** UK police opened an investigation into the virtual gang rape of an under-16-year-old girl. While playing an immersive game in the metaverse, several avatars assaulted her avatar in a virtual room. A police officer reported the victim-survivor suffered psychological trauma 'similar to that of someone who has been physically raped'.[42]

Research on the psychological impacts of embodied harassment in immersive experiences is still emerging. However, research indicates that acts such as invading another user's space in social VR can evoke feelings of discomfort or feeling threatened, as if someone was physically approaching too closely.[43] Research also suggests that when a user perceives interactions, such as touch or invasion of personal space directed at their avatar, they may perceive this interaction as happening to their own physical body.[44]

Experiences of embodied harassment must not be diminished. Although these adverse experiences occur in virtual environments, the effects are real. As the cases above illustrate, victim-survivors have reported impacts comparable to those they would have experienced had the incident occurred in the physical world.

---

[40] T Basu, *The metaverse has a groping problem already*, 16 December 2021, https://www.technologyreview.com/2021/12/16/1042516/the-metaverse-has-a-groping-problem/
[41] N J Patel, *Reality or fiction?*, 22 December 2021, https://medium.com/kabuni/fiction-vs-non-fiction-98aa0098f3b0
[42] C Vallance, *Police investigate virtual sex assault on girl's avatar*, 3 January 2024, https://www.bbc.com/news/technology-67865327
[43] G Freeman et al., *Disturbing the peace: Experiencing and mitigating emerging harassment in social virtual reality*, April 2022, page 12, https://doi.org/10.1145/3512932
[44] K Schulenberg et al., *'Creepy towards my avatar body, creepy towards my body': How women experience and manage harassment risks in social virtual reality*, October 2023, page 10, https://doi.org/10.1145/3610027

Negative online social experiences

Negative online experiences that occur on 2D platforms also persist in immersive environments. The hyper-realistic nature of immersive technologies may intensify the impact of negative interactions.

In immersive environments, harassing behaviours such as stalking and unwanted contact can feel more intense than harassment on traditional social media. Where perpetrators were previously limited to the confines of a 2D screen, immersive technologies can enable them to appear as though they are physically following other users or invading their personal space. This means the onus is often wrongly on users to safely remove themselves from abusive situations and makes it more difficult for them to do so. The real-time and often ephemeral nature of these interactions may also create barriers to reporting harmful behaviour or seeking remedial action.

In November and December 2022, eSafety conducted a survey of over 5,000 Australian adults about their online experiences, including 259 metaverse users. Our research found that 71% of metaverse users reported having at least one negative online experience in the metaverse in the last 12 months. The most common types of negative experiences included:

- being left out by others (23%)
- being called offensive names (21%)
- receiving repeated unwanted messages or contact (16%)
- being provoked to respond to something said or to start an argument (15%)
- being challenged about cultural identity (14%)
- being sent unwanted inappropriate content (13%).

Almost 1 in 2 (49%) respondents who reported a negative experience in the metaverse said the experience had a moderate to extreme impact on their mental or emotional wellbeing.[45]

[45] eSafety, *The metaverse: A snapshot of experiences in virtual reality*, page 19.

Other related risks

While this paper focuses on online risks within eSafety's remit, the following section explores related risks which are interconnected and can exacerbate the online safety harms eSafety regulates.

An overarching consideration is the impact of immersive technologies on human rights. The misuse of immersive technologies can have negative impacts on the protection and promotion of human rights under several instruments. This includes undermining the right to privacy due to the misuse of personal data[46] and facilitating violations of freedoms from discrimination.[47] This section will explore the relationship between immersive technologies and the risks of privacy and security, and barriers to access, lack of inclusivity and bias, which are related to online safety risks. Addressing these risks is important to support a holistic and multidimensional understanding and approach to user safety.

Privacy and security risks

Immersive technologies present a range of privacy and security risks. Upholding privacy and security, alongside safety, is critical for user agency, dignity, and empowerment in digital spaces. eSafety's position is that safety, privacy, and security are not mutually exclusive and each can be maintained through thoughtful and intentional design.

Protecting privacy in immersive technology use is a vital component of user safety. Inadequate privacy measures can contribute to some of the online safety risks associated with immersive technologies. For example, the absence of enforceable community guidelines or mechanisms that promote user privacy may lead to invasions of privacy. Users may believe they are participating in private or sensitive settings, only for another user to record their behaviours and share it without consent.

Immersive technologies can have different privacy implications based on the type and amount of personal information, including sensitive information, collected – and how that data is used, stored, and shared.

[46] *International Covenant on Civil and Political Rights*, art 17.
[47] See e.g. *International Covenant on Civil and Political Rights*, arts 2, 14, 24, 25-26

Immersive technologies can collect a wide range of data types. This can lead to misuse and can have far reaching implications beyond simply identifying 'who' a user is; the range of data types involved means 'sensitive information'[48] may be discerned, such as a person's racial or ethnic origin, religious beliefs, or political opinions, including at a granular level.

