Alliance For Universal Digital Rights

VULNERA THE INTERNATIONAL OBSERVATORY

Human Vulnerability in the Metaverse

Carlotta Rigotti (Leiden University)
Gianclaudio Malgieri, PhD (Leiden University)

Executive Summary

There is much hype and enthusiasm around the development of the metaverse. An interoperable, persistent, and synchronous network of 3D virtual, real-time rendered worlds, the metaverse is expected to be so real and immersive that it will blur the line between our physical and virtual lives. Technology companies promise that it will bring about a newly empowered, just, and equitable society, by removing barriers to social participation and enhancing human capabilities. But despite significant investment, the metaverse remains an intangible concept for many people – one that could enable unknown and unpredictable behaviours and vulnerabilities. Now is the time, therefore, to consider what the metaverse should be, how significant it might become, and how it will affect both the individual and society.

This working paper was written to assess the impact of the metaverse on human vulnerability. The paper questions why the metaverse is being developed, how it will be created and accessed, who will be creating it, and how it is defined. It also explores the nuances and definitions of vulnerability and positions human vulnerability within the context of the metaverse to consider its impact.

Key Findings

→ By helping users satisfy their human needs and exercise fundamental rights, the metaverse could have a positive impact on the daily lives of people who may experience vulnerability, such as people who are physically disabled in the offline world. It could also be a valuable tool for medicine, science, education, art, and social movement.

→ But social inequalities would be reinforced and accelerated by the metaverse, due to the digital divide. Certain categories of people – especially those in the Global South, rural areas, and many women – will have limited opportunity to become meta-users because of the cost of the hardware and software required to access it, or indeed because they have no access to the internet at all.

→ Individuals who are from a marginalized group in the offline world are likely to face the same subordination in the metaverse, if they choose an avatar that reflects their personal characteristics. But equally, it will be problematic if an individual chooses an avatar that has different physical characteristics to their own, in order to conform to socially accepted views of bodily appearance and privilege. Although less likely to experience social subordination, such 'conformist' avatars could erode personal autonomy, self-determination, and diversity in society.

→ The metaverse will create a new channel for abuse, including sexual assault. It is recognized that technology can facilitate abuse, aggravate harm caused to victims, allow for the commission of new forms of violence, and enable abuse. It is likely that the metaverse will bring image-based abuse to the next level, given the embodied and hyper-realistic nature of its content. Similarly, sexual misconduct in the metaverse will cause more trauma than its occurrence on existing online platforms. Moreover, there is currently no definition for sexual offences in the metaverse, and therefore a lack of clarity in the law for dealing with it.

→ The metaverse could turn into a new space for social marginalization, subordination, and oppression for certain categories of people. Consequently, vulnerable groups will withdraw or disengage from it, which will augment a lack of diversity in the virtual world.

→ Governments are not protecting the freedoms and equality of its citizens in the metaverse but are shifting the burden of responsibility to technology companies, based on a 'laissez-faire' attitude. Meta-users will therefore become dependent on companies for removing or mitigating their vulnerabilities. They may accept unfair or undesirable terms and conditions, or choices that are favourable to the technology companies, in order to enjoy the virtual world. In addition, some users may become vulnerable to changes in the service, if they depend on its existence to participate in society; for example, people with physical disabilities who might rely on the metaverse to move around.

→ Meta-users will be vulnerable to mental manipulation based on AI driven emotional recognition, such as eye tracking or behavioural surveillance. This would especially impact children, people with cognitive impairments and those with psychological vulnerabilities.

Conclusion and Recommendations

We believe a measured approach is needed when considering the potential impact of the forthcoming metaverse, rather than the unconditional enthusiasm displayed by its creators. This is not to deny the potential advantages of the new technology, but rather to caution against the new risk of meta-vulnerability. We hope that Big Tech companies take concerns about meta-vulnerability seriously and we urge governments and regulators to prepare for the significant impact the metaverse will have on the fundamental rights of individuals, especially those who are vulnerable or marginalized.

We recommend that governments and regulators:

→ Ensure that businesses developing the metaverse follow the UN Guiding Principles on Business and Human Rights. Technology companies should be required to assess the impact of the metaverse on the human rights of people who are vulnerable or marginalized and conduct meaningful consultation with affected groups and other stakeholders.

→ Establish guidelines for developing robust models to measure the impact on the human rights of vulnerable groups in the metaverse.

→ Require technology companies to involve vulnerable groups in the participative design of the metaverse.

→ Set standards of best practice for the ‘vulnerability-sensitive’ design of this new technology, based on the lessons learnt from the participatory design process.

→ Clarify whether the current laws which prohibit sexual violence are applicable in the metaverse and address any gaps by enacting new laws and policies.

→ Review and, if necessary, restrict the use of AI-driven emotional recognition of users in the metaverse.

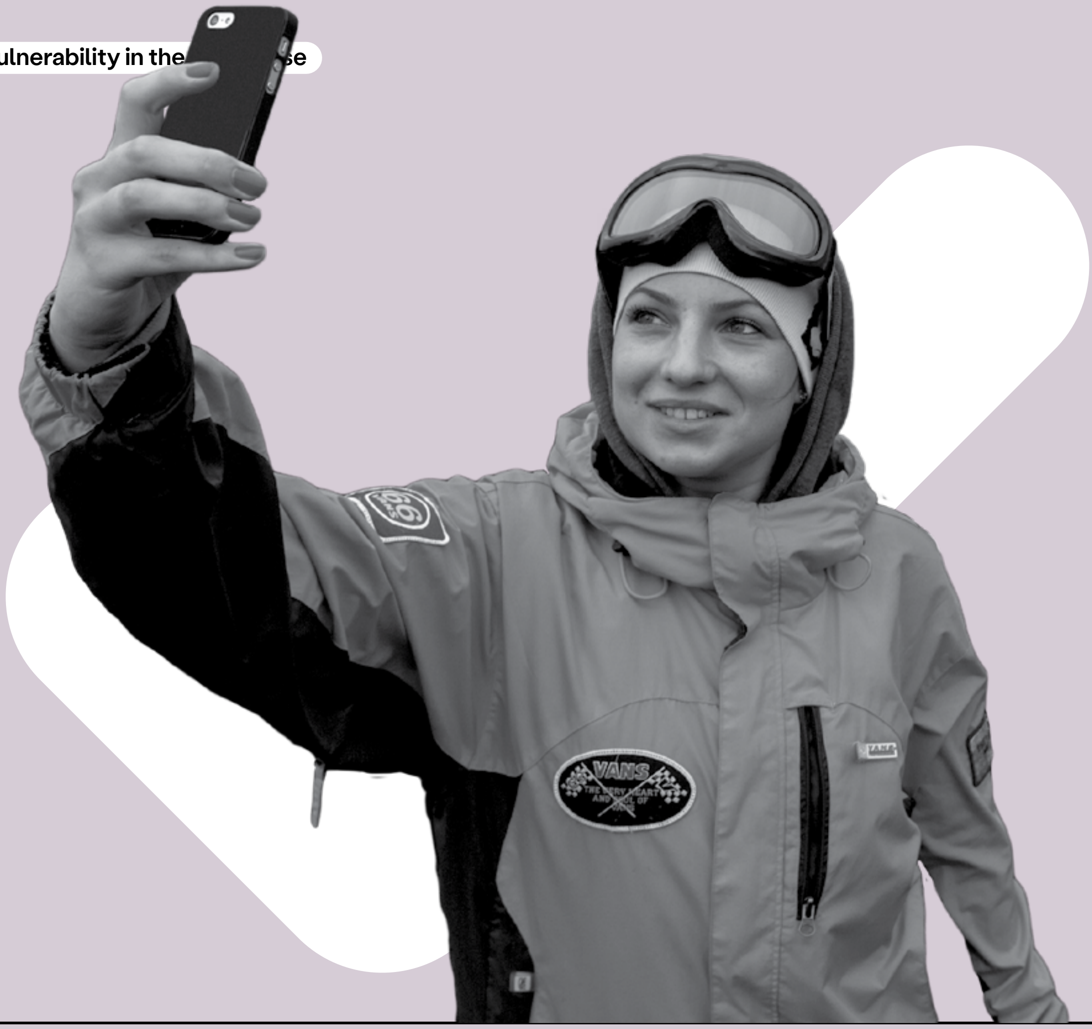
Contents

Acknowledgments

This report is a publication by The Alliance for Universal Digital Rights and Vulnera, The International Observatory on Vulnerable People in Data Protection. It was written by Dr. Carlotta Rigotti and Dr. Gianclaudio Malgieri, both from Leiden University.

The report was published with the support of Equality Now staff, including:
Emma Gibson
Amanda Manyame
Tsitsi Matakaire

Layout and design is by Peter Wilbourne.



1. Introduction

From crossing oceans to exploring space, and even creating the internet, humankind has always been driven by a desire to seek out new worlds; a pursuit that perhaps stems from the want to reset social systems and start life over with a clean slate. The same could be said about the appeal of the metaverse, which, according to the companies engaging in its creation, is expected to be so real and immersive, that it will blur the distinction between our physical and virtual worlds. People across the globe have already discovered the feasibility of moving their lives online, whether that has been for learning, training, working, communicating, or entertaining; a development accelerated by the emergency lockdown measures during the Covid-19 pandemic. This embrace of virtual living has destigmatized the act of spending time in worlds other than the physical one.

> “Now is the time to consider and discuss what the metaverse should be.”

