

Liability for Wrongful Behaviour in the Metaverse

by Lorena Arismendy Mengual *

Abstract: Although the Metaverse presents various potential legal issues including cybersecurity problems, jurisdictional conundrums, an obscure characterization of digital property, and personal data protection just to name a few. This paper specifically focuses on those issues arising from avatar misconduct in online virtual worlds. It is argued that harm suffered by a person may be caused by or through an avatar and that this argument hinges on whether avatars are recognised as legal persons and the lack of said recognition. Currently, avatars do not have legal personhood, making it an essential task to identify a

liable party, prove harm and causation, and establish a basis for imputation to successfully bring a claim for compensation before national courts of Law. These issues must be addressed under existing regulations, highlighting the need for new approaches to handle these situations effectively when adequate. Civil liability for tortious conduct is examined under civil law systems to shed some light on whether actions by or through avatars in online virtual environments can be translated into this framework, leading to the enforceable legal consequence that is compensation for harm suffered.

Keywords: Metaverse, Online Platforms, Avatars, Civil Liability, Fault, Damage

© 2024 Lorena Arismendy Mengual

Everybody may disseminate this article by electronic means and make it available for download under the terms and conditions of the Digital Peer Publishing Licence (DPPL). A copy of the license text may be obtained at <http://nbn-resolving.de/urn:nbn:de:0009-dppl-v3-en8>.

Recommended citation: Lorena Arismendy Mengual, Liability for Wrongful Behaviour in the Metaverse, 15 (2024) JIPITEC 229 para 1

A. Introduction

- 1 In the dynamic landscape of the Metaverse, avatars are the digital embodiment of users, facilitating interactions and representing their virtual identity. As technology continues to evolve, the regulatory framework and the legal implications of interactions through avatars remain a subject of interest and debate. Therefore, this research paper explores the framework of non-contractual civil law claims for damage caused through an avatar in virtual worlds, where such claims are brought under a fault-based liability regime.
- 2 Although the Metaverse presents various potential legal issues, including cybersecurity problems, jurisdictional challenges, an obscure characterization of digital property, and personal data protection, this paper focuses on those issues arising from avatar misconduct in virtual worlds, considering that the Metaverse is currently dominated by so-called “walled gardens.”
- 3 With virtual interactions becoming more prevalent and present in our daily lives, and with the advent of web4.0, the infringement of rights and other legally protected interests multiplies exponentially. Consider, for instance, real cases of alleged virtual

* Dr. Lorena Arismendy Mengual is a Private Law Professor at CUNEF University; ORCID: <https://orcid.org/0000-0002-9969-9186>; The author expresses gratitude to the Institute of European and Comparative Law (IECL) at Oxford University for their insights during the IECL Seminar Series in Trinity term 2024. Acknowledgment is also given to Keble College for its funding and support during the completion of this research.

sexual violence or assault already taking place.¹ Although criminal offences are outside the scope of this paper, these reports are a reminder that misconduct in digital environments, especially when immersive, should be taken seriously to better understand the path towards adequate relief for any compensable harm caused in a digital reality. It is argued in this study that despite their digital nature, avatars can be the source of real-world harm and thus are subject to real-world legal principles. By exploring cases where harm inflicted through avatars must meet with established legal regimes, insights are provided into the institution of civil liability and compensation.

- 4 This paper used a normative juridical method to approach these issues, incorporating statute, case, analytical, and comparative approaches. The study relied on secondary data gathered through literature reviews. Primary and secondary legal materials were used, and information was analysed using mostly a descriptive-qualitative method. After considering a deemed necessary differentiation between the concept of the Metaverse with that of online virtual worlds, an in-depth exploration of avatars in the Metaverse and online virtual worlds is addressed, particularly focusing on the impact of wrongful behavior performed by users when it is harmful. The question of AI-equipped avatars and prohibited AI practices is also addressed. The discussion then raises questions about legal personhood and the appropriate remedial responses to address the harm caused via avatars as well as key issues in fault-based liability regimes, including fault, causation, and compensation.

1 The feelings of disorientation and confusion experienced by a SumOfUs researcher –only two hours into navigating in Meta Horizon Worlds in 2022–, are as real as the result of other types of socially acceptable scarring experiences. In this case, while using a female-looking, the researcher’s avatar was lured into a private room during a virtual party, where another user allegedly, non-consensually approached her in such a way that she described as *rape*. The case exemplifies that the digital reality perceived in online virtual worlds by real-world users can trigger legally relevant responses to harmful conduct when performed by digital means of an avatar. The full report can be consulted here: <https://www.eko.org/images/Metaverse_report_May_2022.pdf>https://www.eko.org/images/Metaverse_report_May_2022.pdf> Accessed: 26 Sept. 2024.

B. Preliminary Remarks

- 5 As an opening remark, it is stressed that it is necessary to distinguish between the Metaverse and metaverses because legal issues arising from these two concepts can differ significantly.
- 6 The Metaverse (singular and with a capital “M”) refers to a unique, interconnected technological landscape which is envisioned as a seamless, persistent online realm where users can interact with digital elements as well as with other users in real-time, which has been described as living in a *digital reality*.² Although there is no consensus on a definitive definition of the Metaverse, it has been argued that the Metaverse is not simply a place or a destination accessible through a virtual reality device. Instead, it represents a complex technological environment that ultimately transforms habits, daily activities, acts and/or legal transactions as they occur in the real world, with equivalent ones in a digital world. Bearing this in mind, the Metaverse can be defined as a unique digital ecosystem which, through the use of different technologies, allows the physical and digital aspects of people’s lives to converge in the same immersive experience in such a way that users perceive a persistent, synchronous, and interoperable environment where they can seamlessly transition between different digital spaces, engaging in social, economic, commercial, labour, cultural, industrial, legal, political, and other activities within a digital reality.³ According to the leading technology scholarly opinion, the Metaverse is currently experiencing a development stage. However, experts anticipate that this evolution could occur relatively swiftly.⁴

2 T R Gadekallu, *et al.* ‘Blockchain for the Metaverse: A Review’ (2022) arXiv preprint arXiv:2203.09738, 7.

3 L M Arismendy Mengual ‘Legal Challenges of the Metaverse: Data Protection, Intellectual Property and Civil Liability’ (2023) 80 Cuadernos de Derecho y Comercio 74.

4 J Dionisio, *et al.* ‘3D virtual worlds and the metaverse: Current status and future possibilities’ (2013) 45(3) ACM Computing Surveys (CSUR) 2-3; D Wang, X Yan and Y Zhou ‘Research on Metaverse: Concept, development and standard system’ (2021) 2nd International Conference on Electronics, Communications and Information Technology (CECIT). IEEE, 983-991; C Hackl *et al* *Navigating the Metaverse* (John Wiley & Sons Newark 2022) 46.

- 7 On the other hand, metaverses (in lowercase and possibly in a plural form) refer to multiple, distinct online platforms showcasing digital worlds. These have been the focus of most scholarly attention over the past few decades, often leading to confusion between this concept and the singular, wider idea of the Metaverse. Online virtual worlds are mainly digital standalone platforms offering an immersive digital world experience where users can interact, create content, and participate in activities. Still, these do not necessarily connect or integrate. Such digital worlds include individual gaming environments, virtual social spaces, or enterprise-focused virtual meeting platforms like Second Life,⁵ Roblox,⁶ Fortnite,⁷ Minecraft⁸ or Meta Horizon Worlds.⁹ As Lastiri emphasizes, the idea of virtual worlds is not new among us. However, they are rapidly gaining more scholarly attention mainly due to the application and impact of blockchain technology in these scenarios.¹⁰
- 8 The legal issues arising from these two concepts are often not the same, as seen throughout this study. For these reasons, the scope of this paper will only address problems arising from avatar interactions in online virtual worlds that nonetheless take place in the Metaverse, notwithstanding some necessary

reference to the Metaverse –as a whole– whenever it is adequate.

