

Marketing In The Metaverse: Transparency Obligations For Restoring The Basic Truth-In- Advertising Principle

Proceedings of the International Congress Towards a Responsible
Development of the Metaverse, 13-14 June 2024, Alicante.

Jacopo Ciani Sciolla
University of Turin



The Chair for the Responsible Development of the Metaverse (MetaverseUA Chair) was created by the University of Alicante (Spain) and financed by Meta Platforms under its [XR Program and Research Funds](#). The Program aims at supporting academic and independent research across Europe into metaverse issues and opportunities. The MetaverseUA Chair is a member of the [European Metaverse Research Network](#). Like all our work, this report has been produced completely independently. The ideas expressed in this paper are the sole responsibility of the author.

How to cite this paper:

Ciani, J., 'Marketing In The Metaverse: Transparency Obligations For Restoring The Basic Truth-In-Advertising Principle' (2024) *Proceedings of the International Congress Towards a Responsible Development of the Metaverse*, Alicante, 13-14 June, 2024.

Abstract:

The immersive nature of the metaverse is populated by virtual influencers, venues and events that create totally unreal consumers experiences. The paper aims to study this phenomenon and its impacts on the end-users from a legal and socio-ethical perspective. As hyper-realistic virtual humans are designed to have human-like features and behaviours, it is increasingly difficult for consumers to distinguish virtuality from real-life. The same issue arises with the use of photo editing softwares to enhance the beauty of images, alter the product's characteristics and make them more appealing. Many warn of the serious consequences coming if we can no longer trust any of the information we consume. Based on the review and evaluation of the applicable hard and soft law, the paper addresses the gaps in existing transparency obligations and reflects on the per se misleading nature of any virtual testimonial or endorsement of products.

Keywords: Marketing, Metaverse, Transparency, Virtual Reality, Deep Fakes, Ai Generated Contents, Ai Manipulated Contents, Right to be Informed, Marking or labelling Ai Generated Contents

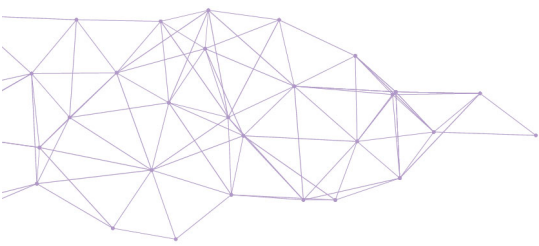


Table of Contents

1. Introduction: the basic truth-in-advertising principle	1
2. Marketing in the Metaverse.....	2
3. Virtual conversational agents.....	3
3.1 Computer v AI generated virtual agents	4
4. Edited and retouched bodies and images	5
5. Ethical concerns.....	5
6. Distinguishing reality from virtuality: the risk of consumer deception and misinformation.....	6
7. A review of the legal framework.....	8
7.1 Transparency obligations under the Artificial Intelligence Act.....	8
7.2 Self-regulation standards.....	10
7.3 Other existing duties of disclosure concerning virtuality.....	12
7.4 Other prospective and forthcoming regulations	12
8. Existing legal framework establishing transparency obligations in advertising	13
8.1 Omission of material information under the Unfair Commercial Practices Directive ...	13
8.2 The duty to disclose the commercial intent of a commercial practice.....	14
8.3 Transparency obligations under the Consumer Rights Directive and the Digital Services Act	15
9. Legislation and practices concerning edited and retouched bodies and images	15
10. Conclusions.....	17

1. Introduction: the basic truth-in-advertising principle

Suppose you meet someone who tells you about a great new product. Would that recommendation factor into your decision to buy the product? Probably. Now suppose the person works for the company that sells the product or has been paid by the company to tout the product. Would you want to know that when you're evaluating the person's glowing recommendation? You bet. That common-sense premise is at the heart of consumer and marketing law as well as advertising self-regulations, all reflecting the basic truth-in-advertising principle that endorsements must be honest and not misleading.

Metaverse advertising, referring to the practice of promoting products, services, or brands within virtual or augmented reality environments, challenges this paradigm as never before.

The immersive nature of the metaverse is populated by virtual influencers, venues and events that create totally unreal consumers' experiences.

This paper aims to study this phenomenon and its impacts on the end-users from a legal and socio-ethical perspectives.

The main objective of the paper is to address the gap and design a legal and ethical benchmark that would set the limits to the scope and use of virtual reality for advertising purposes.

I begin in Sect 2 by introducing the main features of marketing in the metaverse, with a focus on the exploitation of hyper-realistic virtual humans, designed to have human-like features and behaviours as well as of photo editing softwares to enhance the beauty of images, alter the product's characteristics and make them more appealing.