The following are examples of the types of data that could be collected by immersive technologies. This list is not exhaustive:

Personal information refers to a broad range of information, or an opinion, that could identify a person. The following types of data collected by immersive technologies may also constitute personal information under the *Privacy Act 1988* (Cth) (Privacy Act) and be subject to privacy protections where they are about an identified or reasonably identifiable individual.

- **Account information** includes information collected when creating or associated with an account. This can include a person's full name, date of birth, residential address, email address, phone number, or IP address.[49]
- **Psychographic data**[50] refers to information about a user's preferences, interests and values based on their behaviours, actions, and reactions during immersive experiences.
- **Spatial data** is information about the movements, locations, and dimensions of individuals and objects within virtual environments. XR devices often capture and process data such as gestures and eye tracking to create immersive experiences. Emerging studies suggest spatial data may be increasingly sensitive. For example, one study examining data from more than 50,000 VR users showed participants were able to be uniquely identified by 100 seconds of their head and hand motions with 94.33% accuracy.[51]

Information collected by immersive technologies may also be classified as sensitive information under the Privacy Act. For example:

[48] Sensitive information has a higher level of privacy protection than other personal information. See: Office of the Australian Information Commissioner, *What is personal information?*, 2024, https://www.oaic.gov.au/privacy/your-privacy-rights/your-personal-information/what-is-personal-information

[49] Ibid.

[50] World Economic Forum, *Metaverse Safety and Privacy*, July 2023, page 15, Privacy and Safety in the Metaverse | World Economic Forum

[51] V Nair et al., *Unique identification of 50,000+ virtual reality users from head and hand motion data*, 17 February 2023, https://rdi.berkeley.edu/metaverse/identification/

- **Biometric data**,[52] refers to physiological or behavioural characteristics that can be used to identify a person. Physiological biometrics may include a person's fingerprints, facial features, or hand. Behavioural biometrics may include a person's signature or keystroke pattern.

The large volumes of personal information, including sensitive information, collected through users' engagement with immersive technologies may be misused for malicious purposes. This could include perpetrating virtual or physical harm towards individuals identified as belonging to marginalised or minority groups or identifying and doxing victim-survivors of family, domestic and sexual violence who may be seeking refuge in undisclosed locations.

Privacy considerations also have implications for safety moderation. Safety measures are essential to prevent harm in immersive environments. Tech companies should also implement privacy-respecting moderation systems that support synchronous enforcement of community guidelines. Moderation and safety measures should be proportionate to the level of risk and balanced with other rights and freedoms. For example, higher-risk environments may warrant more active moderation than lower-risk spaces. Additionally, 'private' spaces could have reduced moderation, with alternative safety measures in place.[53]

Barriers to access, lack of inclusivity, and bias

Failing to design immersive technologies with inclusivity and accessibility in mind risks deepening inequality, reinforcing social division, and further marginalising individuals and communities. This may impact online safety. If immersive technologies become more embedded in everyday digital experiences – or even essential for participation in aspects of daily life – barriers to access could prevent people from marginalised groups from safely engaging in these environments.

There are several risk factors to consider during the design, development, and deployment of immersive technology products:

- **Lack of accessibility and inclusivity in design:** XR devices are typically designed based on normative male bodies - typically a cisgender, able-

[52] Under the current Privacy Act, biometric information is considered sensitive information. The Australian Government response to the Privacy Act review supported the proposal that Australia's independent privacy regulator, the Office of the Australian Information Commissioner, should continue to develop practice-specific guidance for new technologies and emerging privacy risk.
[53] World Economic Forum, *Metaverse Safety and Privacy*, page 31.

bodied man, among other factors.[54] This impacts the ability for individuals who differ from this to use XR devices and may alienate these groups from immersive experiences. For example, people living with disability may be unable to use XR devices without compatible assistive technology.

- **Lack of inclusive representation in immersive platforms**: A lack of diverse representation may discourage users from participating in immersive environments, reducing their sense of belonging and inclusion.
- **Extension of bias and discrimination into immersive worlds**: Existing prejudices and discriminatory behaviours can carry over into immersive environments and be further compounded. Users from marginalised communities may also experience discrimination in the virtual world, particularly those who choose to design their avatars as representations of themselves (for example, reflecting their gender or ethnicity).
- **Lack of digital literacy**: Structural barriers to developing technical skills or knowledge may prevent some users from engaging with immersive technologies. This may disproportionately affect people without access to digital education, including those from economically disadvantaged communities, those living in rural, remote or regional areas, and elderly users. The immersive nature of these technologies may also require specific skills and competencies in addition to digital literacy.

Opportunities

Immersive technologies offer a range of opportunities for inclusivity and accessibility, health, and education. Some of these benefits also pose opportunities for human rights fulfilment. For example, the right to education may be enhanced when immersive technologies are integrated into classrooms[55] – a benefit we explore in this section.