C. Avatars in the Metaverse and Online Virtual Worlds

- 9 Avatars, as manifestations of digital data, serve as central components of virtual interaction. Designed for immersive experiences, they represent users (whether human or otherwise) rather than static elements within digital landscapes. In online virtual worlds, the user is their avatar; which acts as a conduit for user behavior. Avatars, lack inherent autonomy –unless powered by AI technologies. Moreover, unlike AI agents, they lack opacity or a “black box” effect.¹¹ This section addresses several relevant aspects of avatars to determine and assess civil liability for wrongful behavior in online virtual worlds; digital identity, some inquiries into a possible legal status, and the overall involvement of online platforms are considered.

I. Digital Identity in the Web3.0 and the Web4.0.

- 10 This section addresses the relevance of the connexion between avatars and their controllers through the notion of identity. It is herein considered that the question of the legal consequences of using avatars also falls within a broader digital identity framework. It is also argued that the human user’s perception of the avatar constitutes a fundamental element in establishing the legal relevance of using avatars. Understanding this perception is crucial for framing the legal discourse on avatars, as it underpins some implications of digital representation and the extent to which virtual actions may translate into legal consequences in real-world contexts.
- 11 Besides being considered a key factor to protect individuals and their digital interactions online –as set forth by the European Parliament and the Council and the Commission joint Declaration on Digital Rights and Principles for the Digital Decade,¹² the

5 <<https://secondlife.com/>> Accessed: 26 Sept. 2024. A good example of this is presented by early Metaverse researchers such as J Kemp and D Livingstone ‘Putting a Second Life “metaverse” skin on learning management systems’ (2006) Proceedings of the Second Life education workshop at the Second Life community convention, 12; A Kaplan and M Haenlein ‘The fairyland of Second Life: Virtual social worlds and how to use them’, (2009) 52 6 Business horizons 563-572; A Davis et al. ‘Avatars, people, and virtual worlds: Foundations for research in metaverses’ (2009) 10(2) Journal of the Association for Information Systems, 1; as well as more current works, cf. I Filipova ‘Creating the metaverse: consequences for economy, Society, and Law’ (2023) 1(1) Journal of Digital Technologies and Law 7-32.

6 <<https://www.roblox.com/>> Accessed: 26 Sept. 2024.

7 <<https://www.fortnite.com/>> Accessed: 26 Sept. 2024.

8 <<https://www.minecraft.net/>> Accessed: 26 Sept. 2024.

9 <<https://horizon.meta.com/>> Accessed: 26 Sept. 2024.

10 M Lastiri Santiago ‘Metaverse in the world of trademark law’ (2024) Uniform Law Review, 2.

11 Infra. Section C.IV. (AI-equipped avatars’ wrongful behavior).

12 For the purposes of this paper, identity is also approached

idea of identity relates to the sense of self, and engagement, for this work, within digital platforms. In this sense, a person will use an avatar to identify him or herself in a given digital world.¹³ This raises a concern for the extension and the applicability of this concept from physical to digital domains.

- 12 The immersive experience provided by online virtual worlds in the Metaverse is partly facilitated by the perception of a Metaverse digital reality through the avatar. The graphical representation of an avatar can range from realistic human likenesses to fantastical creatures, abstract shapes, or even inanimate objects. The level of detail can vary considerably depending on the platform where it is created, from highly detailed, lifelike models to simplistic, cartoonish designs. This study argues that this variety raises essential legal questions about the extent to which avatars are treated under existing law.

- 13 Customisation possibilities, although seemingly trivial and unimportant, are not without potential legal consequences, the extent of which remains to be determined. Some examples can be emphasised:

1. The visible appearance of the avatar may qualify as a protected work under European intellectual property protection rules to which the user may be entitled if it meets the necessary originality criteria, provided the platform in question has not reserved such rights under their Terms of Service.¹⁴

2. Changes made by platforms to avatars' appearance can directly impact users' rights to the digi-

tal goods they have acquired to customize them, which could constitute grounds for a contractual liability claim. In this sense, and regardless of the motivation, users might invest in their avatars, e.g., by purchasing digital clothing items or accessories with their real money. This makes it questionable for online platforms to unilaterally alter the overall look of avatars, as recently exemplified by Niantic's decision to update and alter avatar appearances in Pokémon Go, which has sparked controversy among users.¹⁵

3. The likeness of avatars may be subject to unauthorised use by another; as Lake argues, users could be recognised in their online communities based on the appearance and popularity of their avatars; which often results in other users being increasingly tempted to exploit the popularity of another's avatar and deceive or mislead others for personal gain.¹⁶ The use of avatars can therefore make an impact on real-world reputation and rights.

4. According to some studies, it is possible for users to develop deep psychological attachments to their digital twins. In this regard, digital identity in online virtual worlds certainly involves the concept of self-presence, which closely relates to the subjective feeling of existing within the digital sphere. While identifying themselves with their avatars, participants of a virtual world can experience what happens to the avatars in the virtual world as happening to themselves.¹⁷ Self-presence is a psychological condition wherein individuals perceive their virtual identity (avatars) as synonymous with their real-world persona.¹⁸

as a key factor to protect individuals and their digital interactions online, as set forth by the European Parliament and the Council and the Commission joint Declaration on Digital Rights and Principles for the Digital Decade, Brussels, 26 January 2022, COM(2022)28 final. Chapter II: Solidarity and inclusion, and Chapter V: Safety, security and empowerment.

- 13 A comprehensive definition and notion of identity can be found in D Parfit 'Personal Identity' (1971) 80(1) *The Philosophical Review*, 3-27.

- 14 For instance, Second Life allows users to retain intellectual property rights in their digital creations, including avatar characters and other types of digital objects, cf. <<https://lindenlab.com/legal/second-life-terms-and-conditions>> Accessed: 26 Sept. 2024; Also: *infra*. Section D. (Harm caused via avatars).

-
- 15 The press release of the event: <<https://pokemongolive.com/rediscovergo>> Accessed: 26 Sept. 2024.

- 16 J Lake 'Hey, You Stole My Avatar!: Virtual Reality and Its Risks to Identity Protection' (2020) 69 *Emory L. J.* 836.

- 17 J M Balkin 'Virtual Liberty: Freedom to Design and Freedom to Play in Virtual Worlds' (2004) 90(8) *Virginia Law Review*, 2048.

- 18 J-A Lee, L Yang and P Hui 'Legal implications of self-presence in the metaverse' (2023) 25(4) *Media & Arts Law Review*, 268. The authors follow Belk's 'extended self' theory and argue that there is a digital equivalence concerning the connection between users and their digital identities. R W

This argument has been supported by several legal, psychological, and sociological studies, which also provide a basis for potential claims for moral damages.¹⁹

- 14 While acknowledging the ongoing nature of these discussions, the above underscores the possibility of compensable harm being caused to human users, and that it may ground a claim for damages where legal interests can be infringed.
- 15 It is also argued here that one must consider that avatars might not be a reliable source for identifying individuals in online virtual worlds. The nature of interactions in virtual worlds complicates identifying individuals responsible for harmful conduct, potentially impeding the pursuit of compensation for damages. This is founded on two main reasons: (i) the Metaverse is currently composed of “walled gardens”, meaning each virtual world operates independently with its own set of rules, systems, and user data. This fragmentation prevents a unified method of identification across different platforms; (ii) a user may create and operate in an online virtual world with multiple avatars rather than a singular one, making it difficult to associate a specific avatar with a particular individual consistently.²⁰
- 16 Otherwise, as far as technology goes and similarly argued by some scholars, other technologies do a better job for accurately and legally identifying users in online virtual worlds.²¹ For example, blockchain-based ID protocols have emerged as a considerable solution. These protocols could make use of the public, transparent and decentralised nature of blockchain technology to establish secure and immutable digital identities, offering interesting authentication and verification mechanisms within virtual environments.²²
- 17 Additionally, digital avatars will likely play a significant role in the so-called Web 4.0 (an autonomous, interconnected, interoperable, immersive network), according to the recent new EU strategy on Web 4.0 and Virtual Worlds.²³ Digital avatars are, in fact, a central part of virtual worlds and the Metaverse envisioned in Web 4.0. The EU strategy, therefore, aims for virtual worlds reflecting EU values and principles, where people’s rights fully apply.²⁴
- 18 The EU’s commitment to ensuring that virtual worlds reflect values and rights indicates a growing acknowledgment of avatars’ legal implications, reinforcing the need to address these issues as technology and virtual worlds continue to evolve. This should necessarily consider the legal implications of avatar conduct in the Metaverse, which will be addressed in the following section.

