D2.2: Mapping of risks and harms of eXtended reality (XR) technologies

Authors: Cristiana Anca Voinov, Lene A. Hagen, Rigmor C. Baraas, Rosemarie Bernabe, Shereen Cox

Editor:

Project title: The Equitable, Inclusive, and Human-Centered XR Project

Project acronym: XR4Human

Grant Agreement no.: 101070155
Lead contractor for this deliverable: University of Oslo, Norway

Funded by the European Union

Deliverable factsheet:

Project Number:	101070155
Project Acronym:	XR4Human
Project Title:	The Equitable, Inclusive, and Human-Centered XR Project
Title of Deliverable:	Mapping of risks and harms of eXtended reality (XR) technologies: a scoping review
Work Package:	WP 2
Due date according to contract:	October 30, 2024
Editor(s):	
Contributor(s):	Cristiana Anca Voinov, Lene A. Hagen, Rigmor C. Baraas, Rosemarie Bernabe, Shereen Cox
Reviewer(s):	Octavia Madeira
Approved by	Rosemarie Bernabe

ABSTRACT:	This report has two parts. Part 1 used a scoping review method to extensively map risks and harms related to the design, development and application of XR technologies (VR, MR, AR), virtual environments, and metaverses. Through the application of qualitative content analysis, thematic categories emerged which were developed into a unique typology of risks comprising of six main themes and 19 sub-themes: 1) health and wellbeing (cognitive, psycho-emotional, socio-relational and physical), 2) autonomy (privacy, self-sovereignty, surveillance, authenticity), 3) epistemic and scientific (epistemic uncertainty, scientific uncertainty, algorithmic imperfection, manipulation), 4) normative (norm violations, rights violations), 5) structural (socio-political, economic/financial, legal), and 6) technical safety and security (security, dual use). Further considerations identified were risks pertaining to special populations (vulnerable groups), unknown risks and long-term use effects. A subsidiary aim was to map risk mitigations. Other key findings, such as geographical distribution of sources, and use cases, among others, are featured, lending to a robust discussion on the state of XR risks, potential motivators for development of XR harm mitigation-strategies, and key knowledge gaps. Finally, the review supports the XR4Human project's proposed code of conduct for developers, as well as a XR ratings system for users.

Funded by the European Union

	Part 2 of this report presents insights on risks and harms of virtual worlds according to European Citizens. This section is based on a literature review of the final report on 2023 European Citizens' Panel on Virtual Worlds. European citizens' views on emerging issues within virtual worlds are comparable with the findings from Part 1 and include health and wellbeing, autonomy, normative standards, structural impacts, and technical security. Key concerns include data privacy, mental health impacts, user consent, and protection of digital identities, aligning with many risks identified in academic literature. Additionally, citizens emphasized the need for sustainable practices in virtual world development and highlighted the importance of ethical AI use, particularly concerning surveillance and data transparency. These public insights contribute valuable context to the XR4Human project, supporting the development of policies that address both expert and citizen perspectives on risks in extended reality technologies.
Keyword List:	Immersive technologies, virtual reality, augmented reality, mixed reality, harm, risks, ethics

Consortium:

	ROLE	NAME	Short Name	Country
1.	Coordinator	UNIVERSITETET I SOROST-NORGE	USN	NO
2.	Partner	XR4EUROPE	XR4Europe	BE
3.	Partner	UNIVERSITETET I OSLO	UiO	NO
4.	Partner	ETHNICON METSOVION POLYTECHNION	NTUA	EL
5.	Partner	NORGES TEKNISK-NATURVITENSKAPELIGE UNIVERSITET NTNU	NTNU	NO
6.	Partner	UNIVERSITEIT LEIDEN	ULEI	NL
7.	Partner	KARLSRUHER INSTITUT FUER TECHNOLOGIE	KIT	DE
8.	Partner	INSTITUTT FOR ENERGITEKNIKK	IFE	NO
9.	Partner	FRAUNHOFER GESELLSCHAFT ZUR FORDERUNG DER ANGEWANDTEN FORSCHUNG EV	Fraunhofe	DE
10.	Partner	OPEN AR CLOUD EUROPE GEMEINNUTZIGE UG	OARC EU	DE
11.	Partner	STICHTING FONTYS	Fontys	NL
12.	AP	UNIVERSITY OF LEEDS	UNIVLEEDS	UK

Revision history:

VERSION	DATE	Revised by	Reason
0.1	7.10.2024		
1.0			

Table of contents

General Summary

PART 1: Mapping of Risks and Harms in Extended Reality (XR) Technologies: A Scoping Review

Part 1 focuses on identifying potential risks and harms associated with XR technologies, encompassing virtual reality (VR), augmented reality (AR), and mixed reality (MR). Through an extensive scoping review, the report establishes a typology of risks spanning six main themes:

Health and Wellbeing Risks: These include cognitive, psycho-emotional, socio-relational, and physical impacts, with concerns about addiction, anxiety, social isolation, and potential physical injuries.

Autonomy Risks: Issues such as privacy, self-sovereignty, surveillance, and authenticity were identified, highlighting concerns over biometric data privacy, virtual identity control, and user manipulation.

Epistemic and Scientific Risks: Risks include algorithmic imperfection, manipulation, and uncertainties in epistemic and scientific areas, emphasizing challenges in transparency and potential biases in data interpretation.

Normative Risks: Risks of norm and rights violations are explored, including moral and ethical boundaries potentially compromised by XR technologies.

Structural Risks: Socio-political, economic, and legal risks are mapped, focusing on the effects on labor markets, regulatory gaps, and potential inequities.

Technical Safety and Security Risks: These include hacking, data leaks, and dual-use risks, stressing the importance of secure digital infrastructure and robust cybersecurity practices.

The findings in Part 1 underscore the importance of comprehensive risk mapping for responsible XR development, feeding into a proposed XR developer code of conduct and user rating system.

PART 2: Risks and Harms According to European Citizens - A Review of the European Citizens' Panel on Virtual Worlds

Part 2 examines to comparability of the risks and harms in Part 1 with the perceptions of risks and harms of European citizens. This section utilized a literature review methodology to extract perceptions of risks and harms from the final report of the 2023 European Citizens' Panel on Virtual Worlds. Convened by the European Commission, the panel gathered input on the future development of virtual worlds. Citizen concerns were generally aligned with the scoping review in Part 1, covering major risk areas such as:

Health and Wellbeing Risks: Concerns were raised about mental health impacts, addiction, and the need for further research into physical and psychological effects.

Autonomy Risks: Citizens prioritized data privacy, digital identity protection, and advocated for regulations on user consent and anonymity.

Normative Risks: Emphasis was placed on EU values in virtual world development, with calls for inclusivity, accessibility, and ethical behavior standards.

Structural Risks: Potential economic impacts, particularly on employment and labor rights in virtual workspaces, were highlighted, along with recommendations for harmonized training and worker protections.

Technical Safety and Security Risks: Citizens stressed the need for secure digital environments, with suggestions for EU certification and collaboration across sectors to ensure safety.

There were notable divergences between the citizens' concerns and the scoping review findings in two areas:

Sustainability: Citizens highlighted the environmental impact of virtual worlds, a topic minimally addressed in Part 1.

AI-Driven Surveillance: While the scoping review identified risks in algorithmic and scientific uncertainties, citizens showed less concern over potential AI biases in security applications.

These insights from the citizens' panel add a public perspective to the academic scoping review, supporting the development of policies and guidelines to mitigate XR risks in alignment with public values.

PART 1: Mapping of risks and harms of eXtended reality (XR) technologies: scoping review

1. INTRODUCTION

1.1 Background and rationale

XR4Human (1) is a European Union funded project that aims to co-create guidance documents for developers and users of extended reality technologies in Europe. The project comprises of several work packages (WPs), one of which – WP2 – is tasked primarily to map, examine, and evaluate the ethical and philosophical issues, and potential risks and harms relevant to the creation, development, and use of eXtended reality technologies (XR), also called immersive technologies. These knowledge mappings will feed into a conceptual and normative framework on how these issues should be addressed. The mapping will also contribute towards XR4Human's goal of creating a code of conduct for XR developers which upholds human-centered values and practices, as well as ratings system for XR users as a guiding tool. It is therefore critical that risks and harms which arise from these technologies are comprehensively considered and inputted within both code of conduct and ratings system.

1.1.1 eXtended reality

XR is an all-encompassing term commonly used to represent:

> technology-mediated experiences enabled via a wide spectrum of hardware and software, including sensory interfaces, applications, and infrastructures. XR is often referred to as immersive video content, enhanced media experiences, as well as interactive and multi-dimensional human experiences. (2)

In essence, XR is a fusion concept capturing "all the realities": augmented reality (AR), virtual reality (VR) and mixed reality (MR). (2) The meaning and usage of XR and its modes/forms have evolved over time (see (3-5)). These terms remain in flux, expanding as new infrastructures, affordances and trends emerge. AR places digital information on top of a user's physical environment; MR is an active blending of digital and physical, wherein changes to, or interactions with, the digital object would occur as if it existed in physical space; VR is a completely digitized interactive environment which can give the user the experience of complete immersion. (2) Immersion is a complex concept which may refer to the physics or technical abilities offered by a device (measured quantitatively), but may also refer to a user's subjective experience of being "present" in a virtual environment (qualitative). (4)

These interactive environments, commonly characterized in literature as virtual environments (VEs) or virtual worlds, are a primary affordance of XR. These can be understood as "a synchronous, persistent network of people, represented as avatars, facilitated by networked computers". (6, p.2) The Metaverse can be considered an expansive, interconnected set of virtual environments which persist over time and are interoperable. (7) Persistence, here, means virtual objects continue to exist even when the user is not in the digital world. (7) Interoperability is the capacity for intra-system and world interaction, allowing for cross-platform and device communication. (7)

1.1.2 Risk and harm

As XR technologies become more widely adopted across various disciplines and applications, more attention has been placed on the associated risks and harms. There have been a number of governmental and non-governmental agencies and individual interest in examining what constitutes a risk or harm in various technologies. The Institute of Electrical and Electronics Engineers (IEEE) has been one of the key standards agencies to play a significant role in examining risks of technologies. As such, the project incorporated industry specific definitions while considering other nuanced understanding of risks and harms in the academic literature. According to the IEEE, a "risk" is "an unwanted event that may or may not occur". (8) Another industry-specific definition considered is that proposed by the International Organisation for standardization (ISO). The ISO notes that a risk is a "combination of the probability of occurrence of harm and the severity of that harm". (9) Given that risks are directly connected to an unwanted event or probability of harm then it would be prudent to examine what constitutes harm.

Referring back to the IEEE, harm is defined as "any unwanted occurrence which ends in some sort of "value damage or loss" to a person. (10) This very broad interpretation presents the opportunity to capture a wide range of unwanted events in the XR literature by various contributors. However, it may also present a significant challenge in contextualising what harms ought to be considered by developers. Harm has been a concept that has been debated extensively in philosophical, sociological, and jurisprudential academic literature. (11-13) Acknowledging the nuanced understanding of harm and how the probability of harm (risks) may be represented in literature, it was important to examine what may already be broadly categorised as risks and harms of XR to facilitate a foundational awareness of various risk/harm descriptors and how they are used in relation to XR.

In 2020, the Cyber XR Coalition published what is described as a framework for XR risk assessment. (14) It was emphasised that the risk categories in the report was a "work in progress", which could be interpreted as an acknowledgment of the challenge of attempting to map risks and harms. The reported risk categories in this framework were human, financial, legal, information, and societal. (14) They note harms such as physical, psychological, misuse of data, privacy violations, manipulated social discourse etc. societal. (14)

Others have similarly sought to identify risks associated with XR, often focusing on a specific type of XR, application, or user. For example, XRSI released a publication focused on the privacy risks of XR and proposed mitigation measures. (15) Likewise, the IEEE embarked on what they described as the "Global Initiative on Ethics of Extended Reality" which yielded several publications on challenges with various applications (use cases) of XR. (16) The European Union, as part of its Web 4.0 and virtual worlds strategy, has also begun to reflect on "opportunities, pitfalls and potential solutions" of XR technologies, with a focus on virtual worlds/metaverses. (17) Noting that the global market for XR, particularly, virtual worlds should reach over 800 billion by 2030, it is imperative that developers are not only aware of the risks and harms but to ensure that the development of the technology considers fundamental rights and to avoid risk of harming to users i.e., worse off in terms of well-being. (12, 17) Being an EU funded project, the XR4Human project aims to contribute to any knowledge gaps in the existing literature by applying the scoping review method for obtaining a comprehensive lay of the land of XR risks and harms. A scoping review is a type of systematic knowledge mapping in which findings are to be summarised and presented as described in the literature. (2, 3). To achieve this, this review is guided by the following aim and objectives.

1.1.3 Aim and objectives

To conduct a comprehensive mapping of risks and harms associated with the design, development, and use of XR technologies, as reported in the literature. This mapping spans across disciplines and methodologies.

The specific objectives are:

1. To identify and describe the various forms of risks and harms associated with the design, development and use of XR technologies.
2. To identify and describe the mitigation strategies/redress proposed in literature for identified risks and harms.

1.1.4 Previous systematic reviews and research gap(s)

A preliminary search in January 2022 identified only a few systematic/scoping reviews reporting harms and risks related to the use of XR technologies. (18-21) These, however, were narrow in disciplinary scope i.e., mainly relating to a specific domain or specific scenarios (health or construction), or one type of XR technology or experience (ex: VR only). (18-21) We posit therefore that a gap exists in the XR literature for reviews that comprehensively map risks and harms related to XR technologies and experiences. This expanded mapping of risks and harms is building on the findings of an initial mapping of ethics guidelines for XR technologies conducted early 2023, where various ethical issues with the use of the technology were identified. (22)

2 METHOD

A scoping review aligned with Arksey and O'Malley (23) seminal review framework and the Johanna Briggs Institute (JBI) Manual for Evidence Synthesis (24) was selected. This method additionally adheres to reporting standard: the Preferred Reporting Items for Systematic Reviews and Meta-Analysis extension for scoping review (PRISMA-ScR). (25)

The scoping method is well suited to the broad nature of our research question. Sources covering risks and harms of XR are not homogenous, thus a scoping approach allows for the inclusion of literature from various domains, traditions, study designs and methodological approaches. (23, 26) The method is also appropriate to finding potential knowledge gaps. (23, 26). Scoping reviews are primarily descriptive endeavours, meaning that "findings" (i.e. risks and harms themselves) are reported as written within the included sources. (24) Nonetheless, the PRISMA-ScR review guidelines note a critical appraisal of literature as an optional part of the scoping process. (25) However, due to limitations in the reviewers' expertise, this review did not involve a critical appraisal of source quality, specifically regarding the empirical rigor or methodological validity of the included sources. Anticipating that many of the selected publications would be highly technical, and would necessitate a primarily descriptive approach, reviewers considered that attempting a deeper interpretation could introduce biases or inaccuracies, making a straightforward presentation of findings the most practical choice. (24)

The following sections will provide an overview of the review protocol, search strategy, source inclusion/exclusion criteria, the process of selecting sources of evidence, as well as data mapping and reporting. (26)

2.1 Team and expertise

The review team comprises of CV (ethicist/methodologist), LH (vision scientist), SC (ethicist/pharmacist), RB (vision scientist), RdcB (research ethicist). Additionally, the WP members acted as experts providing guidance throughout the research. The members of the WP include philosophers, technology experts and practitioners, psychologists, and other social and natural scientists.

2.2 Review protocol

Reviewers CV, LH, SC prepared a protocol based on JBI standards for scoping review protocols. (27) This protocol served as the roadmap for how to approach this review. However, while the protocol itself was not registered, the data charts which present extracted findings will be uploaded to the XR4Human Community in Zenodo.org,

2.3 Information sources and search strategy

The search strategy comprised of an inductive search using databases followed by citation chaining, and manually inputted sources. This design allows for an iterative process that brings in more sources as we became familiar with appropriate literature, widening the breadth of our mapping. (26)

In January 2023, the search string was constructed after several meetings with the WP team. The search string (Table 1.0) was submitted to the University of Oslo's librarian who executed a search of the databases, IEEE Xplore and Scopus. These databases provide a wide range of technical literature from multidisciplinary sources. This resulted in 2281 records which were then uploaded to the management software Rayyan. (28) Duplicates were removed, leaving 1939 sources considered for eligibility. See Table 2.0. for details.

Citation chaining of included literature from the database search also yielded some information sources. Reviewers CV, SC and LH searched the reference lists of the included database documents (i.e. database documents read in-full and deemed appropriate for data extraction). A total of 111 sources were considered for screening in Rayyan. Finally, some publications were identified for screening through the input of XR4Human consortium members.

Databases	Search string	Number of retrieved documents
Scopus	(TITLE-ABS-KEY ((safe* OR harm* OR danger* OR abus* OR hazard* OR secur* OR "adverse effect" OR "adverse vents" OR risk OR risking OR risks OR benefit* OR privacy OR wellbeing OR "well-being")) AND TITLE-ABS-KEY (ethic* OR bioethic* OR philosoph* OR integrit* OR moral* OR "human centred" OR "human centered") AND TITLE-ABS-KEY ((interactive W/2 (media* OR experience*)) OR metaverse OR xr OR xr4* OR (immersive W/2 (technolog* OR environment*)) OR (virtual W/2 (realit* OR environment*)) OR ((virtualis* OR virtualiz*) W/2 (technolog* OR network*)) OR (augmented W/2 realit*) OR (mixed W/2 realit*) OR (extended W/2 realit*)))	1768
IEEE Xplore	((((interactive NEAR/2 (media OR experience OR experiences)) OR metaverse OR xr OR xr4* OR (immersive NEAR/2 (technology OR technologies)) OR (immersive NEAR/2 (environment OR environments)) OR (virtual NEAR/2 (reality OR realities OR environment*)) OR (virtuali* NEAR/2 (technology OR technologies OR network*)) OR (augmented NEAR/2 (reality OR realities)) OR (mixed NEAR/2 (reality OR realities)) OR (extended NEAR/2 (reality OR realities))) AND (safe OR safety OR harm OR harming OR harms OR danger* OR abuse OR hazard* OR security OR secure OR "adverse effect" OR "adverse events" OR risk OR risking OR risks OR benefit OR benefits OR privacy OR wellbeing OR "well-being") AND (ethic OR ethical OR ethics OR ethically OR bioethic OR bioethical OR philosophi* OR integrity OR integrities OR moral OR morals OR morality OR moralities OR "human centred" OR "human centered")))	513

Table 1.0 Search string used for search in the databases Scopus and IEEE Xplore, with number of retrieved documents for each database.

Database	Number of retrieved records
Scopus	1768
IEEE Xplore	513
Number of records before deduplication:	2281
Number of records after deduplication:	1939

Table 2.0 Overview of number of retrieved records from database search and after deduplication.

2.4 Inclusion and exclusion criteria

See Table 3 for an overview of the type of sources considered and the eligibility criteria used. All included sources referred in either title or abstract to eXtended reality (XR) technologies (VR, AR, MR, Metaverse/virtual worlds/environment) or synonyms (i.e. "immersive technologies", etc.). To reflect how XR technologies have evolved and been redefined over several decades of literature, the approach was to identify and chart how sources employ XR terminology (AR, MR, VR), along with the applications of the technology (use cases) and the accompanying output devices/platforms (see Table 4: Data charting categories).

Having established earlier that sources would have multiple and sometimes divergent understandings of what constitutes a risk and harm, or sometimes use synonymous terms, a broad inclusion strategy for these terms was employed. Selected publication must have referred to harm, risk, or their various synonyms or words associated with risk and harm (i.e. well-being, security, etc.) in the title or abstract.

Sources that did not identify any risk/harm coming from the use and/or application of XR were excluded. Additionally, sources that only mentioned XR very generally (as part of "digital technologies", for example) but did not engage more specifically with XR-related risks/harms were excluded. Further inclusion criteria required full texts to be published in the English Language. The database search ranged from 2000 to January 2023, although some 2023 and 2024 publications were manually inputs. Practical reasons motivated these preferences as the overlapping language for all reviewers was English, and while discussions about XR risks and harms pre-date year 2000, we considered that technological advancement is such that those publications would likely be "out-of-date".

Strictly peer reviewed publications were not required, as doing so would have limited the range of sources. However, due to time constraints, sources such as blogs, vlogs, conference summaries and podcasts were excluded. There were no context or population constraints, as the aim was to identify a broad range of risks and harms with no specific population or context in mind. Additionally, there were no restrictions on research methodologies or discipline. The final inclusion/exclusion criteria were developed through an iterative process, with significant refinements made following the pilot phase.

Screening	
Sources considered	**Types:** Academic /scientific literature, book chapter, guidelines, recommendations, policy documents, regulations, framework, official standards, can be either journal articles or grey literature (report, working paper, policy paper, etc.) from government, non-governmental organisations, or industry.
	Authors: Academic researchers, institutions, government and affiliated agencies, non-governmental organisations, professional societies, non-profit organisations.
	Must be full text.
	Language: English only.
Sources excluded	Conference summaries, videos, images, podcasts, blog articles, news articles, syllabi, books.
Eligibility	
Sources included	Explicitly referred in title and/or abstract to "immersive technologies" or extended reality (XR) or virtual reality (VR) or mixed reality (MR) or augmented reality (AR) or virtual environments (VEs) or virtual worlds or Metaverse. Publications that mentioned specific XR-related artifacts/devices/platforms were also included.
	Referred in title and/or abstract to harm or risk. "Risk" and "harm" encompasses a wide umbrella of meaning, including synonyms (i.e. danger, abuse, etc.), hence words highly associated with risk and harm such as "well-being" or "safety", etc were considered.
	Date range: 2000-2023 (database literature, citation chaining), 2000-2024 (manual input)
Sources excluded	Did not mention immersive technologies, XR or its constituents (VR, AR, MR), virtual environments or virtual worlds, or the Metaverse.
	Did not identify any form of risk or harm or synonym arising from the design/use/implementation of XR in the title or abstract.
	Mentioned XR very generally or broadly but did not engage with the technology specifically (for example, as part of "digital technologies").
	Languages outside of English.

Table 3.0 Overview of the type of sources considered and the Eligibility criteria used.