Technology is neither inherently good nor bad; its impact depends on how it is used. Whether immersive technologies create opportunities or challenges will be determined by the values and intentions embedded in their design and deployment.

[54] Gray, *Governing Social Virtual Reality*, page 50.
[55] *International Covenant on Economic, Social and Cultural Rights*, art 13.

Education and learning

Immersive technologies offer new opportunities to enhance educational experiences. For example, VR could enable virtual excursions to locations in other cities or countries, or to reconstructed historical sites.[56] Teachers could also use AR to overlay visual content onto the physical world to help students visualise complex and abstract concepts.[57]

Research suggests that immersive learning methods may offer increased educational benefits compared to less immersive methods, particularly in subject areas which are highly abstract, conceptually challenging, or have a procedural focus.[58] Immersive technologies could also support experiential learning by allowing individuals to practice, model, and develop skills in practical, virtual scenarios.[59] However, many studies into the effects of immersive learning have largely focused on a limited range of subject areas, such as science, technology, engineering and mathematics (STEM) disciplines.[60]

Spotlight: Consultation with eSafety's Youth Council

eSafety's Youth Council, comprised of members aged 13 to 24, provides a platform for young people to share their insights and lived experiences about online safety. Their insights contribute to policy and program design on online safety education and issues.

In April 2024, eSafety consulted the Youth Council about its views on immersive technologies and their impact on young people. The Youth Council identified several potential educational benefits:

- Equitable access to education, particularly for students in rural, regional and remote areas.
- Truly immersive remote learning, where students could sit in virtual classrooms and interact with teachers and peers in dynamic, 3D environments.

---

[56] UNICEF, *The metaverse, extended reality and children*, May 2023, page 4, https://www.unicef.org/innocenti/reports/metaverse-extended-reality-and-children
[57] Ibid page 10.
[58] D Hamilton et al., *Immersive virtual reality as a pedagogical tool in education: A systematic literature review of quantitative learning outcomes and experimental design*, 11 July 2020, page 21, https://doi.org/10.1007/s40692-020-00169-2
[59] World Economic Forum, *Experiential learning and VR will reshape the future of education*, 23 May 2022, https://www.weforum.org/stories/2022/05/the-future-of-education-is-in-experiential-learning-and-vr/
[60] Hamilton, *Immersive virtual reality as a pedagogical tool in education*.

- Tailored learning experiences, where students could choose virtual classrooms aligned with their individual learning preferences, giving them greater autonomy over their education.
- Exposure to 'trial' experiences, helping students explore unfamiliar activities. For example, medical students could practice surgeries on a virtual body before operating in the physical world, and young people could learn to drive in a virtual setting before driving in the physical world.

The Youth Council also discussed their views on risks and harms associated with immersive technologies, including:

- Regulatory challenges posed by the novel nature of immersive technologies.
- Overuse of immersive technologies, with suggestions that regulation and usage guidelines should be developed to prevent overuse.
- Over-reliance on immersive technologies for social interaction, which Youth Council members believed could negatively affect user's ability to engage in offline social interactions.

Inclusivity and accessibility

When designed with inclusivity and accessibility in mind and in consultation with people with diverse lived experiences, immersive technologies can provide pivotal opportunities for all individuals. For example:

- A study by researchers from the University of South Australia and the University of New South Wales suggests VR headsets may be viable tools to help support people with intellectual disability to live more independently.[61]
- VR and AR can provide social and emotional enrichment for residents in aged care. Aged care residents, particularly those with mobility impairments, often experience physical limitations that can contribute to isolation. VR and AR devices offer residents the opportunity to 'travel' and participate in activities outside their aged care homes.[62]

[61] A Franze et al., *Immersive virtual reality is more effective than non-immersive devices for developing real-world skills in people with intellectual disability*, 21 August 2024, https://doi.org/10.1111/jir.13177

[62] D Johnson, *How immersive virtual reality is helping aged care residents travel the world without leaving home*, 15 March 2024, https://www.abc.net.au/news/2024-03-15/immersive-reality-helps-aged-care-living-with-disability-travel/103585628

- The Apple Vision Pro includes accessibility features which can support users with vision and mobility impairments. A *New York Magazine* article highlights the experiences of a woman living with vision impairment using the device, which has allowed her to access experiences she was previously unable to. However, the article also notes there are still improvements to be made to make the Apple Vision Pro more accessible.