Belk, ‘Possessions and the Extended Self’ (1988) 15(2) *Journal of Consumer Research*, 139–168; R W Belk, ‘Extended Self in a Digital World’ (2013) 40(3) *Journal of Consumer Research*, 478; Although current, this notion is not new, *vid.* J W Penney, ‘Privacy and the New Virtualism’ (2008) 10 *Yale J.L. & Tech.* 221.

19 *Infra.* Section D. (Harm caused via avatars) of this study; F G Lastowka and D Hunter, ‘The Laws of the Virtual Worlds’, (2004) 92 *Cal. L. Rev.* 1, 73; Balkin (n17) 2043; S Triberti et al ‘Changing Avatars, Changing Selves? The Influence of Social and Contextual Expectations on Digital Rendition of Identity’ (2017) 20(8) *Cyberpsychology, Behavior and Social Networking* 501–507.

20 These ideas are emphasised in L M Arismendy Mengual ‘A legal status for Avatars in the Metaverse from a Private Law perspective’ (2024) 2 *InDret* 109. A clear explanation on online virtual worlds governance from a technological perspective can be found in T R Gadekallu (n2) at 2, 8, 10, 13–14.

21 See for instance the proposal for a registration system for online personas. J Bryson et al. ‘Of, for, and by the people: the legal lacuna of synthetic persons’ (2017) 25 *Artificial Intelligence and Law* 273–291.

22 See for instance N Schreier R Renwick and T Ehrke-Rabel (2021). ‘The Digital Avatar on a Blockchain: E-Identity, Anonymity and Human Dignity’ 2(3) *Austrian Law Journal* 202–218.

23 More information on this initiative is available here: <https://ec.europa.eu/commission/presscorner/detail/en/ip_23_3718> Accessed: 26 Sept. 2024.

24 According to the recent *Trend Report of Virtual Worlds (Metaverse)* published on 24 May 2024 by the Directorate-General for Communications Networks, Content and Technology, see: <https://blockchain-observatory.ec.europa.eu/publications/trend-report-virtual-worlds-metaverse_en?prefLang=et> Accessed 28. May 2024.

II. Wrongful Behavior in Online Virtual Worlds. A Brief Assessment of the Impact of the Terms of Services

- 19 For this study, wrongful behavior—regardless of the perpetrator—is outlined narrowly. Hence, this paper focuses on addressing harmful misconduct, and finds an obligation to compensate, also disregarding criminal offences.²⁵
- 20 Wrongful conduct in virtual platforms often involves actions that, if committed in the physical world, would fall under established civil liability fundamentals such as damage —mainly²⁶— fault or negligence as well as causation. As virtual worlds become more immersive and realistic, the lines between virtual and physical worlds blur. As stressed above, actions taken in online virtual worlds can have real-world consequences, both psychological and patrimonial. Therefore, it is critical to identify a clear legal framework to address wrongful conduct in these spaces, as the impact can be just as severe as in the physical world, as is argued herein.
- 21 In online virtual worlds, covenants set out in End User License Agreements (EULA) or in the Terms of Service (ToS) are not a mere formality for access to metaverses. Indeed, these agreements are legally binding for the parties involved and cover a range of matters, including detailed regulations of behavior, rights, and obligations of users. They fundamentally establish the limits of what is allowed in each virtual world while also granting the platforms extensive

moderation powers over user conduct.²⁷

- 22 This provides a context for online interactions among users within an online virtual world but does not set out rules for harm that may be caused, e.g., to a third party. Moreover, as will be explored in this study, even if users agree to the ToS, this does not exempt anyone who causes compensable damage to another person from tort liability. Even with these agreements, online virtual world platform owners still face significant difficulties in ensuring user safety within virtual worlds. The sheer volume of user interactions makes it difficult for platform owners to oversee and manage every occurrence of misconduct, which they are also not obliged to do.²⁸ Another section of this work considers the scope of their obligation in light of the recent EU Digital Services Act (DSA).²⁹

III. Misconduct by or Through an Avatar; Queries on Legal Personhood and Remedial Responses. A Matter of Control

- 23 This section addresses the question of whether avatars themselves can bear liability for harmful interactions in online virtual worlds, considering it a prerequisite to the attribution of liability in any legal system worldwide.
- 24 Legal personhood is typically ascribed to natural persons and legal entities in most legal systems. Avatars, as digital proxies for users, do not possess independent legal personhood today.

25 This clarification is deemed necessary, as for scholars from common law systems, “wrong” typically refers to torts, breaches of contract, and breaches of confidence or confidentiality (whether contractual or not). More broadly, according to the English legal perspective, a “wrong” involves a breach of duty. This broader definition arises from the fundamental effort not to limit the concept of wrongs exclusively to torts or fault. P Birks. *Introduction to the Law of Restitution* (Clarendon Press, Oxford, 1985) 313; A Burrows *English Private Law* (3rd edn Oxford University Press Oxford 2013).

26 Notwithstanding that, in civil law, there are several possibly applicable liability regimes. See *infra*. Section E. (Main issues in fault-based liability regimes. Considerations on fault, causation, fault, and compensation).

27 J Langenderfer ‘End-User License Agreements: A New Era of Intellectual Property Control’ (2009) 28(2) *Journal of Public Policy & Marketing*, 202-211; J M Balkin ‘Law and Liberty in Virtual Worlds’ In *The State of Play* (2nd edn. New York University Press New York 2020) 86-118.

28 This has expressly been stated by the Court of Justice of the European Union. See for instance: Case 70/10 *Scarlet Extended SA v Société belge des auteurs, compositeurs et éditeurs SCRL (SABAM)*, ECLI:EU:C:2011:771. Recital 30 of the Regulation (EU) 2022/2065 (Digital Services Act) –Further addressed in Section C.V. (Other liable agents? Views on secondary or vicarious liability)– confirms the same stance.

29 *Infra* Section C.V. (Other liable agents? Views on secondary or vicarious liability).