In Sect 3 I introduce virtual influencers, their scope, and their weight on influencer marketing.

Sect 4 is devoted to the issues related to the editing and retouching of pictures, hence bodies in the realm of influencer marketing.

In Sect 5, I consider the ethical dilemmas underpinning this phenomenon.

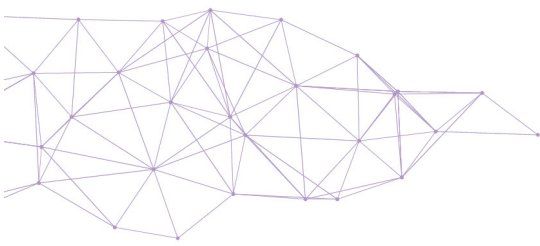
In Sect. 6 I describe the difficulty for consumers to distinguish virtuality from real-life.

In Sect. 7, I evaluate hard and soft law applicable against consumer's deception. Sect 8 also delves on the existing legal framework, with a view on transparency obligations in advertising.

In Sect 9 the legal instruments concerning edited and retouched bodies and images are analysed.

In Sect. 10, I use the conclusions from the previous sections to suggest that brands need to be transparent about using digital identities and photoshopped images in their communications through disclaimers and refrain from engaging in marketing communications referring to any testimonial or endorsement of products that would be per se non-genuine.

Overall, the proposed paper will provide a deeper understanding of the phenomenon, which often operates in a blurred area between legitimate attempts at persuasion and illegitimate manipulation techniques.



2. Marketing in the Metaverse

While there is no universally agreed-upon definition, for the purpose of this paper we will use the word metaverse to refer to a “convergence of our physical and digital lives” through a network of interoperable virtual spaces or worlds. It includes virtual reality experiences, augmented reality experiences, interactive, persistent digital spaces, or all of the above, and differs from both the “physical world” and from most conventional 2D experiences of “the Internet”. The metaverse currently does not identify a single shared virtual space, but is decentralized across various platforms and therefore can only come into full existence once there is a true interoperability between these different platforms. In this paper, the word metaverse will interchangeably be used to talk about the metaverse in this sense, but also about what presently could be called a metaverse embryo.

Whatever its form and definition, experts expect that the metaverse will grow quickly and be immensely valuable. According to McKinsey, the potential impact of the metaverse varies by industry, although implications for all are expected. For instance, they estimate it may have a \$144 billion to \$206 billion impact on the advertising market².

Indeed, from a functionalist perspective, the metaverse is most of all “an advertising channel”³. Brands may exploit the metaverse to enhance engagement with consumers, for example by promoting events, creating virtual stores (to sell virtual or physical goods), or by launching virtual products before their physical counterparts, as prototypes, for market testing, or in order to increase influence on consumers.

New brands may come into existence in the metaverse and then move (or not) to the physical world.

It’s not only about using old school methods in a new environment. Shifting from digital advertising to metaverse marketing allows brands to tap into a space that continues to exist and where things continue to happen even when the user is not connected. Furthermore, interoperability across platforms allows to move a virtual good or avatars themselves from one platform to another. Another turning point for the metaverse is represented by new cutting-edge hardware like Virtual Reality (VR) or Augmented Reality (AR) headsets, sensors, and other interfaces. VR headsets provide an immersive virtual reality for the wearer, while AR headsets create virtual interaction with elements from the physical world. In a nutshell, the metaverse will evolve our current experience of the Internet from a 2D perspective to a 3D one.

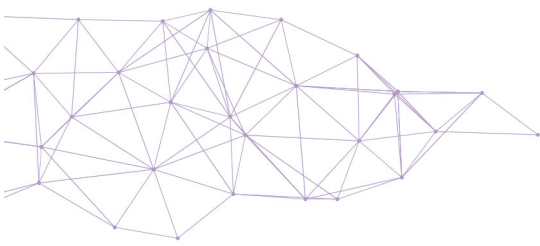
Brands appreciate that establishing a virtual presence in digital worlds allows them to engage with a largely younger audience. To give some examples: Coca-Cola, Samsung, and Volkswagen have invested in virtual billboards within video games like Football Manager and Hyper Scape. Chipotle established a virtual restaurant in Roblox for a Halloween campaign, giving a voucher for a real-life burrito to any visitor wearing a costume. Other uses of brands in the metaverse relate to events: for example, the Metaverse Fashion Week held in Decentraland in March 2022 attracted brands to exhibit their digital products, which in turn could be sold inside or outside the metaverse.