Case study: Convergence of immersive technologies with neurotechnology

Neurotechnology is a rapidly expanding field of technologies dedicated to understanding and interacting with the brain or nervous system activity. Definitions vary, but the Australian Human Rights Commission (AHRC) has adopted the following definition of neurotechnology: 'devices and procedures used to access, monitor, investigate, assess, manipulate and/or emulate the structure and function of the neural systems of natural persons'.[63]

Some neurotechnology devices aim to restore or improve brain function. Brain-computer interfaces (BCIs), for example, enable communication between a person's brain and an external device. BCIs can have profound impacts for people living with disability. They may have the potential to restore motor function for individuals with neuromuscular or paralysing disorders by re-establishing neural and muscular connections.[64] In November 2024, Canada approved Neuralink to launch a clinical trial to study the potential for BCIs to enable individuals with quadriplegia to control external devices through thinking.[65]

Possibilities for the convergence of neurotechnology and immersive technologies are growing. XR devices already use sensors to track real-time user behaviour. Future innovation may see this sensor-based technology integrated with neurotechnological devices such as BCIs.[66]

---

[63] Australian Human Rights Commission, *Protecting cognition: Background paper on neurotechnology*, 12 March 2024, page 8 https://humanrights.gov.au/our-work/technology-and-human-rights/publications/protecting-cognition-background-paper.

[64] For example, see M Sebastián-Romagosa et al., *Brain Computer Interface Treatment for Motor Rehabilitation of Upper Extremity of Stroke Patients – A Feasibility Study*, October 2020, https://doi.org/10.3389/fnins.2020.591435

[65] Reuters, *Elon Musk's Neuralink receives Canadian approval for brain chip trial*, 22 November 2024, https://www.reuters.com/business/healthcare-pharmaceuticals/elon-musks-neuralink-receives-canadian-approval-brain-chip-trial-2024-11-21/

[66] Varjo, *Insight session: Bridging neuroscience and XR for superpowering the human experience*, 26 January 2023, https://varjo.com/webinar/insight-session-bridging-neuroscience-and-xr-for-superpowering-the-human-experience/

This convergence raises important privacy concerns regarding the collection and user of neural data. There is emerging discourse about ensuring 'neurorights' – an umbrella term referring to several rights-related risks to the human mind, such as mental privacy and mental integrity – are preserved in neurotechnology and that users are protected from misuse.[67] In March 2024, the AHRC published its Background Paper on Neurotechnology and Human Rights, highlighting how neurotechnology may pose challenges to existing human rights protections.

Therapeutic benefits

Immersive technologies could serve as complementary tools to aid therapeutic interventions. Virtual reality therapy (VRT) is the use of VR environments to create controlled, interactive experiences in which patients can address mental health conditions. Studies suggest that VRT can be effective in supporting treatment for anxiety and depression.[68]

Immersive technologies may also help raise awareness and understanding of gender-based violence, particularly for those working in and around family, domestic and sexual violence. Teesside University, in collaboration with Police and Crime Commissioners, produced an immersive VR film titled *Through the Eyes of Another*, for judiciaries and other professionals working in North-Eastern UK family courts. Guided by a domestic violence advisor, the film depicts in immersive form a victim-survivor's experience of domestic violence and their navigation of the family justice system.[69]

Emerging research suggests immersive technologies could be used to challenge implicit biases and harmful stereotypes, such as towards race or gender.[70] The first round of eSafety's Preventing Tech-Based Abuse of

[67] Australian Human Rights Commission, *Protecting cognition: Background paper on neurotechnology*, page 10.
[68] N Baghaei et al., *Virtual reality for supporting the treatment of depression and anxiety: Scoping review*, 23 September 2021, https://doi.org/10.2196/29681
[69] L Hobson, BBC, *Virtual reality used to understand domestic abuse*, 16 September 2024, https://www.bbc.com/news/articles/cp8n67wxgjro; P Jacques, Police Professional, *VR gives judges a new perspective on impact of domestic abuse*, 12 September 2024, https://policeprofessional.com/news/vr-gives-judges-a-new-perspective-on-impact-of-domestic-abuse/
[70] D Banakou, P D Hanumanthu and M Slater, *Virtual embodiment of white people in a black virtual body leads to a sustained reduction in their implicit racial bias*, November 2016, https://doi.org/10.3389/fnhum.2016.00601; S Schulze et al., *The effects of embodiment in virtual reality on implicit gender bias*, June 2019, https://doi.org/10.1007/978-3-030-21607-8_28; L Wu and K B Chen, *Examining the effects of gender transfer in virtual reality on implicit gender bias*, May 2024, https://doi.org/10.1177/00187208221145264

Women Grants program provided funding for Monash University to deliver a project aimed at preventing TFGBV. The project will involve developing VR scenarios to be used within men's behaviour change programs to build awareness of TFGBV and build empathy towards TFGBV victim-survivors.

Regulatory challenges and approaches

Challenges

The novel characteristics of immersive technologies, such as their synchronous or ephemeral nature, may pose challenges for online safety regulation not otherwise seen in 2D online media. Immersive technologies may also present further challenges for law enforcement, non-government organisations, and other bodies investigating harms such as online CSEA. For example, immersive environments may pose difficulties for users to capture evidence of harm and for regulators, platforms and law enforcement to detect harm, due to the fleeting nature of immersive interactions and the absence of URLs to pinpoint locations of harm.

Online safety policies and regulations must be capable of addressing the emerging harms associated with immersive technologies.