- 25 The need for proper regulation of avatars has been explored for over a decade,³⁰ coinciding with the launch and popularisation of virtual reality platforms. Scholars advocating for legal personhood of avatars mainly argue over a *de lege ferenda* viability of a recognition analogous to that already granted to juristic persons.³¹ A similar argument has been made regarding AI agents, suggesting they should be given legal personhood due to their autonomous agency and complex capabilities. However, it should be useful to remark that this argument has not changed their legal status; AI systems have not been recognised with any form of independent or intermediary legal personhood despite their increasing autonomy from programmers or designers. Moreover, following the implementation of the Regulation (EU) 2024/1689 (AI Act),³² there has been no such recognition. This can be considered a clear indication of how the European Union intends to approach the issue. Other minoritarian approaches consider, e.g., a *tertium genus* affairs parallel to the e-personality posed by the Committee on Legal Affairs of the European Parliament for AI agents.³³
- 26 Minding this situation, a further case for avatar legal personality has risen under the premise that avatars may be endowed with self-learning capabilities or may be able to make autonomous decisions (which may be wrongful) and should be treated separately from their controller, at least for all civil liability purposes. This is the case of the so-called increasingly smart avatars.³⁴
- 27 To be sure, it was argued that avatars can identify a user, the avatar being a direct representation of the user in the online virtual worlds, also possibly allowing a person to create a psychological bond with this digital immersive element. While this is herein deemed accurate, for the purposes of online virtual worlds interaction dynamics and civil liability law, avatars must be considered tools and mere conduits for another's actions, whether legal or not. It should also be considered that said control upon avatars can be asserted by an AI agent. Therefore, the avatar entails a virtual or digital character that is completely controlled by another, making control a key aspect to further analyse the legal consequences of using them.
- 28 By following this reasoning, responsibility for wrongful behavior using avatars should primarily fall on the controlling entity (human or otherwise); hence, as it stands, harmful behavior in online virtual worlds can only be committed *through* an avatar, rather than by the avatar itself.³⁵
- 29 Moreover, an avatar in an online virtual world more closely resembles a digital good or service rather than a person. To determine its nature in this regard, it is helpful to consider whether the avatar remains within the online virtual world's cloud and can only be accessed while logging in and utilizing the online platform. In such cases, avatars are more akin to digital services than digital goods.³⁶ In this sense, and
-
- 30 A M Franks 'Unwilling avatars: Idealism and discrimination in cyberspace' (2011) 20 Colum. J. Gender & L. 224.
- 31 T Day, 'Avatar Rights in a Constitutionless World', (2009) 32 Hastings Comm. & Ent. L.J., 151; B C Cheong 'Avatars in the metaverse: potential legal issues and remedies', (2022) International Cybersecurity Law Review, 5; It has been argued that avatars bear a closer resemblance to companies than AI agents do. See for instance L M Arismendy Mengual (n20) 112-117'.
- 32 Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828. Accessed: 23 Sept. 2024.
- 33 Said notion of e-personality was intended for overly sophisticated and autonomous AI agents. European Parliament (EP) 'Motion for a European Parliament Resolution' CLA 2015/2103(INL), 27 January 2017 <https://www.europarl.europa.eu/doceo/document/A-8-2017-0005_EN.html> Accessed: 26 Sept. 2024; This proposal was clearly not accepted by renowned scholars and experts, as argued in the Open Letter to The European Commission Artificial Intelligence and Robotics, Available in: <<http://www.robotics-openletter.eu>> Accessed: 26 Sept. 2024.
- 34 W Barfield and A Williams 'Chapter 1: The law of virtual reality and increasingly smart virtual avatars', in *Research Handbook on the Law of Virtual and Augmented Reality* (Edward Elgar, Cheltenham 2018), 2-43.
- 35 This statement will hold true as long as no legal personhood is endowed to avatars or to AI-agents.
- 36 M P García Rubio 'Non Conformity of Goods and Digital Content and its Remedies' in *European Perspectives on the*

according to article 2.2. of Directive (EU) 2019/770,³⁷ digital service means a service that allows the consumer to create, process, store or access data in digital form; or a service that allows the sharing of or any other interaction with data in digital form uploaded or created by the consumer or other users of that service means data which are produced and supplied in digital form. From this broad and ambiguous definition, the avatar can accurately be perceived as a digital service that can be supplied by a digital service provider (labelled trader under this Directive), which in today's Metaverse would be the centralised digital world or the platform on which it is created. The Directive above does not provide many useful elements for this discussion, as it excludes many activities that already occur or are planned to be undertaken in online virtual worlds, such as gambling, health, and financial services (as specified in Article 3.5 of the Digital Content Directive). Additionally, the Directive is designed solely to regulate B2C (business-to-consumer) relationships, focusing on interactions between the trader and consumers.³⁸

IV. AI-Equipped Avatars' Wrongful Behavior

- 30 This section addresses the essential question of scenarios where AI-powered avatars might influence or exploit users with the objective to or the effect of materially distorting human behavior, leading to real-world harm. The issue is approached particularly in light of Article 5(1)(a) and (b) of the AI Act.
- 31 It is herein maintained that if avatars are powered by AI systems—whether fully or partially—they should not be legally distinguishable from other AI agents. As a result, all regulatory frameworks applicable to

AI agents would also apply to avatars in online virtual worlds and the Metaverse. Bearing this in mind, the rapid evolution of AI technology raises important questions about the potential for avatars to engage in increasingly autonomous wrongful actions. This development calls for a closer examination of its legal implications, including liability issues and potential remedies.

- 32 Although the AI Act does not resolve the issue of civil liability, it establishes important fundamental lines, such as the concept of artificial intelligence systems—as per article 3(1) of the AI Act³⁹ that will be ultimately applicable to AI powered avatars. This is also the case of the risk-based classification of AI systems that structures the regulation. It means that the deployment of an AI-powered avatar may entail different levels or risks; namely, unacceptable, high-risk, limited-risk, minimal-risk or no risk whatsoever.⁴⁰ This approach becomes particularly relevant when determining whether certain AI practices involving avatars fall under prohibited activities.
- 33 At the core of determining if an AI system practice that has been used or deployed in the market is prohibited, a joint assessment of provisions 5(1)(a) and (b) of the AI Act reveals three key aspects:
- The objective or the effect of the AI agent's deployment is to materially distort the behavior of a person or group of persons. Consequently, the provider or the deployer's intention—or lack thereof—to cause harm is ultimately irrelevant.
 - The AI system causes individuals to take decisions they would not have otherwise taken.
 - The AI system's intervention causes or it is reasonably likely to cause significant harm to a person, or a group of persons.
- 34 From this point on, the regulation differs depending on whether it addresses manipulative or exploitative practices that cause harm. Article 5(1)(a) specifically requires (i) that the AI system deploys subliminal

Common European Sales Law Studies in European Economic Law and Regulation (Springer New York 2015).

37 Directive of the European Parliament and of the Council of 20 May 2019 on certain aspects concerning contracts for the supply of digital content and digital services regulates contracts for the supply of digital content and services.

38 Further explanation on this topic is presented by J M Carvalho 'Sale of Goods and Supply of Digital Content and Digital Services – Overview of Directives 2019/770 and 2019/771' (2019) 8(5) *Journal of European Consumer and Market Law*, 194–201.

39 Cf. Recital 12 of the AI Act.

40 Critical considerations on the lack of a clear methodology for risk assessment under the AI Act are presented by C Novelli et al 'AI Risk Assessment: A Scenario-Based, Proportional Methodology for the AI Act' (2024) 3(1) *Digital Society* 13–26.

techniques beyond a person's consciousness (e.g. imperceptible audio, image, video stimuli),⁴¹ or that it purposefully uses manipulative or deceptive techniques; (ii) the person's ability to make an informed decision must be appreciably impaired. On the other hand, Article 5(1)(b) addresses the exploitation of vulnerabilities –due to age, disability, a specific economic or social situation (e.g., extreme poverty, ethnic or religious minorities⁴²)– of a specific group of persons.

- 35 Many of the concepts introduced by the regulation as it stands foster greater uncertainty rather than providing much-needed clarity. For instance, it is difficult to identify a specific AI-powered avatar that employs subliminal techniques⁴³ and that is able to imperceptibly manipulate another user to effectively change their behavior 'beyond a person's consciousness' under Article 5(1)(a) –also, what does the latter term mean, and how can it be proven by the victim.⁴⁴ It should also be noted that the final

text of these provisions suggests that manipulative or exploitive practices without harm are nonetheless acceptable.

- 36 Avatar-related misconduct, particularly when involving AI-driven behaviors –even if not fully autonomous legal entities– can cause real-world harm and should be subject to the same regulatory scrutiny as other AI systems. This raises a question in scenarios where the avatar is partially user-controlled yet also possesses AI autonomous capabilities.⁴⁵ Some noteworthy criteria have been raised in this regard and are also echoed here: The identification of a liable agent –who primarily operates the technology– may heavily depend on whether the service provider (ensuring the necessary technical framework for its operation) has a higher degree of control than the owner or user of an AI-equipped product or service, v.gr., the avatar. Also, the view that a person using an avatar with a certain degree of autonomy should not be held less accountable for any resulting harm than if that harm had been caused by a human auxiliary can also be herein supported.⁴⁶

41 Recital 29 of the AI Act.

42 Recital 29 of the AI Act, cf. Directive (EU) 2019/882 of the European Parliament and of the Council of 17 April 2019 on the accessibility requirements for products and services.

43 Although prohibitions against subliminal techniques date from more than five years, cfr. Article 9(1)(b) of the Audiovisual Media Services Directive –Directive (EU) 2018/1808 of the European Parliament and of the Council of 14 November 2018 amending Directive 2010/13/EU on the coordination of certain provisions laid down by law, regulation or administrative action in Member States concerning the provision of audiovisual media services in view of changing market realities. A discussion on this topic can be found in M Franklin, et al. 'Missing Mechanisms of Manipulation in the EU AI Act' (2022) The International FLAIRS Conference Proceedings 2022.