Despite the significant investment geared towards its forthcoming emergence *(Bobrowsky, 2021)*, the metaverse is still an intangible concept for many people. But the fact that it is based on past and present technologies, such as online platforms, blockchain, and virtual and augmented reality, gives us the opportunity to better understand and influence its design and functionality. Now is the time, therefore, to consider and discuss what the metaverse should be, how significant it might become, and how it will affect both the individual and society. Although the metaverse has the potential to present people with many opportunities and freedoms, we should be wary of its potential to replicate, exaggerate and amplify existing stereotypes, power imbalances and misogynistic behaviours that exist in the world as we know it today. It is possible that the metaverse will enable and potentially reinforce and accelerate human vulnerabilities, including social asymmetries of power that lie outside state control. By considering the impact of the metaverse now, meta-users will have the chance to shape their own personhood within the new virtual world and the opportunity to divert any sexism, racism, and other discrimination based on social inequalities that they already experience in their physical lives.[1]

With this in mind, this working paper will pose the question: "How will the metaverse affect human vulnerability?". Section 1 will develop a working definition of the metaverse, providing a brief overview of its necessity, infrastructure, creators, and timing. Section 2 will set the scene and discuss the traditional formulations and understandings of vulnerability. Section 3 will transpose human vulnerability in the metaverse and identify the following three new forms of vulnerability – that the metaverse will entrap meta-users in bodily appearances and performances that conform to socially accepted views of normality, that it will cause harm to people who choose an avatar with personal characteristics that generally lead to social marginalization and stigmatization in the offline world, and that it will keep users in a permanent state of dependency on private platforms. Finally, Section 4 will provide conclusions and recommendations.

[1] See, for example, the webpage of Meta with information about diversity and inclusion in its design process of the metaverse: https://about.fb.com/news/tag/diversity-and-inclusion/ accessed: 10 March 2023

2. Understanding the Metaverse

Neil Stephenson first popularized the term 'metaverse' in his 1992 novel Snow Crash, which was named after a software failure mode on early Macintosh computers *(Ball, 2022; Burrows, 2022)*. Drawing inspiration from the massively multiplayer online games that already existed in the 1990s, the American author imagined the metaverse as a virtual world that took over the internet and enabled people to transpose their lives fully online. The metaverse could be accessed through the creation of an avatar and the use of a personal or a public terminal, while some remained constantly connected via portable terminals. The novel is set in a world governed by private companies, following worldwide economic collapse that precedes the creation of the metaverse *(Ball, 2022; Burrows, 2022; Stephenson, 2008)*.

Far from being a novelty, this piece of science-fiction is part of a long-running tradition of fictional works that have inspired technological innovation. Take, for instance, the ancient myth of Pygmalion, which foresaw the modern use of sex robots *(Liveley, 2021)*, and the series of Matrix movies, which sparked debate about the nature, use, and governance of reality and virtuality *(Edwards et al., 2020)*. In the case of the metaverse, Herman Narula goes back much further than the publication of Snow Crash in 1992 and gives the example of the Egyptian pyramids, which were used to represent the eternality of the post-mortem world *(Narula, 2022)*. More recently, the metaverse has been represented in popular culture *(Burrows, 2022)*, with episodes of the TV series Black Mirror and the Steven Spielberg movie Ready Player One being classic examples.

While popular culture and academic literature largely focus on how each aspect of our lives are transposed in the metaverse *(Dwivedi et al., 2022)*, most private companies engaging in its creation take a more limited interest, in line with their own sector-based worldviews. For example, the Match Group – which owns online dating services like Tinder, Hinge, and OKCupid – recently claimed that it might provide "augmented features, self-expression tools, conversational AI and a number of what we would consider metaverse elements" *(Ball, 2022)*.[2] This then prompts the question, will there be only one or many 'metaverses'? Such a question is difficult to answer. According to Florian Buchholz et al., it depends on whether one would similarly argue that there is only "one internet". As cyberspace varies in different purposes, fora, and forms of interaction, with many servers and networks behind it, the metaverse is likely to reproduce the same diversity *(Buchholz et al., 2022)*. In any case, it currently appears that Big Tech companies conceive a number of interconnected metaverses, covering a wide range of virtual worlds that differ in several ways, including function, access, and governance models *(Ball, 2022)*.

Against this backdrop, the following sections will seek to solve the problem of terminology that most studies on the metaverse are currently facing *(Hackl et al., 2022)*. In doing so, we will position the metaverse within the broad analysis of its possible purposes, infrastructures, methods of access through avatar creation, and community. Later, we will adopt a working definition and explain the reasons behind our demand for scholarly discussion and regulatory intervention.

[2] Although it appears that the Match Group stepped back from its metaverse dating plans in August 2022 (Samantha Delouya, 2022), Sangeeta Singh Kurtz and Lakshmi Rengarajan report that many other companies still see a future and continue investing in this technological innovation (Sangeeta Singh Kurtz & Lakshmi Rengarajan, 2023).

2.1 Why should we build a metaverse?

The aim of the forthcoming metaverse is to enhance human capabilities in the physical world, and to create brand new possibilities. In the case of the former, the metaverse is designed to solve difficulties that we face in our physical lives *(Dwivedi et al., 2022)*, and could help us perform virtual tasks that are otherwise more complex or impossible to perform offline. For instance, in a recent article published in Nature Machine Intelligence, Wang et al. welcome a wide number of meta-applications, including virtual comparative scanning that allows the creation of a patient's digital twin and a more comprehensive analysis of their pathologies, even with regard to disease prevention *(Wang et al., 2022)*. Several Big Tech companies even promote the metaverse as a tool for empowering certain groups of vulnerable people, because it will allow them to overcome physical and socio-economic obstacles *(World Economic Forum, 2022)*.

Alternatively, the metaverse could open up brand new possibilities, most of which are beyond our current understanding and imagination *(Burrows, 2022)*. In this scenario, its design is likely to involve stand-alone applications. This means that the metaverse will no longer be a reproduction of the offline world, but will rather offer its own opportunities, meanings, and value *(Dwivedi et al., 2022)*. Cathy Hackl et al. emphasize the digital economy that will characterize the metaverse, which will eventually allow users to earn, pay, exchange, and invest. In sum, the ultimate goal of the metaverse will be to access goods and services in new and immersive ways *(Hackl et al., 2022)*.

"The aim of the forthcoming metaverse is to enhance human capabilities in the physical world, and to create brand new possibilities."

2.2 How will we create and access the metaverse?

Although some architects claim responsibility for the societal function of all design – Zaha Hadid Architects has even launched its own metaverse project *(Schumacher, 2022)* – most of the literature about creating and accessing the metaverse focuses on the hardware we will need *(Ball, 2022; Burrows, 2022)*. In pop culture, the main character of Snow Crash uses goggles with fibre optic cables running down a plastic tube to access the metaverse, while the characters of Black Mirror have retinal implants or special lenses *(Burrows, 2022)*. Many Big Tech companies are indeed investing in existing technologies in order to improve displays, reduce weight, and increase battery life. They are also investing in new technologies, such as designs to fit a supercomputer into the frame of everyday accessories like glasses and bodysuits *(Ball, 2022)* or to conceive features that have not yet been considered *(Burrows, 2022)*.

In creating a virtual world that differs from the current cyberspace, the metaverse will inhabit a three-dimensional (3D) space. As such, it will be possible to represent a world as we currently experience it offline – it will be richly detailed, with a mix of audio and video and the sense of being live, rather than static, or outdated. But the metaverse is not just about VR and AR technologies[3], as is sometimes held in popular magazines, grey literature, and scholarship *(Hackl et al., 2022)*.

Blockchain is another technology that is mentioned when discussing the infrastructure of the metaverse.[4] Several authors believe that blockchain is structurally required for the metaverse to become a reality, serving many functions including governance protocol, incentive mechanism, global payment rail, trustless participation, and a global immutable ledger *(Ma & Huang, 2022)*. Cathy Hackl et al. consider blockchain technology a means of empowerment that gives meta-users the ability to design and govern their community *(Hackl et al., 2022)*. Indeed, one of the main benefits of blockchain technology is its alleged incorruptibility; the larger and more decentralized a network is, the more difficult it is for data to be overwritten or disputed, because the majority of the decentralized network would have to agree to it *(Ball, 2022)*. However, it appears that decentralization is more expensive, and time-consuming, meaning the meta-user is likely to be unsatisfied with the synchronous shared experience it is supposed to offer *(Ball, 2022)*. In addition, Huynh-Thee et al. make it clear that blockchain is only one of the state-of-the-art methods that has exploited artificial intelligence (AI) to drive the creation of the metaverse *(Huynh-The et al., 2022)*. It is important to highlight, however, that if blockchain becomes an integral part of the metaverse, its complexity will create limitations for people experiencing vulnerabilities, such as digital illiteracy. The main obstacle for such people will be difficulty in understanding the complex functionality of blockchain, which will impair a user's autonomy and informational self-determination – for example, the user might find it difficult to exercise their right to be forgotten on blockchain technology *(Fink, 2018)*.

In conclusion, the metaverse is a mostly intangible experience composed of a persistent network of virtual worlds, data, and supporting systems. However, physical devices and AI systems are still the gateway to accessing and creating these experiences, and therefore Big Tech companies and other private actors will have to carefully consider how they define this new world *(Ball, 2022)*. Moreover, because a decent online connection is the *conditio sine qua non* to access the metaverse, the contemporary digital divide will be a key limitation for many people who wish to become a meta-user. According to

[3] By VR, we mean a technology creating immersive and interactive, simulated environments. AR refers to a technology enhancing the real world by placing data, interactive digital objects, or other digital media on top of the physical world.

[4] According to us, the term 'blockchain' could be generally understood to mean a digital ledger that is shared across a public or private computing network and is mathematically encrypted so that no data alteration is allowed post recording into the blockchain. When a decision needs to be made, it is up to the participating nodes to reach a consensus (Tran & Krishnamachari, 2022).

the United Nations Development Programme, nearly 37% of people across the world still have no access to the internet, especially in the Global South, rural areas and when belonging to certain social groups, including women *(UNDP, 2022)*.