Striking a balance between tech-agnostic and tech-specific approaches

Immersive technologies are continually evolving, and concepts like a fully interconnected metaverse remain speculative. Additionally, immersive technologies will also continue to converge with other emerging technologies, such as generative AI.[71] To avoid falling behind as new forms of immersive technologies emerge, legislation and regulatory frameworks should have the ability to capture both current and future iterations of immersive technologies. This can be achieved through tech-agnostic legislation or regulation, which does not focus on specific types of technology, but focuses on the experiences and issues enabled by technology; for example, types of online harms.[72] Tech-

[71] B Heller, *Revisiting code as law: Regulation and extended reality,* September 2023, pages 57-58, https://dx.doi.org/10.2139/ssrn.4559458.
[72] E Hine et al., *Safety and privacy in immersive extended reality: An analysis and policy recommendations*, 3 July 2024, page 2, https://doi.org/10.1007/s44206-024-00114-1; See also, Principle 5 in A van der Spuy et al., Digital Futures for Children Centre *Guiding principles for*

agnostic legislation and regulation can provide broader frameworks which can account for rapidly evolving spaces.[73]

At the same time, the uniqueness of immersive technologies may require some level of targeted and specific approaches to address the specific risks and harms posed by immersive technologies,[74] distinct from other technology. Overall, regulatory approaches should strike a balance between being tech-agnostic and tech-specific.[75]

Identifying which actors bear responsibility

There are several types of actors involved in the development and use of immersive technologies. These include, but are not limited to, hardware and device (and operating system) developers, service providers of immersive platforms, users who create events or worlds in immersive environments, and users who use immersive technology products and engage in immersive environments.

Given the range of actors involved, an appropriate level of accountability should apply throughout the ecosystem. However, it may be difficult to determine which actor is responsible for identifying or mitigating specific risks, or who should be held liable for harmful or malicious use. Interoperability may further complicate this issue. As the metaverse continues to evolve, new categories of actors may emerge.[76]

How do immersive technologies fit within eSafety's regulatory remit?

The Act provides eSafety with a range of powers and functions to address online safety issues, including those related to immersive technologies. The Act is designed to be broadly technology neutral and flexible, focusing on specific harm types rather than types of technology. As such, where the relevant legislative thresholds for eSafety's regulatory schemes are met, they are capable of capturing harms that occur on immersive technologies. The Act

---

addressing technology-facilitated child sexual exploitation and abuse, December 2024 https://eprints.lse.ac.uk/126219/.

[73] B Egliston et al., *Who will govern the metaverse? Examining governance initiatives for extended reality (XR) technologies*, 29 January 2024, page 15 https://doi.org/10.1177/14614448231226172

[74] Ibid page 16.

[75] Gray, *Governing Social Virtual Reality*, page 82.

[76] C Wallace et al., *Whitepaper: The Metaverse and Standards*, May 2023, page 12, https://www.standards.org.au/documents/h2-3061-metaverse-report

empowers eSafety to regulate eight sections of the online industry, including those supporting or providing immersive experiences.

eSafety takes a multi-faceted approach to online safety that involves prevention, protection, and proactive and systemic change. The following section outlines how the risks associated with immersive technologies may intersect with eSafety's regulatory schemes.

Complaints schemes

eSafety has schemes that allow Australians to report cyberbullying of children, adult cyber abuse, image-based abuse (sharing, or threatening to share, intimate images without the consent of the person shown) and illegal and restricted content that is class 1 or class 2.[77]

eSafety provides support to people who make complaints under these investigations schemes by offering guidance, assisting in or requiring the removal of certain content, and minimising the risk of further harm.

It is possible that experiences facilitated by immersive technologies may add complexity for the collection of evidence in investigations. Under the Act, eSafety's existing suite of remedies for complaints focuses primarily on the removal of material. This may present limitations on how eSafety can respond to complaints of harm in immersive environments – which allow for synchronous, ephemeral virtual experiences and may have limitations for capturing content in real time or retrospectively – where that harm has not been captured in material form.

Industry codes and standards

The Act provides for industry bodies or associations to develop codes to regulate certain types of seriously harmful online material, and for eSafety to register and enforce the codes. If the Commissioner is not satisfied that a Code meets the statutory requirements of the Act for registration, then eSafety may develop an industry standard for that section of the online industry instead.

---

[77] The definition of Australian adult in the *Online Safety Act* is 'an adult who is ordinarily resident in Australia.'
The definition of Australian child in the *Online Safety Act* is 'a child who is ordinarily resident in Australia.
The image-based abuse scheme requires either the end-user who shares or threatens to share an intimate image or the person shown in the image to be ordinarily a resident in Australia.

Phase 1 Codes and Standards

There are currently six industry codes in operation. They apply to social media services, app distribution services, hosting services, internet carriage services, equipment providers, and search engine services.

In addition to these codes, two industry standards were registered in June 2024 and came into effect on 22 December 2024. They apply to relevant electronic services and designated internet services.