2.4 Source selection process

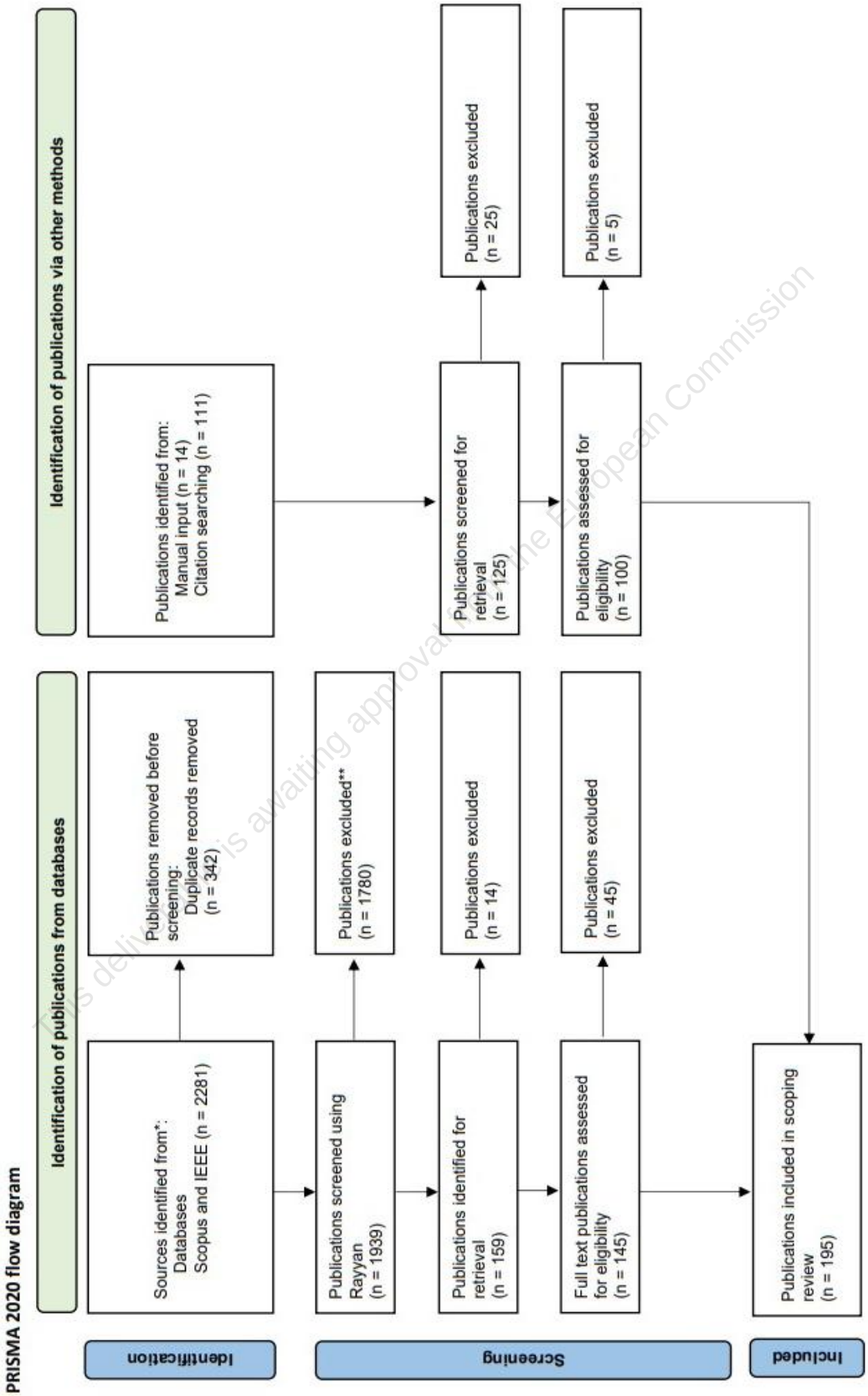
Reviewers (SC, CV, LH, and RB) conducted an initial screening of each database source according to the aforementioned inclusion criteria using the software Rayyan. (28) This review was a "first pass", focusing on title and abstract. Disagreements were settled during our weekly meetings. Reviewer RdcB had the task of independently reviewing conflicts. After which, we were left with 159 database sources, which were then uploaded to NVivo 14 (29) for data extraction.

To ensure that sources aligned with answering the research question, and to further refine both inclusion criteria and the draft charting table (section 2.5), twenty-five pilot articles were selected to read in full by reviewers CV, SC, and LH. (25) This led to clarifications of eligibility criteria, refinement of what to categorise and how to define these categorisations in the data chart, as well as the addition of new categories for charting. After the pilot stage, it was agreed that there would be one reviewer per source, instead of the recommended two, complemented by weekly discussions about areas of uncertainty/ambiguity. (24) This was a practical decision, as the expected number of included sources after citation chaining and manual input would be high. This approach both reinforced the appropriateness

of including the source and facilitated continued refinement of the data items and charting process. Of the 159 database publications read in full, 145 were assessed as suitable for final inclusion.

Citation chaining commenced on these included articles mid-2024, while manual input was ongoing. The same criteria used to include/exclude database sources were used to inform inclusion of chained sources. This process yielded 111 eligible texts. Combined with manual input of 14 publications, the total number of publications screened from non-database source was 125, of which 25 were excluded for not meeting eligibility criteria. The final number of included publications from database, citation chaining and manual input was 195 (see Figure 1.0. Prisma flow chart).

Figure 1.0 The PRISMA flow diagram (30) illustrating the selection process of publications, with the number of publications identified, screened, assessed for eligibility and included in the review. The search process comprised of database search, citation searching, and manual input. The final number of publications included in the scoping review was 195.



Funded by the European Union

2.5 Data charting, coding and theme identification

2.5.1 Data charting

As per Aromataris et al. (24), the approach to data charting was primarily descriptive and to a lesser extent analytical. Source data was classified into several categories and presented in a table or chart format (the "charting" method). The purpose of data charting is to organise information. This aids in the process of synthetising findings and identifying commonalities or differences between sources. Charts were discussed between reviewers at weekly meetings; individual charts were also accessible to each team member. These meetings enabled a deeper understanding of how categories were defined as well as the identification of additional items to be charted.

Categories were title, author, year published, type of publication, country, method (s) used, units of analysis, output device or platform mentioned, use cases, risks or harms discussed, and recommended strategies for mitigation of risks or redress for harms. Unit of analysis refers to the specific experience described by the authors (ex: VR, AR, MR, Metaverse), use cases refers to the specific areas of application and device refers to the hardware components or systems used to enable the experiences being analysed as described by the source This helps narrow down the type of experience the risks and harms are referring to by connecting the device or application. (See Table 4)

Category	Description	
Title of publication	Extracted as written in the publication	
Author	The names of the authors in the order as written in publications	
Year of publication	As reported	
Type of publication	Journal article, book chapter, conference paper, report, pre-print	
Country	The country or countries where the authors' research institution or organization is/are located, as given in the publication	
Methods	Empirical	Experimental, quasi experimental, survey, observational, case study, mixed method
	Non-empirical	Reviews, theoretical, philosophical, commentaries/opinion pieces, book reviews, conceptual, policy, historical
Unit of analysis (Type of XR experience mentioned)	XR, AR, MR, VR, virtual worlds, virtual environments (VEs), Metaverse(s)	
Output device or platform mentioned	"Wearables" such as head-mounted displays (HMD), glasses or goggles, haptic gloves, vests and suits. Virtual environment platforms (VEs; can be desktop), mobile devices and platforms, desktop, 360 video, CAVE. If the source is about a *particular* VE platform or device, this will be specified (ex: sources analysing VE platform, Pokemon GO, etc.).	
Use cases	General application, industry/manufacturing, commerce, healthcare, education/training, entertainment, journalism, governance/law, urban planning, socializing, research	
Meets inclusion criteria (Include)	Yes	
	Reasons	
Risks noted	Risk is to be coded only when there is mention of the likelihood of harm to individual or society	
Harms noted	Harm is to be coded if there are explicit examples of harm that occurred during exposure to/with use of XR, and if there was an explicitly stated harm	
Mitigations measures/recommendations	Measures proposed by authors on how to address various risks or harms noted in the publications	

Table 4.0: Data charting categories as presented in section 2.5.1.

2.5.2 Coding, analysis, and theme identification

In addition to data extraction and presentation through the charting method, qualitative descriptive content analysis was used as an analytic tool to synthesize findings. (31) Due to the large volume and wide range of risks and harms mapped, it was critical to present the data specific to risks and harms in a way that would be understandable for the reader. Descriptive content analysis is a method which uses codes to group and classify text, allowing themes to emerge inductively. (23)

Reviewers CV, SC and LH used software NVivo 14 to each assign codes to their identified risks and harms, placing codes with conceptual or practical similarities together. (26, 27) During pilot testing it was recognized that the way in which "risks" and "harms" were framed (i.e. whether from a quantitative, qualitative, or philosophical point of view) vary widely between source materials, and many sources used synonyms or associated words instead of "risks" and "harms" explicitly. Given all these things, some interpretation was necessary. (26) For conceptual clarity, expansive definitions for risks and harm were established a priori (as described in Section 1.3). If sources indicated that an unwanted event may occur, the described event was classified as a risk. If sources indicated an actual event had taken place, then the classification was as a harm. In other words, harms were coded as events which have taken place. Probable, implied or speculative occurrences were coded as a risk. In this way, the reviewers acknowledge that there is conceptual slippage between risks and harms (i.e. sources might refer to a "risk of harm") but maintain a level of difference between what has occurred and what may, which contributes to future analyses on XR vulnerabilities.

Of note, however, the codes remained as close in meaning to the source as possible (26, Ch.11.2.8) (ex: sources describing risk of "cybersickness" were assigned the code "cybersickness" and grouped together). This process led to the emergence of several themes (ex: "physical risks", under which cybersickness and other risks pertaining to physical well-being were classified). The reviewers then compared themes and reflected on the appropriateness of each code to each developed theme and the establishment of new themes. Eventually, this led to the collective development of a novel typology of risks comprising six main risk themes (health and wellbeing; autonomy; epistemic and scientific; normative; structural; technical safety and security) and 19 sub-themes (described further in Section 3).

3. FINDINGS

The following sections will present the data, which will feed into subsequent discussion. This will include an overview of the characteristics of the source material (as charted), as well as a report addressing the primary research question: "what are the risks and harms of XR technologies?".

3.1 Source characteristics

3.1.1 Characteristics of publications

The findings of this review were extracted from the included 195 publications with date range from 2000-2024. Most sources were published between 2016 and 2022. Since the database search covered publications from 2000 through January 2023, and only manually selected publications were added for 2023 and 2024, Figure 2 reflects only the period from 2000 to 2022. The manually inputted sources for 2023 and 2024 are not included, as they would not represent a comprehensive output for those years. The types of sources included were journal articles, book chapters, conference papers, pre-prints, and reports.

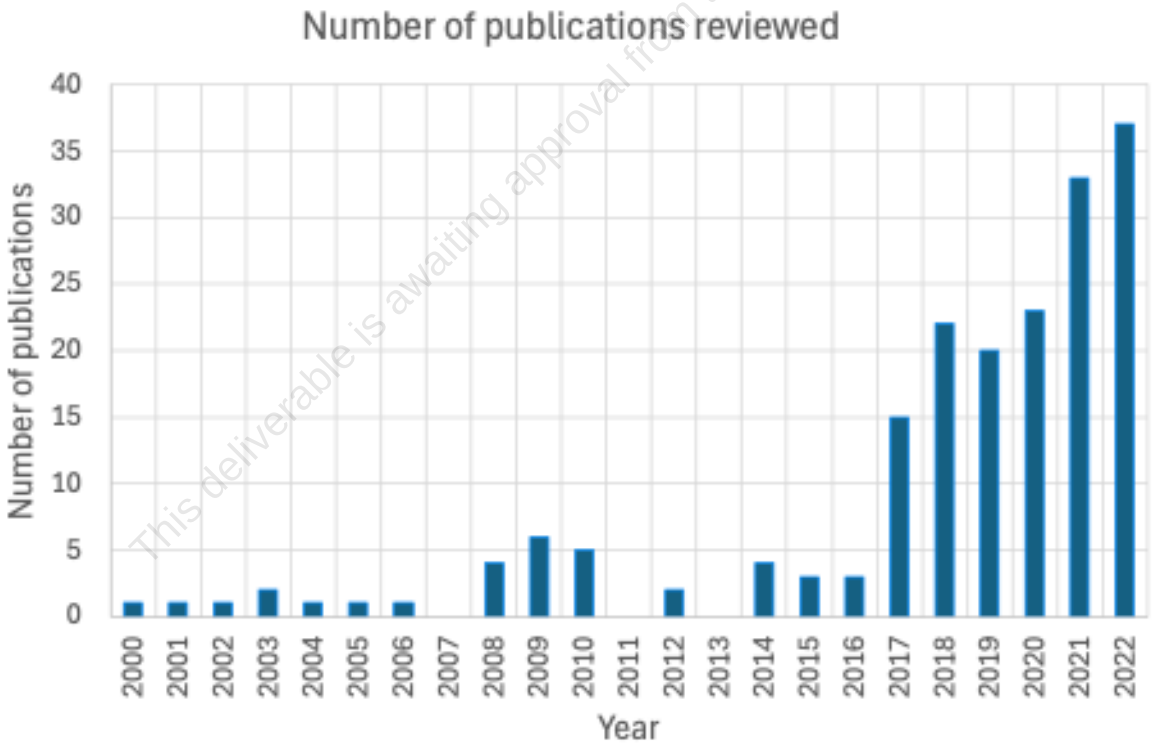


Figure 2.0 Number of publications from year 2000 to 2022 included in the review. Since database search was performed in January 2023, the number of publications from year 2023 (n=12) and 2024 (n=5) were not included in the bar chart.

3.1.2. Geographic distribution of authors

The sources represented a broad geographic range of authors. However, the number of countries listed surpass the number of sources due to multi-authored publications from various affiliated countries. Most authors originated from North America, largely the United States (n=90). European sources comprised of 92 publication sources, with the majority from UK (n=24) and Germany (n=20) respectively. There were 19 sources from Asia and 16 from Oceania (Australia and New Zealand). The minority sources were from the Middle East (n=4) and South and Central America (n=1). We did not identify any sources from African or Caribbean countries. Some of these sources may have been filtered out due to English-language constraints or excluded source types.

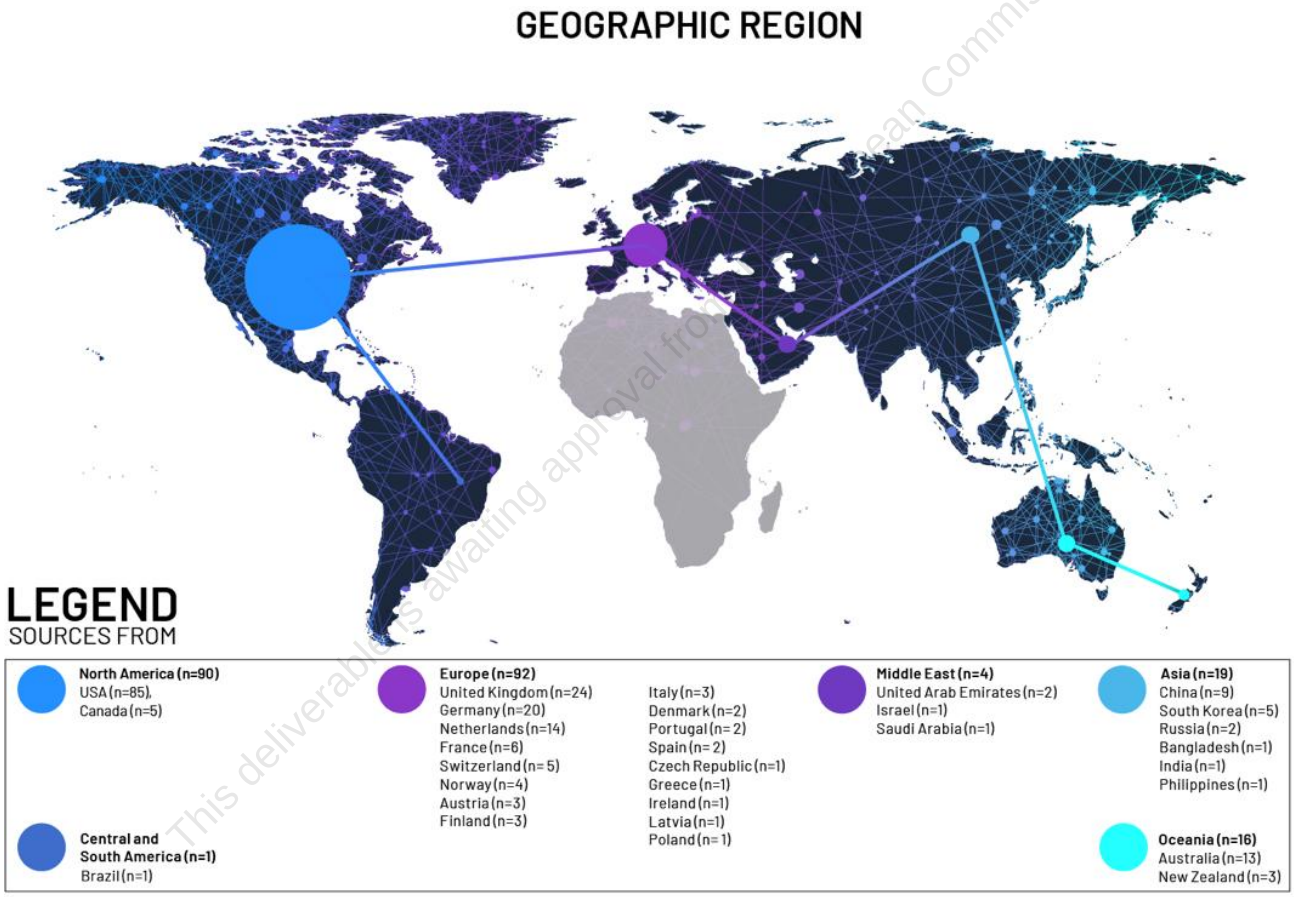


Figure 3.0 Geographic distribution of source authors. Note that the number of countries listed (n=222) exceeds the number of publications reviewed (n=195) due to multi-authored publications with affiliations from multiple countries.

3.1.3 Distribution of use cases

The sources reflected a variety of use cases, i.e. the domain which the source was applicable to (Figure 4.0). The majority written applied to the general user (charted as general use/application (n=94), the other publications focused span healthcare (n=27), entertainment (n=18), socializing (n=16), education (n=11), research (n=10), governance (n=9), and journalism (n=5). The use cases least mentioned were industry, commerce, and urban planning.

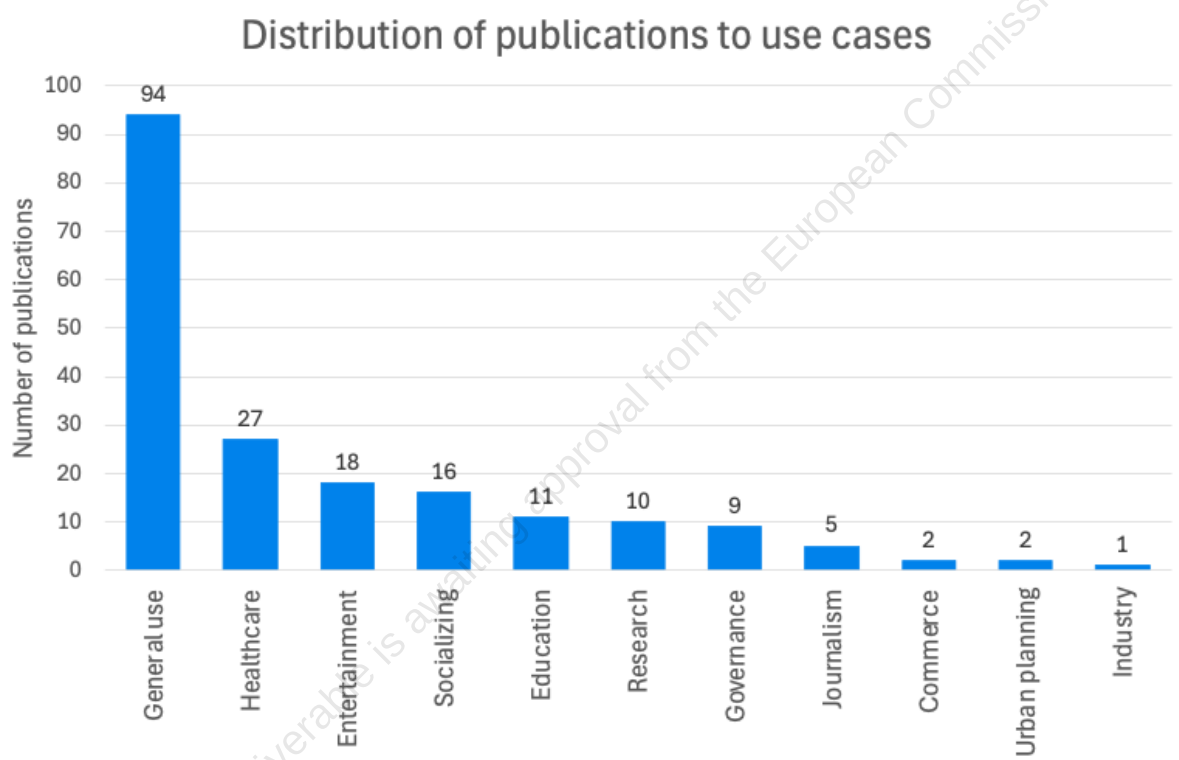


Figure 4.0 Distribution of all included publications (n=195) to use cases; the specific areas of application as presented in the literature.

Use cases	Number of publications	References
General use	94	(18, 19, 32-123)
Healthcare	27	(124-150)
Entertainment (gaming)	18	(151-168)
Socializing	16	(169-184)
Education/training	11	(185-195)
Research	10	(196-205)
Governance, law/legal	9	(206-214)
Journalism	5	(215-219)
Commerce	2	(220, 221)
Urban planning	2	(222, 223)
Industry e.g., manufacturing, logistics	1	(224)

Table 5.0 Use cases, number of publications, and corresponding references.

3.1.4 Distribution of units of analysis

In terms of charted unit of analyses (Figure 5.0), 84 publications pertained to virtual reality only (VR), 46 to virtual worlds, metaverses or virtual environments, and 29 to AR. Additionally, there were 18 non-specific publications on XR or immersive technologies. MR saw a relative underrepresentation at 6 publications.