2.3 Who is creating the metaverse?

Since the words we use and the meanings we attach to them reflect our thinking and inform our behaviour, we should consider who is driving the creation of the metaverse and, consequently, setting the scene for its aims and governance. As is the case with the novel Snow Crash, it appears that the metaverse will be run by profit-driven Big Tech companies *(Ball, 2022; Burrows, 2022)*, such as Meta, Microsoft, and Google *(Meghan Bobrowsky, 2021)*. This implies that the metaverse will mostly rely on a profitable business model based on obtaining and retaining users *(Burrows, 2022)*. Some Big Tech companies expect to cooperate with each other *(Mark Zuckerberg, 2021)* so that meta-users benefit from a seamless experience *(Burrows, 2022)*, where their avatar might use a dating app, while simultaneously doing the grocery shopping in a virtual supermarket and listening to a podcast on Spotify.

Conversely, Matthew Ball points out that, due to its complexity, the creation of the metaverse infrastructure could engage diverse actors, especially start-ups, in addition to the Big Tech companies *(Ball, 2022)*. The key problem with this prediction, however, lies in the war for supremacy among technology firms, and the possible acquisition of start-ups by Big Tech companies, which has already occurred with AR and VR technology *(Ball, 2022)*.

Lastly, Narula stresses the pivotal role that individual users could play in the definition and governance of the metaverse, thereby mitigating the risk of private monopoly. According to the British Indian author, meta-users will inevitably advance and expand the parameters of the metaverse, and the experiences available within it, in ways that we can neither predict nor control. Similarly, small businesses and entrepreneurs who access the metaverse will drive its design when offering new goods and services that ease and enhance users' experiences of these worlds *(Narula, 2022)*.

2.4 How do we define the metaverse?

This working paper uses the term 'metaverse' to refer to the broad, technical definition by Ball, namely that the metaverse is: "[a] massively scaled and interoperable network of real time rendered 3D virtual worlds that can be experienced synchronously and persistently by an effectively unlimited number of users with an individual sense of presence, and with continuity of data, such as identity, history, entitlements, objects, communications, and payments" *(Ball, 2022, epub)*. This definition encapsulates most of the features of the metaverse that we have thus far discussed.

In unpacking each element of this definition, Ball briefly explains that virtual worlds refer to any computer-generated, simulated environment that are in immersive 3D and real-time rendered; the last feature being necessary for the individual to migrate their daily life from the physical world to the online one. The metaverse is likely to be composed of a huge number of virtual worlds, where people can enjoy as many or more diverse experiences as they can in the physical world. Because the individual is expected to transform into an avatar and carry their belongings with them, the metaverse should be interoperable, persistent, and synchronous. This means that technological development must allow the exchange and use of information sent from one virtual world to another, manage the persistence of data across various worlds and over time, as well as allowing synchronous and shared experiences. Thus, if a virtual world goes offline, resets, or shuts down, an individual's history, achievements, social relations, and goods, reliably endure and therefore a large number of people can experience the same event, at the same time, and in the same place *(Ball, 2022)*.

It is important to clarify that, at the time of writing, we expect the metaverse to turn into a real world where people will go through a continuum of offline and online experiences until the line between the two worlds is blurred. This scenario might never become a reality, but it is what the Big Tech companies are aiming for *(Ball, 2022)*, and we should therefore take these ambitions seriously and use them as the starting point for our analysis. For this reason, when distinguishing the metaverse from offline reality, we will refrain from using the common binary 'real' versus 'virtual'. Rather, we will refer to the 'physical' or 'offline' experiences that we have in the offline world and compare them to the 'virtual' or 'online' ones that we will have in the metaverse.

2.5 When and why should we discuss the metaverse?

One could argue that the emerging virtual world that we are discussing is unlikely to come to fruition or is so far in the future that discussing it or being concerned by it now is unnecessary. However, Cathy Hackl et al. make it clear that the metaverse is already in its infancy *(Hackl et al., 2022)*. Indeed, Florian Buchholz et al. give the example of the 2022 launch of Volvo's XC40 Recharge vehicle, which took place in the metaverse – or 'Volvoverse' – in order to reduce the carbon footprint of the launch event. Meanwhile, like Meta and its Horizon Workrooms, Accenture created digital twins of its offices so that employees could join meetings and experience 'normal' workdays during the Covid-19 pandemic *(Buchholz et al., 2022)*. And Patrik Schumacher expands on the design of the Liberland Metaverse, which Zaha Hadid Architects and other organizations are developing, "with the intention to establish a libertarian micronation on a seven square kilometre small, uninhabited, disputed piece of land on the Danube, between Croatia and Serbia" *(Schumacher, 2022, 11)*.

Ball explains that a proto-metaverse has already grown from text-based chat, multi-user dungeons and other virtual worlds in the last 70 years and will continue to evolve and grow, thereby offering more realism, diversity of experience, participants, cultural influence, and value to virtual worlds *(Ball, 2022)*. This means that we cannot draw a line between a pre-metaverse era and a post-metaverse one, because technological innovation has traditionally been an iterative process in which several changes occur and converge.

Given the ambitions for the advance of the metaverse, the aim of moving some of our daily activities there, and the fact that we are unable to precisely predict how it will change human life, we consider it necessary to prepare for what might lie ahead. We therefore believe that our analysis is timely; the metaverse is expected to be a new world, and the ideas on which it is based should be moderated by human-made rules and lessons learnt from offline society, as well as our learnings from our online experiences to date.

3. Understanding the Nuances of Human Vulnerability

Because this paper positions human vulnerability in the metaverse, this section is designed to help the reader understand how human vulnerability has so far been understood offline.

Generally, the term 'vulnerability' is used to describe social marginalization, economic insecurity, precarious employment conditions or violence caused by wars and similar situations. Early definitions and conceptualization of vulnerability stressed its links to fragilities, harms, and the fact of being wounded, as its etymology likewise suggests. Indeed, the Latin word 'vulnus' means wound *(Mackenzie et al., 2013)*. The term served almost as a synonym of dependency, helplessness, pain, violence, and weakness. As Robert E. Goodin affirmed: "to be vulnerable is to be susceptible to harm to one's interests" *(Goodin, 1985; see also Schroeder & Gefenas, 2009)*. Many other scholars and institutions followed this view of conceptualizing vulnerability around exposure and the likelihood of being harmed in the context of autonomy, dignity or integrity *(Mackenzie et al., 2013)*.

However, vulnerability is a condition situated in opposition to actual harm or injustice, rather it indicates potentiality *(Gilson, 2014)*. Well-established views also stress that vulnerability is a condition that should be avoided – it is something negative or a risk that should be mitigated *(Malgieri & Niklas, 2020)*. This approach is often criticized by feminist scholars who explore the positive sides of vulnerability, showing it as a precondition of empathy, social connectedness, and intimacy *(Cole, 2016)*. Therefore, vulnerability is not only a limitation but also something that allows us to act and feel, with Erinn Gilson formulating vulnerability as an "openness to being affected and affecting" *(Gilson, 2014, p. 36)*.

Nonetheless, many problematic dichotomies and uncertainties affect the application of vulnerability in institutional environments. One of these dichotomies is between the particular and universal character of vulnerability.

"Vulnerability is not only a limitation but also something that allows us to act and feel."

In more traditional approaches, vulnerability is a distinctive characteristic of particular weaker individuals and groups, based on specific situations or socio-economic contexts *(Fineman, 2012)*. Typical examples of such groups are racial minorities, asylum seekers, and people with disabilities. This is the predominant way the concept of vulnerability is used in circumstances like research, social policy, or policing *(Hewer, 2019; Nicole L. Asquith et al., 2017)*. However, several authors have criticized this way of understanding vulnerability, since it might bring about stigmatizing effects and harmful regulations for marginalized groups *(Cole, 2016)*.

For these reasons, some scholars reformulate the understanding of vulnerability as a universal human condition, which can change in different situations, periods, and spaces. They portray vulnerability as a general feature of human existence; a characteristic of every human being *(Martha Albertson Fineman, 2008; Nussbaum, 2006; Butler, 2004)*. Some disagree with this approach, arguing that the universal concept of vulnerability ignores the structural violence, injustice, and exploitation that particular groups of people experience *(Cole, 2016; Cooper, 2015)*. But proponents of a universalized notion of vulnerability show this can be a way to deflect the failures of existing diversity and equality policies, and anti-discrimination laws *(Fineman, 2008)*. Another area of dispute about vulnerability concerns the organizational, legal, and political responses to vulnerability. Martha Albertson Fineman calls on institutions to recognize human vulnerability. She criticizes existing systems of rights and laws that depend on formal equality and embrace an individualistic, self-sufficient, and rationalist liberal subject. In a similar way, for Goodin, the idea that there are some members of society who are more vulnerable is a rationale for why we need a welfare state, which could help to address inequalities in relation to access to essential goods and services *(Goodin, 1985)*.

In response to these criticisms, Florencia Luna formulated the concept of vulnerability as layers *(Luna, 2019)*. According to this scholar, layers of vulnerability are not fixed attributes of specific individuals or groups, but are features constructed by status, time, and location. In this sense, the concept of layering provides an opening to a more intersectional approach and stresses its cumulative and transitory potential *(Luna, 2009)*. We can summarize this universal focused theory as follows *(Luna, 2009; Luna, 2019)*: all individuals are vulnerable and labels should not be ascribed to particular groups, but some individuals have more layers of vulnerability based on particular contexts and relational balances. The intensity of legal protection needed by vulnerable individuals is proportional to the quantity and quality of layers of vulnerability. The identification and assessment of layers of vulnerability should be based on several criteria, including an analysis of the origins of vulnerability, such as stimulus conditions – including whether some layers are "cascade vulnerability", meaning layers that have a cascade effect on other sources of vulnerability – and of its impact (that is, probability and intensity of harms). Lastly, this theory on layered vulnerability suggests that each vulnerability layer has its own mitigation measures, including avoiding exacerbating layers, eradicating layers, and minimising layers of vulnerability through different strategies, such as protections, safeguards, and empowerment.