The industry codes and standards relating to Class 1A and 1B material set out mandatory requirements currently in force across eight sections of the online industry. This coverage across the tech stack enables eSafety to take a holistic regulatory approach, which requires all providers to play a role in improving their online safety in relation to online harms such as CSEA. The registered codes and standards are available on the eSafety website.

The various layers that comprise an immersive experience will have different requirements under the codes and standards. For example, providers that manufacture XR headsets may be subject to requirements set out in the Equipment Online Safety Code (Class 1A and Class 1B Material) (**Equipment Code**), such as providing annual compliance reports and enabling users to report potential Code breaches.

If another provider or the same provider operates a service that enables end-users to download apps to use in immersive environments, then that service is likely to be subject to the **App Distribution Services** Online Safety Code (Class 1A and Class 1B Material)**.**

Providers offering XR services and apps may be subject to the Online Safety (**Designated Internet Services**— Class 1A and Class 1B Material) Industry Standard 2024. If they have a messaging or chat functionality, they are likely to be subject to the Online Safety (**Relevant Electronic Services** – Class 1A and 1B Material) Industry Standard 2024.

The Minimum Compliance Measures required under a code or standard vary in accordance with a service's risk tier and categorisation.

More information on identifying the applicable code or standard, and how to comply with the relevant obligations, is available in eSafety's **regulatory guidance**:

- Phase 1 Codes (Class 1A and Class 1B Material) Regulatory Guidance (Updated December 2024)
- Phase 1 Standards (Class 1A and 1B Material) Regulatory Guidance

Phase 2 Codes

Development of the Phase 2 industry codes formally commenced in July 2024 with the publication of eSafety's Phase 2 position paper and the issuing of notices to industry to develop the Phase 2 industry codes.

The material to be considered under the Phase 2 industry codes is 'class 1C' and 'class 2' material, which includes online pornography and other high-impact material, as defined by reference to the National Classification Scheme. This includes themes such as suicide and simulated gambling.

The aims of the Phase 2 industry codes are to prevent children from accessing or being exposed to age-inappropriate material online (including material like online pornography) and to provide all end-users with effective information, tools and options to limit access and exposure to such material.

As with Phase 1 Codes and Standards, it is likely the various layers that comprise an immersive experience will also have different requirements under the Phase 2 industry codes.

The industry associations submitted seven industry codes to eSafety for review on 28 February 2025, and the eighth code (for app distribution services) was submitted to eSafety on 28 March 2025.

At the time of publishing, eSafety has not endorsed the draft codes and is currently undertaking an assessment of whether the draft codes submitted meet the statutory requirements for registration.

Basic Online Safety Expectations

The Act also empowers eSafety to require social media services, relevant electronic services (such as messaging, gaming, and dating services), and designated internet services (other apps and websites) to report on the reasonable steps they are taking to comply with the Australian Government's Basic Online Safety Expectations (the Expectations). This is to make sure these services are transparent, accountable, and safe for people to use.

eSafety has the power to compel information from online providers and publish information about the steps they are taking to meet the Expectations

in transparency reports. Future reports could cover immersive platforms and functionalities.

Social media age of access

The *Online Safety Amendment (Social Media Minimum Age) Act 2024* (Cth) will require age-restricted social media platforms to take reasonable steps to prevent children under 16 from having an account.

The Minister may make legislative rules specifying services that are or aren't covered – the eSafety Commissioner will advise through separate independent assessment.

The age restriction requirements will take effect on a date to be determined by the Minister, no later than **10 December 2025.**

Prevention and education

In addition to its regulatory schemes, eSafety also provides scaffolded, age-appropriate and contextualised programs and resources for children, parents and carers. It also provides professional learning for educators and supports the delivery of best practice online safety education. eSafety collaborates with mental health professionals, child protection services and other frontline workers when developing resources for specific at-risk groups.

By understanding the benefits and risks of immersive technologies, people can better manage their online experiences and contribute to a safer, more positive online environment.

Evolving domestic context

Forthcoming policy developments may have implications for the use, governance and regulation of immersive technologies. Recognising the importance of regulatory coordination, eSafety will continue to work with other Australian Government departments, agencies and regulators where our remits intersect.

Safe and Responsible AI

Immersive technology developers are increasingly incorporating AI capabilities into their products and services. Through its Safe and Responsible AI initiative, the Australian Government is taking an integrated approach to harness the benefits of AI, supporting innovation, while building public trust.

The Government has consulted on proposed mandatory guardrails to shape the use of AI in high-risk settings. These proposed guardrails focus on mitigating serious risks by ensuring AI systems are tested, transparent, and subject to clear accountability. Service providers that develop or deploy AI-enabled immersive technologies which meet the definition of 'high-risk' may be subject to these proposed mandatory guardrails.

The Australian Government has also released the Voluntary AI Safety Standard (VAISS) to support businesses to develop and deploy AI safely.