44 The issue closely relates to the so-called dark patterns introduced in Recital 67 of the DSA regarding 'practices that materially distort or impair, either on purpose or in effect, the ability of recipients of the service to make autonomous and informed choices or decisions'; An in-depth assessment of the lack of clarity posed by Articles 5.1.a and 5.1.b of the AI Act can be found in M Leiser 'Psychological Patterns and Article 5 of the AI Act: AI-Powered Deceptive Design in the System Architecture and the User Interface' (2024) 1(1) Journal of AI law and Regulation, 6-14; Also: H Zhong et al. 'Regulating AI: Applying Insights from Behavioural Economics and Psychology to the Application of Article 5 of the EU AI Act'. (2024) 38 Proceedings of the AAAI Conference on Artificial Intelligence, 20001-20009.

V. Other Accountable Agents? Views on Secondary or Vicarious Liability

- 37 Before further examining the legal configuration of a claim for damages arising from misconduct in online virtual worlds, it is important to consider the overall involvement of these worlds, which may lead to liability risks due to their platform operations.

45 While users can generate scripts or pre-defined tasks for their avatars when offline, no evidence of user's avatars AI-driven operation was found in this research. However, technological advancements suggest this could happen in the future. For example, Somnium Space (a virtual world platform < <https://somniumspace.com/> > Accessed: 26 Sept. 2024) is working on integrating AI into its avatars through a feature called the "Live Forever" mode. It is intended to allow users to have their movements, conversations, and behaviors recorded as data, which would then be used to create an AI-driven avatar that continues to exist and interact even when the original user is offline or deceased. < <https://somniumtimes.com/2024/04/04/live-forever-in-somnium-space-again/> > Accessed: 26 Sept. 2024.

46 These aspects reflect the opinion of the Expert Group on Liability and New Technologies New Technologies Formation Liability for artificial intelligence and other emerging digital technologies (Publications Office, 2019) 23.

- 38 It can be argued that platforms should bear part of the costs associated with addressing illegal content or interactions online, as they are the primary beneficiaries of the increased internet traffic generated by such content. However, solutions to this issue are not straightforward; regulatory obligations in this area may threaten users' freedom of expression and access to information online. Additionally, they may reinforce market asymmetries by favoring larger, well-resourced players over smaller competitors, potentially causing more harm than good.⁴⁷
- 39 As it has been until now, the opportunities to benefit from conditional exemptions and immunity that effectively mitigate online platforms' exposure are in place under the Directive 2000/31/EC (E-commerce Directive).⁴⁸ Therefore, online platforms mostly enjoy a so-called "safe harbor" provision for claims other than intellectual property rights infringements.⁴⁹ Essentially, this provision shields online platforms from legal liability regarding user-transmitted content as long as they remove illegal content promptly upon notification. This safeguard extends to various online services like social media platforms, search engines, e-commerce sites, and hosting providers.⁵⁰ However, as a matter of definition, it is problematic to consider that online virtual worlds match such characterization as per recitals 17 and 18 of the E-commerce Directive.⁵¹ This makes said guideline ultimately disconnected from the discussion.⁵²
- 40 In this regard, the DSA mostly maintains the conditional liability exemptions for online intermediaries from the e-Commerce Directive. Still, it provides more detailed rules on notice-and-action mechanisms for illegal content. However, the DSA does not directly establish a new framework for civil liability related to online platforms and services. Indeed, it does not harmonize or create new rules across the EU for platforms' civil liability. Rather, it leaves this aspect largely to existing national laws, simply updating the liability exemption conditions that platforms must merely meet.⁵³
- 41 Even if it applied to the subject matter, the underlying rules about acceptable content or behavior
-
- 47 M Mariniello 'Online Content and Platform Liability' in *Digital Economic Policy: The Economics of Digital Markets from a European Union Perspective* (Oxford University Press, Oxford 2022) 217. The author builds the argument upon Sartor's analysis and classification of online platforms and their liability risks. Cf. G Sartor *Providers liability: from the eCommerce Directive to the future: in-depth analysis* (European Parliament Brussels 2017).
- 48 Directive 2000/31/EC of the European Parliament and of the Council of 8 June 2000 on certain legal aspects of information society services, in particular electronic commerce, in the Internal Market. I C Ballon 'Chapter 49. The Liability of Platforms (including Website Owners, App Providers, eCommerce Vendors, Cloud Storage and Other Internet and Mobile Service Providers) for User Generated Content and Misconduct' *E-Commerce and Internet Law* (2d edn Thomson Reuters West 2019).
- 49 Provided in Articles 12-15, but also outlined in recital 42-46 and 52 of the E-commerce Directive, Cf. Copyright Directive (Directive (EU) 2019/790), Article 17 on the use of protected content by online content-sharing service providers, under which IP right holders have the option to seek an injunction against intermediaries whose services are being used by a third party to infringe copyright.
- 50 G Sartor (n 47). As for the types of online platforms addressed above, a wide range of economic activities which take place on-line, including the selling of goods online on e-commerce platforms (e.g., Alibaba or Amazon) the provision offering on-line information or commercial communications (e.g. advertisement funded by sponsorship revenue), the offering of online search engine tools (e.g., Google or Bing), the transmission of information or the hosting of information through internet intermediaries, etc. G Pearce and N Platten 'Promoting the Information Society: The EU Directive on Electronic Commerce' (2002) 6(4) *European Law Journal* 363.
- 51 Á Carrasco Perera and C Álvarez López 'Operadores y responsabilidad civil en el metaverso' (2022) *Publicaciones GA_P* 6. Nonetheless, the authors considered that the E-commerce Directive's abstract rules of imputation could be usefully transferred in part to the metaverses model. Therefore, platforms that serve as "mere conduits" or provide hosting services, without intervening in content, are generally not held liable for third-party content they do not editorially control. These platforms are not initially obligated to filter content, except in instances of intellectual property infringements. See note 41.
- 52 Although certain activities within online virtual platforms could fall under digital services regulation, they do not neatly align with the current definitions of online platforms under European law.
- 53 M Husovec 'Rising Above Liability: The Digital Services Act as a Blueprint for the Second Generation Of Global Internet Rules' (2023) 38(3) *Berkeley Technology Law Journal* 118.

remain to be set by the parties. Private individual remedies like claims for damages or injunctive relief against platforms do not directly arise from the obligations set out in the DSA; therefore, injured parties will still need to rely on national tort law and liability provisions when seeking compensation for harm, which, as argued here, primarily rests upon the user.⁵⁴ Furthermore, if platforms were to be made vicariously liable for user content or interactions, the model would need to be constructed in accordance with national regulations.⁵⁵

- 42 These further stress that the subject matter must be examined under national rules. While it is clear that the avatar controller might be liable, a secondary liability regime for virtual worlds requires further exploration. However, this aspect will not be addressed in this discussion due to its broad scope.⁵⁶

D. Harm Caused Via Avatars

- 43 This section addresses the idea that harm in virtual platforms can manifest in various forms, such as reputational damage, psychological distress, or economic loss, considering that avatar usage does not inherently contest the existing range of compensable harm. Nonetheless, while some wrongdoings depend on tangible or physical harm in the real world—such as killing or causing physical injuries—other types of wrongful actions or omissions can transcend physical boundaries. The central question is whether these actions, when carried out through avatars, can be translated into the framework of civil liability law and result in an enforceable duty to compensate the injured party. For example, an avatar might be used to post defamatory content, leading to reputational harm, or engage in virtual harassment, resulting in psychological trauma.⁵⁷ Patrimonial losses may occur through fraudulent schemes, e.g., involving vir-

tual currencies or digital goods, or as a consequence of using avatars to make false statements to induce others to act to their detriment.