In summary, discussions about human vulnerability can prompt debate and new ideas. In this research, we adopt a contextual, relational, and risk-based notion of vulnerability, where human vulnerability – although inherent in all human beings – should be identified as the inevitable inequality of resilience among individuals *(Fineman, 2017)*, and the risk to their fundamental rights and freedoms in specific contexts and power relations. The higher the risk of a person suffering adverse effects to their fundamental rights, and their incapability to mitigate that risk or face its consequences, determines their vulnerability. Since individuals rely on social infrastructures to satisfy their fundamental rights and freedoms – for example, social media potentially enables freedom of expression and association, caregivers enable freedom of movement, an app enables freedom to work or to conduct business – an exclusive or quasi-exclusive dependence on these specific entities might enhance their vulnerability.

4. Positioning Human Vulnerability in the Metaverse

Having established the context of this working paper, to understand how we define the metaverse and human vulnerability, we will move on to our core analysis exploring how the metaverse is expected to affect human vulnerability from both a positive and a negative perspective. While literary narratives and Big Tech companies imagine a newly empowered, just, and equitable society *(Narula, 2022)* as a result of the metaverse, we nonetheless consider it necessary to approach this new technology with caution. After all, the metaverse is expected to raise generations of users who will no longer be able to distinguish between their online and offline experiences. This can already be seen with the increasing number of tools available to create an avatar – a 'digital twin' that allows users to represent and express themselves in the metaverse. People can customize their avatars by choosing various physical characteristics, such as skin pigmentation, hair colour, and body shape. Such customization options are available with the likes of Meta or can be taken a step further through websites such as Ready Player Me *(Ready Player Me, 2023)*[5], which allows users the option to create a full-body 3D avatar based on a selfie. Although most avatars are lifelike, it is sometimes possible for the user to go beyond human anatomy. For example, avatar app Genies allows the user to choose whatever form one self-identifies with, including "a Zen Teacup" and "a Psychotic Bunny" *(Genies, 2023)*[6].

"The metaverse will give rise to new forms of vulnerabilities."

Though disembodying, and possibly mitigating, some traditional vulnerabilities of humankind, it is our opinion that the metaverse will give rise to new forms of vulnerabilities. For instance, in creating their own avatar to access the metaverse, the individual will have to choose between two risks of vulnerability. On the one hand, they could choose an avatar that does not look like them but instead conforms to socially accepted views of bodily appearance and performance. For example, young women might represent themselves as old men in the metaverse, to increase their employability. While benefiting from the social privilege of being an older male, because such an avatar would be less likely to experience social subordination, the choice would have a detrimental impact on diversity in society. On the other hand, the individual could resist social conformity and create an avatar that reflects their personal characteristics but runs the risk of reproducing long-lasting asymmetries of social power. For example, instead of choosing older male avatars, the young women of the previous example could go beyond the stereotypical, social construction of femininity and even opt for gender fluidity. Regardless of which avatar one chooses to create, a universal meta-vulnerability is likely to arise from the users' dependence on Big Tech companies to access the metaverse, in order to self-determine and participate in society. At present, States are refraining from accepting full responsibility for the regulation of the metaverse, which means that its governance and responsibility to meet users' needs are left to private companies.[7] This hesitancy to intervene could be for a number of reasons – States may fear stifling technological innovation or disadvantaging the competitiveness of their national market. They may experience lobbying from large companies,

[5] Ready Player Me is an avatar platform for developers, creators and 'residents' of the metaverse, which can be integrated with other apps and games (see https://readyplayer.me/it).

[6] Genies is a consumer app that allows users to create an avatar to be used in other apps, such as Giphy, iMessage and Instagram (see https://genies.com/).

[7] With some (albeit minimal) exceptions, see the declaration of intent of the European Union (European Commission, 2022).

their legislators may lack competence, or they may be sceptical about regulating digital contexts that are too new to be assessed.

4.1 Mitigating human vulnerability through disembodiment

The metaverse could have many positive impacts on the daily lives of people experiencing different forms of vulnerabilities in the offline world. It can be a valuable tool in fields such as medicine, science, education, art, and social movements *(Access Now, 2022)*. More precisely, the metaverse could help users satisfy their human needs,[8] as well as exercise their human or fundamental rights – such as the freedom of movement, freedom of expression, and freedom of association – especially in contexts where such human needs are more challenging to achieve.

For example, the metaverse offers virtual mobility, which could open up accessible alternatives for activities that usually require physical mobility and prevent people with mobility issues from equally participating in society *(European Commission, 2022; Zuckerberg, 2021)*. The metaverse will also provide people with new solutions for remote working, learning, training, socializing, and entertainment, and will take intimacy and other human-machine interaction even further than the world wide web *(Council of the European Union, 2022)*.

According to some of the promises made by metaverse providers, the metaverse will be a tool to guarantee diversity and enable people to have social interactions that they would not otherwise have. This could be due to health issues such as immunological deficiencies that prevent people from meeting others in person, limited mobility, psychological reasons including difficulties with social interaction, or economic reasons such as being unable to afford travel costs to conferences, project meetings or even to meet family members abroad. Since research has traditionally tied vulnerability to the embodiment of humankind *(Matambanadzo, 2012; Fineman: 2013; Herring: 2016)*, in this sense, the metaverse could prove to be a useful tool for mitigating human vulnerabilities and enhancing human resilience.

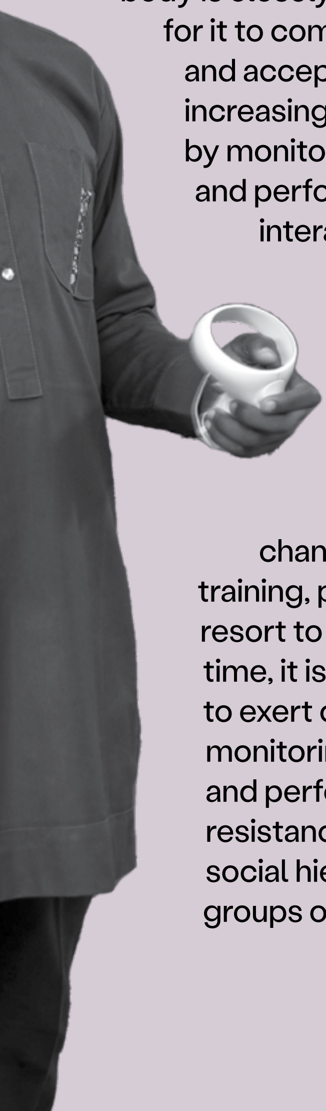
However, as the sections below will explain, rather than producing 'disembodiment', the metaverse will produce a shift from physical to virtual embodiment, and all the related risks that this might bring. This could include new or augmented forms of harm, privilege-based conformism, increasingly powerful emotional surveillance and manipulation, and enhanced dependency on proprietary platforms.

In addition, it is worth stressing that, although the metaverse might hold much promise for mitigating some existing forms of human vulnerability, its deployment could likewise reinforce and accelerate social inequalities. As nearly half of the global population is still offline and there is a gender gap in global internet use (UN, 2020) *(Lee, 2021)*, it is likely that certain categories of people will have limited opportunities to access the metaverse due to the cost of hardware and software that allows access to it, or because they do not have access to the internet at all.

[8] To identify these human needs, it is possible to refer to the capability theory of Martha Nussbaum (Nussbaum, 2003).

4.2 Channelling human vulnerability

Generally speaking, the body is a material reality; it is visible, tangible, and takes up space. In being incomplete at birth, it develops and changes throughout our lives. While a naturalistic approach to the definition of the human body considers it a naturally specific entity whose biological constitution permanently determines our being *(Shilling, 2012)*, we embrace its understanding as a social construct too. In our opinion, the way the human body is seen, experienced and treated establishes an individual's sense of their own body and the way in which it embodies social privileges and subordinations *(Shilling, 2012)*. In this context, the use of the plurals is an intentional one, arising from a number of intersecting memberships that each human body normally mirrors at once *(Natalie Boero & Katherine Mason, 2021, drawing on the intersectionality theory of Kimberlé Crenshaw)*.



The physiological development of the human body is closely associated with the pressure for it to comply with social expectations and accepted views. As a result, people increasingly try to control their image, by monitoring their bodily appearance and performance to facilitate social interaction and recognition. The aim is to present themselves as 'normal' people worthy of gaining social acceptance and joining society *(Shilling, 2012; Lorber & Yancey Martin, 2013)*. This is often done, for example, when we change our bodies through physical training, put ourselves on a diet, or resort to cosmetic surgery. At the same time, it is possible for the individual to exert control over their identity by monitoring their bodily appearance and performance as a form of social resistance, in order to stop perpetuating social hierarchies that favor certain groups over others, based on their embodied, personal characteristics *(Lorber & Yancey Martin, 2013)*. Although training, dieting, and undergoing cosmetic surgery is often linked to social conformity, it can also be an expression of nonconformity if used to reach a result other than a socially accepted one, similar to how people change their bodies through tattooing and piercing. The chance to conform to or resist socially accepted views of the human body also arises from the deployment of technology. Since the late 1980s onwards, cyberfeminism has considered the internet a possible means to abandon the corporeal presence of existing, gendered bodies, and to empower the individual *(Haraway, 1991; Stone, 1995; Plant, 1997)*. A minority response, however, immediately adjusted this social expectation, by pointing out that cyberspace and other technologies were likely to reproduce the same asymmetries of power because "the new cannot be spoken except in relation to the old" *(Katherine N. Hayles, 1999, 158. See also Balsamo, 1999; Gonzalez, 1999)*.

To conclude, because the human body shapes the way society sees and treats us, we believe that our embodiment has serious implications for the way our vulnerability is defined. Accordingly, the following sections will discuss the chance for our avatar to level out embodied differences or maintain our sense of otherness, and the consequences this personal choice might have for both the individual and society.

4.2.1 Risking conformity with socially accepted views on bodily appearance and performance

According to many of the Big Tech companies that are promoting the launch of the metaverse, the virtual world will give the user the freedom to decide how they self-identify and are seen, and will eventually remove some socio-economic and physical barriers to self-determination and equal participation in society *(Meta, 2022)*. While recognizing the good intentions behind this ambition, we do not believe that the technology will effectively contribute to the eradication of social marginalization, stigmatization, and subordination within society. Rather, the metaverse risks entrapping people in bodily appearances and performances that conform to socially accepted views of normality, to the detriment of diversity in society. This phenomenon has previously been described in analogous contexts as "cosmetic vulnerability", which is based on aesthetic consumer desire *(Garcia Sanchez, 2016)*.