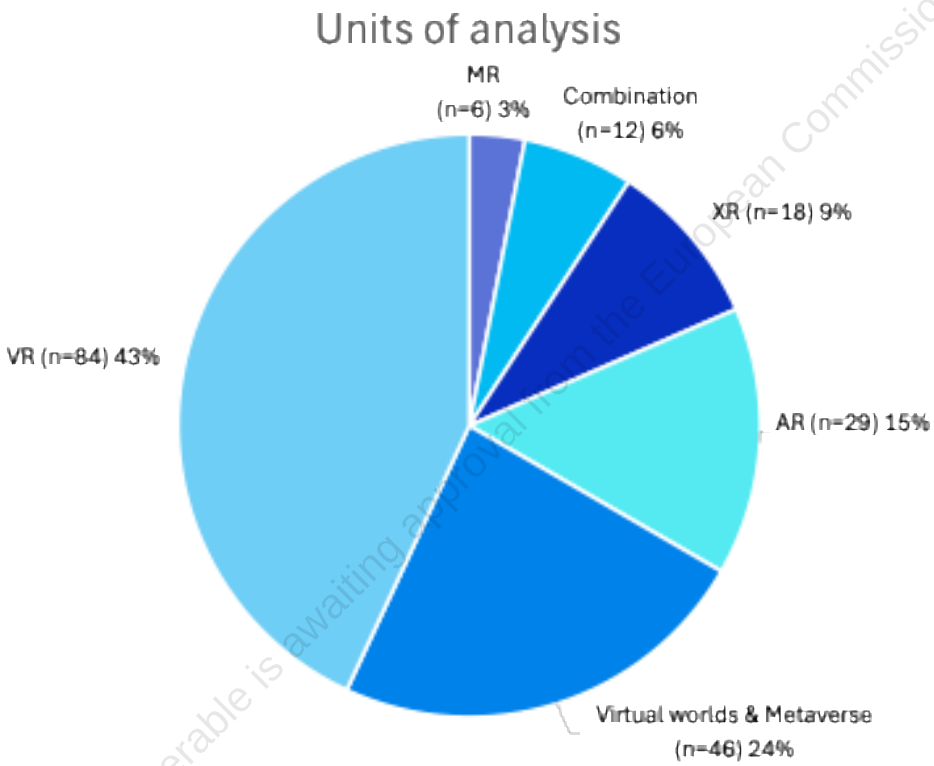


Figure 5.0 Distribution of units of analysis with number of publications in parentheses. Publications with more than one unit of analysis (eg. AR and MR) are categorized as "Combination".

We observed conceptual differences in how terms such as "virtual reality", "virtual environments/worlds" and the "Metaverse" were utilized between 2000-2024. The types of output devices and platforms discussed changed over time as well. Publications from earlier years (2000-2012) discussed virtual environments or virtual worlds, with particular attention to desktop platforms such as Second Life, Habbo, LambdaMOO, among others, as well as to avatars. (99, 110, 113, 167, 172, 173, 206, 212, 214) Not only are these platforms discussed as virtual environments or world, or sometimes "metaverses", some authors referred to them as VR technologies. (18, 99, 113, 129, 167, 205) We also observed an uptick in discussions regarding virtual environments where the Metaverse becomes a term used more frequently in recent years (see discussion section for more). (54, 97, 102, 125, 191, 208, 209) After 2012, publications appeared to shift focus towards various output devices such as wearables, particularly head mounted displays (HMDs). (96, 178, 179)

There were inconsistencies regarding the meaning of or how virtual reality was understood over time. We note a shift where the authors start to consider "immersion" (quantitatively understood) as a critical component of what constitutes VR. Therefore, while many sources discuss HMDs as primary facilitators of virtual reality, bracketing desktop platforms as non-VR, there are some recent sources that still use desktop platforms as an example of VR. (107, 145, 181) These sources use rather broad definitions, where VR is as "close to perfect simulation of the real-world including representations of oneself and other persons" where the most convincing of such simulations are "fully immersive". (145, p.574) Thus, by this author and the like, desktop simulations would be a less convincing type of virtual reality, but nevertheless VR. Critically, many publications did not define their unit of analysis. (96, 149, 180, 192, 204)

Additionally, there is inconsistency in meanings of Metaverse, virtual worlds, and virtual environments, some sources referring to these interchangeably. (150) "Virtual environments are usually defined through their immersive aspects, e.g., a virtual environment is a 'immersive, interactive, multi-sensory, viewer-centered, three-dimensional computer-generated environment (Cruz-Neira in (90, p.9)). How "immersive" is interpreted here, and whether that would limit virtual environments to devices which offer quantitatively immersive experiences (like HMDs) was unclear. Another states: "virtual environments (VEs) attempt to create a replica of an object of experience in a way for a human to understand" (123, p.550) indicating a more expansive understanding of virtual environments.

AR applications in the form of wearables (glasses or goggles), as well as mobile devices and their associated platforms, were the third most discussed unit of analysis. Of note, the mobile platform Pokemon GO was the AR application with the highest number of mentions in the literature. (103, 104, 151, 155, 157, 161, 168) While HMDs are frequently associated with VR, some of the more recent discussions have focused on AR and MR applications that utilize passthrough technology. (47, 72, 83, 93)

Finally, disciplinary scope and aims should be mentioned. The sources come from a wide range of disciplines, and similarly purport a variety of aims. These sources went about achieving their aims through various methods, from offering a critical discussion, to philosophical argumentation, to mixed methods and proof of concepts/experimental work. Discussions and evidence of risks and harms are not siloed to a narrow disciplinary or methodological regime. Though these findings do not have an accompanying table, they are available in reviewer's data charts.

Funded by the European Union

3.2 Typology of risks and corresponding harms

The main purpose of this review was to map risks and harms of XR. The process of coding and subsequent iterative classification of codes allowed for reviewers CV, SC, and LH to develop a typology of risks. What emerged after this process included six major themes and 19 sub-themes: 1) health and wellbeing (cognitive, psycho-emotional, socio-relational and physical), 2) autonomy (privacy, self-sovereignty, surveillance, authenticity), 3) epistemic and scientific (epistemic uncertainty, scientific uncertainty, algorithmic imperfection, manipulation), 4) normative (norm violations, rights violations), 5) structural (socio-political, economic/financial, legal), and 6) technical safety and security (security, dual use).

A typology is practical for the presentation of such a large dataset, with large variation in *types* of risks and harms. This means that certain themes will be inherently more concrete (like health & well-being risks), while other's more abstract (autonomy risks, normative risks, epistemic risks). (See Figure 6) Additionally, the codes are not mutually exclusive and can (and should) be considered as overlapping between thematic categories. The typology should therefore operate mainly as a revisable, heuristic mapping for risks. A substantial number of codes were associated with each risk/harm category and are listed in tables 6-11 accompanying the thematic elaborations.

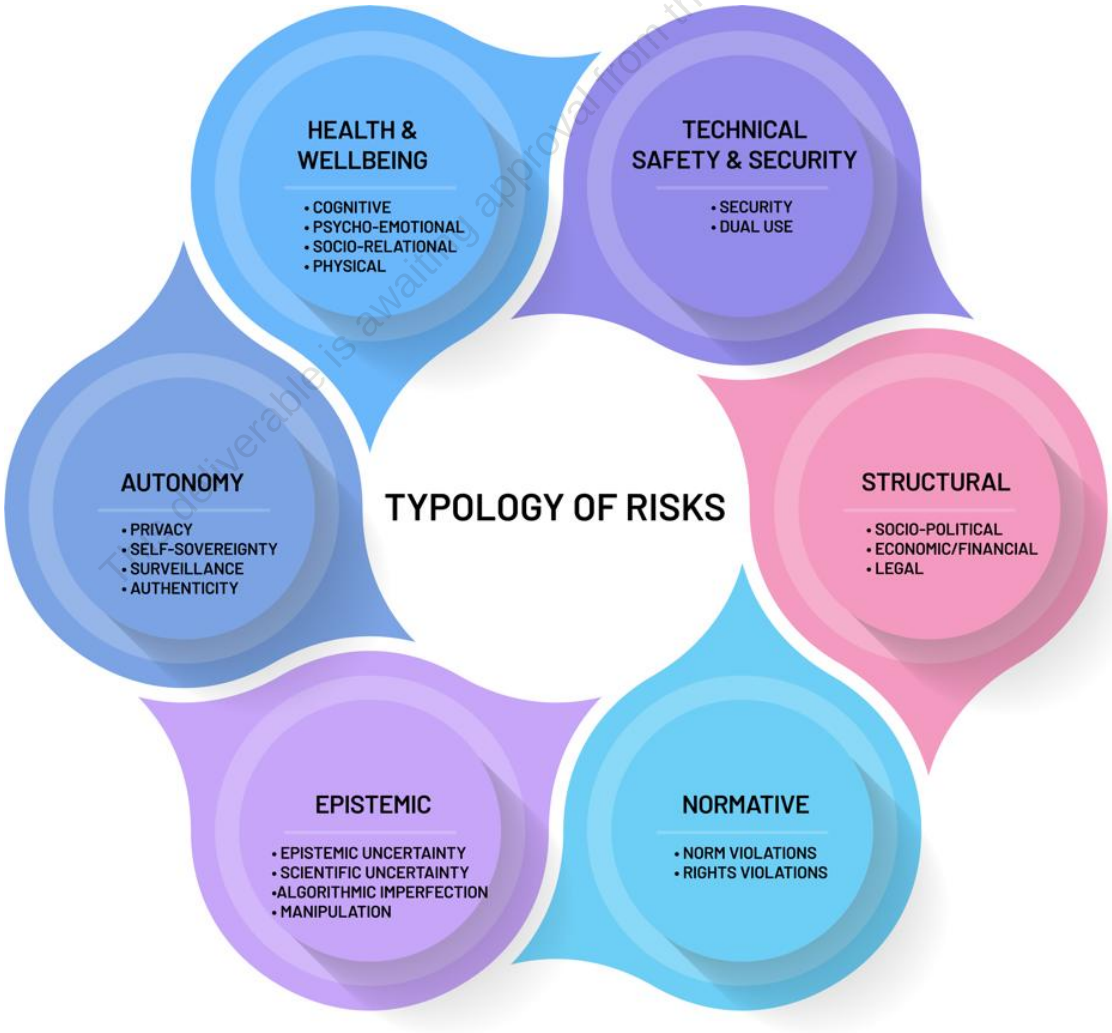


Figure 6.0 Risk typology developed by the reviewers from the iterative classification of codes for risk and harm. The typology includes six major themes and 19 sub-themes. See further details in Section 3.2.

The following section is a reporting of the risks and harms based on the typology. Note that written sections will not present all the codes, but rather the codes that the reviewers had in common or were frequently discussed. Later discussion will then reflect upon and use evidence to contextualize both the charted findings as well as the findings classified within the typology.

3.2.1 Health and Well-being risks

Risks to health and well-being encompass "a state of complete physical, mental and social well-being and not merely the absence of disease or infirmity". (225)

3.2.1.1 Cognitive

Cognitive impacts are those to the brain's conscious activity: to think, understand, learn, remember, and make decisions about, among other things. Here, the affordances of XR – particularly immersive VR – offered several unique risks. Immersive VR is able to convince the brain of new environments which can make a user feel like sensations akin to those experience in outside of VR, and may even influence behaviour. (123) These behavioural changes can be limited to the virtual environment itself, where the participant takes on associated behavioural traits of their avatar representation ("proteus effect"), or may extend virtual activities into extra-virtual spaces.(90, 105, 120, 145)

The combination of immersion and presence (a user feeling like they are in a virtual environment) among other phenomenon, can contribute to the unique difficulty of separating virtual and physical "realities". (67, 72, 123, 129, 191, 207) This is not only a problem for immersive XR, but also for AR which may potentially confuse digitally over-laid information with real information. (119) This may mean that some persons may confuse the two realities, or even obtain false memories from a VR experience, a problem especially for children who are not as capable as adults in distinguishing between imaginary and real. (19, 115, 141, 204)

In addition to its immersive and presence-inducing properties, XR offers a sense of embodiment, where the user believes their virtual body as their own. (143) This can affect the way a user may perceive 3D space, impeding or inhibiting spatial understandings. Additionally, immersive XR can have a distorting effect in how we think about our non-virtual bodies. (112) These include changes in our body image, leading to practices of self-objectification, issues in how persons remember their body, as well as perception that the virtual body is "more real" than the biological body which may contribute to various disordered behaviours and emotional difficulties. (93, 112, 142)

Related to how persons may understand VEs is how they interpret virtual entities, avatars, or objects. Users may anthropomorphize, or add human characteristics, to virtual entities, which can lead to interacting with an avatar as if it were "capable of empathy, individual perspective, and rationality. As a result, they may place undue confidence, trust or expectations in these agents". (90, p.37)

Several authors noted the addictive potential of XR; the more immersive the artifact, the more potentially addictive. (90, 102, 141, 142, 183, 205, 218) Authors differed on the exact mechanism motivating this, but it was generally posited that a user's sense of presence – of being in the virtual environment – in combination with compelling and "real"-feeling experiences have compelling force. (90, 178, 193) The virtual environments themselves can be voluminous and complex to the point of overwhelming users' ability to process it, resulting in information overload. (112, 204) Sources also describe XR as capable of reducing

cognitive performance, contributing to poor decision making within intra-virtual task assignments, difficulty concentrating, as well as problems with attention. (102, 112, 119, 143, 165, 194)

3.2.1.2 Psycho-emotional

Psycho-emotional risks are those pertaining to the psychological and/or emotional states in a person. Here again, XR's unique virtual environment-building capabilities invites several considerations. These include the potential triggering or intensification of certain emotional states like anger (aggression) and mental health conditions like anxiety, as well as dissociative conditions such as depersonalisation and derealisation. (90, 115, 141, 221) Related to this is a feeling of agential uncertainty, where "the risks of no longer perceiving oneself as the author of one's own action". (145, p.579) Uses of the technology may serve to desensitize persons to certain experiences, or to decrease empathy. (19, 108)

Further, immersion and embodiment can make experiencing offensive, immoral or violent virtual content or interactions feel real, which have been shown to have a traumatic effect on users. (95, 102, 137, 145, 148, 193, 203, 216) XR's highly realistic graphic and haptic capabilities may make violations of personal space feel more distressing. (193) Some authors point to this feeling of "realness" potentially facilitating aggression in real life, or to XR as a space where radicalisation and violence can be potentialised. (97, 178) This realism may also facilitate a preference for virtual experiences over extra-virtual ones, which may lead to a tendency to escape into VEs as a form of real-life avoidance. (102, 112, 205)

3.2.1.3 Socio-relational

Socio-relational risks arise from or pertain to interpersonal relationships, relations between group members, or family. Following psycho-emotional impacts, escapism may also negatively affect real-world relationships. (112) Time spent with the technology may serve to isolate and disconnect users from families and friends, a reduction in physical or social contact, and difficulty forming relationships. (109, 111) Further, social interactions within XR can be fraught with harmful behaviours, including harassment and assault in the form of defamation, griefing, groping, modifying user avatars, stalking, trolling, and virtual graffiti. (72, 175, 180, 182, 193, 211, 213, 218) These environments may also serve as conduits for control or exclusion, where users feel the need to conform or self-sensor in order to fit in. (111) In addition to personal relationships, professional relationships may also be impacted. XR has the potential to increase remote work, which may lead to heightened workplace surveillance, as well as difficulties balancing between work and life. (90)

3.2.1.4 Physical

Authors note several impacts on the body from the design or use of XR. A primary risk is that of cybersickness, sometimes called simulation sickness, where users of immersive XR experience a series of negative symptomatology including, but not limited to: "nausea, vomiting, eyestrain, disorientation, ataxia, and vertigo". (205, p.10); (38-40, 44, 46, 48, 49, 53, 124, 126, 127, 166, 197) Those using HMDs face additional issues with reduced user stability, where balance and orientation may be affected. (48, 194)

Use of these devices also have impacts on the visual system, from discomfort, blurred vision, and eye strain (asthenopia), to worsening of stereoacuity and other oculomotor effects. (46, 53, 188) Some of the adverse effects may be a consequence of the vergence-accommodation conflict, an unnatural disruption between vergence and accommodation that occur in a typical stereoscopic HMD. (46, 48, 121) HMDs that are not properly aligned in front of the eyes may also lead to perceptual distortions and worsening of the discomfort. (53) Uncomfortable and poor fit of the headset is more common in those with smaller head size and shorter interpupillary distances, such as in females and children (see section 3.3. for more on this). (53, 188) To have the correct prescription and the possibility to wear own prescription when using the devices, are of importance for comfortable use. (121, 195) Wearing own pair of glasses under the headset is, however, not possible for all headsets. (37, 121) Some people lack stereovision and are not able

to see in 3D, while others have reduced stereovision (~30% of the population). (121) These conditions will negatively affect the experience of, and interaction with, the virtual environment. (121, 147)

There are several ergonomic effects of XR use, from fatigue due to hardware weight, strain from requiring more head movement, but also movement limitations. (119, 194) The possibility of injury is frequently mentioned across units of analysis and output devices. These injuries may be minor but at times be life-threatening or even result in death. (163) Injuries may arise due to several issues: from hardware or game design to malfunction and manipulation. (47, 168) When it comes to HMDs, visual and auditory occlusion may lead to self-injury or collisions with bystanders or objects. (216) HMDs also have a small field of view which contributes to inhibited spatial understanding, compromising a user's ability to navigate spaces safely. (93)

Game design in AR, particularly with mobile AR games like Pokémon GO, are frequently cited as facilitators of "distraction"-related injuries, including accidental death when a user pays closer attention to the AR display than their physical environment. (139, 155, 156, 164, 168, 193) It has also been reported that AR-playing individuals have been killed for trespassing. (193) Some other effects include headache, auditory occlusion, seizures and infections secondary to sharing devices. (67, 133, 205, 224) Finally, authors write about how the cognitive, psycho and socio-appeal of virtual environments can also have negative effects in how persons take care of themselves, selecting to participate in virtual worlds instead of managing their bodily needs. (112)

HEALTH & WELLBEING RISKS			
COGNITIVE	**PSYCHO-EMOTIONAL**	**SOCIO-RELATIONAL**	**PHYSICAL**
Addiction Anthropomorphizing virtual objects Avatar proteus effect Creation of false memories Difficulty concentrating Difficulty distinguishing real and virtual realities Embodiment illusion and body memory Incongruent perception/understanding of self to 3D space and physical surroundings Information overload Manipulation Negative effect on child cognitive development Negative effects on users with pre-existing cognitive conditions Perceptual changes and unconscious influence on cognition leading to: • real life immoral actions • real world behaviour transfer from virtual • dangerous behaviour Poor decision making from visual fatigue Problems with attention Lack of awareness over surroundings Reduced cognitive performance Unknown cognitive effects from general or long-term use to wider population and vulnerable groups	Agential uncertainty Aggression Anxiety Attachment to avatar/virtual beings Body dysmorphia Brainwashing Deception Depersonalisation and derealisation Desensitization & decreased empathy Exacerbation of pre-existing conditions Detachment and Escapism Exposure trauma Intensification of experience Nudging Violent propensity Reduced inhibitions Re-traumatisation Self-consciousness Self-objectification Technostress Withdrawal symptoms Unknown psycho-emotional effects from general or long-term use to wider population and vulnerable groups	Abuse of power Assault Bystander haptic proxy Controlling vulnerable persons Defamation Harassment False attribution towards specific groups based on XR experience Impersonation of others Impoverished inter-personal relationships Griefing Groping Manipulative advertising Radicalization and extremist recruitment Remote work gender disparities Reputational damage Self-censorship Social engineering Social isolation or asocial Social exclusion Socially unacceptable use of device Stalking Surveillance & Sousveillance Threat misidentification Third party avatar modification Trolling Virtual graffiti Work-life imbalance Unknown socio-relational effects to wider population and vulnerable groups from general or long-term use	Auditory occlusion Cybersickness and discomfort Dizziness and postural instability Ergonomic issues Fatigue Strain Headache Infections - (microorganisms) Increased reaction time Falls, injuries, life threatening situations & death for user and bystander. Manipulated and limitations of movements Nausea Neglect of well being Obesity Seizures Sensorimotor after-effects Unspecified adverse effects on minority users Sweating Visual disturbances and asthenopia Unknown physical effects from general or long-term use to wider population and vulnerable groups

Table 6.0 Health and well-being risks

3.2.2 Autonomy Risks

This category notes risks related to the ability to "govern oneself, to be directed by considerations, desires, conditions, and characteristics that are not simply imposed externally… but are part of what can somehow be considered one's authentic self". (226, p.1) This definition is selected for its broad scope and is reflective of issues brought about by authors which have been grouped into sub-themes: risks to privacy, to self-sovereignty, risk of surveillance, and to authenticity.