- 44 A nuanced approach is necessary when harm has been caused due to an AI-powered avatar's intervention. Under the AI Act, manipulative or exploitative practices discussed above will only be prohibited when they are harmful. However, the Act sets a special severity standard for harm that is not generally required. It does so by reference to an undefined notion of "significant harm". Despite the examples provided in Recital 29 of the AI Act (v.gr., important adverse impacts on physical, psychological health or patrimonial interests), the term remains ambiguous. Even though it should not be excessively controversial to identify and measure important physical injuries or financial losses,⁵⁸ further difficulties arise when trying to determine the extent of significant psychological distress. As rightly argued by Zhong, the challenge lies in the fact that the AI Act does not clearly define what level of psychological distress counts as "significant harm." Additionally, proving that a specific AI practice caused someone's psychological harm is difficult, given that many factors can affect a person's mental state. Effectively addressing this issue may require detailed, case-by-case discussion.⁵⁹
- 45 Harm is undoubtedly a fundamental element of civil liability,⁶⁰ compelling plaintiffs to prove they suffered actual damage due to the wrongful behavior exerted upon them, in this case, using an avatar. Additionally, civil liability for damages in the Metaverse, as well as in virtual worlds according to current law will only arise if a natural or legal person experiences harm in the physical world.⁶¹

54 M Husovec (n 45) 114-116.

55 Ibid. 118.

56 An in-depth assessment on this topic can be consulted in M Husovec *Injunctions against intermediaries in the European Union: accountable but not liable?* (Cambridge University Press Cambridge 2017).

57 Criminal liability, however, is ultimately out of the scope of this work.

58 Which is not to say that it is without controversy.

59 H Zhong 'Implementation of the EU AI act calls for interdisciplinary governance' (2024) AI Magazine, 2.

60 Cf. Article VI-2:101 Draft Common Frame of Reference (DCFR).

61 Á Carrasco Perera and C Álvarez López (n 51) 4. However, this raises practical issues, particularly when users exploit anonymity or pseudonymity to evade accountability. Courts must then determine how to trace and identify the individuals behind such avatars to attribute liability suitably. See also *infra*. Section D (Harm caused via avatars).

46 Some authors have considered that actions of any kind, if committed with the consent of other players or allowed by the rules of a virtual world, do not result in real-world liability, even if they cause personal or economic harm. This is allegedly because by agreeing to the end-user agreement of the virtual world, users consent to abide by that world's specific rules and laws.⁶² This statement requires some considerable nuance.⁶³ It highlights the idea that legal scholarship has been concerned with this issue for a long time, suggesting that avatar misconduct is not an overly novel concern.

47 What sets online virtual worlds apart from traditional social networks' user interaction experience, in terms of Web 2.0 practice,⁶⁴ is the immersive nature they offer with increasingly sophisticated technology deployed. This immersion leads to various interactions being experienced more vividly and intensely as they occur from a first-person perspective. Consequently, any misconduct directed at an avatar becomes drastically relevant, as it may target the user directly. For example, unprecedented counts of sexual assault and rape in online virtual worlds have been brought up, especially since early 2000 with the launch of Second Life (a paradigmatic example of an online virtual world).⁶⁵ Although this

article does not focus on the criminal nature of such conduct,⁶⁶ it does argue that damage (e.g., mental distress) can indeed result from these types of interactions which we have referred to as wrongful behavior. Furthermore, the damage –whether moral or patrimonial– should be compensated for not being a reasonable consequence of agreeing to their Terms of Service, nor should it always be considered a risk assumed by the victim when participating in an online virtual world.⁶⁷ Additionally, if the rule of law allows for a legal right to be asserted upon avatars, treating them as the object matter of a service contract, then a user could potentially suffer patrimonial losses due to the unauthorised interference by another party with one's avatar.⁶⁸

rape" incident involving a Belgian user. The original press report: <<https://www.demorgen.be/nieuws/hoer-second-life~ba9ce069/>> Accessed 26 Sept. 2024; O Bellini 'Virtual Justice: Criminalizing Avatar Sexual Assault in Metaverse Spaces' (2024) 50(1) Mitchell Hamline Law Review 3.

62 F. E. Marx 'Iniuria in cyberspace' (2010) 31(1) *Obiter*, 150.

63 In accordance with most legal systems, when executing a contract, parties may owe duties not only to each other but also to third parties. If the actions of a contracting party harm a third party, the injured third party might have grounds for a civil liability claim. This is particularly true if the harm resulted from negligence or other fault in the performance of the contract, although contracts often include terms that allocate risks between the parties, which means that each participant might assume certain risks explicitly stated in the contract. For example, indemnity clauses, limitation of liability clauses, and waivers can all influence the extent of liability.

64 (n 4) 48; M V Rijmenam, *Step into the Metaverse: How the Immersive Internet Will Unlock a Trillion-Dollar Social Economy* (John Wiley & Sons Inc Hoboken 2022).

65 Second Life was launched in 2003, and although the most well-known instance of this type of allegation involves a SumOfUs researcher in 2022 mentioned above (note 1), it was certainly not the first. For example, in 2007, it was reported that the Brussels public prosecutor had directed patrol detectives from the Federal Computer Crime Unit to enter Second Life to investigate a "virtual

66 To add some insights into this topic, it has been argued that even though rape (namely, a physical act) is not possible in an online virtual world. "Virtual rape" might constitute another type of criminal offence. For instance, Marx argued that under South African Law it could be covered in terms of section 86 of the Electronic Communications and Transactions Act 25 of 2002. The provision, basically prohibits unauthorised access and interference with data, reads as follows: "86(1) (...), a person who intentionally accesses or intercepts any data without authority or permission to do so, is guilty of an offence. 86(2) A person who intentionally and without authority to do so, interferes with data in a way which causes such data to be modified, destroyed or otherwise rendered ineffective, is guilty of an offence. 86(3) A person who unlawfully produces for use ... a device ... which is designed primarily to overcome security measures for the protection of data ... or performs any of those acts with regard to a password, access code or other similar kind of data with the intent to unlawfully utilise such item to contravene this section, is guilty of an offence. 86(4) A person who utilises any device or computer program mentioned in subsection (3) in order to unlawfully overcome security measures designed to protect such data or access thereto, is guilty of an offence"; Other authors propose metaverse sexual assault to be regulated a crime as a matter of statute. F. Marx (n 62) 150-151. See also: Bellini (n 65) 99-107.

67 That is, unless the specific nature and purpose of the online virtual world is to allow users to perform these types of sexually explicit acts to one another in the digital environment. However, consent would undoubtedly persist as a standard for said interactions.

68 Currently, there is a significant gap in the legal

- 48 Rather than devising a catalogue, we will address some relevant examples of potential tort law claims arising from alleged avatar misconduct in metaverses that leads to real-world harm.

1. Right of publicity v. IPR infringement: If a real-world person's likeness is used to create a digital avatar for commercial purposes, no doubt said person would be entitled to compensation due to the avatar exploiting these traits for commercial purposes.⁶⁹ However, there is significant room for exploring the issue of commercially exploiting the likeness of a user's digital avatar for personal gain.⁷⁰ It could be argued that the right of publicity for an avatar cannot be infringed, as an avatar is not a legal person. However, it would be incorrect to assume that an avatar's possible uniqueness or distinctiveness is not currently protected under the law. This protection, under IP rights rules, likely depends on the technical capabilities of the online virtual world to create unique avatars that meet the criteria of originality and distinctiveness required for intellectual property rights protection. Furthermore, it is essential to consider whether the platform reserves economic rights over the avatar's appearance, potentially making it its own asset, as highlighted in the recent Ada case ongoing in China.⁷¹ Those mentioned above

conceptualization of damage inflicted upon digital property. This loop presents an interesting ground for further scholarly inquiry and analysis. While the digital realm has witnessed an exponential expansion, accompanied by a surge in the creation and ownership of virtual assets, the legal impact of harm to such assets remains mostly unexplored and largely unaddressed in case law. Cf. *Bragg v. Linden Research, Inc.* 487F.Supp.2d593 (E.D. Penn.2007) where the very interesting issue of digital property confiscation was not judged due to the parties reaching a private settlement.

69 For instance, in *Hart v. Electronic Arts, Inc.*, No. 11-3750 (3d Cir. 2013), a case involving a football player's likeness used in a video game, in which the plaintiff sued for the violation of his right of publicity. Initially dismissed by the district court, the Third Circuit Court of Appeals overturned the decision, asserting that the games didn't sufficiently alter Hart's identity, thus upholding his rights.