"The Metaverse risks entrapping people in bodily appearances and performances that conform to socially accepted views of normality, to the detriment of diversity in society."

Cathy Hackl et al. explain that, as is the case with social media profiles, the creation of an avatar in the metaverse will become extremely valuable to our sense of self and social acceptance *(Hackl et al., 2022)*. In this regard, the US authors stress that although "avatars give people the freedom to be anything they want", even objects like "a lamp, books, Legos" *(Hackl et al., 2022, epub)*, we should not expect meta-users to go to these extremes. Instead, it will be more likely for people who feel trapped in their bodies, or those who experience social subordination due to personal characteristics, to go beyond these socio-physical constraints *(Hackl et al., 2022)*. For instance, as an avatar, a transgender person will not be bound by the sexual organs or sex and gender representations that they feel do not represent them, nor experience social marginalization due to a sense of otherness. Additionally, people who carry out socially stigmatized jobs and consequently face socio-economic obstacles will be able to separate their lives more safely.

Critically, Giddeon Burrows stresses the social pressure meta-users will feel to 'normalize' among other participants in order to make their social interaction and integration easier, to the extent that "a non-white person in a predominantly white-designed virtual metaverse may choose a lighter color of skin for their avatar, in order to fit in" and "[a] woman may choose to present as a man, so as to avoid harassment" *(Burrows, 2022, epub)*. Incidentally, the British author questions whether it should be considered cultural appropriation or, more radically, a new form of colonialism, if a white person, or someone from another privileged group, was to adopt another skin color or action that did not belong to their own culture or identity. Furthermore, in imagining that an individual could have a different avatar for different contexts or settings, such as for work or for entertainment, Narula is critical of this continuous metamorphosis, which he believes would prevent self-determination. "In a world where you have to create two hundred different user profiles for two hundred different websites [...], there is little incentive to invest in or feel protective of any of the myriad digital persons that you [are] forced to maintain" *(Narula, 2022, epub)*.

As mentioned in the previous section, the act of creating a 'conformist avatar' will form part of the long-running search for compliance with socially accepted views of bodily appearance, which has so far included physical training, dieting, cosmetics, cosmetic surgery, and social media profiles, to name a few examples. In the last decade, Instagram filters have provided many of us with the opportunity to craft how we appear to others in the social sphere, often in accordance with stereotypical ideas of body appearance and performance *(Caldeira et al., 2018)*. In this regard, it is worth observing that the body normativity we have come to expect from the process of avatar creation will no longer be limited to the traditional categories protected by anti-discrimination law – such as gender/sex, race, age, and disability – but will also cover other personal characteristics, such as breast size, height, and fashion style, which is a product of living in a culture that emphasizes and fetishes these traits.

Eventually, the individual might be pragmatic in the way they create a conformist avatar, using the new technology to address their bodily dissatisfaction and social subordination. However, since we consider the metaverse to be a continuum off the offline world, the individual might worry about how other people will react to a conformist avatar that does not correspond to their physical, and perhaps socially subordinated, identity. This could be in the same way that the process of undergoing cosmetic surgery is shrouded in secrecy *(Northrop, 2012),* or how deception is commonly used to initiate online and offline dating relationships *(Sharabi & Caughlin, 2019)*. While we would not blame users for creating conformist avatars, we nonetheless caution against them. In our opinion, creating a conformist avatar will be unlikely to solve the problem of social subordination that arises from possessing certain personal characteristics in the offline world, but rather that it might erode the personal autonomy and self-determination of the individual, as well as diversity in society.

At an individual level, we expect the creation of conformist avatars to restrain and censor people, by prompting them to airbrush their personal characteristics in order to abide by socially accepted views of bodily normalcy. This will therefore mean that everyone will look and behave in the same way, thereby blurring the line between the individual and society. In this scenario, we do not argue that people should disregard others' opinions. But it is worth stressing the importance of an individual's ability to critically reflect on their personhood and their decision-making processes. In fact, as Jeremy Weissman points out in his work on conformity and control in social media that "it is through our ability for reason and self-determination that we find what is right for ourselves, develop our individuality, fulfil our highest potential, and in turn become happier people. But if the pressure of public opinion becomes too great, with too little ability to shield our lives from that pressure, that inner voice of self-reflection can become squelched in a potentially mindless conformity toward social conventions" *(Weissman, 2021, 19)*.

At a societal level, the infinite creation of conformist avatars will lead to the loss of diversity in a society that, on the contrary, should remain open to fluidity and constant change. We believe that society should not be regarded as a machine that is built by default and intended to be fixed, but rather something that is designed to continuously evolve, taking into account its inward and competing forces, much like a living being. Diversity in society, in fact, encourages us to seek out new information and ideas, thereby resulting in better problem-solving and decision-making processes. Overall, it can improve individual wellbeing and social coexistence.

In short, we voice concern about the creation of avatars that reproduce bodily appearances and performances in order to conform with socially accepted views of normalcy. The creation of such avatars could make us vulnerable to conformity and privilege. However, opting for a conformist avatar is ultimately a personal choice and would not be intrinsically wrong. Indeed, everyone should be able to decide the best self-identification and representation for themselves, so that they can control how they are seen and treated. Rather, it is our aim to raise awareness about the advantages and disadvantages that this personal choice might have at a macro and micro level, especially when considering how to eradicate the root causes of social subordination. In other words, we aspire for the meta-user to gain knowledge and control of their virtual body so they can proactively respond to the long-running asymmetries of power and give rise to a more diverse and inclusive society.

4.2.2 Enabling harm

People could be proud of their otherness and decide to show it in the metaverse *(Burrows, 2022)*. More precisely, they could create an avatar with personal characteristics that usually marginalize and stigmatize them within society.

They could opt for an avatar that crosses the socially accepted views of bodily normativity and eventually mix personal characteristics together, in reaction against gender binarism, heteronormativity, and other social categories. Said otherwise, it is possible for the metaverse to facilitate the establishment of a fluid society, outside normative rules of physical appearance and behaviours. In designing the metaverse, many Big Tech companies also promote the right and freedom to diversity. For example, Meta plans to provide its users with more than one quintillion different combinations of personal characteristics, such as hairstyle and skin pigmentation, and accessories including glasses and wheelchairs *(Meta, 2022)*. We nonetheless believe that their utopia is more likely to turn into a dystopia, where the metaverse reproduces old asymmetries of power and therefore new channels for harm.

In this regard, the very idea of 'virtual rape' was first coined in 1993, after the avatar of a user of the virtual world called "LambdaMoo" was denounced for forcing other avatars to have sex with him *(MacKinnon, 2006; Strikwerda, 2015)*. In 2007, the Belgian Federal Police announced the criminal investigation of a virtual rape incident that had taken place four years before on the online multimedia platform Second Life. Although little was known about what precisely happened, the criminal investigation did not turn into a charge, and law enforcement agents discovered that taking control of another person's avatar and forcing it to engage in sexual conduct was common in Second Life *(Danaher, 2018)*. Since 2007, many other instances of sexual misconduct have been reported to occur in proto-metaverses *(AccessNow, 2022; Weissman, 2021; Danaher, 2018; Esparza, 2018; Strikwerda, 2015)*.

Overall, it is not the first time that technology has been used as a channel for abuse, and researchers seem little surprised by this new form of harm *(Wiederhold, 2022)*, which can be traced back to the contemporary, fourfold classification of online and technology-mediated violence *(GREVIO, 2022)*. According to this classification, specific technologies could first facilitate abuse, such as is the case with intimate partner violence, committed via the use of spyware and other tracking devices. Second, it appears that information and communication technologies (ICTs) aggravate the harm caused to the victim. This is evident in image-based sexual abuse – such as the non-consensual creating, taking, or sharing of intimate content, including threats to share it – where the victim feels continuous abuse arising from each new, non-consensual distribution and/or viewing of the intimate content. Third, technology seemingly allows for the commission of new forms of violence, including deepfakes. Lastly, it is said that some technologies increasingly enable abuse to take place. As has already been the case with the use of social media being central to various forms of sexual violence against women and girls, we expect the metaverse to do the same. For instance, it is likely that the metaverse will bring image-based sexual abuse to the next level, given the embodied and hyper-realistic nature of its content.

"Their utopia is more likely to turn into a dystopia, where the metaverse reproduces old asymmetries of power and therefore new channels for harm."

At present, existing definitions of sexual offences may not be suitable for the metaverse. However, a growing body of literature describes them with reference to the non-consensual engagement of an avatar in the simulation of sexual conduct *(Danaher, 2018)*. It could be argued that if the virtual world goes offline, resets, or shuts down, it is almost as though the harm never existed. But an increasing number of studies report that some people feel that they had been harmed by sexual assault in the virtual world *(Wiederhold, 2022;*

“We expect these ‘vulnerable’ people to withdraw or disengage from the creation of this new, virtual society as a means of self protection.”

Danaher, 2018). It is far from being a novelty to claim that online and technology-facilitated violence can cause harm, with some victims even dividing their lives before and after the cyber-abuse *(McGlynn et al., 2021)*. Briefly, most victims report a negative impact to their mental wellbeing, such as depression, and panic and anxiety attacks, as well as physical symptoms *(Bates, 2017; Champion et al., 2022; Ruvalcaba & Eaton, 2020)*. Additionally, online and technology-facilitated violence negatively affects the professional and economic life of the victim, who often miss work or school, incur heavy expenses in seeking legal redress and finding specialist support, as well as refraining from online engagement for professional purposes *(UNESCO, 2020)*. Lastly, victims are said to experience a profound sense of isolation due to victim-blaming responses and mistrust of family, friends, and colleagues *(Hearn & Hall, 2022; McGlynn et al., 2021)*. According to Brenda K. Wiederold, sexual misconduct in the metaverse will more negatively impact people than those occurring on online platforms. This is because “[w]hen a user enters a virtual environment, the virtual world becomes their world, their avatar becomes their body. Because of this, if someone is sexually assaulted in such an environment the trauma can easily move to the real world” *(Wiederhold, 2022, 479)*. Simply put, because the metaverse experience feels so immersive and real, the harm can feel very real too.