Online Safety Act Review

In 2024, Ms Delia Rickard PSM conducted an independent review of the Act. The Final Report was provided to the Australian Government on 31 October 2024. This was tabled in parliament on 4 February 2025.

The independent review examined the operation and effectiveness of the Act and considered whether additional protections are needed to combat online harms, including those posed by emerging technologies.

A key recommendation of the review is the introduction of a digital duty of care model. In November 2024, the Australian Government announced its intention to legislate a Digital Duty of Care.

Children's Online Privacy Code

The *Privacy and Other Legislation Amendment Act 2024* (Cth) enables the Office of the Australian Information Commissioner (OAIC) to develop a Children's Online Privacy Code (the Code).

The Code applies to businesses or organisations covered by the Privacy Act (APP entities) if:

- they are a provider of a **social media service, a relevant electronic service or designated internet service**
- the service is likely to be accessed by children, and

- the entity is not providing a health service.

The Code will specify how online services accessed by children must comply with the Australian Privacy Principles and may impose other additional requirements on services.

The Code will be in place by **10 December 2026**.

More information about privacy reforms is available here.

Safety by Design measures

eSafety's Safety by Design initiative encourages technology companies to take proactive steps to embed user safety and invest in risk mitigation at the forefront – and throughout all stages – of product design, development, and deployment.

Services have a critical opportunity to embed safer design practices to minimise the risk of harm from immersive technologies now, before they become ubiquitous like 2D social media services. By proactively elevating user safety now, we can avoid the need to retrofit measures in response to harms once they become prevalent.

Services can take practical steps by following the three Safety by Design principles: service provider responsibility, user empowerment and autonomy, and transparency and accountability. Safety by Design principles, although voluntary, can be used by industry as a way to support compliance with regulatory requirements.

Service provider responsibility

The burden of safety should never fall solely on the user. Immersive technology product and service providers should mitigate harm by identifying and assessing the potential safety risks associated with their product or service and, and by engineering out potential abuse and misuse.

Service providers can take the following key steps in relation to immersive technologies:

- Conduct risk and impact assessments to identify and remediate potential harms that could be enabled, facilitated, or exploited through their immersive device or immersive service.

- Nominate individuals or teams accountable for creating, implementing, operating, and evaluating user safety policies that include immersive products and platforms, and promoting a culture of community safety throughout the organisation.
- Develop, and fairly and consistently enforce, robust community guidelines and terms of service to protect users from online harms and ensure they adequately capture and address immersive platforms.
- Provide a range of proactive safety features to prevent harm, along with reactive features to respond to harm, in all online spaces, including immersive ones. Relying solely on reactive safety features places the burden of managing safety on users.
- Establish processes to detect, surface, flag and – where appropriate – remove harmful content or conduct in immersive environments, with the aim of preventing harms before they occur.
- Implement clear protocols for internal and external triaging and escalation pathways in response to user safety concerns arising from content and experiences on immersive services.
- Implement clear protocols for engaging with law enforcement, support services, and illegal content hotlines. Service providers of immersive platforms should also understand and fulfil their obligations related to mandatory reporting requirements for harms relating to children, in accordance with their jurisdiction's requirements.
- Embed age-appropriate design, supported by robust age assurance measures. Immersive devices and platforms which are designed to be accessible to children should prioritise their rights, safety, and best interests. Providers of immersive services which apply a minimum age of 18 because they are not suitable for children, but are still able to be accessed by children, should implement protections to prevent children from encountering or being exploited to produce harmful material or being exposed to inappropriate interactions in immersive environments. Services should use robust age assurance measures to ascertain which users are children and apply age-appropriate safety and privacy settings. They can also create age-appropriate, designated rooms and experiences tailored for children. For further reading on age assurance, refer to eSafety's age assurance issues paper.

- Conduct ongoing safety evaluations of the immersive product or service and continually improve systems to ensure they preserve user safety.

User empowerment and autonomy

Services should ensure their products and platforms uphold the dignity of users and align with their best interests. This involves understanding that online abuse and harms can be intersectional, and that immersive technologies can be misused to reinforce and compound social inequalities.

To ensure that products and services incorporate features and designs that respect user empowerment and autonomy and support safe online interactions, service providers should implement the following steps in relation to immersive technologies:

- Consult with marginalised and underrepresented individuals and communities at all stages on the development of immersive products and platforms. This should include, but is not limited to, women, children, young people, older people, people living with disability, LGBTQI+ people, people from low-economic backgrounds, people living in regional and remote communities, First Nations people, and people from culturally and linguistically diverse communities.
- Employ technical interventions to educate and empower users to safely use immersive technologies. This could include prompts that advise users when their behaviour in immersive environments may violate community guidelines, or alerts to draw attention to available safety features.
- Implement intuitive, easy to navigate safety features. While responsibility for harm should not be placed on the user and sits with the perpetrator of abuse, safety features can equip users with greater choice and control over their immersive experience, including how to respond after being harmed. Although negative and harmful interactions in immersive environments can be fleeting, their impacts can be long-lasting. Onerous safety and reporting tools place an unnecessary burden on users and can make it difficult to exit harmful situations or report harmful behaviour, especially in immersive experiences. Safety tools should also be supported by appeals mechanisms and appropriate governance frameworks to prevent misuse.