70 J Lake (n16) 835-878.

71 This case marked the first time a court of law has addressed whether digital avatars are entitled to copyright protection on their own. The court concluded that Ada, being

could unfold much-needed discussions about the extent of user ownership in relation to intellectual property rights in virtual environments. However, it should be remarked that there is enough evidence today that the unauthorised use or reproduction of copyrighted or trademarked materials in the Metaverse could lead to infringement claims and compensation.⁷²

2. Other personality rights infringements: These types of issues allow us to consider that users employing avatars to make false statements that harm the reputation of a real-world person, or disclosing private information could face a lawsuit for compensation if harm is caused to a user.⁷³ However, it appears unlikely that a defamation claim could be brought if an avatar is the target of malicious or false comments affecting the avatar's reputation, or if an avatar's private affairs are publicly disclosed without their consent.⁷⁴ Such cases would more likely fall under the right to freedom of speech of the individual making the statements, whatever their nature. However, alleging that no protection can be granted in the latter scenarios would be inaccurate. One may consider, for example, that an unfair competition practice took place, or that the controller committed passing off, ultimately damaging the goodwill of a company.⁷⁵

controlled by humans and supported by various aiding technologies, cannot be considered the author. Rather her developer was recognised as the intellectual property rights holder of the videos featuring the avatar's image online. Hangzhou Internet Court (2022) Zhejiang 0192 Minchu No. 9983, Civil Judgement.

72 See for instance: *Hermès International, et al. v. Mason Rothschild* (2023), 1:22-cv-00384 (SDNY), also known as the *Metabirkin* case.

73 Cheong (n31); Lee (n18); Also: B M Chin, 'Regulating Your Second Life: Defamation in Virtual World' (2007) 72(4) Brooklyn Law Review 1303, 1333.

74 It is remarked that to withhold such rights, an individual must be considered a legal person.

75 Should all criteria be met in a case for these alleged causes of action, there is no fundamental or substantive reason to deny injunctive relief or compensation to the victim.

E. Main Issues in Fault-Based Liability Regimes. Considerations on Fault, Causation, and Compensation

- 49 Unsurprisingly, modern regulation has not specifically covered the responsibility of individuals for damage caused through avatars, nor can it evolve as fast as technology does. Even though creating a special ad hoc liability regime is always an option for legislators, the potential application of existing civil liability regimes should be assessed to avoid unnecessary efforts, especially if current frameworks can be effective. These existing regimes include the possessor's liability for damage caused by objects, often found in civil codes, product liability regimes, and liability arising from fault or negligence as a general rule by *default*.
- 50 Legal systems typically have strict rules regarding liability for objects or vicarious liability.⁷⁶ This first approach should result in the liability of the real-world controller of the avatar.⁷⁷ However, the rules provided in most Civil Codes do not seem to directly apply to damage caused by the emerging technology covered in this study. This is mainly because the damage an avatar can cause is akin to that caused by

a person, such as infringing a person's trademark by selling goods that use another's trademark without the legitimate IP right holder's authorization—rather than that caused by an inanimate object, e.g., falling on top of someone's head.⁷⁸

- 51 These parameters appear unsuitable, primarily because they are tailored to ownership or possession of tangible property, whereas avatars possess an intangible, digital nature. For instance, considering the avatar as an animal or a building under Article 1905 or 1907 of the Spanish Civil Code (CC) is likely unrealistic.⁷⁹ Similarly, Article 1908 CC outlines owner liability in specific scenarios—such as machine explosions, excessive smoke, falling trees, and sewage emissions—that are framed to address real-world physical harm and do not readily apply to avatar-related issues. Furthermore, Article 1910 CC holds the head of the household liable for damage caused by objects falling from a house or part of a house. This provision diverges from issues likely to arise in the Metaverse and in online virtual worlds, as examined above.
- 52 Another possibility would be the integration of avatars within the harmonised legal framework of defective product liability.⁸⁰ However, as argued by legal scholarship, the current definition of a product predominantly encompasses tangible goods, making it a major inconvenience—and overall, not feasible—to categorize avatars as such due to their inherently digital nature. While the European Commission recognizes the inclusion of certain intangible assets as products, the applicability of this classification to avatars remains ambiguous. Furthermore, it does not seem legally suitable. This ambiguity exposes a

76 Spanish Civil Code is considered as the main reference for the purposes of this work. However, other Civil Codes or legal systems are certainly noteworthy. In general, these regimes, often approach the issue by imposing strict liability for harm or damage caused by things, (e.g., dangerous objects, defective products, or sources of special danger under one's custody or control) which will rest upon the person have factual control over the object: For instance, the French Civil Code article 1243 al. 4 (covers strict liability for harm caused by corporeal things within one's keeping) and article 1266 (which allows courts to prescribe measures to prevent harm or stop unlawful nuisances caused by things); the German Civil Code Section 833 (covers strict liability for damage caused by dangerous things or activities); the Italian Civil Code article 2051 (covers strict liability for damage caused by dangerous things in one's custody, unless the custodian proves unavoidable circumstances); the Dutch Civil Code article 6:173 (covers strict liability for damage caused by dangerous substances or objects) and article 6:174 (on liability for damage caused by defective products); the Swiss Code of Obligations article 58 (on liability for damage caused by things under one's custody, unless it is proven that all due care was taken), among others.

77 Notwithstanding the problems posed by the pseudo-anonymity of avatars in the Metaverse.

78 Should such an analogy be possible, it has been argued that "if someone can be held liable for the wrongdoing of some human helper, why should the beneficiary of such support not be equally liable if they outsource their duties to a non-human helper instead, considering that they equally benefit from such delegation?". Expert Group on Liability and New Technologies New Technologies (2019) (n46) 25.

79 Liability for animals being more similar to avatar issues for that matter.

80 Council Directive 85/374/EEC of 25 July 1985 on the approximation of the laws, regulations and administrative provisions of the Member States concerning liability for defective products.

disparity between conceptual ideals and practical implementation.⁸¹ In this regard, the recently approved Directive (EU) 2024/2853 of the European Parliament and of the Council of 23 October 2024 on liability for defective products and repealing Council Directive 85/374/EEC (PLD), aims to encompass digital content or services appears to offer a more promising approach, which warrants further attention and observation.⁸²

- 53 In general, when no specific rule exists, liability for damage caused (perhaps by using an avatar) usually depends on the defendant's fault, whether it be intentional or negligent conduct.⁸³ It should also be considered that users may have varying levels of control over their avatars' actions, ranging from direct manipulation to mere indirect influence; the avatar may in fact be partially AI-powered.
- 54 As argued by the Expert Group on Liability and New Technologies of the European Commission, whether a legal system differentiates between objective or subjective wrongdoing, or even if it separates the basis of liability for misconduct into wrongfulness and fault, identifying the duties of care the perpetrator should have fulfilled and proving that the perpetrator's conduct failed to meet these duties are both crucial.⁸⁴ Courts may need to establish a clear duty of care guidelines and reasonable standards of behavior expected from avatar controllers (users or

AI deployers) based on their capabilities. However, it remains true that the plaintiff must prove fault.⁸⁵

- 55 Bearing this in mind, it should be noted that running an avatar by a human controller involves that the duty of care can be measured according to duties of care intended for human conduct.⁸⁶ For example, where users deliberately direct their avatars to engage in harmful behavior, said users should be held liable at fault.
- 56 Should an avatar be AI-equipped, the deployer's lack of compliance with their duties of care can be assessed by specific requirements set by the AI Act.⁸⁷ The opaque nature of AI—v.gr., the difficulty in understanding and explaining how decisions are made, due to the technology's inherent complexity—makes it particularly difficult for the claimant to establish not only fault but also causation. Furthermore, the amount of agency users possess over their avatars can vary depending on the platform or virtual environment due to technical limitations or the design of online virtual worlds' dynamics.⁸⁸ As supported here, traditional compensation for damage may not always be considered a reasonable outcome if no physical or economic injury was inflicted on another. While measures undertaken by online virtual worlds, such as confiscating virtual assets or banning avatars, could serve as deterrents, they may prove insufficient in some cases. Remedial responses for

81 Especially considering that the EU aims to approach emerging technologies from this perspective. Expert Group on Liability and New Technologies New Technologies (2019) (n46) 6, 27.