As we previously clarified, in our opinion, there should be no major distinction between virtual and physical life. It is therefore necessary to position each sexual misconduct within the continuum of violence that women and other categories of people are exposed to in all spheres of their lives *(Kelly, 1988)*, including the metaverse. In this context, the commission of harm in the metaverse is not only a wrongdoing that negatively affects the individual victim, but is also a much broader, socio-cultural one.[9] Rather than being a means for self-determination, social interaction and participation, we are concerned that the metaverse might turn into a new space for social marginalization, subordination, and oppression for certain categories of people due to their personal attributes. Accordingly, we expect these ‘vulnerable’ people to withdraw or disengage from the creation of this new, virtual society as a means of self-protection. As was the case of ‘conformist’ avatars that embody normalized and privileged characteristics, the disengagement of certain groups will augment the lack of diversity and inclusion in the metaverse, and all the negative consequences that will arise from this.

With legal redress still lacking, it appears that some Big Tech companies and other private actors have already imagined strategies to minimize the risk of harm in the metaverse. For example, Ball reports that the user might be required to grant explicit permission to interact in certain spaces, while platforms might automatically block certain capabilities, thereby creating “no-touch zones” *(Ball, 2022)*. Also, it is likely for most of the policing in the metaverse to be done by specific AI applications and/or rely on the reporting of other users *(Burrows, 2022)*. If companies developing the metaverse apply universal technical solutions to put in place virtual boundaries, there may be a danger that these solutions could be applied in a blunt rather than a personalised way, thus restricting the personal autonomy of the metaverse

[9] More precisely, a type of harm manifesting in the normalization and legitimacy of a socio-cultural milieu that draws on social subordination and is conducive to discrimination and violence. On this point see: (Vera-Gray, 2020).

user, who will be unable to draw their own line between legitimate and illegitimate sexual conduct through the provision of their explicit consent. We wonder whether technology companies who govern the metaverse will make design choices that both shift their burden of responsibility for sexual misconduct to users, and at the same time prevent users from exercising self-determination. In conclusion, we see ourselves and our avatars increasingly exposed and vulnerable to the ever-expanding power of private companies.

4.2.3 Depending on the private sector to self-determine and participate in society

The final form of vulnerability we expect the metaverse to carry with its emergence is the structural dependence of "empowered individuals" on the Big Tech companies who, as the designers and rulers of this new technology, will be responsible for the satisfaction of user's needs. As Narula observes: "For now [...] it is worth noting that companies such as Google and Meta have already assumed more power than many nation-states, and in some ways have started to act like autonomous countries" *(Narula, 2022)*. This means that the State will no longer remove socio-economic obstacles that impact the freedom and equality of its citizens but will instead shift this burden of responsibility to the private sector, based on a laissez-faire approach.

As previously observed, human vulnerability is a contextual and relational concept that arises from the encounter between the personal characteristics of the individual – including sex, gender, race, age, and disability – and the dependence on Big Tech companies to access the metaverse, in order to self-determine and participate in society. Though broadly referring to the digital world, Mireille Hildebrandt stresses that there is a "framing power" able to "reconfigure choice architectures in line with whoever pays for them" *(Hildebrandt, 2021)*. This means that Big Tech companies are likely to nudge the user to accept unfair or undesirable terms and conditions in exchange for the satisfaction of their needs and the removal or mitigation of the human vulnerabilities they experience in the physical world.

Generally, power comes to be understood as the "capacity of A to motivate B to think or do something that B would otherwise not have thought or done" *(Forst, 2015)*. This interpretation is in line with EU consumer law definitions[10] and it is, thus, not limited to brute force, but also includes (undue) influence *(Véliz, 2020.)* In rational choice theory, consumers and citizens are modelled as 'actors' who must choose from a "choice set" of possible actions to be able to achieve desired outcomes. This choice set is influenced by "incentive structure". An actor's incentive structure encompasses (his or her beliefs about) the costs associated with different actions in the choice set, and the estimated likelihoods that different actions will lead to desired outcomes *(Dowding, 1996)*. While such a structure is inherently transactional in consumer law, it is worth observing that a wider consideration of choice determinants can be found in private law already; for example, defective consent protects individuals, allowing them to conclude

"Big Tech companies are likely to nudge the user to accept unfair or undesirable terms and conditions in exchange for the mitigation of the human vulnerabilities they experience in the physical world."

[10] Particularly, see Article 5 of the Unfair Consumer Practice Directive (2005).

contracts that encompass provisions that they would not have accepted in a transparent environment *(e.g., Schermaier, 2005)*.

The network effect and virtual 'lock in' within the metaverse – arising from the market dominance of a few Big Tech companies – are compelling "incentive structures" that can easily frame the "choice set" of users in a way that is more favourable to the Big Tech companies.

Max Weber defines power as the ability "to carry out one's own will despite resistance" *(Weber, 1978, p. 53)*. In the present scenario, even if we have personal resistance to accessing the metaverse and even if there is no brute force compelling us to consent, the (market and technological) architecture makes us act in a way that our powerful counterparts prefer. In other words, the choice of architecture and the incentive structures – namely, enjoying a virtual reality that is considered essential in our daily lives – push us to accept undesirable terms and conditions. As it has been argued, consumers can experience psychological dependency on suppliers *(Micklitz et al., 2010; Strycharz & Duivenvoorde, 2021)*, something we, therefore, also expect from the metaverse.

Ultimately, Véliz observes that power – like energy – can transform itself from one kind to another; for example, economic power can become political power *(Véliz, 2020)*. In our example, market power – market dominance and the consequent lock in of customers – can become knowledge power (customers will accept spending, with little or no resistance, more time on the metaverse). This could create an increasingly personalized environment (choice architecture) where the user is more and more dependent on the platform, in a self-reinforcing loop that creates vicious circles *(See Zuboff, 2017)*.

This analysis of power is further amplified in the metaverse. Indeed, as explained above, the metaverse and the technological capabilities, or 'affordances', it offers may help users overcome specific vulnerabilities (for example, disabilities that make it hard to move), thereby strengthening their resilience against vulnerabilities. However, users' dependence on these technologies may also mean that they become vulnerable to changes in the contractual terms or functionalities of the service, since they might depend on its existence to participate in society. While the technology may offer a form of liberation, using online platforms may also induce conformist behaviour, based on the possibilities the provider offers *(Hildebrandt, 2021)*. In other words, the affordances, or design choices, offered by online platforms may lead to new vulnerabilities *(Helberger et al., 2022)*, to the detriment of diversity in society. As well as the earlier example of avatars created with the personal characteristics of privileged groups – such as maleness, or heterosexuality – it is worth referring to the possibility of vulnerability caused by emotional manipulation based on AI-driven emotion recognition of users (through, for instance, eye tracking or behavioural surveillance). This may have a particular impact on children or consumers with cognitive impairments or temporary psychological vulnerabilities *(Malgieri, 2023; Malgieri and Niklas, 2020)* which could warrant the review and, if necessary, the restriction of the use of AI-driven emotional recognition of users in the metaverse.

Because the most used online platforms are, often, offered by private-held businesses, it is not a given that they will fully consider the vulnerabilities that users may experience when using their services. While EU legislation increasingly considers fundamental rights in regulating online platforms *(see, e.g., Article 14(4) of the DSA)*, the (legal and technological) design of online platforms is hardly value sensitive.

5. Conclusion: Policy Recommendations

This working paper has attempted to position human vulnerability within the context of the metaverse. While the creators of this new technology display unconditional enthusiasm for its potential to tackle intersecting social subordinations, we felt the need to take a more measured approach in our assessment, given the unknown and unpredictable impacts and outcomes of the metaverse. In doing so, we recognize that we have mostly focused on the limitations, rather than the ambitions, of the metaverse. Our analysis, however, is not to deny the many advantages that the metaverse might bring, but rather an attempt to caution against the new risk of meta-vulnerability. Against this backdrop, we hope that the Big Tech companies will take concerns about meta-vulnerability seriously and not treat it as a tick box exercise.

We believe that governments and regulators need to prepare for the significant impact the metaverse is likely to have on society and its implications on the fundamental rights of individuals, especially those who are vulnerable or marginalized. We therefore make the following recommendations:

Governments and regulators must ensure that the businesses developing the metaverse follow the UN Guiding Principles on Business and Human Rights. This means that technology companies should be required to assess the impact of the metaverse on users' human rights, especially those at risk of being made vulnerable or marginalized. . Technology companies must also conduct meaningful consultation with affected groups and other stakeholders, such as CSOs.[11] Given that there are currently no clear guidelines for assessing the impact of social media platforms or the metaverse on fundamental human rights, governments and regulators need to establish guidelines for developing robust assessment models to measure the impact on the human rights of vulnerable groups in the metaverse.

To ensure good user experience and to tackle potential underrepresentation or marginalization, technology companies must be required to involve vulnerable groups in the participative design process of the metaverse. Governments and regulators must also set standards of best practice for the 'vulnerability-sensitive' design of this new technology, based on the lessons learnt from the participatory design process. Governments and regulators must clarify whether the current laws that prohibit sexual violence are applicable in the metaverse and address any gaps by enacting new laws and policies. Finally, the use of AI-driven emotional recognition, such as eye tracking or behavioural surveillance, of users in the metaverse must be reviewed and if necessary restricted.

[11] The Alliance for Universal Digital Rights (AUDRi) has developed a set of digital principles to inform global efforts for a digital future in which everyone can enjoy equal rights to safety, freedom and dignity: https://audri.org/digital-principles/

Glossary

AR: Augmented Reality. It refers to the technology that overlays computer-generated sensory information, such as images, sounds, or 3D models, onto the real world, enhancing the user's perception and interaction with their surroundings.