3.2.2.1 Privacy

Privacy is both a state and a protective mechanism, creating a space in which people may "direct their lives as they see fit irrespective of social and political pressures". (227, p.118) Privacy was one of the most frequently cited concerns, across all units of analysis, devices and platforms. (33, 47, 56, 62, 66, 67, 70-72, 74, 78, 80, 82, 85, 191, 208, 209, 215) Privacy discussions centre around data – how it is ascertained, aggregated, analyzed, shared, controlled, and what that means for the average user. XR is changing and expanding in terms of data collection possibilities, alongside acquiring "greater contextual awareness" from analysing bystander and environmental data. (193, p.4) This means that XR is growing better equipped to make predictions about personal location habits (where a user goes, when, and how often). (193)

Mobile AR platforms and devices, despite not collecting unconscious biometric data like gait or eye tracking, nevertheless pose significant issues when it comes to environmental or spatial data and are prevalent as geolocating and potential recording devices. Privacy, therefore, is a significant problem for AR, where Global Positioning systems are heavily relied upon, as well as user information. (164)

Moreover, the discussions point to a uniqueness of XR technologies, particular immersive XR, in exacerbating the privacy issues already present with other forms of technology. Some consider digital technologies to be inherently privacy-reducing, especially in the linking of social networks (SNs) and VR where data capture and interpretation will take on novel dimensions. (178, p.6) Indeed, enhanced data capture is a notable feature of XR technologies. (56, 90, 103, 104, 107, 111, 178)

XR makes it easier to classify and track both users and bystanders, whereby bystanders cannot necessarily consent or opt out. (104, 105, 107) The ease with which users and bystanders can be de-anonymized due to highly capable environmental sensors and efficient machine learning is also a concern. (95, 107, 139) Additionally, some authors discuss how users may lack an understanding of these risks, not fully comprehending the privacy and terms of use agreements offered by companies.(56, 103) These threats are further complicated by the increasing incorporation of artificial intelligence (AI), which combined with sensor data may lead to hitherto unknown privacy threats. (55, 91) Authors describe violations of informational, associational, emotional, and physical privacy. (104, 212) XR can more easily reveal who one associates with, user's emotional or affectional states, as well as where they are. (52, 178) It may do this through unexpected or novel means such as through reflections on a user's pupil which indicate their location. (118)

3.2.2.2 Self-Sovereignty

Self-sovereignty can be understood as the ability to govern oneself, and one's information, deciding upon and pursuing a self-selected course of action without interference. XR has unique implications on user's ability to self-govern. Data, again, is a critically discussed issue. (95, 107, 167, 221) Not only does XR have access to intimate data, it can share that data with third parties. (183) This results in an increased vulnerability of data: more of it, ease with which it can be de-anonymized, and then shared and analysed. (119, 178) Importantly, many users are unaware of what has been shared, and do not have much control over what is, nor any way to retrieve or change this data. (90, 105, 192) This data may be available for others for the entirety of one's lifetime. (88)

Then, there is the risk of data inaccuracies and misuse. (56, 95) XR, supported by AI, can infer personal and intimate user characteristics and protected traits (i.e. sexual preference, etc.) based on information extracted without the user's knowledge or consent. (107) This type of data analysis is referred to as "biometric psychography", wherein XR is able to gain a "sophisticated, longitudinal understanding" of a person by tracking not only their behaviour but also their mental and effective states (ex: arousal, etc.) and cognitive processes. (107, p.11) Authors note that this inference-making capability is not always accurate, but can nevertheless be used to potentially discriminate against the user, or for targeted marketing efforts. (105, 111) It can also be used to alter behaviour in a way that benefits a company or brand. (103, p.289)

Self-sovereignty, as discussed by authors is related to control. "VR," writes Marloth et al. (145) is "understood as artificially created and completely controllable". It is difficult to determine who, exactly, is in control within a virtual environment setting – the user, the developer/designer, or the technology itself. Though authors like Marloth et al. imply that there is an inherent lack of control in artificially created spaces where the user has fundamental design input. (145) The Metaverse is also discussed in terms of control, where it is pointed out that though users have an expectation of being exposed to other actors (i.e. lack total control), there are nevertheless social and interpersonal expectations within worlds about issues like anonymity, personal reputation, and who can watch and when. (212, p.96)

3.2.2.3 Surveillance

To be surveilled is to be monitored or observed. (228) While some forms of surveillance enhance (public health surveillance), sometimes to be surveilled serves to reduce autonomy (and other rights). (228) Surveillance is closely related to the topics discussed above – privacy and self-sovereignty. Multiple sources refer to XR as a vehicle for surveillance, across units of analyses, devices, and platforms. (56, 60, 61, 65, 104, 215) Surveillance is discussed as an inherent aspect/feature of virtual environments/Metaverse. (222, p.5)

AR-related literature discusses issues of persistent and pervasive devices and platforms. (119) Pervasive AR is "always on" and a "disappearing interface" which makes issues of private use but also workplace use problematic. (119, p.87) Increased workplace surveillance is cited as a risk of XR technology in general, made possible by the unique affordances of immersive wearables such as HMDs. Employers can potentially monitor employees through biometric sensors, and make potentially scientifically unsound conclusions about focus and procrastination. (90, p.41) It is not only the workplace that is monitored. Some authors proffered XR as instrumental to government spying (big brother). They suggest that criminal investigators may use XR to identify individuals at a protest, or may obtain a journalist's 360 video. (218, p.340) Finally, the risk of citizen spying (little brother, sometimes called sousveillance) both in the real world and in virtual environments or the Metaverse, is discussed. (98, 111, 212, 218) Wearables, especially AR smart glasses, are considered efficient tools for tracking and spying on persons without their knowledge. (111, p.130)

Funded by the European Union

3.2.2.4 *Authenticity*

Authenticity concerns what it means to be oneself, to act oneself, or to be "expressive of one's own self-identity". (229, p.1) Authors raise the question of whether an avatar is, in fact, an autonomous agent acting of their own authentic desires, or whether the controlled and constructed nature of virtual environments prevent authenticity. (111, 184) Maslen and Salvulescu note:

> "immersive technology might render the actions of the virtual agent inauthentic or, at least, not continuous with those of the "real life" agent. The concern is not that the agent herself becomes inauthentic, but rather that her activities in the virtual environment are not her own thus, the question is not so much whether the use of immersive technologies renders the self inauthentic but, rather, whether one can be authentically oneself in virtual environments. If I visit you as my robot avatar, have you really met me, and are the actions I take mine?" (184, p.4)

Other sources discuss a reduction in authenticity in terms of a pressure to conform under constant monitorization (authenticity as relational to issues of surveillance and self-sovereignty). (141, 178) Conformism in the Metaverse is talked about as a flattening of the individual in favour of a collective or group identity. (222, p.4) Other analyses posit that VR can reinforce certain social interactions in a way that may lead to homogenisation. (90)

AUTONOMY RISKS			
PRIVACY	**SELF-SOVEREIGNTY**	**SURVEILLANCE**	**AUTHENTICITY**
Biometric psychography **Breach of privacy** • **Associational** • **Informational** • **Physical** **De-anonymising data of user/bystander** **Location registration** **Lack of or limited awareness of privacy risks** **Third party identification of user/bystander** **Traceability of avatars** **Tracking of user/bystander** **Unintentional sharing of sensitive (intimate) data** **Violation of the principle of purpose limitation** **Unspecified AI influenced novel privacy threats** **Unknown privacy threats**	**Confusing choice architecture** **Compromised consent** **Inability to opt-out** **Lack of or limited control of virtual environment** **Lack of transparency** **Manipulation of persons to act as robots** **Misuse of data** **Problems with data control** **Poor comprehension of terms of use and privacy agreements** **Over-trust**	**Enhanced data capture** **Sousveillance (Inter-Citizen)** **Surveillance (Government/Workplace)** **Live mapping** **World scraping**	**Conformism and homogeneity** **Instrumentalization of user and bystander for data** **Inauthenticity of avatar or user** **Reduced authenticity**

Table 7.0 Autonomy risks

3.2.3 Epistemic and Scientific Risks

Epistemic risks pertain to how knowledge or beliefs are made (230); how XR informs the actions, choices and judgements of users. This theme is divided into: epistemic and scientific uncertainty, algorithmic imperfect and manipulation.

3.2.3.1 Epistemic Uncertainty

XR has several issues pertaining to epistemic and scientific uncertainty. Some authors claim that immersive XR is inherently epistemically uncertain. (112) This is because virtual environments are generated, and these generated environments then interact with a user's senses in various ways. Inhabiting a virtual environment, therefore, is inherently risky when it comes to knowing what is real/true. One author notes that the visual experience of a person is generally assumed to correspond with "ground truth" or an objective physical reality. However, XR is inherently unstable, where generated data can be controlled, rearranged and manipulated such that the sensory experiences of users are altered. (112, 142, 145)

Within XR, the issue of what is true and how it may be known is also one that relates to distinguishing between real and simulation – a problem first discussed under Health and Wellbeing risks. Authors not only consider this as a cognitive and psychological problem, but also a risk relevant to knowledge acquisition, belief, and truth. This is not only an issue for VR, but AR and MR, where illusions generated by a computer may transform into a belief (119, p.88). Pervasive AR, further, aims to further diminish the real from the virtual, which authors warn may mean "perceptual artifacts" become increasingly apt at moulding person's attitudes, beliefs and knowledge. (90, 119)

There are also risks related to avatar identity and potential user harms. It is difficult to determine whether a virtual identity and a real-world identity cohere. There is no guarantee that someone online is who they claim to be, (97-99, 213) nor can it be guaranteed that a user is interacting with a real individual or bot, a problem which will be exacerbated by AI. (97, 98) This may have several consequences in terms of eroding trust. (97)

3.2.3.2 Scientific Uncertainty

Several authors discuss the relationship between XR and science. These discussions revolve around two main points: authors who question the evidence supporting XR as a therapeutic, educational, or training tool, and authors questioning the epistemic validity of experiments using XR, i.e. that an experiment is actually measuring what it is claims to measure. (194, 201, 202). First, while there is much research taking place on the effects of XR, there is still a lack of knowledge within certain domains and subject matters; particularly the physiological, ethical and social dimensions and psychological impacts on special populations or vulnerable groups. (19, 35, 139, 198). As such, there is a risk of overstating the possibilities and benefits of XR technologies. (19, 185) Regarding scientific validity – specifically how generalisable (quantitative) experimental finding can be – studies may not be adequately accounting for gender differences. (201) Additionally, some authors question how field of view impacts results (194) and whether "virtual" experiences can adequately inform us about "material" (i.e. non-virtual) experiences at all. (90, p.35) Authors also caution that XR could be used with clinical populations without sufficient training or might be applied without the participant fully understanding associated risks. (149). Finally, there is a risk of engendering the participant or patient with a sense of false hope at the possibilities afforded by XR. (139, 140, 145)

3.2.3.3 *Algorithmic Imperfection*

XR technologies are reliant on data to function. Collecting, processing, and transmitting this data is, in turn, dependent on algorithms and other automated systems (ex: machine learning). (231) Algorithms, however, can be "black boxes" whose inner processes can be unknown even to those developing them, and whose results may be flawed. (107)

What this means for a user in XR – whose data can include environmental or biometric components – is that the nature of these errors has implications on personal identity, where an algorithm makes an inference that is incorrect (ex/ identifying someone as having a mental illness, etc.). (107, 192)

Authors further question the training datasets for XR technologies and XR learning analytics, as well as the predictive power of such analytics, often trained on male, white, neurotypical and able-bodied engineers. (192) In other words, these algorithms are subject to bias. (107, 192)

3.2.3.4 *Manipulation*

Knowledge manipulation or the manipulation of virtual environments include the "usual suspects" of internet manipulation: misinformation and disinformation, fake news, deepfakes, filter bubbles, gatekeeping, cyberbalkanization, among others. (86, 91, 92, 178, 218, 219) What further complicates these manipulatory practices common to all digital platforms, authors claim, is that the immersive XR experiences facilitate these practices more easily and convincingly. (47, 91, 92, 219) Consider a VR deepfake, where persons may "be portrayed as carrying out actions and saying things they did not do. This is already powerful enough in photos and videos, but in XR could be even more dangerous because plausibility includes the automatic attribution of realness to virtual humans". (112, p.7) AI, which offers enormous generative potential, used in conjunction with VR or other immersive technologies, can make malicious AIVR design and disinformation a greater possibility. (219)

Immersive virtual environments may use cognitive or psychological manipulation to achieve various objectives. (94, 218) Marketing agencies may take advantage of persons by designing VEs towards the goal of furthering sales. (119, 221) An HMD user can be physically steered into various directions through subtle but imperceptible changes in spatial knowledge cues ("human joystick attack"); they could be made dizzy or confused. (116, 219) The wealth of knowledge available about a user may fuel a nefarious actor's ability to take advantage of them, perhaps for financial or social gain. (219) Users may be scammed, experience extortion or "sextortion". (92)

Manipulative potential may be scaled up such that there is "risk of harm to our own processes of knowledge creation and reasoning". (92, p.93) For example, pervasive AR, where digital objects become indistinguishable from "real" objects, may be used to facilitate a false reality. This effect on "objectivity" includes fabricating journalistic or social events. Scaled up, malicious actors may introduce "paranoviruses" as a mode of social control, were networked devices are used to provoke group panic, paranoia, or to facilitate propaganda. (96, p.736)

EPISTEMIC & SCIENTIFIC RISKS			
EPISTEMIC UNCERTAINTY	**SCIENTIFIC UNCERTAINTY**	**ALGORITHMIC IMPERFECTION**	**MANIPULATION**
Avatar Identity Biased immersive Journalism Difficulty distinguishing between real and simulation Superrealism Inherent uncertainty Immersive experiences become beliefs	Diminished objectivity Difficulty to test all aspects of technology Questionable ecological validity Questionable empirical evidence Lack of or limited assessment tools for usability of device Lack of or limited research on long-term cognitive and psychological impact Overstating benefits Clinical and research issues: • Contraindication to participate in XR • Consent issues • Hasty use • Paternalism • Therapeutic misconception • Unsafe application of XR in treatment • Lack of training • Lack of established standards for medical or research	Algorithmic bias Algorithmic miscategorization Facial recognition incongruency Misuse of augmented intelligence	Cyberbalkanization Dark patterns Deepfakes Digital deception Fake news Filter bubbles Gatekeeping Fabricated visual or auditory information Malevolent creativity Misinformation and disinformation Misleading or manipulative design Paranoviruses

Table 8.0 Epistemic risks

3.2.4 Normative Risks

To make a "normative" decision is to determine which actions or outcomes are good or permissible and which are bad, or impermissible. (232) In this way, constructing a "norm" means determining rules for behaviour, sometimes based out of convention or etiquette, other times rooting in moral or ethical reasoning. Human rights may be considered a type of explicitly recognized (legally, socially) moral norm.

3.2.4.1 Norm Violations (Moral and Conventional)

Authors generally refer to potential norm violations as arising from or being influenced by virtual environments. Some argue that the immersive and highly realistic experiences of some virtual spaces/worlds may encourage immorality in general, or unethical behaviours, as these environments expose people to what may be considered immoral content. This is, of course, subjective, as perceptions of what is moral vary widely from society to society and even within societies. However, there are some things that may be deemed morally objectionable by a broad segment of society. An example is the simulation of underage sex. (163, 193, 221) There is also some speculation that certain exposures, for example, violent content, might alter normativity. (110, 203)

Further, the felt "realness" of XR experiences may serve to motivate what some call "empathy content":

> "VR as an 'empathy machine' puts the device to work as an effective curative to circumvent cruel or indifferent thoughts about already immiserated people and replace them with intense sensory experiences that invoke feelings of presence... 'Seeing as' a refugee... making racial empathy pleasurable puts the under commons to work providing empathy content". (108, p.53)

Nash discusses empathy content as a risk of "improper distance", where immersive virtual environments or virtual videos facilitate "forms of self-focus and self-projection rather than a more distance position that allows for recognition of distance between self and other.... it presents distant spaces and others in an aesthetic mode, inviting a contemplation of the scene as a tableau vivant or spectacle rather than a painful reality". (217, p.125) In other words, moral engagement with another person becomes about how the user comes to feel or experience the feelings of other people – centering the user rather than the person whose "painful reality" is being portrayed. There is moral power, Nash argues, in VR's ability to *simulate* experience as opposed to merely showcasing events (like traditional television). This runs the risk of no longer recognising the other *as* other, but rather foregrounds the emotional experience of user themselves.

Avatar representations within VEs and AR filters may also have an impact on norms. The homogenization of society intra-virtually may influence a loss of diverse, culturally specific norms,(129) as well as may alter norms of beauty. (112) When analysing normativity within virtual environments themselves, Deng and Ruan argue that an avatar is a symbolic representation of a person's past and present selves, in addition to being an "extension" of a user's personal identity – highlighting that personal dignity in virtual environments can indeed be violated. (167, p.338) The avatar, considered by some to be an extension of a person is also referred to as having a prominent "moral" effects on users represented by their avatars. (90, p.49) Other examples of dignity violations, authors argue, are the presence of deepfakes of the dead in VEs. (90, 99)

Some authors discuss the "moral danger" of XR, suggesting that normative accountability in XR becomes challenging in a virtual environment where consequences for actions may be absent : "if I feel I have nothing to lose, why treat other people with dignity and sensitivity?" (115, p.1542) Virtual Reality Social Networks (VRSNs) may facilitate a societal and personal shallowness, wherein information is presented instantly, which according to some, means little time is spent contemplating and "digesting" that information. (178, p.19)

Other potential violations occur when inappropriately using devices. Playing AR mobile games like Pokémon GO can risk crossing cultural or ethical boundaries. The game, for example, has been played in churches or at the Holocaust Museum. (163, 168)

3.2.4.2 Rights Violations

Rights violations, potential and actual, are reported across units of analyses, platforms, and virtual environments. Of note are discriminatory practices in the design or deployment of XR, un-inclusive design and accessibility issues, and the possibility of XR deepening existing societal inequalities. (139, 195, 201) Some authors suggest that XR devices are often modelled off white, able-bodied young men, accommodating their physical attributes and thus using them as standards for device-builds. (192) In addition to the devices themselves, XR platforms and virtual environments, are said to be built to reflect the needs, desires, and implicit biases of the same demographic group. Thereby, negatively impacting the XR experience of other races or genders despite their initial interest in the device and/or platform.(100, 114, 193)

Persons living with disabilities or cognitive impairments may also encounter use limitations, such as some HMDs requiring a non-sitting position, or an overt focus on visual or auditory cues to navigate

platforms. (139) Cisgendered women, or those with a smaller head size, are not accommodated in HMD hardware design, increasing the likelihood of experiencing simulation sickness. (114, 201)

Hosfelt et al. (105) and Lee & Hu-An (193) argue that player characters/representations (ethnic etc) and roles are not adequately represented in XR. Further, avatar representations themselves may be built around gendered or racialised stereotypes and often assume the player or user is a heterosexual male, thus catering to their perceived preferences (ex: females portrayed as damsels in distress or objectified (usually sexual)). (90, 193) Thus, there are concerns that platforms and virtual environments, programmed with biased data, will serve to perpetuate real-world inequalities and the marginalisation of certain groups, especially as XR grows to increasingly rely on AI. (100, 105, 223) Pokémon GO, for example, appears to be systemically biased toward urban places with few minorities as its geographic mapping of in-game points of interest is predominantly concentrated in these communities. (161)

A final consideration is that not all people can afford XR or have access to it. (110, 223) The increasing use of XR in school or work contexts may potentially engender inequalities, where those with digital literacy or economic potential face a competitive advantage, turning XR into a gatekeeping technology. Gatekeepers are those who may hinder career progress through making XR an implicit (or explicit) precondition of applying. (90)

NORMATIVE RISKS	
NORM VIOLATIONS (Moral & Conventional)	**RIGHTS VIOLATIONS**
Crossing respected legal, social, cultural and ethical boundaries **Changing beauty norms to align with avatars** **Dignity violations** **Deepfakes of the dead** **Exposure to immoral or offensive content** **Encouraging immorality and unethical behaviour** **Legitimation of immersive technology as weapon of war** **Loss of cultural diversity** **Moral damage** **Private spaces become non-existent** **Reversal of the relationship between man and technology** **Shift in values** **Surveillance as new normal** **Inappropriate behaviour due to unclear norms in virtual environment** **Unknown norm violations from general or long-term use to wider population and vulnerable groups**	**Accessibility issues** **Discriminatory or un-inclusive design** **Expansion of digital divide** **Gatekeeping** **Stigmatisation** **Marginalisation** **Replication of real-world inequities in virtual environments** **Exposing people to simulated acts without their consent**

Table 9.0 Normative risks

3.2.5 Structural Risks

This category describes risks to high-level structures: socio-politics, law, and economics/ finance. It looks at large scale arrangements of people, entities or systems, and how they may be affected by XR technologies.

3.2.5.1 Socio-political Risks

This is a combination category addressing risks to political and social systems. One large-scale social harm is the datafication of society, closely related to the algorithmisation of persons. (182, 223) To "datafy" is to transform life into digitally analysable units. (192) These phenomena are further facilitated by the large-scale data acquisition capacities of XR, and the nature of virtual environments as surveillable, analysable, and controllable. (192, 223)

Critics argue that datafication is inherently limiting, bracketing or setting boundaries on knowledge in a way which affects social relations on both a personal level and within large scale systems. (178, 191, 192) Algorithmisation is, similarly, described as the reductionist application of algorithms to analysing life. (223) Datafication and algorithmisation are considered by some as a form of "data violence" which may harm persons in unintended or unforeseen ways, predicated on implicitly (and sometimes explicitly) flawed data systems and engineering. (192, p.4) These processes, authors claim, are highly beneficial to big tech companies, reinforcing power and influence and producing new forms of surveillance capitalism i.e., a form of capitalism which relies on personal information as value. (193, 223) The concentrated power of technology companies may contribute to democratic decay, and also facilitate governance via technocracy. (105, 222) It is argued that these technologies are increasingly privatizing social interactions, shifting them to virtual environments and platforms – spaces where "inherently biased viewpoints and ideological framing" are present. (193, p.8)

3.2.5.2 Economic/Financial Risks

This category describes risks which apply to how resources are produced, used and managed. Several authors discuss the potential for XR technologies to increase the cost of living, as well as contribute to processes of gentrification. This is due to XR aiding in the shift of work to digital spaces, which may mean local workforces are no longer relied upon. (90) Additionally, certain XR affordances such as augmented sensing in combination with artificial intelligence have been shown to diminish the need for human labour, which can further limit opportunities for work. (88) Healthcare, too, may face increasing costs due to the implementation of XR systems in clinical settings. (158, 172, 189)

Another risk is that the XR market may become monopolized, which not only contributes to consolidating power and political impact, but also limits competition on an open market. (90, p.52) The monopolisation of platforms has effects on competition, and serves to facilitate the "interests of the few". (105, p.1) Virtual environments also present a range of resource management difficulties, particularly concerning virtual assets. (97)

3.2.5.3 Legal Risks

A legal risk is any risk that pertains to the lack of or breaking of existing laws. There are several identified by authors, stemming from actions taking place within virtual environments themselves, but also from using the technology in unlawful ways such as using geolocational data to target someone for a crime, or trespassing for game purposes. (163, 164, 209, 220) Virtual environments are described as "lawless", underscoring that governance (legal but also social) is at its infancy. (214, p.3) Thus, legal responsibility – whether the developer, designer, or user – for unethical or illegal actions is unclear, and regulation is in its infancy. (90, 164) This lawlessness may facilitate frauds and scams, identity theft, pyramid and ponzi schemes, copyright violations, intellectual property theft, and unregulated banking and gambling. (90, 97, 214) XR users can potentially experience violations of already existing data protection rules such as the General Data Protection Regulations (GDPR), however, it is not well understood how to apply the regulations to XR. (112)

STRUCTURAL		
SOCIO-POLITICAL	**ECONOMIC/FINANCIAL**	**LEGAL**
Algorithmisation **Authoritarianism** **Creation of panopticon** **"Data violence"** **Datafication** **Democratic decay** **Limited potential for socio-political solutions to risks** **Perpetuation of biases or power-structures** **Perpetuation of colonising ideals** **Privatization of social and political thought** **Social control** **Surveillance capitalism**	**Attention economy** **Creation of market monopoly** **Difficulty managing virtual assets** **Gambling** **Gentrification** **Higher cost of living** **Increased cost of healthcare** **Unemployment** **Harvesting in game currency eg gold farming** **Replacement costs for personal glasses**	**Copyright violations** **Defamation of character** **Exploitation of legal rights** **Fraud and scams** **Geolocating user to target for crime** **Identity theft** **Intellectual property theft** **Lack of governance/regulations** **Potential violations of data protection law** **Pyramid and ponzi schemes** **Taking advantage of lawless VR** **Trespassing in real life for game purposes** **Unclear legal responsibility** **Unconventional uses of VR as evidence in court** **Unregulated banking and gambling** **Unspecified breaking of law**

Table 10. Structural risks

3.2.6 Technical Safety and Security

This category pertains to risks stemming from the technical or digital limits and vulnerabilities of XR. These can be understood as risks inherent to XR's hardware or software – technical systems which can be disrupted or exploited.