82 In light of the new legal framework: "Products in the digital age can be tangible or intangible. Software, such as operating systems, firmware, computer programs, applications or AI systems, is increasingly common on the market and plays an increasingly important role for product safety". Recital 13; "Product" includes electricity, digital manufacturing files and software." Article 4.1 PLD.

83 C V Dam 'Liability for Movable Objects' *European Tort Law*, (2nd edn Oxford University Press Oxford 2013) 402-403; Under Spanish law, the general rule is that set forth in article 1902 CC.

84 Expert Group on Liability and New Technologies New Technologies (2019) (n46) 23.

85 This would involve identifying the duties of care that the perpetrator should have fulfilled and proving that their conduct failed to meet said duties. Moreover, these duties are determined e.g., by statute, or they must be assessed by the Court afterwards, based on the individual's conviction about a reasonable course of action in specific the circumstances. Ibid.

86 However, if avatars were to gain legal personhood, they could potentially be held directly at fault, raising questions about what negligence would entail in the context of avatar behavior.

87 For instance, if the deployer or the user runs a high-risk AI-driven avatar, non-compliance with requirements outlined in Chapters 2 and 3 of the AI Act can adequately be set as the standard for fault. Cfr. Recital 26 of the Proposal for a Directive of the European Parliament and of the Council on adapting non-contractual civil liability rules to artificial intelligence, COM/2022/496 final.

88 Barfield and Williams (n34) 2-43.

victims of harmful behavior committed through avatars may include several well-known forms of relief. Monetary compensation is the most common remedy aimed at covering the actual damages suffered by the plaintiff. Injunctive relief, such as court orders to remove defamatory content, may be sought to prevent ongoing or future harm.

- 57 If an innovative way of thinking about these issues can be noted here, minimum capitalization requirements for avatars to ensure they can pay for damages, along with the potential need for platforms to have liability insurance or victim compensation funds for major incidents, are also to be considered.
- 58 Provided that the avatar controller (the user or the AI-equipped avatar deployer) could be held liable for their online misconduct, the plaintiff must prove a factual and legal causal link between an avatar's actions and the harm suffered. Several aspects on a technological and juridical level should be considered for this purpose. The development of the Metaverse ought to be directed towards a reasonable balance between the anonymity rights of users. This clashes with the need to identify wrongdoers to attribute fault by tracing them to their real-world operators, thereby obscuring the causal chain.⁸⁹ However, this is not a new problem arising from using avatars but rather from using Metaverse technologies.
- 59 Taking all of the above into account, and although it is not a perfect analogy, the discussion on proving causation may benefit from the action taken by the EU, especially in their recent Proposal for an AI Liability Directive. It is intended to apply to non-contractual civil law claims for damages caused by an AI system (possibly an AI-equipped avatar), where such claims are brought under fault-based liability regimes.⁹⁰ The proposal above is funded by a common problem with avatars in the Metaverse (as well as in some online virtual worlds, whether AI equipped or user driven); namely, that current national liability

rules, particularly when based on fault, are deemed inadequate for addressing liability claims for damage caused by digital AI. The inadequacy allegedly arises from the fact that (i) victims must demonstrate a wrongful action or omission by the individual responsible for the damage, and (ii) due to the complex structure of AI, it may be excessively difficult or costly for victims to identify the responsible party and satisfy the prerequisites for a successful liability claim. As contemplated in the proposal, victims could face substantial expenses and significantly prolonged legal proceedings when seeking compensation, potentially dissuading them from pursuing it as a legitimate remedy.⁹¹ It was pointed out that avatars do not share the same complex architecture of AI systems (opacity or the so-called "black box effect", for that matter) per se. Nonetheless, in a scenario where an avatar is operated by a human user, it's important to note that Metaverse technologies—including AI—render it excessively expensive and complicated for victims to pursue such claims. Thus, even though addressing the use of an avatar does not substantively equate to addressing that of a high-risk AI agent (as per Article 6 of the AI Act), the Proposal for an AI Liability Directive suggests adopting a presumption of causal link in instances of fault that is herein deemed also a suitable answer. The reason to apply this specific measure is that it might be analogously excessively costly for a person to prove it—not due to the use of an avatar, but rather because of the Metaverse technologies deployed that may obscure the causal chain, making that same consideration useful for the subject matter.

- 60 However, as a bottom line, existing laws should apply to avatar misconduct without great need for further adaptation. Current regulations are, therefore, sufficient and should be enforced in cases of wrongful behavior conducted via avatars.

89 As emphasised by technology scholars. Cf. L-H Lee et al 'All One Needs to Know about Metaverse: A Complete Survey on Technological Singularity, Virtual Ecosystem, and Research Agenda' (2021) arxiv.2110.05352 1-47.

90 COM/2022/496 final (n87).

91 Explanatory Memorandum. Proposal for a Directive of the European Parliament and of the Council on adapting non-contractual civil liability rules to artificial intelligence, COM/2022/496 final.

F. Conclusion

- 61 In virtual worlds, a person's online identity is represented and made visible to others through their avatar. An avatar –when used by a natural or legal person– serves as a means of interaction among users, acting as their virtual representation and identity within an online virtual world. Avatars may also be the conduit for AI systems to be deployed, which is a foreseeable outcome of the advancement of technology. Despite the importance of avatars, rightly so, the idea of regulating them as legal entities has not gained significant traction. Instead, as the technology stands today, avatars can be better considered as digital data, assets or as the subject matter of a contract concluded by a real person (natural or legal), or by an AI agent within a virtual platform. This type of arrangement allows the user to assume a digital form, navigate and undertake several actions, all of which are carried out through the essentially controllable vessel, the avatar.
- 62 This paper argues that damage caused by using an avatar is a real-world problem and is subject to real-world law. Furthermore, as a prerequisite, any harm considered must be experienced in the real world to successfully ground a claim for compensation, which will be assessed by national rules governing this issue.
- 63 In respect to compensable harm, and as explained herein, while it is true that the prejudice suffered by a person constitutes an actionable claim, alleged damage asserted upon one's avatar may also ground a compensation claim wherein an actual tort can be committed. For example, as argued above, even if an avatar's right to publicity cannot be protected, the intellectual property rights of that avatar's controller can certainly be.
- 64 This is grounded in the principle that the user will be liable provided the usual requisites are met. It also suggests that only some minimal nuances are necessary to conduct an analysis of non-contractual civil law claims for damages caused through an avatar, where such claims are brought under a fault-based liability regime. This is also true for instances when an avatar is run by an AI-system that can be manipulative or exploitative of others, leading to significant harm; arguably some nuances are necessary to address the issue under Union or national rules.
- 65 For instance, wrongdoing that causes harm is deemed relevant only if the harm extends beyond physicality in a digital environment. In these cases, fault-based civil liability remains adequate for addressing avatar misconduct. The standard of the duty of care closely resembling that of a real-world individual's behaviour when an avatar is controlled by a user, but can be assessed otherwise if the avatar is controlled by an AI-agent. Additionally, difficulties that arise, e.g., in proving the causation link, do not inherently result from avatar misconduct, but from the wider context of Metaverse technologies, which also entail the use of AI technology. As such, existing legal frameworks can address these challenges without requiring entirely new principles.
- 66 Without prejudice to the above, and bearing in mind the inevitable advance of technology, we may soon need to consider appropriate legal consequences for harmful occurrences that take place entirely in the interoperable, persistent digital realm and are not covered by current legislation. This could comprise the infringement of the right to an avatar's own image caused by another avatar or the protection of a trademark created by an avatar in the Metaverse without a physical world counterpart –should we reach a point in which this is possible, and they are granted legal personhood. More advanced avatars acting independently may be held directly at fault themselves, whereas less autonomous avatars merely mirroring user inputs may shift more fault to the human operator. Nevertheless, it is necessary to ground these highly hypothetical remarks in our current reality, where one can assert that such considerations only prompt a speculative exercise.