Examples of existing safety tools in immersive platforms which prioritise user empowerment and autonomy include:

- **Intuitive muting, blocking, and reporting functions**: For example, a mute or block function triggered by the user covering their ears[78] may be more intuitive in an immersive environment than requiring the user to navigate a menu. In RecRoom, users can point to another user to mute, block, or report them.[79] Services should ensure their intuitive safety features are accessible for users living with disability.

- **In-built recording mechanisms**: The transient nature of interactions in immersive environments can make it difficult for users who experience harm or abuse to capture the conduct for reporting. Requiring users to provide detailed accounts of their experience can be retraumatising. Users should not be expected to manually initiate a recording during an experience of abuse. Instead, platforms can implement in-built recording mechanisms that are triggered when a user wants to submit a report. For example, Horizon Worlds continually records the last few minutes of a user's interactions on a rolling basis. When a user wants to make a report, the previous two minutes of interaction are saved. Recordings are not viewed unless they are submitted as part of a report and are otherwise deleted.[80] Platforms that provide in-built recording mechanisms should implement robust privacy-preserving measures to reduce the risk of recordings being misused.

- **Personal boundary features**: For example, Horizon Worlds has implemented a personal boundary feature that places a few feet of distance between users. When another user invades that personal space, their hands, feet, or whole avatar automatically disappears.[81] This can be a critical feature for preventing harms such as embodied harassment. Users should be able to customise these boundaries to suit their requirements – for example, by turning them on or off and applying different settings for different people.

[78] Gray, *Governing Social Virtual Reality*, page 64.
[79] Rec Room, *Comfort and Safety*, n.d., https://recroom.com/comfortandsafety
[80] Meta, *Notice of monitoring and recording to improve safety in Horizon Worlds*, n.d., https://www.meta.com/au/legal/quest/monitoring-recording-safety-horizon/
[81] Meta, *Use personal boundary in Meta Horizon Worlds*, n.d., https://www.meta.com/en-gb/help/quest/articles/horizon/safety-and-privacy-in-horizon-worlds/personal-boundary-horizon-worlds/

Transparency and accountability

Services should share information with users, regulators, and researchers about how their immersive devices and platforms operate.

Services can promote transparency and accountability by undertaking the following key actions in relation to immersive technologies:

- Embed user safety considerations – including the lived experiences of marginalised and underrepresented communities – into the training, roles, functions, and workplace practices of all individuals who work with, for, or on behalf of the immersive product or platform.
- Provide clear and accessible information about user safety policies, terms and conditions, community guidelines, and safety processes, including in relation to immersive platforms. Ensure this information is easy to understand and regularly updated.
- Publish regular transparency reports detailing measures taken to improve user safety, as well as reported abuse or misuse on their services, including in relation to immersive environments. These reports should include meaningful analysis of metrics and effectiveness of safety features.
- Innovate and invest in new technologies that enhance user safety. This includes sharing and collaborating on safety tools, best practices, processes, and technologies, including in relation to immersive products and platforms. Services should also seek, consider, and implement user feedback on safety issues.

Looking ahead

While immersive technologies present a myriad of potential benefits, they also have the capacity to amplify existing harms and introduce novel safety risks. We are already seeing these harms manifest; for example, in victim-survivors' experiences of embodied harassment and children's exposure to age-inappropriate experiences on immersive platforms. These harms may proliferate as immersive technologies evolve and become more widely adopted.

To mitigate risks and harness benefits, all stakeholders in the digital ecosystem must play a role in ensuring safety is prioritised, underpins, and guides discussions about immersive technologies.

Consultation with underrepresented and marginalised communities is essential to ensure diverse voices are represented during the design, development, and deployment of immersive technologies. Lived experiences must be embedded in all stages, including policy and product design.

Equally important is building digital literacy though a strengths-based approach to promote safer engagement and exploration within immersive environments. This is especially critical for individuals and communities who may face greater risk in immersive settings.

Understanding the risks, harms, and benefits associated with immersive technologies, ensuring regulation remains responsive to evolving harms, and implementing Safety by Design principles are all essential to making immersive technologies safer for all users.

Further information and resources

- If you or someone in your care is experiencing serious online abuse or harm – whether or not immersive technology is involved – you can make a report to eSafety. You can also speak to a mental health professional through an expert counselling and support service.
- Report incidents of child sexual exploitation and abuse material to the Australian Centre to Counter Child Exploitation. If you believe a child is in immediate danger, call 000 or contact your local police station.
- eSafety's Gift Guide provides information on how to be safety conscious with popular tech gifts, including immersive technologies.