3.2.6.1 Security risks

This subcategory identifies issues related to the design, maintenance and management of hardware, software, data, and services provided through XR technologies. These can be targeted attacks upon the system or user, exploits such as malware, as well as issues with negligent or unsatisfactory design. (111, 223) These targeted attacks may take on various methods; some may attack infrastructure directly by way of controlling the application or system (sometimes without user knowledge). (43, 219) Other attacks may occur by observing user interactions with devices, such as by way of shoulder surfing, or using sensors to deduce password inputs. (105, 107) Perhaps most worrisome, some authors purport, is the ease with which wearables such as HMDs can be overtaken and manipulated such that the user is physically manipulated (more below). The human joystick attack has been shown to unwittingly and unknowingly move HMD users from one physical location to another one selected by the manipulator. (123, p.550)

Other technical security risks may be brought about as a function of poor code or design. These can lead to various breaches, data leaks, and difficulty authenticating information. (97, 106, 119) The complex relationship between intra-devices also means that systems may share security problems between them. (123)

3.2.6.2 Dual Use Risks

According to EU Regulation 2021/831 "Dual Use" could be understood as any technology which can be used to "meet many civilian needs, but .....can also be used for defense, intelligence and law enforcement". (233) However, this review treats dual use more expansively, as authors describe dual use scenarios as any XR that, though designed nominally for a person's benefit, may be used to "threaten or harm people, animals or the environment". (234)

Immersive VR, accessible via wearables like HMD, were often the target of analysis. Authors write about the potential "double-edged sword" nature of XR technologies, which offer great societal promise by way of their unique immersive capabilities (among other potentialities), yet can simultaneously be used for nefarious purposes. (205, p.21)

Attributes of VR include its purported ability to increase empathy, which authors warn could, instead, decrease empathy,(216, p.6) or facilitate other behavioural or emotional changes. (140) These impacts on empathy may be applied to not only the general population but could be applied to military contexts, which may desensitize actors in delicate combat situations to further a nation-state's goals. (90, 122, 216)

TECHNICAL SAFETY AND SECURITY	
SECURITY	**DUAL USE**
Attacks **- Chaperone attack** **- Data scraping bot** **- Disorientation attack** **- Hack controls app in real time** **- Hacking** **- Human joystick attack** **- Implanting virtual eavesdropping** **- Keystroke inference attack** **- Overlay attack** **- Shoulder surfing** **- System becomes owned by attacker** **- Using sensors to extract information from user** **Exploits** **- Malicious apps altering digital object** **- Malware** **Negligence** **- Breaches via 3rd party apps** **- Data leaks** **- Difficulty authenticating information** **- Shared security problems between software or hardware** **Unspecified security threats** **Unknown security vulnerabilities**	**Empathy increaser to empathy decreaser** **Illicit applications** **Unspecified harms due to dual use**

Table 11.0 Technical safety and security risks

3.3 Further considerations

3.3.1 Special Populations (Vulnerable Groups)

There are some consistencies across categories which should be noted. First, that risks to certain populations or groups are made worse/heightened in some way. (93, 94) These populations were generally referred to as having a heightened susceptibility to experiencing harm, the causes of which were multi-factorial (physical, cognitive, social, etc.). In the category of health and well-being, these were often children, the elderly, disabled persons, as well as people with pre-existing conditions. (94, 140) Notable risks for children and young people pertain to how reality is experienced and understood, particularly when using HMDs. (94, 140) Here, a heightened risk of false memory adoption is noted, concern over whether children can discern properly between biological and digital realities after being immersed, as well as overall impact on cognitive development. (94, 140) Further, those with pre-existing cognitive or psychiatric conditions also experience heightened use-risks, such as those with reality-perception disorders or dissociative disorders becoming further destabilised.(147) Autistic users may experience specific psychological and physiological effects. (146) Those with vestibular or oculomotor conditions, or damage to the central nervous system, face a higher risk of cybersickness. (142, 143) Older users, too, face injury and usability limitations from deteriorated vision and hearing, as well as unstable gait. (147, 148)

3.3.2 Unknowns and Long term uses

Common to all categories and across units of devices, platforms and virtual environments were discussion pertaining to what is simply unknown about these technologies. For health and wellbeing, a combination of lacking and potentially problematic research means cognitive, psycho-emotional, socio-relational and physical risks of XR on persons is unknown (with special populations like children also being mentioned). (19, 140, 141, 145, 180, 218) The infancy of the tech produces unknowns at the level of governance and responsibility, as well as security vulnerabilities which may yet to be discovered (novel ways of attacking XR devices). (106, 222) It is also unclear what XR can do to the normative sphere, how it will change moral and social conventions, or to socio-political structures. (119) These unknowns deepen temporally – it is unclear how long-term use will further deepen these risks. (93)

3.4 Mitigation measures

A subsidiary objective of this review was to identify corresponding harm-reducing or risk mitigation strategies proposed in the literature. During the review process, however, it was noted that the understanding of some of the more technical measures required highly specialized knowledge outside the combined competencies of the reviewers. Therefore, the coded measures are reported based on our interpretation of the literature and, as much as possible, reported in a general (non-specific) format to facilitate ease of comprehension. The main themes are organised into four domains, namely: society (government and non-governmental), industry (design, development and deployment), research (with XR technologies), and individual (users and bystanders) (see Table 12).

3.4.1 Society (Government and non-governmental organisations)

At the societal level, several authors emphasised the need for the interrogating of existing laws to identify how and when these may be applicable to current identifiable legal issues arising with the use of XR technologies. (63, 97, 136) The recommendations at the societal level include governance measures and a call for more research-knowledge building in relation to implementation and impact of XR technologies. Governance considerations include the examination of existing laws and development of policies to address known crimes related to breach of privacy, property and ownership, and trespassing. (172, 173, 192, 210) There is a general appeal across the literature for the creation of or modernizing of laws, policies to address new and emerging legal challenges such as those that arise with property and commercial transactions in virtual environments. (66, 173) Corporate restrictions or limitations are also proposed to prevent data monopoly, emotional analysis, managing virtual content and types of games that may be considered dangerous and inappropriate. (223) Crime management measures include incorporation of digital forensics and cross border information sharing. (210) Rules and guidelines were recommended to address norm (moral and conventional) violations with regards to bias and discrimination as well as respect for cultural diversity, thereby ensuring inclusivity. (80) Codes of conduct and etiquette would be essential for social interactions when using the technology. (63)

Governance measures are particularly emphasized in relation to the Metaverse as it becomes more expansive in various activities including the purchase of property, commercial and other social activities. The risk of new types of crime in the metaverse beyond the scope of current legislative measures are significant considerations in a comprehensive report issued by Interpol and the World Economic Forum. (177, 210) Finally, building and/or facilitating equitable and safe XR should be prioritised, thereby ensuring ethical and human centered XR technologies. (121) There is an appeal for further longitudinal studies (knowledge building) on the technologies to understand the long-term impact and unknowns. (35) Some authors posited the use of design fiction to speculate on future impacts and for investment in digital equity and connectivity. (80, 219)

3.4.2 Industry

The industry mitigation measures were focused on improvements with software and hardware. Software recommendations were extensive and highly technical. To simplify the reporting, measures were organised as structural, privacy and well-being, and technical security. Structural software considerations include intelligent systems designs as well as ethics by design, including users in the design process, empowering users to define their virtual representations. (19, 37, 69, 194) There is a need for implementing mechanisms in virtual environments for mitigating harassment and promoting a culture that values transparency and accountability were reported. (170, 208, 210) Interoperability is an important structural measure. (97) Measures to protect privacy and well-being include minimizing tracking and collection of biometric information, enabling protocols for identity protection, permission requests, and augmented security protocols. (97, 105, 167) To minimize or mitigate against collisions and dangerous situations, some authors recommended spatial awareness sensors and warning message. (158, 208) Warning lights may also be placed on output devices such as AR glasses for example to signal use of facial recognition to bystanders. (63) Hardware design was also highlighted as a key consideration for industry. The authors asserted that developers should also consider designing lightweight and ergonomic headsets that are adjustable to facilitate different head size and inter-pupillary distance as well as ensure comfortable use with prescription glasses. (37, 121) Technical security measures include the use of AI based or dynamic security control. (170) Some authors emphasised encryption and preventive game design. Within the virtual environments such as the Metaverse, privacy enhancing technologies, block chain technology, shadow avatars and privacy bubbles were presented as plausible measures of protection. (208, 210)

3.4.3 Research with XR technologies

Proposed measures within the research domain include (a) adopting a review system tailored to the unique challenges of conducting research with XR technologies, (b) designing experiments with inclusion/exclusion screening for persons likely to experience cybersickness or may have contraindications or vulnerabilities, and finally (c) safety measures to facilitate gradual acclimation to the technology, ensuring researchers are adequately trained on how to use the technology, assessment of post-experience well-being for research participants. (146, 200) Privacy measures include use of the technology offline during experiments and minimising tracking or data and biometric extraction. (105, 136, 167) Safe physical environs to prevent both foreseeable and unintended harm to the participants could include a practical measure such as grab bars that would prevent users (especially the elderly) from falling when feeling disoriented. (200) The adoption of an Institutional Review Board (IRB) model within organisations where there is awareness of the possibility of dual use of XR for unintended purposes, ensuring informed consent by users, and development of XR specific code of ethics. (122, 208) Some authors argued for the application of the equivalence principle when assessing the ethics of a virtual experience when conducting research. They posit that "if would be wrong to subject a person to an experience then it would wrong to a virtually real analogue of that experience". (203, p.260)

3.4.4 Individual

Individual mitigation measures aimed to empower and protect users and bystanders. The main recommendations centred on the responsible use of XR technology. Maslen and Savulescu contends:

> ".....moral responsibility is a function of the intentions and foresight the agent had with respect to the harm risked or caused, and the avoidability of that harm, together with whatever duties of care exist..... Such principles would extend to telepresence. Whether a manufacturer is liable for some portion of the harm, or whether responsibility attaches fully to the agent who assumed the risks associated with using the technology, will in large part depend on the respective duties of care in the situation, the user's expectations relating to the functioning of the technology, any abnormal use of the technology, and the degree of foreseeability of the harm". (184)

Users are encouraged to balance virtual time with real life and take breaks in play and going into nature. (158) Risks awareness is important for both users and bystanders. (66, 80, 109)

Domains	Proposed Mitigation measures			
Society *Government & non-governmental organisations*	**Governance**	**Laws, standards, and policies**	**Data protection** • implement existing data standards (user and sharing) • clear and comprehensible privacy agreements • use of explicit consent forms **Modernize laws for virtual environments** • property rights • trespass **Corporate restrictions/limitations** • data monopoly • emotional analysis • type of content & games • tracking and sharing of sensitive data • virtual product placements	
		Crime management	Digital (NFT) forensics, cross border information sharing	
		Rules and guidelines	• address bias and discrimination • respect for cultural norms and inclusivity • Codes of conduct and etiquette	
	Equitable and safe XR	Ensure ethical and human centered XR technology Use design fiction to speculate on future Digital equity and connectivity		
Industry *Design, Development & Deployment*	**Software**	**Structural**	**Privacy & wellbeing**	**Technical security**
		Ethics by design Empower users to define own virtual representation Interoperability Mechanisms for mitigating harassment Valuing and enhancing transparency and accountability	Access restriction Augmented consent protocols Limit audio inputs Minimizing tracking and biometrics Motion cues Protocols for identity protection Permission requests Spatial awareness sensors in headset to prevent collision User controlled profiles Warnings/notification of notification of danger/ tracking Warning lights to signal use of facial recognition to bystander	AI-based security protocol Block chain technology Dynamic security control Encryption Intelligent system design Preventive game design Privacy bubbles Shadow avatars Trustworthy XR architecture
	Hardware	Design lightweight and ergonomic devices for comfortable use with prescription glasses Design adjustable headsets for different head size and inter-pupillary distance		
Research *with XR Technologies*	**Review system**	Apply the equivalence principle (TEP) Dual use awareness Institutional Review Board model in organization Informed consent XR-specific code of ethics		
	Experimental design	Control of simulation environment Inclusion/exclusion criteria • Screening for person with cybersickness, • Contraindications • Vulnerabilities/vulnerable groups Transparency		
	Safety	Gradual acclimation Minimizing tracking and biometric extraction Offline use of technology during experiments Post experience well-being Use of grab bars Training researchers		
Individual *Bystander, User*	**Responsible use of technology**	Balance time virtual with real life Breaks in play User awareness of bystander Risk awareness Ensure correctable eye and vision problems are corrected		

Table 12.0 Mitigation measures

4. DISCUSSION

The review gives a mapping of risks and harms of XR technologies across various use cases. Following a scoping method in conjunction with a qualitative descriptive content analysis, we identified risks and harms spanning across six major categories of risk: health and well-being, autonomy, epistemic and scientific, normative, structural, and technical safety and security. The following discussion will contextualise some of the charted items and findings (including the risk typology), point to potential knowledge gaps, and offer some reflections on the scoping review process.

4.1 Summary of results

First, the increase in XR publications over the last few years, the geographical distributions of said publications and how the distribution relates to knowledge-building and mitigation strategy are discussed. Then, the significance of the emergent risk typology will be addressed alongside a comparison of similar typologies.

The increasing number of publications on XR risks and harms over the last decade reflects the change in academic and societal interest in these issues. This might feed strategic efforts towards redress, and hopefully, more knowledge will bring about change. It is uncertain, however, what that change will look like or whose voices will be heard in this changemaking. The review's majority of sources originate from economically advantaged countries, with no sources and therefore an underrepresentation of views from regions/countries that are economically disadvantaged (ex., Africa, Latin America, and the Caribbean). How this underrepresentation may affect or steer the global discussions on how risks and harms are to be considered and mitigated is unclear.

Most of the sources where from the USA (46%). A country with advanced XR infrastructure accompanied by a robust startup ecosystem, accounting for 37% of global XR revenue share. (235) China and Japan are also projected to be key builders and investors in XR, (235) but only 6% of the sources where from China. There were no sources form Japan. Having no contributed works, however, does not necessitate that scholarship and knowledge building on XR risk is not happening. Sources may have been published non-English languages. Thus, there may be discussions taking place that account for risks and harms to which non-native speakers are not privy.

This review includes a comprehensive mapping of risks and harms related to XR technologies, with over 200 codes mapped. The listing of this large number of actual/potential harms in a reportable format may be overwhelming for the reader, but it is also a starting point for further research on risks and harms. The listed risks/harms are not connected to particular use cases, output devices or platforms. This connection can be made, through cautious inference, as 70% of the publication use cases were for general application, healthcare, and entertainment (gaming) as well as the main types of XR being VR and AR.

Although use cases such as education, research, governance, commerce, and industry were not significant in number of publications reviewed, the relevance of risks within these areas of application is of importance as these areas are projected to grow, especially for industry training. (236) Social and entertainment applications of XR were amongst the most frequent use cases reviewed. With the increasing convergence of social networks and XR – Meta's Horizon, for example – it is expected that novel risks and harms will emerge. (237) Military applications were notably absent from our findings. Given the sensitive and security risks associated with use of XR in military settings, this absence is explainable.

The mapping is not only comprehensive in terms of breadth of risks/harms, but also covers risks/harms across domains and units of analyses (explicitly including virtual environments and the metaverse). As far as can be found, only Cyber-XR offers a risk typology which includes all units of analysis (VR, MR, AR, but not VEs) and covers all domains (i.e. for general use). (14) There are several areas of

alignment between their typology and ours. This alignment is encouraging, as it reinforces our findings and subsequent categorisation of risk themes. For example: Cyber XR has five risk categories: human, financial, legal, information, and societal. (14, p.56-57) They do not, however, explicitly identify epistemic risks as a major concern. Instead, they grouped what herein is outlined as structural, normative, and epistemic under societal and legal risks, identifying issues such as disinformation, deep fakes and other epistemic risks as "manipulated social discourse". (14, p.56) Notably, Cyber-XR classify "lack of ground truth" and "inability to establish Truth" under societal and legal risks, respectively. (14, p.57)

While these things may indeed produce legal and social problems, they point to an inherent issue with XR: XR makes it more difficult to distinguish between digital information and non-digital information. Thus, the epistemic risk category is important to bring forward as an independent/unique conceptual category. It is critical that the knowledge ascertained and interpreted from XR technologies are assessed in their own right – especially given the nature of immersive virtual environments as changeable and experientially convincing, and the increasing integration of AI. Further, epistemic and scientific uncertainty is highly relevant to those conducting research with or on XR technologies.

The risk typology also distinguishes social risks from socio-political risks, signifying a separation of scale – social risks operating on a micro, inter-personal level whilst socio-political risks operate at a macro level. This separation is important to future scholarship which take on micro, meso, and macro-level analyses. Finally, normative risks were separated from risks to autonomy, while acknowledging the overlap between the two as fundamentally rights-based issues. The reason being that the abundance of literature pertaining to issues of self-sovereignty, surveillance, and privacy, especially, and as such these risks could be considered within a self-contained theme: autonomy.

4.1.1 The development of XR risk-strategies: potential motivators

The uptick in discussions and subsequent publications on XR technologies and its risks may be a function of the relatively increasing affordability and public visibility due to expanding range of applications (use cases). The early 2010s saw computer storage advancements, faster processing, powerful graphics chips, improved form factor and the release of the Oculus Rift in 2012. (238, 239) Oculus was soon acquired by Meta (then Facebook), "leading to a significant increase in the popularity of VR devices for home use". (238, p.2) Facebook's 2021 rebrand into Meta Platforms also signifies that large technology companies are serious about XR – Meta's own branding echoing the "metaverse". As of 2024, Meta owns nearly three quarters (74%) of the global XR market share. (240)

Technology companies' efforts towards designing and developing XR devices and platforms can have significant effects and influence on research and public discourse. This is especially salient given the relatively "high degree" of independence under which the XR industry operates. (241) For example, there was an increase in AR-related literature, which coincided with the release of Pokémon GO in 2016; with usership at the time in the tens of millions. (242)Thus, the role of technology companies in *mitigating* risks should not be underestimated. Whether the motivation to make critical design and governance changes comes from within company bounds, or whether it requires extra-regulatory power remains to be seen.

It is important to consider that regulatory bodies and civil society groups (NGOs or advocacy groups) are critical to the governance of XR risks as they too, alongside industry, are stakeholders – and that strategic efforts are indeed being made. Several included publications within this review propose regulatory strategies: The World Economic Forum, Interpol, and IEEE, for example. (97, 100, 107, 139, 177, 181, 210) Co-governance practices, where strategies for redress are collectively established between industry, civil society and governments, are growing increasingly popular. (241) The XRSI (Extended Reality Safety Initiative), an organisation which provides research and standards have advised both industry and governments, for example. (243)

The way technology changes over time also affects how conceptualisation of XR is popularized and communicated, which, in part, may explain the shift in meaning from "virtual reality" as desktop experiences to wearable displays with immersive affordances. "The immersive experience is what most people think of when it comes to VR and is one of the most marketable aspects of VR technology". (238, p.19) It would seem VR's affordances as a technology which can convincingly transport a person into a new space and have them feel as if they were there, embodied and present, is of interest to scholars and others in various domains. Indeed, VR comprised 43% whereas AR comprised 15% of sources within this review.

AR, whose goal is not to completely immerse but rather to augment or overlay digital information, may seem like less of a risk. However, the potential for AR as a "pervasive" technology, facilitated through the ubiquity of cellphones and mobile platforms, should not be underestimated. (119, 244) As Girginova et al point out: "Globally, many more people in 2022 experienced AR technologies on a daily basis than VR technologies. Yet, VR research dominates academic literature". (245)

Additionally, scholars predict an uptake of wearable AR which allows for an integration of everyday activity with digital information through optical displays; this "always-on" state means that the volume of user data tracked, stored, and interpreted will increase. (119) The network sharing capacities of these platforms makes it easier to create large scale 3D spatial maps – potentially against people's knowledge or will. (119) The wearability of XR and its ability to be integrated more seamlessly into everyday life – the latter being something immersive VR, by nature, cannot do – may bring about epistemic, structural, autonomy-related risks yet to be identified (among others). This is something to consider for MR as well, which was the least discussed unit of analysis, comprising 3% of sources. It may be that the relative newness of MR, or its conceptual ambiguity and practical similarities to AR which subsumes discussions about MR under AR. (246) With the 2024 introduction of HMD products like the Apple Vision Pro – marketed as a "spatial computing device" that "just about everyone calls" a mixed reality headset – an increase in company interest in MR/AR technologies is likely, as well as more public consumption. (247, 248)

A final thing to consider is that though units of analysis have changed in meaning and expanded over the years in terms of capabilities or affordances, risks identified in earlier sources are still reflected in recent literature. (99, 100) At least two decades have passed and problems persist, and become amplified, in immersive XR: harassment (notably of women and minorities), privacy violations, for example, among others. Is this a function of lack of effort, inadequate or insufficient mitigation techniques, or is it simply 'the nature of the beast', an inherent feature of a technology which is controlled by someone else?

4.1.2 Identified knowledge gaps

4.1.2.1 The need for interdisciplinary research

There are several research opportunities to consider which have not already been mentioned: first, further research into the mental, social and physical well-being of persons using these technologies – particularly in special populations and over the long term. These should be addressed not only from within a rigorous disciplinary perspective, but also through interdisciplinary approaches. An analysis of the more abstract dimension of XR risks and harms – epistemic, normative, etc. – should also be considered. A knowledge mapping by Girginova et al. (245) point to the relative underrepresentation of non-STEM research within the XR space. Thus, the social dimensions of XR, they argue, are important to understand if an ethical product is to be made. (245) Reviewers add that non-empirical dimensions -- philosophical, legal – are critical as well. Indeed, XR is an industry which operates on global computing infrastructure, therefore regulating at more local levels become difficult and requires unique legal analysis. Further opportunities for research involve how to world-build and govern virtual environments themselves.