Avatar: In the context of virtual reality or online environments, an avatar refers to a digital representation or embodiment of a user. It is typically a customizable character that represents the individual in the virtual world and can be used to interact with others and navigate the digital space.

Global Immutable Ledger: A global immutable ledger refers to a decentralized and tamper-proof record-keeping system that maintains a transparent and unchangeable history of transactions or information. It typically utilizes blockchain or distributed ledger technology, where each transaction is verified and recorded in a permanent and transparent manner, ensuring the integrity and immutability of the data. In simpler words, it is a network or infrastructure that connects banks, financial institutions, and payment platforms around the world, enabling smooth and secure cross-border transactions. Just like a rail system transports goods across vast distances, a global payment rail facilitates the movement of money across borders, making it easier for businesses and individuals to engage in international transactions.

Global Payment Rail: A global payment rail refers to a network or infrastructure that facilitates the transfer of funds or value across borders and between different financial institutions or payment systems. It enables the seamless and efficient movement of money internationally, supporting various payment methods and currencies.

Governance Protocol: A governance protocol refers to a set of rules, processes, and mechanisms established to manage and make decisions within a decentralized system or network. It outlines how decisions are proposed, discussed, and implemented, ensuring the smooth operation and evolution of the system while involving stakeholders in the decision-making process.

Incentive Mechanism: An incentive mechanism is a system designed to motivate and reward individuals or participants within a network or ecosystem. It provides incentives, such as tokens or rewards, to encourage desired behaviors, contributions, or actions that align with the goals and objectives of the system.

Metaverse: The metaverse is a term used to describe a collective virtual shared space that encompasses multiple interconnected virtual worlds. It is often visualized as a virtual reality space where users can interact with each other and computer-generated environments in real-time.

MR: Mixed Reality. It refers to the merging of virtual and real-world environments, allowing users to interact with both physical and digital elements simultaneously. It combines elements of both Augmented Reality (AR) and Virtual Reality (VR). In MR, digital content is overlaid onto the real world in a way that it appears to coexist and interact with the physical surroundings. This technology enables users to see and manipulate virtual objects while maintaining a sense of presence in the real world. MR experiences often involve the use of specialized headsets or devices that overlay virtual content onto the user's view of the physical environment.

Trustless Participation: Trustless participation refers to the ability of individuals to engage in a decentralized system or network without the need to trust or rely on a centralized authority. It is achieved through cryptographic protocols and consensus mechanisms that ensure the security, integrity, and transparency of the system, eliminating the need for intermediaries and fostering a trustless environment.

VR: Virtual Reality. It refers to the technology that creates a simulated, computer-generated environment or experience. Users typically wear a headset that immerses them in a virtual world, allowing for interactive and immersive experiences.

XR: Extended Reality. XR is an umbrella term that encompasses the combination of AR, VR, and MR. It refers to technologies that blend the real and virtual worlds to create immersive and interactive experiences.

References

AccessNow. (2022). Virtual Worlds, Real People Human Rights in the Metaverse. https://www.accessnow.org/cms/assets/uploads/2021/12/Virtual-Worlds-Real-People-1c-1082766984.pdf

Anne Balsamo. (1999). Reading cyborgs writing feminism. In Jenny Wolmark (A c. Di), A Reader in Feminist Theory, Cyborgs and Cyberspace (pp. 145–156). Edinburgh University Press.

Ball, M. (2022). The Metaverse: And How It Will Revolutionize Everything (First edition). Liveright Publishing Corporation, a division of W.W. Norton & Company.

Bates, S. (2017). Revenge Porn and Mental Health: A Qualitative Analysis of the Mental Health Effects of Revenge Porn on Female Survivors. Feminist Criminology, 12(1), 22–42. https://doi.org/10.1177/1557085116654565

Butler, J. (2004). Precarious Life: The Powers of Mourning and Violence. Verso.

Champion, A. R., Oswald, F., Khera, D., & Pedersen, C. L. (2022). Examining the Gendered Impacts of Technology-Facilitated Sexual Violence: A Mixed Methods Approach. Archives of Sexual Behavior, 51(3), 1607–1624. https://doi.org/10.1007/s10508-021-02226-y

Chris Shilling. (2012). The Body and Social Theory. Sage.

Cole, A. (2016). All of Us Are Vulnerable, But Some Are More Vulnerable than Others: The Political Ambiguity of Vulnerability Studies, an Ambivalent Critique. Critical Horizons, 17(2), 260–277. https://doi.org/10.1080/14409917.2016.1153896

Council of the European Union. (2022). Metaverse — Virtual World, Real Challenges. https://www.consilium.europa.eu/media/54987/metaverse-paper-9-march-2022.pdf

Danaher, J. (2018). The law and ethics of virtual sexual assault. In W. Barfield & M. Blitz, Research Handbook on the Law of Virtual and Augmented Reality (pp. 363–388). Edward Elgar Publishing. https://doi.org/10.4337/9781786438591.00021

Donna Haraway. (1991). A Cyborg Manifesto: Science, Technology, and Socialist-Feminism in the Late Twentieth Century. In D. Haraway (A c. Di), Simians, Cyborgs, and Women: The Reinvention of Nature (pp. 149–182). Routledge.

Dowding, K. M. (1996). Power. Open University Press.

Dwivedi, Y. K., Hughes, L., Baabdullah, A. M., Ribeiro-Navarrete, S., Giannakis, M., Al-Debei, M. M., Dennehy, D., Metri, B., Buhalis, D., Cheung, C. M. K., Conboy, K., Doyle, R., Dubey, R., Dutot, V., Felix, R., Goyal, D. P., Gustafsson, A., Hinsch, C., Jebabli, I., … Wamba, S. F. (2022). Metaverse beyond the hype: Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. International Journal of Information Management, 66, 102542. https://doi.org/10.1016/j.ijinfomgt.2022.102542

Edwards, L., Schafer, B., & Harbinja, E. (2020). The Future's Already Here: It's Just Unevenly Edited. In L. Edwards, B. Schafer, & E. Harbinja, Future Law: Emerging Technologies, Regulation and Ethics (pp. 1–10). Edinburgh University Press. https://doi.org/10.1515/9781474417631-005

European Commission. (2022). Virtual worlds (metaverses) – A vision for openness, safety and respect. https://ec.europa.eu/info/law/better-regulation/have-your-say/initiatives/13757-Virtual-worlds-metaverses-a-vision-for-openness-safety-and-respect_en

Fineman, M. A. (2008). The Vulnerable Subject: Anchoring Equality in the Human Condition. Yale Journal of Law and Feminism, 20, 23.

Fineman, M. A. (2019). Vulnerability and Social Justice. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.3352825

Fink, M. (2018). Blockchain Regulation and Governance in Europe, Cambridge University Press.

Florencia Luna. (2009). Elucidating the Concept of Vulnerability: Layers Not Labels. International Journal of Feminist Approaches to Bioethics, 2(1), 121–139.

Forst, R. (2015). Noumenal Power. Journal of Political Philosophy, 23(2), 111–127. https://doi.org/10.1111/jopp.12046

Frank Rudy Cooper. (2015). Always Already Suspect: Revising Vulnerability Theory. North Carolina Law Review, 93, 1340–1379.

García-Sánchez, E. (2016). Cosmetic Vulnerability: The New Face of Human Frailty. In A. Masferrer & E. García-Sánchez (Eds.), Human Dignity of the Vulnerable in the Age of Rights: Interdisciplinary Perspectives (pp. 189–217). Springer International Publishing. https://doi.org/10.1007/978-3-319-32693-1_9

Genevieve Liveley. (2021). Beyond the Beautiful Evil? The Ancient/Future History of Sex Robots. In Classical Literature and Posthumanism. Bloomsbury.

Gennet, É., Andorno, R., & Elger, B. (2015). Does the new EU Regulation on clinical trials adequately protect vulnerable research participants? Health Policy, 119(7), 925–931. https://doi.org/10.1016/j.healthpol.2015.04.007

Gideon Burrows. (2022). Your Life in the Metaverse.

Gilson, E. C. (2014). The Ethics of Vulnerability: A Feminist Analysis of Social Life and Practice (1. publ). Routledge.

GREVIO. (2022). The digital dimension of violence against women as addressed by the seven mechanisms of the EDVAW Platform. https://rm.coe.int/thematic-report-on-the-digital-dimension-of-violence-against-women-as-/1680a933ae

Hackl, C., Lueth, D., & Bartolo, T. D. (2022). Navigating the metaverse: A guide to limitless possibilities in a WEB 3.0 world. Wiley.

Hearn, J., & Hall, M. (2022). From physical violence to online violation: Forms, structures and effects. A comparison of the cases of 'domestic violence' and 'revenge pornography'. Aggression and Violent Behavior, 67, 101779. https://doi.org/10.1016/j.avb.2022.101779

Hewer, R. M. (2019). A Gossamer Consensus: Discourses of Vulnerability in the Westminster Prostitution Policy Subsystem. Social & Legal Studies, 28(2), 227–249. https://doi.org/10.1177/0964663918758513

Huynh-The, T., Pham, Q.-V., Pham, X.-Q., Nguyen, T. T., Han, Z., & Kim, D.-S. (2022). Artificial Intelligence for the Metaverse: A Survey. https://doi.org/10.48550/ARXIV.2202.10336

Jennifer Gonzalez. (1999). Envisioning Cyborg Bodies: Notes from Current Research. In Jenny Wolmark (A c. Di), Cybersexualities: A Reader in Feminist Theory, Cyborgs and Cyberspace (pp. 264–279). Edinburgh University Press.

Judith Lorber & Patricia Yancey Martin. (2013). The Socially Constructed Body: Insights From Feminist Theory. In P. Kivisto (A c. Di), Illuminating Social Life: Classical and Contemporary Theory Revisited (pp. 226–244). SAGE Publications, Inc. https://doi.org/10.4135/9781506335483

Katherine N. Hayles & Jenny Wolmark. (1999). The Life of Cyborgs: Writing the Posthuman. In Cybersexualities: A Reader in Feminist Theory, Cyborgs and Cyberspace (pp. 157–173). Edinburgh University Press.