A high personalization capacity serves an important role in the business model of the metaverse. This concerns in particular the use of conversational agents like virtual assistants, chatbots, avatars etc. who, even in their virtual form, represent brands,

¹ McKinsey, ‘What is the Metaverse and what does it mean for businesses’ (www.mckinsey.com) <www.mckinsey.com/business-functions/mckinsey-digital/our-insights/what-is-the-Metaverse-and-what-does-it-mean-for-business> accessed 21 June 2024.

² McKinsey, Value creation in the metaverse. The real business of the virtual world, June 2022, 6.

³ McKinsey.



sometimes promoting their products and services but also acting as ambassadors of principles and rights. Even if the language in advertising distinguishes conversational agents in different categories, they will be grouped for the purpose of this paper under the “virtual influencers” umbrella-term.

3. Virtual conversational agents

Traditionally, brands collaborate with real-life influencers (i.e., humans living in a physical world) who can make their own decisions regarding sponsored collaborations with brands and form opinions about the products and services they promote.

Virtual influencers are non-human digitally created characters sharing social media content and engaging in interactive communications to obtain influential status among consumers.

Virtual influencers can have different forms. Some authors developed a taxonomy based on their similarity to human appearance, also known as anthropomorphism⁴, and their placement on the reality-virtuality continuum⁵, ranging from unimaginable characters to hyper-realistic characters that can be nearly impossible to distinguish from humans⁶. Like real-life influencers, hyper-realistic human virtual influencers share content about their personal and social lives, which often features them in the physical world performing human tasks, including attending fashion shows and commercial photoshoots. Consider Lil Miquela who claims to be a 19-year-old AI robot with a passion for social justice, fashion, music, and friendship. Currently, Miquela has over 190,000 monthly listeners on Spotify and gives interviews at major events⁷. She has been featured in campaigns by Calvin Klein and Prada.

Accordingly, virtual influencer marketing can also be classified as mixed reality because it allows to mix objects from both physical and virtual worlds and makes the boundaries between the real and virtual world blurred.

It is worth noting that some digital characters appear to exist only in a virtual world, while some others are avatars of real-life celebrities⁸. For example, a digital double of supermodels Naomi Campbell and Kendall Jenner starred in Burberry's TB summer monogram collections. In September 2023, Meta launched 28 AI-powered chatbots featuring Kendall Jenner (Billie), Paris Hilton (Amber), and Snoop Dogg (Dungeon Master). Currently, they are only available for testing in the US but AI shall make celebrities, shortly, omnipresent, since they can penetrate every market and format at any time. Most of the time AI clones of celebrities, grabbing user's attention on YouTube, are just scams relying

⁴ Martin Mende and others, 'Service Robots Rising: How Humanoid Robots Influence Service Experiences and Elicit Compensatory Consumer Responses' (2019) 56(4) Journal of Marketing Research 535 <<http://dx.doi.org/10.1177/0022243718822827>> accessed 21 June 2024.

⁵ Sarah Hudson and others, 'With or without you? Interaction and immersion in a virtual reality experience' (2019) 100 Journal of Business Research 459, <<http://dx.doi.org/10.1016/j.jbusres.2018.10.062>> accessed 21 June 2024. .

⁶ Simone Lykke Tranholm Mouritzen, Valeria Penttinen and Susanne Pedersen, 'Virtual influencer marketing: the good, the bad and the unreal.' (2023) ahead-of-print European Journal of Marketing.

⁷ Louisa Savageaux, 'Virtual Influencers: Harmless Advertising or Dystopian Deception?' [2022] De Pauw The Prindle Institute for Ethics.

⁸ Jbid Arsenyan and Agata Mirowska, 'Almost human? A comparative case study on the social media presence of virtual influencers' (2021) 155 International Journal of Human-Computer Studies 102694 <<http://dx.doi.org/10.1016/j.ijhcs.2021.102694>> accessed 21 June 2024.

on AI voice cloning paired with decontextualized video of the celebrity⁹. Scarlett Johansson recently claimed that OpenAI would have developed an AI personal assistant voice sounding uncannily similar to the actress'one.

3.1 Computer v AI generated virtual agents

Within this wide category, experts distinguish between conversational agents created with computer-generated imagery technology (CGI influencers) and AI influencers that rely on AI technologies in creating content and interacting with consumers.

Both of them are very realistic graphical simulations of a person and disambiguating them may be at least at first glance troublesome. Still, their capabilities are very different.