4.1.2.2 Ramifications of AI/XR intersection

Attention should be paid to risks related to AI/XR. Though we did not explicitly code AI-related risks, anticipating that XR will increasingly rely on AI, some authors highlighted risks related to generating virtual worlds and the analysis and interpretation of large swaths of extracted information. Hirzle et al., (249) reports in their scoping review of XR and AI literature that AI can create XR worlds by "creating a realistic replication of the real world, modifying the real world, or generating a synthetic world". (249) They also note that AI may be used to understand and support interactions with users, facilitate interactions with intelligent virtual agents, and enable user authentication and identification. (249) Noting that the AI/XR intersection comes with not only the usual AI-related issues (bias, transparency, etc.) but unique challenges, reviewers posit that AI/XR combinations being used in simulations will augment many of the potential harms discussed earlier. AI/XR may contribute not only to epistemic and scientific uncertainty risks, but could be used to manipulate behaviour, cognition, emotions, etc. Computer vision and generative AI can facilitate real time surveillance and interactions that are tailored to specific users/bystanders. Reviewers therefore posit that all risk categories are of relevance for deeper reflection in relation to AI/XR intersecting and as such echo Hirzle et al's call for more discussion and research on the ethical and societal impacts of AI/XR combinations. (249)

4.1.2.3 Environmental risks, impact and opportunities

In addition to sources detailing AI/XR, attention should be paid to the unique environmental risks associated with XR. Some purport that XR reduces the need to travel for work or leisure, in addition to minimising office-space use which lowers an individual's carbon footprint. (250) XR may also promote immersive strategies for building relationships with nature through storytelling, thereby communicating environmental issues in potentially novel ways. (251) However, only one source within this review identified XR's potential *negative* effects on the environment. Adomaitis et al. (90) point out that the production and maintenance of XR infrastructure necessitates large amounts of material resources and energy. Additionally, they note, "information is physical". (90, p.50) Thus, it is not only the production of hardware as such, but the large energy-consumptive *data* stores and cloud services required by XR which have environmental costs - this includes the data stores necessary to maintaining Metaverse infrastructure. Considering these things, it is surprising that such a low number of references to environmental risks are reported.

It can be assumed that XR's likely integration of AI will serve to further problematize its impacts on the environment by creating additional energy needs. Nascent environmental impacts of generative AI include "expanding demand for computing power, larger carbon footprints, shifts in patterns of electricity demand, and an accelerated depletion of natural resources". (252, p.1); see also (253) Additionally, energy intensive practices are expected to increase with demand for high resolution images and video. (254) It is also unclear how increasing consumption will further drive electronic waste. (254)

Funded by the European Union

4.2 Establishing risks/harms within the XR4Human project

The task of effectively mitigating risks for emerging technologies within XR is challenging. XR is a speculative space and predicting with exactitude what the long-term effects of XR and its constitutive experiences is not possible. A first step is to understand the nature of the problem – the kinds of risks which need addressing in the first place, and what might be overlooked within and across disciplines (environmental risks, for example).

Further, this mapping buttresses XR4Human's proposed code of conduct for developers, its rating system, and is used to support normative frameworks proposed by the project. This review's typology offers a way to assess by way of comparison how potential principles within the code of conduct may or may not adequately consider various risks categories. Though non-binding codes and guidelines are not without critique (see 255), they should be considered a necessary component of a collective governance strategy undertaking between government, industry, and CSGs. These principles should be engaged with as a tool for active developer self-reflection in the making of XR technologies.

4.3 Limitations and challenges with scoping review process

There are several potential limitations. Indeed, undertaking a review of 195 sources from a range of disciplinary domains and methodologies did not come without challenges. First, registering a scoping review protocol would have been preferred, as it reinforces the values of open science and transparency, as well as potentially yielding valuable community inputs. Though a protocol was made, time constraints directed team efforts towards proceeding with the review.

Secondly, though we believe to have a robust selection of sources, and an effective search strategy, we may have missed valuable inputs by constraining our grey literature sources. We reasoned that the popularity of immersive technologies as a topic of speculation assumes a vast number of newspaper articles or popular magazine references – untenable to include for a small research team. The database search ending in January 2023 also may mean the most recent discussions on risks and harms were not included. Though, we do have resources from citation chaining as well as manually inputted articles between 2023–2024 which may temper this. We hope to augment this review with an up-to-date search in the future.

An additional limitation stems from having one reviewer reading texts in-full instead of the recommended two. (24) However, pilot articles were read together, and from there we established clear criteria under consensus. We also met regularly to discuss discrepancies, uncertainties, and when data items were unclear.

Finally, there may be some limitations to the typology itself. Coding was indeed a challenge, both from the perspective of interpreting what sources considered a "risk" or "harm", but also what type of XR or experience they applied to. This means the risks and harms which apply to, for example, a VR desktop platform might not apply to immersive devices or vice versa. It would benefit scholarship to map risks and harms according to conceptual specificity. We hope to do this in the future.

Additionally, multiple risks/harms cut across several thematic categories, even if they might not be categorised as such. There are also sub-categories within our typology with great conceptual overlap, such as cognitive risks and psycho-emotional risks. There are some practical reasons for this separation,

as combining sub-categories would have made for a lengthy category. Further (non-descriptive) critical analysis should be done on the inter-sections of risks and harm.

5. CONCLUSION

The aim of this article was to capture risks and harms related to the application and use of XR technologies, an umbrella term referring to a spectrum of technologies (AR, MR, VR) offering different ways of interacting with digital information. Facilitating virtual environments or metaverses is a key component of XR, and thus risks related to such terms were also mapped. Another task included mapping any mitigative efforts described. The report selected a scoping review method in conjunction with the PRISMA guidelines for scoping review to do this. Qualitative content analysis was selected to group identified risks into several thematic categories and sub-categories which emerged as coding progressed, alongside the categorisation of several ways of preventing or alleviating these potential risks.

These categories include risks to health and wellbeing (cognitive, psycho-emotional, socio-relational, physical), to autonomy (privacy, self-sovereignty, surveillance, authenticity), epistemic and scientific risks (epistemic uncertainty, scientific uncertainty, algorithmic imperfection, manipulation), normative risks (norm and rights violations), structural (socio-political, economic/financial, legal), and technical safety and security risks (security, dual use). Further considerations include how special populations (i.e. children, etc.) may face novel or heightened risks, and how there are a large swath of unknowns due to technological infancy and limited long-term research.

XR operates by gathering a large amount of data, both in breadth and depth, about its users (conscious, environmental and depending on the device, unconscious). This data capture is enhanced compared to other technology. In addition to this, the more immersive forms of XR can convince users that they are experiencing a virtual environment from a first-person perspective – environments which are felt as "real", even if they are digitally rendered. These features, among others, position the technology uniquely in terms of risks and harms. Mitigations at multiple levels – individual, industry, and societal – as well as within specific domains like research, are proposed to combat these unique risks.

The number of publications on the ethical issues of eXtended reality technologies are rapidly increasing and evolving in number, scope, and context. Consequently, the research gap exploring ethically relevant issues is narrowing with time, though several issues should be further explored including XR's impact on the environment as well as the unique epistemic and scientific risks manifested via AI/XR combinations. A further task for scholars is a mapping on risks related to specific definitions of XR, perhaps one which captures how meaning, and thus risk, has or has not changed over time. We hope to contribute to this scholarship in the future.

6. REFERENCES

1. XR4Human. The Equitable, Inclusive, and Human-Centered XR Project. Responsible development and uptake of XR technologies. n.d. [cited 2024 October]. Available from: https://xr4human.eu.
2. XR Safety Initiative (XRSI). The XRSI Definitions of Extended Reality (XR); XR Safety Initiative Standard Publication XR-001 2024 [cited 2024 October]. Available from: https://xrsi.org/wp-content/uploads/2020/03/XRSI-Standard-XR-Definitions-XR-001v002.pdf.
3. Carroll J, Hopper L, Farrelly AM, Lombard-Vance R, Bamidis PD, Konstantinidis EI. A Scoping Review of Augmented/Virtual Reality Health and Wellbeing Interventions for Older Adults: Redefining Immersive Virtual Reality. Frontiers in Virtual Reality. 2021;2.
4. Mann S, Havens JC, Iorio J, Yuan Y, Furness TA. All Reality: Values, taxonomy, and continuum, for Virtual, Augmented, eXtended/MiXed (X), Mediated (X,Y), and Multimediated Reality/Intelligence. 2018.
5. Rauschnabel PA, Felix R, Hinsch C, Shahab H, Alt F. What is XR? Towards a Framework for Augmented and Virtual Reality. Computers in Human Behavior. 2022;133:107289.
6. Bell M. Toward a Definition of "Virtual Worlds". Journal of Virtual Worlds Research. 2008;1(1).
7. X Reality Safety Initiative (XRSI). The Metaverse: XRSI; 2024 [cited 2024 October]. Available from: https://xrsi.org/definition/the-metaverse.
8. Hansson SO. "Risk". 2023. In: The Stanford Encyclopedia of Philosophy [Internet]. Stanford University: Metaphysics Research Lab. Summer 2023. Available from: https://plato.stanford.edu/archives/sum2023/entries/risk/.
9. International Organization for Standardization (ISO). Medical devices — Application of risk management to medical devices (ISO14971:2019). 2019.
10. IEEE. IEEE Standard Model Process for Addressing Ethical Concerns during System Design. IEEE Std 7000-2021. 2021:1-82.
11. Canning V, Tombs S. From Social Harm to Zemiology: A Critical Introduction. 1st ed. London: Routledge; 2021.
12. Zhou YK. What it means to suffer harm. Jurisprudence. 2022;13(1):26-51.
13. Unruh CF. A Hybrid Account of Harm. Australasian Journal of Philosophy. 2023;101(4):890-903.
14. The CyberXR Coalition. Immersive Technology Standards for accessibility, inclusion, ethics, and safety. Special Edition Cyber XR-2020-1.0. 2020.
15. Initiative XS. The XRSI Privacy Framework, version 1.0. 2020.
16. IEEE Standards Association. The IEEE Global Initiative on Ethics of Extended Reality n.d. [cited 2024 October]. Available from: https://standards.ieee.org/industry-connections/ethics-extended-reality/.
17. European Commision. Towards the next technological transition: Commission presents EU strategy to lead on Web 4.0 and virtual worlds 2023 [cited 2024 October]. Available from: https://ec.europa.eu/commission/presscorner/detail/en/ip_23_3718.
18. Nichols S, Patel H. Health and safety implications of virtual reality: a review of empirical evidence. Appl Ergon. 2002;33(3):251-71.
19. Kenwright B. Virtual Reality: Ethical Challenges and Dangers [Opinion]. IEEE Technology and Society Magazine. 2018;37(4):20-5.
20. Prabhakaran A, Mahamadu A-M, Mahdjoubi L. Understanding the challenges of immersive technology use in the architecture and construction industry: A systematic review. Automation in Construction. 2022;137:104228.
21. Costello PJ. Health and Safety Issues associated with Virtual Reality-A Review of Current Literature. 1997.
22. Cox S, Kadlubsky A, Svarverud E, Adams J, Baraas R, Bernabe R. A scoping review of the ethics frameworks describing issues related to the use of extended reality [version 1; peer review: 3 approved with reservations]. Open Research Europe. 2024;4(74).

23. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. International Journal of Social Research Methodology. 2005;8(1):19-32.
24. Aromataris E, Lockwood C, Porritt K, Pilla B, Jordan Z. JBI Manual for Evidence Synthesis 2024 [cited 2024 October]. Available from: https://synthesismanual.jbi.global.
25. Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, et al. PRISMA Extension for Scoping Reviews (PRISMA-ScR): Checklist and Explanation. Ann Intern Med. 2018;169(7):467-73.
26. Peters M, Godfrey C, McInerney P, Munn Z, Tricco A, Khalil H. Scoping Reviews 2020. In: JBI Manual for Evidence Synthesis [Internet]. JBI. Available from: https://synthesismanual.jbi.global.
27. Peters MDJ, Godfrey C, McInerney P, Khalil H, Larsen P, Marnie C, et al. Best practice guidance and reporting items for the development of scoping review protocols. JBI Evid Synth. 2022;20(4):953-68.
28. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan—a web and mobile app for systematic reviews. Systematic Reviews. 2016;5(1):210.
29. Lumivero. NVivo 14. 2024.
30. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. Bmj. 2021;372:n71.
31. Pollock D, Peters MDJ, Khalil H, McInerney P, Alexander L, Tricco AC, et al. Recommendations for the extraction, analysis, and presentation of results in scoping reviews. JBI Evid Synth. 2023;21(3):520-32.
32. Wierzbowski M, Pochwatko G, Borkiewicz P, Cnotkowski D, Pabiś-Orzeszyna M, Kobyliński P. Behavioural Biometrics in Virtual Reality: To What Extent Can We Identify a Person Based Solely on How They Watch 360-Degree Videos? 2022 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct) 2022. p. 417-22.
33. Bye K, Hosfelt D, Chase S, Miesnieks M, Beck T. The ethical and privacy implications of mixed reality. ACM SIGGRAPH 2019 Panels; Los Angeles, California: Association for Computing Machinery; 2019. p. Article 4.
34. Haas EHAd, Lee L-H. Deceiving Audio Design in Augmented Environments : A Systematic Review of Audio Effects in Augmented Reality. 2022 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct). 2022:36-43.
35. Madary M, Metzinger TK. Real Virtuality: A Code of Ethical Conduct. Recommendations for Good Scientific Practice and the Consumers of VR-Technology. Frontiers in Robotics and AI. 2016;3.
36. Jung J, Lee H, Choi J, Nanda A, Gruenefeld U, Stratmann T, et al. Ensuring Safety in Augmented Reality from Trade-off Between Immersion and Situation Awareness. 2018 IEEE International Symposium on Mixed and Augmented Reality (ISMAR); 16-20 Oct. 2018. p. 70-9.
37. Thomas D, Holmquist LE. Is Functionality All That Matters? Examining Everyday User Opinions of Augmented Reality Devices. 2021 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW); 27 March -1 April 2021. p. 232-7.
38. Lopes P, Tian N, Boulic R. Exploring Blink-Rate Behaviors for Cybersickness Detection in VR. 2020 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW); 22-26 March 2020. p. 794-5.
39. Vinson NG, Lapointe J-F, Parush A, Roberts S. Cybersickness induced by desktop virtual reality. Proceedings of Graphics Interface 2012; Toronto, Ontario, Canada: Canadian Information Processing Society; 2012. p. 69–75.
40. Souchet AD, Xie W, Lourdeaux D. Distinguishing Visual Fatigue, Mental Workload and Acute Stress in Immersive Virtual Reality with Physiological Data: pre-test results. 2022 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW); 12-16 March 2022. p. 720-1.
41. Zielasko D. Subject 001 - A Detailed Self-Report of Virtual Reality Induced Sickness. 2021 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW). 2021:165-8.
42. O'Hagan J, Gugenheimer J, Bonner J, Mathis F, McGill M. Augmenting People, Places & Media: The Societal Harms Posed by Everyday Augmented Reality, and the Case for Perceptual Human Rights. Proceedings of the 22nd International Conference on Mobile and Ubiquitous Multimedia; Vienna, Austria: Association for Computing Machinery; 2023. p. 225–35.
43. Yang Z, Li CY, Bhalla A, Zhao BY, Zheng H. Inception Attacks: Immersive Hijacking in Virtual Reality Systems. arXiv. 2024.
44. LaViola JJ. A discussion of cybersickness in virtual environments. SIGCHI Bull. 2000;32(1):47–56.

45. Pase S. Ethical considerations in augmented reality applications. Proceedings of the International Conference on e-learning, e-Business, Enterprise Information Systems, and e-Government. 2012.
46. Saredakis D, Szpak A, Birckhead B, Keage HAD, Rizzo A, Loetscher T. Factors Associated With Virtual Reality Sickness in Head-Mounted Displays: A Systematic Review and Meta-Analysis. Frontiers in Human Neuroscience. 2020;14.
47. Wang X, Lee L-H, Bermejo Fernandez C, Hui P. The Dark Side of Augmented Reality: Exploring Manipulative Designs in AR. International Journal of Human–Computer Interaction. 2023;40(13):3449-64.
48. Park S, Lee G. Full-immersion virtual reality: Adverse effects related to static balance. Neurosci Lett. 2020;733:134974.
49. Mittelstädt JM, Wacker J, Stelling D. VR aftereffect and the relation of cybersickness and cognitive performance. Virtual Reality. 2018;23:143 - 54.
50. Miller MR, Herrera F, Jun H, Landay JA, Bailenson JN. Personal identifiability of user tracking data during observation of 360-degree VR video. Sci Rep. 2020;10(1):17404.
51. Adams H, Narasimham G, Rieser J, Creem-Regehr S, Stefanucci J, Bodenheimer B. Locomotive Recalibration and Prism Adaptation of Children and Teens in Immersive Virtual Environments. IEEE Trans Vis Comput Graph. 2018;24(4):1408-17.
52. Abraham M, Saeghe P, Mcgill M, Khamis M. Implications of XR on Privacy, Security and Behaviour: Insights from Experts. Nordic Human-Computer Interaction Conference; Aarhus, Denmark: Association for Computing Machinery; 2022. p. Article 30.
53. Stanney K, Lawson BD, Rokers B, Dennison M, Fidopiastis C, Stoffregen T, et al. Identifying Causes of and Solutions for Cybersickness in Immersive Technology: Reformulation of a Research and Development Agenda. International Journal of Human–Computer Interaction. 2020;36(19):1783-803.
54. Nair V, Garrido GM, Song DX, O'Brien JF. Exploring the Privacy Risks of Adversarial VR Game Design. Proceedings on Privacy Enhancing Technologies. 2023;4:238-56.
55. Bozkir E, Özdel S, Lau KHC, Wang M, Gao H, Kasneci E. Embedding large language models into extended reality: Opportunities and challenges for inclusion, engagement, and privacy. Proceedings of the 6th ACM Conference on Conversational User Interfaces. 2024:1-7.
56. Selinger E, Altman E, Foster S. Eye-Tracking in Virtual Reality: A Visceral Notice Approach for Protecting Privacy. Privacy Studies Journal. 2023;2:1-34.
57. Sharples S, Cobb S, Moody A, Wilson JR. Virtual reality induced symptoms and effects (VRISE): Comparison of head mounted display (HMD), desktop and projection display systems. Displays. 2008;29(2):58-69.
58. Wen J, Ming K, Wang F, Huang B, Ma J. Cyber-I: Vision of the Individual's Counterpart on Cyberspace. 2009 Eighth IEEE International Conference on Dependable, Autonomic and Secure Computing; 12-14 Dec. 2009. p. 295-302.
59. Kruijff E, Swan JE, Feiner SK. Perceptual issues in augmented reality revisited. 2010 IEEE International Symposium on Mixed and Augmented Reality. 2010:3-12.
60. Heimo OI, Kimppa KK, Helle S, Korkalainen T, Lehtonen T. Augmented reality - Towards an ethical fantasy? 2014 IEEE International Symposium on Ethics in Science, Technology and Engineering; 23-24 May 2014. p. 1-7.
61. Denning T, Dehlawi Z, Kohno T. In situ with bystanders of augmented reality glasses: perspectives on recording and privacy-mediating technologies. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems; Toronto, Ontario, Canada: Association for Computing Machinery; 2014. p. 2377–86.
62. Roesner F, Kohno T, Molnar D. Security and privacy for augmented reality systems. Commun ACM. 2014;57(4):88–96.
63. Verbeek P-P. Designing the Public Sphere: Information Technologies and the Politics of Mediation. In: Floridi L, editor. The Onlife Manifesto: Being Human in a Hyperconnected Era. Cham: Springer International Publishing; 2015. p. 217-27.
64. Motti VG, Caine K. Users' Privacy Concerns About Wearables. In: Brenner M, Christin N, Johnson B, Rohloff K, editors. Financial Cryptography and Data Security; Berlin, Heidelberg: Springer Berlin Heidelberg; 2015. p. 231-44.
65. Franks MA. The Desert of the Unreal: Inequality in Virtual and Augmented Reality. UC Davis Law Review. 2017;51.

66. Adams D, Bah A, Barwulor C, Musabay N, Pitkin K, Redmiles EM. Ethics emerging: the story of privacy and security perceptions in virtual reality. Proceedings of the Fourteenth USENIX Conference on Usable Privacy and Security; Baltimore, MD, USA: USENIX Association; 2018. p. 443–58.
67. Lebeck K, Ruth K, Kohno T, Roesner F. Towards Security and Privacy for Multi-user Augmented Reality: Foundations with End Users. 2018 IEEE Symposium on Security and Privacy (SP). 2018:392-408.
68. Hill RW. Ethics of Immersive Technologies. In: Abbas AE, editor. Next-Generation Ethics: Engineering a Better Society. Cambridge: Cambridge University Press; 2019. p. 39-53.
69. Valluripally S, Gulhane A, Mitra R, Hoque KA, Calyam P. Attack Trees for Security and Privacy in Social Virtual Reality Learning Environments. 2020 IEEE 17th Annual Consumer Communications & Networking Conference (CCNC); Las Vegas, NV, USA: IEEE Press; 2020. p. 1–9.
70. Happa J, Glencross M, Steed A. Cyber Security Threats and Challenges in Collaborative Mixed-Reality. Frontiers in ICT. 2019;6.
71. Kudina O, Verbeek P-P. Ethics from Within: Google Glass, the Collingridge Dilemma, and the Mediated Value of Privacy. Science, Technology, & Human Values. 2019;44(2):291-314.
72. Gugenheimer J, McGill M, Huron S, Mai C, Williamson J, Nebeling M. Exploring Potentially Abusive Ethical, Social and Political Implications of Mixed Reality Research in HCI. Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems; Honolulu, HI, USA: Association for Computing Machinery; 2020. p. 1–8.
73. Maloney D, Zamanifard S, Freeman G. Anonymity vs. Familiarity: Self-Disclosure and Privacy in Social Virtual Reality. Proceedings of the 26th ACM Symposium on Virtual Reality Software and Technology; Virtual Event, Canada: Association for Computing Machinery; 2020. p. Article 25.
74. Kumarapeli D. [DC] Privacy in VR: Empowering Users with Emotional Privacy from Verbal and Non-verbal Behavior of Their Avatars. 2021 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW); 27 March-1 April 2021. p. 715-6.
75. Lu F. [DC] Glanceable AR: Towards an Always-on Augmented Reality Future. 2021 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW); 27 March -1 April 2021. p. 717-8.
76. Nakul E, Jeannin R, Lopez C. A new device to restore sensory congruency in virtual reality and to prevent cybersickness. 2021 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW); 27 March -1 April 2021. p. 156-60.
77. Dao E, Muresan A, Hornbæk K, Knibbe J. Bad Breakdowns, Useful Seams, and Face Slapping: Analysis of VR Fails on YouTube. Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems; Yokohama, Japan: Association for Computing Machinery; 2021. p. Article 526.
78. Guzman JAd, Seneviratne A, Thilakarathna K. Unravelling Spatial Privacy Risks of Mobile Mixed Reality Data. Proc ACM Interact Mob Wearable Ubiquitous Technol. 2021;5(1):1-26.
79. Bajorunaite L, Brewster S, Williamson JR. Virtual Reality in transit: how acceptable is VR use on public transport? 2021 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW) 2021. p. 432-3.
80. Dick E. Risks and Challenges for Inclusive and Equitable Immersive Experiences. Information Technology & Innovation Foundation [Internet]. 2021. Available from: https://www2.itif.org/2021-challenges-immersive.pdf.
81. Jung S, Li C, McKee R, Whitton MC, Lindeman RW. Floor-vibration VR: Mitigating Cybersickness Using Whole-body Tactile Stimuli in Highly Realistic Vehicle Driving Experiences. IEEE Transactions on Visualization and Computer Graphics. 2021;27(5):2669-80.
82. Giaretta A. Security and Privacy in Virtual Reality - A Literature Survey. ArXiv. 2022;abs/2205.00208.
83. Cheng K, Tian JF, Kohno T, Roesner F. Exploring user reactions and mental models towards perceptual manipulation attacks in mixed reality. Proceedings of the 32nd USENIX Conference on Security Symposium; Anaheim, CA, USA: USENIX Association; 2023. p. Article 52.
84. O'Hagan J, Saeghe P, Gugenheimer J, Medeiros D, Marky K, Khamis M, et al. Privacy-Enhancing Technology and Everyday Augmented Reality: Understanding Bystanders' Varying Needs for Awareness and Consent. Proc ACM Interact Mob Wearable Ubiquitous Technol. 2023;6(4):Article 177.
85. Gopal SRK, Shukla D, Wheelock JD, Saxena N. Hidden reality: caution, your hand gesture inputs in the immersive virtual world are visible to all! Proceedings of the 32nd USENIX Conference on Security Symposium; Anaheim, CA, USA: USENIX Association; 2023. p. Article 49.