Kelly, L. (1988). Surviving Sexual Violence. Polity Press ; B. Blackwell.

Luna, F. (2019). Identifying and evaluating layers of vulnerability – a way forward. Developing World Bioethics, 19(2), 86–95. https://doi.org/10.1111/dewb.12206

Ma, W., & Huang, K. (2022). Blockchain and Web3: Building the Cryptocurrency, Privacy, and Security Foundations of the Metaverse. Wiley.

Mackenzie, C., Rogers, W., & Dodds, S. (2013). Introduction: What Is Vulnerability, and Why Does It Matter for Moral Theory? In C. Mackenzie, W. Rogers, & S. Dodds (A c. Di), Vulnerability: New Essays in Ethics and Feminist Philosophy (pp. 1–30). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199316649.003.0001

MacKinnon, R. (2006). Virtual Rape. Journal of Computer-Mediated Communication, 2(4), 0–0. https://doi.org/10.1111/j.1083-6101.1997.tb00200.x

Malgieri, G., & Niklas, J. (2020). Vulnerable data subjects. Computer Law & Security Review, 37, 105415. https://doi.org/10.1016/j.clsr.2020.105415

Mark Zuckerberg. (2021). Founder's Letter, 2021. https://about.fb.com/news/2021/10/founders-letter/

Martha Albertson Fineman. (2008). The Vulnerable Subject: Anchoring Equality in the Human Condition. Yale Journal of Law and Feminism, 20(1), 1–23.

Martha Albertson Fineman. (2012). Beyond Identities: The Limits of an Anti-discrimination Approach to Equality. Boston University Law Review, 92, 1713–1770.

McGlynn, C., Johnson, K., Rackley, E., Henry, N., Gavey, N., Flynn, A., & Powell, A. (2021). 'It's Torture for the Soul': The Harms of Image-Based Sexual Abuse. Social & Legal Studies, 30(4), 541–562. https://doi.org/10.1177/0964663920947791

Meghan Bobrowsky. (2021, 9 November). Big Tech Seeks Its Next Fortune in the Metaverse. The Wall Street Journal. https://www.wsj.com/articles/big-tech-seeks-its-next-fortune-in-the-metaverse-11636459200

Meta. (2022). Helping to build a diverse, equitable and inclusive metaverse. https://www.metacareers.com/life/helping-to-build-a-diverse-equitable-inclusive-metaverse

Micklitz, H.-W., Stuyck, J., & Terryn, E. (2010). Cases, Materials and Text on Consumer Law. Hart. https://cadmus.eui.eu/handle/1814/13750

Micky Lee. (2021). Feminist Scholarship on the Global Digital Divide. A Critique of International Organizations and Information Companies. In Dal Yong Jin (A c. Di), The Routledge Handbook of Digital Media and Globalization (pp. 66–76). Routledge.

Mireille Hildebrandt. (2021). The Issue of Bias. The Framing Powers of Machine Learning. In M. Pelillo & T. Scantamburlo (A c. Di), Machines We Trust: Perspectives on Dependable AI (pp. 43–60). The MIT Press.

Narula, H. (2022). Virtual Society (First Edition). Currency.

Natalie Boero & Katherine Mason. (2021). Toward a Sociology of the Body. In N. Boero & K. Mason (A c. Di), The Oxford Handbook of the Sociology of Body and Embodiment (pp. 1–20). Oxford University Press. B

Nicole L. Asquith, Isabelle Bartkowiak-Théron, & Karl A. Roberts. (2017). Police Encounters with Vulnerability. Springer Berlin Heidelberg.

Northrop, J. M. (2012). Reflecting on Cosmetic Surgery: Body Image, Shame and Narcissism. Routledge.

Nussbaum, M. (2003). CAPABILITIES AS FUNDAMENTAL ENTITLEMENTS: SEN AND SOCIAL JUSTICE. Feminist Economics, 9(2–3), 33–59. https://doi.org/10.1080/1354570022000077926

Nussbaum, M. Craven. (2006). Frontiers of Justice: Disability, Nationality, Species Membership. Harvard Univ. Press; /z-wcorg/.

Peroni, L., & Timmer, A. (2013). Vulnerable groups: The promise of an emerging concept in European Human Rights Convention law. International Journal of Constitutional Law, 11(4), 1056–1085. https://doi.org/10.1093/icon/mot042

Plant, S. (1997). Zeroes + ones: Digital women + the new technoculture (1st ed). Doubleday.

Robert E. Goodin. (1985). Protecting the Vulnerable: A Reanalysis of Our Social Responsibilities. University of Chicago Press.

Ruvalcaba, Y., & Eaton, A. A. (2020). Nonconsensual pornography among U.S. adults: A sexual scripts framework on victimization, perpetration, and health correlates for women and men. Psychology of Violence, 10(1), 68–78. https://doi.org/10.1037/vio0000233

Ryan Esparza. (2018). "The Way I Felt": Creating A Model Statute to Address Sexual Offenses Which Utilize Virtual Reality. Criminal Law Practitioner, 4(5), 24–40.

Samantha Delouya. (2022, agosto 3). Match Group says it's stepping back from its metaverse dating plans, citing the economy and uncertainty about «what will and won't work» on the new platform. Business Insider. https://www.businessinsider.com/tinder-match-group-metaverse-dating-crypto-economic-slowdown-uncertainty-meta-2022-8?international=true&r=US&IR=T

Sangeeta Singh Kurtz & Lakshmi Rengarajan. (s.d.). AI to IRL: The Future of Dating.

Schermaier, M. J. (2005). Mistake, misrepresentation and precontractual duties to inform: The civil law tradition. In R. Sefton-Green (A c. Di), Mistake, Fraud and Duties to Inform in European Contract Law (1 ed., pp. 39–64). Cambridge Univ. Press.

Schroeder, D., & Gefenas, E. (2009). Vulnerability: Too Vague and Too Broad? Cambridge Quarterly of Healthcare Ethics, 18(2), 113–121. https://doi.org/10.1017/S0963180109090203

Schumacher, P. (2022). The metaverse as opportunity for architecture and society: Design drivers, core competencies. Architectural Intelligence, 1(1), 11. https://doi.org/10.1007/s44223-022-00010-z

Sharabi, L. L., & Caughlin, J. P. (2019). Deception in online dating: Significance and implications for the first offline date. New Media & Society, 21(1), 229–247. https://doi.org/10.1177/1461444818792425

Sofia P. Caldeira, Sander De Ridder, & Sofie Van Bauwel. (2018). Exploring the Politics of Gender Representation on Instagram: Self-representations of Femininity. DiGeSt. Journal of Diversity and Gender Studies, 5(1), 23. https://doi.org/10.11116/digest.5.1.2

Stephenson, N. (2008). Snow Crash. Bantam Books.

Stone, A. R. (1995). The war of desire and technology at the close of the mechanical age. MIT Press.

Strikwerda, L. (2015). Present and Future Instances of Virtual Rape in Light of Three Categories of Legal Philosophical Theories on Rape. Philosophy & Technology, 28(4), 491–510. https://doi.org/10.1007/s13347-014-0167-6

Strycharz, J., & Duivenvoorde, B. (2021). The exploitation of vulnerability through personalised marketing communication: Are consumers protected? Internet Policy Review, 10(4). https://policyreview.info/articles/analysis/exploitation-vulnerability-through-personalised-marketing-communication-are

Tran, D. A., & Krishnamachari, B. (2022). Blockchain in a Nutshell. In D. A. Tran, M. T. Thai, & B. Krishnamachari (A c. Di), Handbook on Blockchain (Vol. 194, pp. 3–54). Springer International Publishing. https://doi.org/10.1007/978-3-031-07535-3_1

UNDP. (2022). Traversing the metaverse whilst managing risks with opportunities. https://www.undp.org/blog/traversing-metaverse-whilst-managing-risks-opportunities

UNESCO. (2020). Online Violence against Women Journalists: A Global Snapshots of Incidence and Impacts. https://unesdoc.unesco.org/ark:/48223/pf0000375136

Véliz, C. (2020). Privacy is Power: Why and How You Should Take Back Control of Your Data. Bantam Press.

Vera-Gray, F. (2020). Rape Porn, Cultural Harm, and the Law. In K. Ross, I. Bachmann, V. Cardo, S. Moorti, & M. Scarcelli (A c. Di), The International Encyclopedia of Gender, Media, and Communication (1ª ed., pp. 1–5). Wiley. https://doi.org/10.1002/9781119429128.iegmc032

Wang, G., Badal, A., Jia, X., Maltz, J. S., Mueller, K., Myers, K. J., Niu, C., Vannier, M., Yan, P., Yu, Z., & Zeng, R. (2022). Development of metaverse for intelligent healthcare. Nature Machine Intelligence, 4(11), 922–929. https://doi.org/10.1038/s42256-022-00549-6

Weber, M. (1978). Economy and Society: An Outline of Interpretive Sociology. University of California Press.

Weissman, J. (2021). The Crowdsourced Panopticon: Conformity and Control on Social Media. Rowman & Littlefield.

Wendy Rogers, Catriona Mackenzie, & Susan Dodds. (2012). Why bioethics needs a concept of vulnerability. International Journal of Feminist Approaches to Bioethics, 5(2), 11–38.

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

World Economic Forum. (2022). New Initiative to Build An Equitable, Interoperable and Safe Metaverse. https://www.weforum.org/press/2022/05/new-initiative-to-build-an-equitable-interoperable-and-safe-metaverse/

Zuboff, S. (2017). The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power (1st ed). Public Affairs.

Alliance For
Universal
Digital Rights

AUDRi.org

VULNERA
THE INTERNATIONAL
OBSERVATORY

brusselsprivacyhub.com/vulnera