Computer generated characters can engage in a dialogue with the human user. The algorithm managing the dialogue is, normally, a large set of 'if-then' rules. Each rule maps an incoming user utterance, occurring in a specified dialogue context, onto a response utterance, and an accompanying new dialogue context. These rules are specified by a human 'script author'. The author defines a set of dialogue contexts, and for each context specifies a set of possible utterance types to expect from the user in that context. The script author can also define emotional gestures (facial expressions, body gestures) to accompany each avatar response utterance. The simulated physical body does not feature in the representation of emotions. There is nothing to simulate the agent's perception or representation of her own emotions. Its internal state is extremely minimal—a set of symbolic contexts supplied by a human author. The emotional system consists of a circuit that classifies the user's current emotion, using evidence from the words and acoustic features of the user's current utterance, and from the user's current facial expression, and then responds to this with an emotional gesture, again using a set of hand-authored rules (for instance, 'if user is happy, be happy'; 'if user is angry, be worried'). The data that train the emotion classifier are assembled by human authors, who label utterances and video images with the relevant emotion categories.

An increasing number of AI systems aim to provide a more 'complete' model of agents, having not only a realistic face and body but also a range of humanlike abilities and emotions driven by a human brain modelling component.

A concrete example is an avatar, called BabyX, simulating an 18-month-old baby produced by the New Zealand company Soul Machines¹⁰. BabyX is already quite a convincing simulation of a real baby: a user can interact with her in various natural ways, and her responses are also quite natural, both at the graphical level and at a more cognitive behavioural level. She can see and hear the user via video and audio feeds; she can see and interact with objects in her own simulated environment, she can learn words and actions. She also manifests various emotional behaviours in response to events she perceives: she can smile, laugh, cry, get cross or frustrated.

BabyX's cognitive model includes a model of episodic memory: she can remember the events and states she experiences, and she can retrieve a sequentially structured series of events and states. She can use her memories of generic event sequences to make

⁹ Jason Koebler, 'Deepfaked Celebrity Ads Promoting Medicare Scams Run Rampant on YouTube' (*404 Media*, 9 January 2024) <www.404media.co/joe-rogan-taylor-swift-andrew-tate-ai-deepfake-youtube-medicare-ads/?mc_cid=f250e2b063&mc_eid=f720a42bfb> accessed 21 June 2024.

¹⁰ 'Benefits of Biological AI | Soul Machines' (*Soul Machines_042324*) <www.soulmachines.com/biological-ai> accessed 21 June 2024.

predictions, she also learns emotional associations of events, and her memory for events is influenced by these associations. Her storage of events is weighted towards events with strong emotional associations¹¹.

Even if at present, BabyX's brain is still quite simple, the authors anticipate that at some point in the future these models could be placed on the scale of moral patients and in a relationship of moral parity with some lower animal, like a rat or mouse¹².

4. Edited and retouched bodies and images

As hyper-realistic virtual humans are designed to have human-like features and behaviours and appear in the physical world, like real-life restaurants and events, it might be particularly difficult for consumers to distinguish virtuality from real-life. The same issue arises with the use of photo editing softwares to enhance the beauty of images, alter the product's characteristics and make them more appealing. Digital photo retouching of advertising imagery is ubiquitous¹³ and often involves correcting perceived "flaws" in the appearance of featured models, including modifying skin tone, minimizing signs of wrinkles or blemishes, or modifying their body size or shape¹⁴. The impact of Photoshop is considerable. Photoshop has made a "once unattainable image of beauty and perfection much less a figment of the imagination and much more a tangible reality, leaving beauty in the hands of its digital creator."¹⁵ The growing recognition that so much of social media is unrealistic has given rise to a form of online social activism where fake and realistic images of oneself are posted side-by-side to show how photo retouching dramatically alter perceived appearance¹⁶.

5. Ethical concerns

AI agents that simulate 'whole humans' have always been a focus for ethical discussion¹⁷. The ethical questions for these agents move from the implications for building computer models that explicitly aim to reproduce the functionality of human or other biological brains¹⁸ and then range from the safety of AI-related products and the risks of misuse to a wide variety of social impacts.

¹¹ Endel Tulving, 'Episodic Memory: From Mind to Brain' (2002) 53(1) Annual Review of Psychology 1, <<http://dx.doi.org/10.1146/annurev.psych.53.100901.135114>> accessed 21 June 2024.

¹² Alistair Knott, Mark Sagar and Martin Takac, 'The ethics of interaction with neurorobotic agents: a case study with BabyX' [2021] AI and Ethics <<http://dx.doi.org/10.1007/s43681-021-00076-x>> accessed 21 June 2024.