86. Turner C. Augmented Reality, Augmented Epistemology, and the Real-World Web. Philosophy & Technology. 2022;35(1):19.
87. Gao B, Mai Z, Tu H, Duh HBL. Effects of Transfer Functions and Body Parts on Body-Centric Locomotion in Virtual Reality. IEEE Transactions on Visualization and Computer Graphics. 2023;29(8):3670-84.
88. Raisamo R, Rakkolainen I, Majaranta P, Salminen K, Rantala J, Farooq A. Human augmentation: Past, present and future. International Journal of Human-Computer Studies. 2019;131:131-43.
89. Rudenko A. Analysis of virtual communities in the context of human security. Man in India. 2017;97(23):73-83.
90. Adomaitis L, Grinbaum A, Lenzi D. TechEthos D2.2: Identification and specification of potential ethical issues and impacts and analysis of ethical issues of digital extended reality,neurotechnologies, and climate engineering. CEA Paris Saclay; 2022.
91. Aliman N-M, Kester L. Facing Immersive "Post-Truth" in AIVR? Philosophies. 2020;5(4):45.
92. Aliman NM, Kester L. VR, Deepfakes and Epistemic Security. 2022 IEEE International Conference on Artificial Intelligence and Virtual Reality (AIVR)2022. p. 93-8.
93. Bailenson JN, Beams B, Brown J, DeVeaux C, Han E, Queiroz ACM, et al. Seeing the World Through Digital Prisms: Psychological Implications of Passthrough Video Usage in Mixed Reality. Technology, Mind, and Behavior. 2024;2(2: Summer 2024).
94. Bonnail E, Tseng W-J, Mcgill M, Lecolinet E, Huron S, Gugenheimer J. Memory Manipulations in Extended Reality. Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems; Hamburg, Germany: Association for Computing Machinery; 2023. p. Article 875.
95. Buck LE, Bodenheimer B. Privacy and Personal Space: Addressing Interactions and Interaction Data as a Privacy Concern. 2021 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW). 2021:399-400.
96. Cook A. Paranovirus: Bridging Realities Using Digital Media Simulations. Proceedings of the Eleventh International Conference on Tangible, Embedded, and Embodied Interaction; Yokohama, Japan: Association for Computing Machinery; 2017. p. 733–8.
97. World Economic Forum. Metaverse Cybersecurity: Building Resilience in the Future Internet. 2024.
98. Falchuk B, Loeb S, Neff R. The Social Metaverse: Battle for Privacy. IEEE Technology and Society Magazine. 2018;37(2):52-61.
99. Ford PJ. A further analysis of the ethics of representation in virtual reality: Multi-user environments. Ethics and Information Technology. 2001;3(2):113-21.
100. Fox D, Thornton IG. The IEEE Global Initiative on Ethics of Extended Reality (XR) Report - Extended Reality (XR) Ethics and Diversity, Inclusion, and Accessibility. 2022:1-25.
101. Gruenefeld U, Stratmann TC, Jung J, Lee H, Choi J, Nanda A, et al. Guiding Smombies: Augmenting Peripheral Vision with Low-Cost Glasses to Shift the Attention of Smartphone Users. 2018 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct); 16-20 Oct. 2018. p. 127-31.
102. Han D-ID, Bergs Y, Moorhouse N. Virtual reality consumer experience escapes: preparing for the metaverse. Virtual Reality. 2022;26(4):1443-58.
103. Harborth D. Human Autonomy in the Era of Augmented Reality—A Roadmap for Future Work. Information. 2022;13(6):289.
104. Harborth D, Pape S. Investigating privacy concerns related to mobile augmented reality Apps – A vignette based online experiment. Computers in Human Behavior. 2021;122:106833.
105. Hosfelt DD. Making ethical decisions for the immersive web. ArXiv [Internet]. 2019; abs/1905.06995.
106. Luo S, Hu X, Yan Z. HoloLogger: Keystroke Inference on Mixed Reality Head Mounted Displays. 2022 IEEE Conference on Virtual Reality and 3D User Interfaces (VR); 12-16 March 2022. p. 445-54.
107. McGill M. The IEEE Global Initiative on Ethics of Extended Reality (XR) Report. Extended Reality (XR) and the Erosion of Anonymity and Privacy - White Paper. 2021:1-24.
108. Nakamura L. Feeling good about feeling bad: virtuous virtual reality and the automation of racial empathy. Journal of Visual Culture. 2020;19(1):47-64.
109. O'Hagan J, Williamson JR, McGill M, Khamis M. Safety, Power Imbalances, Ethics and Proxy Sex: Surveying In-The-Wild Interactions Between VR Users and Bystanders. 2021 IEEE International Symposium on Mixed and Augmented Reality (ISMAR); 4-8 Oct. 2021. p. 211-20.

110. Papagiannidis S, Bourlakis M, Li F. Making real money in virtual worlds: MMORPGs and emerging business opportunities, challenges and ethical implications in metaverses. Technological Forecasting and Social Change. 2008;75(5):610-22.
111. Royakkers L, Timmer J, Kool L, van Est R. Societal and ethical issues of digitization. Ethics and Information Technology. 2018;20(2):127-42.
112. Slater M, Gonzalez-Liencres C, Haggard P, Vinkers C, Gregory-Clarke R, Jelley S, et al. The Ethics of Realism in Virtual and Augmented Reality. Frontiers in Virtual Reality. 2020;1.
113. Spence EH. Meta Ethics for the Metaverse: The Ethics of Virtual Worlds. Proceedings of the 2008 conference on Current Issues in Computing and Philosophy: IOS Press; 2008. p. 3–12.
114. Stanney K, Fidopiastis C, Foster L. Virtual Reality Is Sexist: But It Does Not Have to Be. Front Robot AI. 2020;7:4.
115. Spiegel JS. The Ethics of Virtual Reality Technology: Social Hazards and Public Policy Recommendations. Sci Eng Ethics. 2018;24(5):1537-50.
116. Tseng W-J, Bonnail E, Mcgill M, Khamis M, Lecolinet É, Huron S, et al. The Dark Side of Perceptual Manipulations in Virtual Reality. Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems. 2022.
117. Wolfendale J. My avatar, my self: Virtual harm and attachment. Ethics and Information Technology. 2007;9(2):111-9.
118. Yao P, Lympouridis V, Zyda M. Virtual Equipment System: Face Mask and Voodoo Doll for User Privacy and Self-Expression Options in Virtual Reality. 2021 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW) 2021. p. 747-8.
119. Regenbrecht H, Zwanenburg S, Langlotz T. Pervasive Augmented Reality—Technology and Ethics. IEEE Pervasive Computing. 2022;21(3):84-91.
120. Jia J, Chen W. The Ethical Dilemmas of Virtual Reality Application in Entertainment. 2017 IEEE International Conference on Computational Science and Engineering (CSE) and IEEE International Conference on Embedded and Ubiquitous Computing (EUC) 2017. p. 696-9.
121. Pladere T, Svarverud E, Krumina G, Gilson SJ, Baraas RC. Inclusivity in stereoscopic XR: Human vision first. Frontiers in Virtual Reality. 2022;3.
122. Khlif M. Virtual Reality consistency with Common Concerns of Humanity: an overview. 2020 Seventh International Conference on Information Technology Trends (ITT); 25-26 Nov. 2020. p. 218-23.
123. Casey P, Baggili I, Yarramreddy A. Immersive Virtual Reality Attacks and the Human Joystick. IEEE Transactions on Dependable and Secure Computing. 2021;18(02):550-62.
124. Reger GM, Smolenski D, Edwards-Stewart A, Skopp NA, Rizzo AS, Norr A. Does Virtual Reality Increase Simulator Sickness During Exposure Therapy for Post-Traumatic Stress Disorder? Telemed J E Health. 2019;25(9):859-61.
125. Wiederhold BK, Riva G. Metaverse Creates New Opportunities in Healthcare. Annual Review of CyberTherapy and Telemedicine. 2022.
126. Bouchard S, St-Jacques J, Renaud P, Wiederhold BK. Side effects of immersions in virtual reality for people suffering from anxiety disorders. JCR. 2009;2(2):127.
127. Tyrrell R, Sarig-Bahat H, Williams K, Williams G, Treleaven J. Simulator sickness in patients with neck pain and vestibular pathology during virtual reality tasks. Virtual Reality. 2018;22(3):211-9.
128. Sarkar U, Lee JE, Nguyen KH, Lisker S, Lyles CR. Barriers and Facilitators to the Implementation of Virtual Reality as a Pain Management Modality in Academic, Community, and Safety-Net Settings: Qualitative Analysis. J Med Internet Res. 2021;23(9):e26623.
129. Botella C, Garcia-Palacios A, Baños RM, Quero S. Cybertherapy: Advantages, limitations, and ethical issues. PsychNology Journal. 2009;7(1):77-100.
130. Rose V, Stewart I, Jenkins K, Ang CS, Matsangidou M. A scoping review exploring the feasibility of virtual reality technology use with individuals living with dementia. International Conference on Artificial Reality and Telexistence Eurographics Symposium on Virtual Environments 2018.
131. Kellmeyer P, Biller-Andorno N, Meynen G. Ethical tensions of virtual reality treatment in vulnerable patients. Nature Medicine. 2019;25(8):1185-8.
132. Szpak A, Michalski SC, Saredakis D, Chen CS, Loetscher T. Beyond Feeling Sick: The Visual and Cognitive Aftereffects of Virtual Reality. IEEE Access. 2019;7:130883-92.

133. Felix RB, Rao A, Khalid M, Wang Y, Colloca L, Murthi SB, et al. Adjunctive virtual reality pain relief following traumatic injury: protocol for a randomised within-subjects clinical trial. BMJ Open. 2021;11(11):e056030.
134. Ara J, Karim FB, Alsubaie MSA, Bhuiyan YA, Bhuiyan MI, Bhyan SB, et al. Comprehensive Analysis of Augmented Reality Technology in Modern Healthcare System. International Journal of Advanced Computer Science and Applications. 2021;12(6):840-9.
135. Schmidt M, Newbutt N, Schmidt C, Glaser N. A Process-Model for Minimizing Adverse Effects when Using Head Mounted Display-Based Virtual Reality for Individuals with Autism. Frontiers in Virtual Reality. 2021;2.
136. Nason EE, Trahan M, Cash D. Evaluation of a Mobile Virtual Reality Intervention for Social Anxiety Disorder: Ethical and Methodological Lessons Learned. Journal of Technology in Behavioral Science. 2023;8(1):79-86.
137. Jicol C, Feltham J, Yoon J, Proulx MJ, O'Neill E, Lutteroth C. Designing and Assessing a Virtual Reality Simulation to Build Resilience to Street Harassment. Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems New Orleans, LA, USA: Association for Computing Machinery; 2022. p. Article 117.
138. Lewis-Evans B. A short guide to user testing for simulation sickness in Virtual Reality. In: Drachen A, Mirza-Babaei P, Nacke L, editors. Games User Research: Oxford University Press; 2018.
139. Evans J. The IEEE Global Initiative on Ethics of Extended Reality (XR) Report - Extended Reality (XR) Ethics in Medicine. 2022:1-31.
140. Fromberger P, Jordan K, Müller JL. Virtual reality applications for diagnosis, risk assessment and therapy of child abusers. Behav Sci Law. 2018;36(2):235-44.
141. Kaimara P, Oikonomou A, Deliyannis I. Could virtual reality applications pose real risks to children and adolescents? A systematic review of ethical issues and concerns. Virtual Reality. 2022;26(2):697-735.
142. Kellmeyer P. Neurophilosophical and Ethical Aspects of Virtual Reality Therapy in Neurology and Psychiatry. Camb Q Healthc Ethics. 2018;27(4):610-27.
143. Kourtesis P, MacPherson SE. How immersive virtual reality methods may meet the criteria of the National Academy of Neuropsychology and American Academy of Clinical Neuropsychology: A software review of the Virtual Reality Everyday Assessment Lab (VR-EAL). Computers in Human Behavior Reports. 2021;4:100151.
144. Ligthart S, Meynen G, Biller-Andorno N, Kooijmans T, Kellmeyer P. Is Virtually Everything Possible? The Relevance of Ethics and Human Rights for Introducing Extended Reality in Forensic Psychiatry. AJOB Neurosci. 2022;13(3):144-57.
145. Marloth M, Chandler J, Vogeley K. Psychiatric Interventions in Virtual Reality: Why We Need an Ethical Framework. Camb Q Healthc Ethics. 2020;29(4):574-84.
146. Schmidt M, Newbutt N. Towards a Conceptual Model for Consideration of Adverse Effects of Immersive Virtual Reality for Individuals with Autism. In: Soares MM, Rosenzweig E, Marcus A, editors. Design, User Experience, and Usability: Design for Diversity, Well-being, and Social Development. Cham: Springer International Publishing; 2021. p. 333-42.
147. Washburn M, Hagedorn A, Moore S. Creating Virtual Reality Based Interventions for Older Adults Impacted by Substance Misuse: Safety and Design Considerations. Journal of Technology in Human Services. 2021;39(3):275-94.
148. Waycott J, Kelly RM, Baker S, Neves BB, Thach KS, Lederman R. The Role of Staff in Facilitating Immersive Virtual Reality for Enrichment in Aged Care: An Ethic of Care Perspective. Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems; New Orleans, LA, USA: Association for Computing Machinery; 2022. p. Article 414.
149. Botero M, Whatley E. Harm, Consent, and Virtual Selves in Full-Body Ownership Illusions: Real Concerns for Immersive Virtual Reality Therapies. Camb Q Healthc Ethics. 2020;29(4):585-91.
150. Gorrindo T, Groves JE. Crime and hate in virtual worlds: a new playground for the id? Harv Rev Psychiatry. 2010;18(2):113-8.
151. Alomar N, Alsaleh M, Alarifi A. Behavioral consequences of Pokémon GO: The exaggerated picture. Computers in Human Behavior. 2019;90:223-45.

152. Joseph B, Armstrong DG. Potential perils of peri-Pokémon perambulation: the dark reality of augmented reality? Oxf Med Case Reports. 2016;2016(10):omw080.
153. Pourchon R, Léger P-M, Labonté-LeMoyne É, Sénécal S, Bellavance F, Fredette M, et al. Is Augmented Reality Leading to More Risky Behaviors? An Experiment with Pokémon Go. In: Nah FF-H, Tan C-H, editors. HCI in Business, Government and Organizations Interacting with Information Systems. Cham: Springer International Publishing; 2017. p. 354-61.
154. Kade D. Ethics of Virtual Reality Applications in Computer Game Production. Philosophies. 2016;1(1):73-86.
155. Ayers JW, Leas EC, Dredze M, Allem JP, Grabowski JG, Hill L. Pokémon GO-A New Distraction for Drivers and Pedestrians. JAMA Intern Med. 2016;176(12):1865-6.
156. Barbieri S, Vettore G, Pietrantonio V, Snenghi R, Tredese A, Bergamini M, et al. Pedestrian Inattention Blindness While Playing Pokémon Go as an Emerging Health-Risk Behavior: A Case Report. J Med Internet Res. 2017;19(4):e86.
157. Judge EF, Brown TE. Pokémorials: Placing Norms in Augmented Reality University of British Columbia Law Review. 2017;50:46.
158. Heidrich D, Oberdörfer S, Latoschik ME. The Effects of Immersion on Harm-inducing Factors in Virtual Slot Machines. 2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR); 23-27 March 20192019. p. 793-801.
159. Yildirim C. Cybersickness during VR gaming undermines game enjoyment: A mediation model. Displays. 2019;59:35-43.
160. Lavoie R, Main K, King C, King D. Virtual experience, real consequences: the potential negative emotional consequences of virtual reality gameplay. Virtual Reality. 2020;25:69-81.
161. Colley A, Thebault-Spieker J, Lin AY, Degraen D, Fischman B, Häkkilä J, et al. The Geography of Pokémon GO: Beneficial and Problematic Effects on Places and Movement. Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems. 2017.
162. Huff C, Johnson DG, Miller K. Virtual harms and real responsibility. IEEE Technology and Society Magazine. 2003;22(2):12-9.
163. Millard DE, Hewitt S, O'Hara K, Packer H, Rogers N. The Unethical Future of Mixed Reality Storytelling. Proceedings of the 8th International Workshop on Narrative and Hypertext; Hof, Germany: Association for Computing Machinery; 2019. p. 5–8.
164. Mukhra R, Baryah N, Krishan K, Kanchan T. "Pokémon Go" and Ethical Considerations-Are Pokémons Poking Us Enough? Journal of Information Ethics. 2019;28(2):117-24.
165. Oberdörfer S, Heidrich D, Latoschik ME. Think Twice: The Influence of Immersion on Decision Making during Gambling in Virtual Reality. 2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR) 2020. p. 483-92.
166. Wang J, Shi R, Xiao Z, Qin X, Liang HN. Resolution Tradeoff in Gameplay Experience, Performance, and Simulator Sickness in Virtual Reality Games. 2022 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW); 12-16 March 2022. p. 542-3.
167. Deng X, Ruan J. Users' privacy in the Second Life Library. 2009 IEEE International Symposium on IT in Medicine & Education. 2009;1:337-40.
168. Makhmutov M, Asapov T, Brown J, Alexander. Safety Risks in Location-Based Augmented Reality Games. 20th International Conference on Entertainment Computing (ICEC); Coimbra, Portugal: Springer International Publishing; 2021. p. 457-64.
169. Blackwell L, Ellison N, Elliott-Deflo N, Schwartz R. Harassment in Social VR: Implications for Design. 2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR). 2019:854-5.
170. Fiani C, Bretin R, Mcgill M, Khamis M. Big Buddy: Exploring Child Reactions and Parental Perceptions towards a Simulated Embodied Moderating System for Social Virtual Reality. Proceedings of the 22nd Annual ACM Interaction Design and Children Conference; Chicago, IL, USA: Association for Computing Machinery; 2023. p. 1–13.
171. Maloney D, Freeman G, Robb A. It Is Complicated: Interacting with Children in Social Virtual Reality. 2020 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW); 22-26 March 2020. p. 343-7.

172. Dudley A, Braman J, Wang Y, Vincenti G, Tupper D. Security, Legal, and Ethical Implications of Using Virtual Worlds. 2010. Available from: http://www.iiis.org/CDs2010/CD2010SCI/SCI_2010/PapersPdf/SA049JX.pdf.
173. Warren I, Palmer D. Crime risks of three-dimensional virtual environments. Trends and issues in crime and criminal justice. 2010(388):1-6.
174. Blackwell L, Ellison N, Elliott-Deflo N, Schwartz R. Harassment in Social Virtual Reality: Challenges for Platform Governance. Proceedings of the ACM on Human-Computer Interaction. 2019;3(CSCW).
175. Freeman G, Zamanifard S, Maloney D, Acena D. Disturbing the Peace: Experiencing and Mitigating Emerging Harassment in Social Virtual Reality. Proc ACM Hum-Comput Interact. 2022;6(CSCW1):1-30.
176. Vondráček M, Baggili I, Casey P, Mekni M. Rise of the Metaverse’s Immersive Virtual Reality Malware and the Man-in-the-Room Attack & Defenses. Computers & Security. 2023;127:102923.
177. World Economic Forum. Metaverse Identity: Defining the Self in a Blended Reality. Switzerland; 2024.
178. O’Brolcháin F, Jacquemard T, Monaghan D, O’Connor N, Novitzky P, Gordijn B. The Convergence of Virtual Reality and Social Networks: Threats to Privacy and Autonomy. Science and Engineering Ethics. 2016;22(1):1-29.
179. Outlaw J, Duckles B. Why women don’t like social virtual reality: A study of safety, usability, and self-expression in social VR 2017 [cited 2024 October].
180. Shriram K, Schwartz R. All are welcome: Using VR ethnography to explore harassment behavior in immersive social virtual reality. 2017 IEEE Virtual Reality (VR). 2017:225-6.
181. Cortese M, Outlaw J. The IEEE Global Initiative on Ethics of Extended Reality (XR) Report--Social and Multi-User Spaces in VR: Trolling, Harassment, and Online Safety. The IEEE Global Initiative on Ethics of Extended Reality (XR) Report--Social and Multi-User Spaces in VR: Trolling, Harassment, and Online Safety. 2021:1-17.
182. Hirskyj-Douglas I, Kantosalo A, Monroy-Hernández A, Zimmermann J, Nebeling M, Gonzalez-Franco M. Social AR: Reimagining and Interrogating the Role of Augmented Reality in Face to Face Social Interactions. Companion Publication of the 2020 Conference on Computer Supported Cooperative Work and Social Computing; Virtual Event, USA: Association for Computing Machinery; 2020. p. 457–65.
183. Maloney D, Freeman G, Robb AC. Social Virtual Reality: Ethical Considerations and Future Directions for An Emerging Research Space. 2021 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW). 2021:271-7.
184. Maslen H, Savulescu J. The ethics of virtual reality and telepresence. In: Prescott TJ, Lepora N, Verschure PFMJ, editors. Living machines: A handbook of research in biomimetics and biohybrid systems: Oxford University Press; 2018. p. 587-95.
185. Zhong J, Zheng Y. Empowering Future Education: Learning in the Edu-Metaverse. 2022 International Symposium on Educational Technology (ISET) 2022. p. 292-5.
186. Rauschenberger R, Barakat B. Health and Safety of VR Use by Children in an Educational Use Case. 2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR); 22-26 March 2020. p. 878-84.
187. Kye B, Han N, Kim E, Park Y, Jo S. Educational applications of metaverse: possibilities and limitations. J Educ Eval Health Prof. 2021;18:32.
188. Yamada-Rice D, Mushtaq F, Woodgate A, Bosmans D, Douthwaite A, Douthwaite I, et al. Children and Virtual Reality: Emerging Possibilities and Challenges. 2017.
189. Iserson KV. Ethics of Virtual Reality in Medical Education and Licensure. Camb Q Healthc Ethics. 2018;27(2):326-32.
190. Gulhane A, Vyas A, Mitra R, Oruche R, Hoefer G, Valluripally S, et al. Security, Privacy and Safety Risk Assessment for Virtual Reality Learning Environment Applications. 2019 16th IEEE Annual Consumer Communications & Networking Conference (CCNC). 2019:1-9.
191. Li Y, Wei W, Xu J. The Exploration on Ethical Problems of Educational Metaverse. In: Zhang L-J, editor. Metaverse – METAVERSE 2022; Cham: Springer Nature Switzerland; 2022. p. 29-38.
192. Carter M, Egliston B. What are the risks of Virtual Reality data? Learning Analytics, Algorithmic Bias and a Fantasy of Perfect Data. New Media & Society. 2023;25(3):485-504.
193. Lee JJ, Hu-Au E. E3XR: An Analytical Framework for Ethical, Educational and Eudaimonic XR Design. Frontiers in Virtual Reality. 2021;2.

194. Flowers BA, Rebensky S. Are you Seeing what I'm Seeing?: Perceptual Issues with Digital Twins in Virtual Reality. 2022 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW). 2022:126-30.
195. Simon-Liedtke JT, Baraas R. Towards eXtended Universal Design. Stud Health Technol Inform. 2022;297:391-9.
196. Southgate E, Smith SP. Designing and conducting research using immersive technologies in schools: Seven observations. 2017 IEEE Virtual Reality Workshop on K-12 Embodied Learning through Virtual & Augmented Reality (KELVAR) 2017. p. 1-3.
197. Almeida A, Rebelo F, Noriega P, Vilar E. Virtual Reality Self Induced Cybersickness: An Exploratory Study. In: Rebelo F, Soares M, editors. Advances in Ergonomics in Design; Cham: Springer International Publishing; 2018. p. 26-33.
198. Southgate E, Smith SP, Scevak J. Asking ethical questions in research using immersive virtual and augmented reality technologies with children and youth. 2017 IEEE Virtual Reality (VR); 18-22 March 2017. p. 12-8.
199. Kim A, Darakjian N, Finley JM. Walking in fully immersive virtual environments: an evaluation of potential adverse effects in older adults and individuals with Parkinson's disease. J Neuroeng Rehabil. 2017;14(1):16.
200. Brown JA, Dinh AT, Oh C. Safety and Ethical Considerations When Designing a Virtual Reality Study with Older Adult Participants. In: Gao Q, Zhou J, editors. Human Aspects of IT for the Aged Population Design, Interaction and Technology Acceptance. Cham: Springer International Publishing; 2022. p. 12-26.
201. Peck TC, Sockol LE, Hancock SM. Mind the Gap: The Underrepresentation of Female Participants and Authors in Virtual Reality Research. IEEE Transactions on Visualization and Computer Graphics. 2020;26(5):1945-54.
202. Ramirez EJ. Ecological and ethical issues in virtual reality research: A call for increased scrutiny. Philosophical Psychology. 2019;32(2):211-33.
203. Ramirez EJ, LaBarge S. Real moral problems in the use of virtual reality. Ethics and Information Technology. 2018;20(4):249-63.
204. Behr K-M, Nosper A, Klimmt C, Hartmann T. Some practical considerations of ethical issues in VR research. Presence: Teleoper Virtual Environ. 2005;14(6):668–76.
205. Rizzo A, Schulteis MT, Rothbaum BO. Ethical issues for the use of virtual reality in the psychological sciences. In: Bush S, Drexler M, editors. Ethical Issues in Clinical Neuropsychology. Lisse, Netherlands: Swets & Zeitlinger Publishers; 2003. p. 243-77.
206. Nelson JW. A Virtual Property Solution: How Privacy Law Can Protect the Citizens of Virtual Worlds. 2010.
207. Lake J. Hey, You Stole My Avatar!: Virtual Reality and Its Risks to Identity Protection. Emory Law Journal. 2019;69(4):833.
208. Fernandez CB, Hui P. Life, the Metaverse and Everything: An Overview of Privacy, Ethics, and Governance in Metaverse. 2022 IEEE 42nd International Conference on Distributed Computing Systems Workshops (ICDCSW). 2022:272-7.
209. Kasiyanto S, Kilinc MR. The Legal Conundrums Of The Metaverse. Journal of Central Banking Law and Institutions. 2022;1(2):299-322.
210. Interpol. Metaverse. A law enforcement perspective. Use Cases, Crime, Forensics, Investigation, and Governance. 2024.
211. Cheong BC. Avatars in the metaverse: potential legal issues and remedies. International Cybersecurity Law Review. 2022;3(2):467-94.
212. Leenes R. Privacy in the Metaverse. In: Fischer-Hübner S, Duquenoy P, Zuccato A, Martucci L, editors. The Future of Identity in the Information Society; Boston, MA: Springer US; 2008. p. 95-112.
213. Lin J, Latoschik ME. Digital body, identity and privacy in social virtual reality: A systematic review. Frontiers in Virtual Reality. 2022;3.
214. Lee CY. Understanding Security Threats in Virtual Worlds. Americas Conference on Information Systems (AMCIS) 2009 Proceedings 2009.
215. Johnson DG. Promises and perils in immersive journalism. In: Turo Uskali, Astrid Gynnild, Sarah Jones, Sirkkunen E, editors. Immersive Journalism as Storytelling. London: Routledge; 2021. p. 71.

216. McIntosh V. Dialing up the danger: Virtual reality for the simulation of risk. Frontiers in Virtual Reality. 2022;3.
217. Nash K. Virtual reality witness: exploring the ethics of mediated presence. Studies in Documentary Film. 2018;12(2):119-31.
218. Pavlik JV. The moral mandate of virtual reality journalism. The Routledge Companion to Journalism Ethics: Taylor and Francis; 2021. p. 337--45.
219. Aliman NM, Kester L. Malicious Design in AIVR, Falsehood and Cybersecurity-oriented Immersive Defenses. 2020 IEEE International Conference on Artificial Intelligence and Virtual Reality (AIVR). 2020:130-7.
220. Snijders D, Horsman S, Kool L, van Est R. Responsible VR. Protect consumers in virtual reality. Den Haag; 2020.
221. Mhaidli AH, Schaub F. Identifying Manipulative Advertising Techniques in XR Through Scenario Construction. Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems; Yokohama, Japan: Association for Computing Machinery; 2021. p. Article 296.
222. Bibri SE. The Social Shaping of the Metaverse as an Alternative to the Imaginaries of Data-Driven Smart Cities: A Study in Science, Technology, and Society. Smart Cities. 2022;5(3):832-74.
223. Bibri SE, Allam Z, Krogstie J. The Metaverse as a virtual form of data-driven smart urbanism: platformization and its underlying processes, institutional dimensions, and disruptive impacts. Computational Urban Science. 2022;2(1):24.
224. Lagorio A, Pasquale VD, Cimini C, Miranda S, Pinto R. Augmented Reality in Logistics 4.0: implications for the human work. IFAC PapersOnLine. 2022;55(10):329-34.
225. World Health Organization (WHO). Constitution of the World Health Organization. Geneva: World Health Organization (WHO); 1946.
226. Christman J. Autonomy in Moral and Political Philosophy. 2020. In: The {Stanford} Encyclopedia of Philosophy [Internet]. Metaphysics Research Lab, Stanford University. Available from: https://plato.stanford.edu/archives/fall2020/entries/autonomy-moral/.
227. Mokrosinska D. Privacy and autonomy: On some misconceptions concerning the political dimensions of privacy. Law and Philosophy. 2018;37(2):117-43.
228. Lyon D. Surveillance. Internet Policy Review [Internet]. 2022; 11(4). Available from: https://policyreview.info/concepts/surveillance.
229. Varga S, Guignon C. Authenticity. 2023. In: The {Stanford} Encyclopedia of Philosophy [Internet]. Metaphysics Research Lab, Stanford University. Available from: https://plato.stanford.edu/archives/sum2023/entries/authenticity/.
230. Cambridge University Press. Epistemic n.d. [cited 2024 October]. Available from: https://dictionary.cambridge.org/dictionary/english/epistemic.
231. Berrick D, Spivack J. Understanding eXtended Reality Technology & Data Flows: Privacy and Data Protection Risks and Mitigation Strategies 2022 [cited 2024 October]. Available from: https://fpf.org/blog/understanding-extended-reality-technology-data-flows-privacy-and-data-protection-risks-and-mitigation-strategies/.
232. Thomson JJ. Normativity. Shafer-Landau R, editor: Open Court; 2008. 713-5 p.
233. European Union. Dual-use export controls. Summary of Regulation (EU) 2021/821. EUR-Lex. 2021 [cited 2024 October]. Available from: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=legissum:4532505.
234. Royal Netherlands Academy of Arts and Sciences. A Code of Conduct for Biosecurity. Report by the Biosecurity Working Group. Amsterdam. 2008 [cited 2024 October]. Available from: https://www.bureaubiosecurity.nl/documenten/knaw-code-of-conduct.
235. Grand View Research. Extended Reality Market Size, Share & Growth Report 2024 [cited 2024 October]. Available from: https://www.grandviewresearch.com/industry-analysis/extended-reality-xr-market-report.
236. Precedence research. Extended Reality Market Size, Share, and Trends 2024 to 2034 2022 [updated 2022 November; cited 2024 October]. Available from: https://www.precedenceresearch.com/extended-reality-market.
237. Meta Horizon. Meta Horizon n.d. [cited 2024 October]. Available from: https://horizon.meta.com.
238. Hamad A, Jia B. How Virtual Reality Technology Has Changed Our Lives: An Overview of the Current and Potential Applications and Limitations. Int J Environ Res Public Health. 2022;19(18).

239. Greengard S. Virtual Reality: The MIT Press; 2019.
240. Counterpoint. Global XR (AR & VR Headsets) Market Share: Quarterly 2024 [cited 2024 October]. Available from: https://www.counterpointresearch.com/insight/global-xr-ar-vr-headsets-market-share-quarterly.
241. Egliston B, Carter M, Clark KE. Who will govern the metaverse? Examining governance initiatives for extended reality (XR) technologies. New Media & Society. 2024;0(0):1-21.
242. Statista Research Department. Number of daily Pokémon GO users in the United States as of August 2016. Statista. 2016 [cited 2024 October]. Available from: https://www.statista.com/statistics/589213/pokemon-go-user-number-us/.
243. X Reality Safety Intelligence (XRSI). Who we are [cited 2024 October]. Available from: https://xrsi.org/who-we-are.
244. Hayes A. Cyborg Cops, Googlers, and Connectivism [Leading Edge]. IEEE Technology and Society Magazine. 2013;32(1):23-4.
245. Girginova K, Cassidy K, Foxman M, O'Donnell M, Rawson K. Introduction and Executive Summary. Social Grammars of Virtuality, no 1 [Internet]. 2023; (1). Available from: https://penn.manifoldapp.org/projects/social-grammars-of-virtuality-no-1.
246. Hudon D. The slow rise of Mixed Reality 2023 [cited 2024 October]. Available from: https://medium.com/stellarx-en/the-slow-rise-of-mixed-reality-c6e3ef5ad3dd.
247. Apple. Apple Vision Pro 2023 [cited 2024 October]. Available from: https://www.apple.com/apple-vision-pro/.
248. Rosenberg L. Mixed Reality or Spatial Computing. 2024 [cited 2024 October]. Available from: https://medium.com/predict/mixed-reality-or-spatial-computing-346e62148026.
249. Hirzle T, Müller F, Draxler F, Schmitz M, Knierim P, Hornbæk K. When XR and AI Meet - A Scoping Review on Extended Reality and Artificial Intelligence. Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems; Hamburg, Germany: Association for Computing Machinery; 2023. p. Article 730.
250. Lo TTS, Chen Y, Lai TY, Goodman A. Phygital workspace: a systematic review in developing a new typological work environment using XR technology to reduce the carbon footprint. Frontiers in Built Environment. 2024;10.
251. Reis CM, Câmara A. Expanding Nature's storytelling: extended reality and debiasing strategies for an eco-agency. Frontiers in Psychology. 2023;14.
252. Noman B, Donti P, Cuff J, Sroka S, Ilic M, Sze V, et al. The Climate and Sustainability Implications of Generative AI. An MIT Exploration of Generative AI. 2024.
253. Huang J, Tong S, Singh P, Shi M, Liu C. White Paper on Global Artificial Intelligence Environmental Impact. AI and Sustainability. 2024;1(1).
254. World Economic Forum. Even though it's virtual, the metaverse does actually impact the environment. Nature and Biodiversity [Internet]. 2022. Available from: https://www.weforum.org/agenda/2022/02/how-metaverse-actually-impacts-the-environment/.
255. Veale M. A Critical Take on the Policy Recommendations of the EU High-Level Expert Group on Artificial Intelligence. European Journal of Risk Regulation. 2020;11(1):e1.

7. ACKNOWLEDGEMENTS

7.1 Authors roles and contribution

1st author - Cristiana Anca Voinov, PhD (candidate), University of South-Eastern Norway

Screening, methodology, full text reading, data charting, thematic analysis, original draft preparation, writing, revising and editing of report.

2nd author – Lene A. Hagen, PhD, Researcher, University of South-Eastern Norway

Screening, full text reading, data charting, thematic analysis, visualisation, original draft preparation, revising and editing of report.

3rd author -Rigmor C. Baraas, PhD, Project Coordinator, University of South-Eastern Norway

Funding acquisition, project administration, conceptualization, screening.

4th author - Rosemarie DLC Bernabe, PhD, Project coordinator, University of Oslo

Funding acquisition, project administration, work package supervision, conceptualization, screening, and report submission.

5th author- Shereen Cox, PhD, Researcher, University of Oslo

Screening, methodology, full text reading, data charting, thematic analysis, original draft preparation, visualization, writing, revising and editing of report.

7.2 Non-author contribution

Vivian Mbanya, researcher, and Hilde Iren Flaatten, librarian, University of Oslo, for development of search string, database search, and deduplication.

Alina Kadlubsky, Open AR Communications - visualisation

Members of Work package 2 – conceptualisation and search terms development

Part 2: Risks and Harms according to European Citizens: a Review of the European Citizens' Panel on Virtual Worlds Final Report

1. Introduction

The European Citizens' Panel on Virtual Worlds was convened by the European Commission, with three sessions taking place between February 24th and April 23rd, 2023. The panel aimed to gather citizen input on the future of virtual worlds (also known as the metaverse) and to formulate recommendations for the European Commission to consider in its policymaking. The panel was part of a wider initiative by the European Commission to engage citizens in shaping key policy areas.

Up to 150 citizens from all 27 EU Member States were randomly selected to participate in the panel. Of the total invited participants, 140 citizens participated in at least one of the three sessions of the European Citizens' Panel on Virtual Worlds. The selection process, carried out by Kantar Public, aimed to create a panel that reflected the diversity of the EU population, with quotas for gender, age, educational level, and urban/rural residence. Notably, the panel included an over-representation of young people (aged 16-25) to ensure their voices were adequately heard.

The panel focused on the question: "What vision, principles, and actions should guide the development of desirable and fair virtual worlds?" The goal was to identify and prioritize key issues related to the development of virtual worlds and to formulate recommendations to guide EU policy in this area. The European Commission defines "virtual worlds" as "Persistent, 3D, real-time, immersive environments, blurring the line between real and virtual, for socialising, working, learning, making transactions, playing and creating."[1] How virtual worlds differ from XR (VR, AR, MR) was not clearly specified in the Virtual Worlds documents; however, virtual worlds are related to XR in the sense that XR technologies may be used to create virtual worlds.[1–3] As such, and in this manner, the findings of the Virtual Worlds European Citizens' Panel consultation are relevant to the goals of D2.2.

This document has aims to achieve the following: the identification of risks and harms according to European Citizens; and to demonstrate convergences and divergences between Parts 1 and 2 of D2.2. Note that though the panel was not primarily concerned with the identification of risks and harms of immersive technologies, by looking at the participants' concerns, we were able to identify what the risks, harms, and burdens are that European citizens who participated in this panel associate with immersive technologies.

2. Methodology

A literature review approach from a Grounded Theory Perspective elaborated by Hennie Boeije[4] was used to process the document, *European Citizens' Panel on Virtual Worlds Final Report*.[5] Using NVivo and adapting the codes from Part 1 of this report, *Mapping of risks and harms of eXtended reality (XR) technologies: a scoping review*, RdlCB and RB individually coded the document, European *Citizens' Panel on Virtual Worlds Final Report*.[5] Disagreements were discussed to achieve resolution. To identify similarities and differences, the major codes from the latter were compared with those of the former.

The major codes from *Mapping of risks and harms*, which were adapted in this study, were as follows:

1. Health and Wellbeing Risks: Cognitive; Psycho-emotional; Socio-relational; Physical

2. Autonomy Risks: Privacy; Self-sovereignty; Surveillance; Authenticity

3. Epistemic and Scientific Risks: Epistemic Uncertainty; Scientific Uncertainty; Algorithmic Imperfection; Manipulation

4. Normative Risks: Norm Violations (Moral and Conventional); Rights Violations

5. Structural Risks: Socio-political Risks; Economic/Financial Risks; Legal Risks

6. Technical Safety and Security: Security Risks; Dual-use Risks

RCB and RB created a new code or sub-code when the need arose.

3. Results

3.1. Concerns of Risks and Harms according to European Citizens

Results show that the concerns of the European citizens cover all six major codes, though divergences were noted on the level of the subcodes.

3.1.1. Health and Wellbeing Risks

Citizens expressed concerns about the potential negative health impacts of virtual worlds, especially on mental and physical health. They advocated for research on the impact of virtual worlds on health, suggesting the creation of indicators to measure these effects. The citizens highlighted concerns about potential addiction to virtual worlds. This reflects the "Addiction" risk identified in the scoping review, which notes the immersive and potentially addictive nature of XR technologies.

Concerns related to mental health were frequently raised. Citizens emphasized the need to address negative psychological impacts, such as anxiety and social isolation. They advocated for the protection of vulnerable groups from manipulation and threats within virtual worlds.

This aligns with the scoping review's emphasis on the psycho-emotional risks of XR, including anxiety, detachment, and potential manipulation.

3.1.2. Autonomy Risks

Data privacy was a paramount concern for the citizens. They called for transparent data usage policies, user-friendly consent mechanisms, and stricter regulations on data collection and usage by companies.

This reflects the scoping review's discussion of privacy risks in XR, particularly concerning the collection of biometric data and the potential for data misuse.

Citizens recommended the development of regulations on digital identity, with an emphasis on security and anonymity. They also discussed the importance of user control over their virtual identities and data.

These concerns echo the scoping review's "self-sovereignty" risk, which addresses the potential for compromised consent and manipulation within virtual environments.

The citizens expressed concerns about potential surveillance within virtual worlds, advocating for police presence and the responsible use of AI in policing.

This relates to the scoping review's discussion of surveillance risks, both from governmental and corporate entities, enabled by the enhanced data capture capabilities of XR technologies.

3.1.3. Normative Risks

The citizens repeatedly stressed the importance of EU values and ethical principles in the development of virtual worlds. They called for ethical guidelines for behaviour in virtual worlds, addressing issues like misinformation and environmental responsibility.

This connects with the scoping review's broader discussion of "norm violations" within XR, including the potential for crossing ethical boundaries and engaging in harmful behaviour.

Citizens strongly advocated for inclusivity and accessibility in virtual worlds, ensuring equal access for all citizens regardless of their background or abilities. They highlighted the risk of replicating real-world inequalities in virtual environments.

This resonates with the scoping review's focus on "rights violations" within XR, including accessibility issues, discrimination, and the perpetuation of real-world biases.

3.1.4. Structural Risks

The citizens addressed the potential impacts of virtual worlds on the labour market, calling for legislative adjustments, harmonized training programs, and the protection of workers' rights within virtual workspaces.

These concerns reflect the scoping review's analysis of "economic/financial risks" associated with XR, including unemployment, gentrification, and the potential for exacerbating existing inequalities.

3.1.5. Technical Safety and Security Risks

The need for a secure and reliable digital infrastructure for virtual worlds was highlighted. Citizens called for EU labels and certifications to ensure the security of virtual world applications.

This aligns with the scoping review's identification of various security risks in XR, including hacking, data leaks, and the potential for malicious attacks.

The citizens repeatedly emphasized the need for regulation and oversight of virtual worlds to mitigate potential harms. They called for collaboration between governments, industry, and researchers to establish standards and regulations, highlighting the EU's potential role as a global leader in this area.

This resonates with the scoping review's finding that many sources identify the lack of clear governance and regulatory frameworks as a key risk in the development and use of XR.

3.2. Divergences between Part 1 (scoping review) and Part 2 (Virtual Worlds consultation)

Overall, the citizens' concerns closely align with the risks and harms identified in the scoping review, demonstrating a shared understanding of the potential challenges posed by virtual world technologies. There were, however, divergences, outlined below:

Sustainability: Citizens emphasized the need for environmentally sustainable virtual world development. They called for actions to address the environmental footprint of virtual worlds and promote the use of green energy in their development. This concern is only briefly mentioned in the scoping review, as there was limited discussion about this aspect in the reviewed material.

Epistemic and Scientific Risks: Utilization of AI was recommended as potentially useful for supporting police presence in virtual worlds. Citizens emphasized that AI for this use would have to be built on sound ethical foundations. Associated risks related to algorithmic imperfection and epistemic uncertainty were not something citizens appeared to be concerned about. This concern was identified as important in the reviewed material and duly discussed in the scoping review.

Reference List:

1. European Commission. *An EU Initiative on Web 4.0 and Virtual Worlds: A Head Start in the next Technological Transition, Accompanying Document*. (2023).
2. European Commission. *European Citizens' Panel Virtual Worlds: Have Your Say -- Information Kit*. (2023).
3. European Commission Directorate-General for Communications Networks, C. and T. *European Citizens' Panel on Virtual Worlds, Final Report*. (2023).
4. Boeije, H. *Analysis in Qualitative Research*. (Sage, London, 2010).

5. European Commission. *European Citizens' Panel on Virtual Worlds Final Report*. https://citizens.ec.europa.eu/system/files/2023-11/ECP%20on%20Virtual%20Worlds_Final%20Report.pdf (2023) doi:10.2775/392615.