¹³ Hany Farid, 'Seeing Is Not Believing' (2009) August 2009 IEEE Spectrum 44.

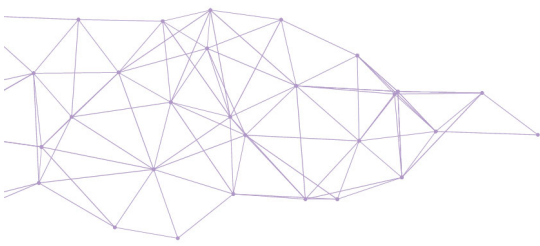
¹⁴ Eric Kee and Hany Farid, 'A perceptual metric for photo retouching' (2011) 108(50) Proceedings of the National Academy of Sciences 19907, <<http://dx.doi.org/10.1073/pnas.1110747108>> accessed 21 June 2024.

¹⁵ Ashley Brown, 'Picture [Im]Perfect: Photoshop Redefining Beauty in Cosmetic Advertisements, Giving False Advertising a Run for the Money' (2015) 87(87) Tex. Rev. Ent. & Sports L.

¹⁶ Marika Tiggemann and Isabella Anderberg, 'Social media is not real: The effect of 'Instagram vs reality' images on women's social comparison and body image' (2019) 22(12) New Media & Society 2183, <<http://dx.doi.org/10.1177/1461444819888720>> accessed 21 June 2024.

¹⁷ Elisabeth Hildt, Kelly Laas and Monika Sziron, 'Editorial: Shaping Ethical Futures in Brain-Based and Artificial Intelligence Research' (2020) 26(5) Science and Engineering Ethics 2371, <<http://dx.doi.org/10.1007/s11948-020-00235-z>> accessed 21 June 2024.

¹⁸ Christine Aicardi and others, 'Ethical and Social Aspects of Neurorobotics' (2020) 26(5) Science and Engineering Ethics 2533, <<http://dx.doi.org/10.1007/s11948-020-00248-8>> accessed 21 June 2024.



The preliminary question, of course, is whether simulated agents have performative equivalence with any class of biological agent to which we accord moral status. Someone proposes that the ethical status of a simulated human should be decided on the basis of its behaviours. Other argue that the difference in algorithms is ethically significant.

A large literature is growing up around the topic of user mistreatment of avatars. This field of study builds on the assumption that the way a human user treats an avatar may be of ethical significance, because of its effects on the user and the capacity to affect their behaviours towards real people¹⁹.

Another research topic concerns harms on users themselves. In this regard, an important issue to consider is that a user may become emotionally invested in the simulated person. Many users are emotionally vulnerable and there's potential harm in their becoming attached to something that cannot truly reciprocate or could say things leading to strong emotional distress²⁰.

At the root of most of the harms arising from the interactions between human being and avatar there is a failure to disclose context-specific information, such as those concerning the virtual nature of the latter.

6. Distinguishing reality from virtuality: the risk of consumer deception and misinformation

Even though virtual agents do not exist in real life, several studies show they are perceived as authentic, regarding their physical appearance, personality and behaviour²¹. This is coherent with the social response theory²², according to which when consumers come across virtual influencers, they engage with them as they do with real-life ones, by applying the same social rules of interactions with humans, though they know virtual influencers are not humans²³. Hence, as long as consumers will respond to virtual influencers as they do to real-life ones, it is not surprising that the former are capable of being preferred to humans. However, as virtual agents are designed to have human-like features and behaviours, they might be particularly difficult to distinguish from real-life ones.

Several studies warn of the risk of consumer manipulation²⁴, or warn of the serious consequences when we can no longer trust²⁵ any of the information we consume. The

¹⁹ John Danaher, 'Welcoming Robots into the Moral Circle: A Defence of Ethical Behaviourism' (2019) 26(4) *Science and Engineering Ethics* 2023, <<http://dx.doi.org/10.1007/s11948-019-00119-x>> accessed 21 June 2024.

²⁰ Jacopo Ciani, 'The legal challenges of the digital afterlife industry and the new AI driven memorial chatbots', *AI Approaches to the Complexity of Legal Systems* (Springer 2024, forthcoming).

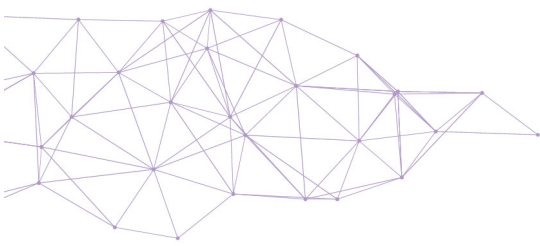
²¹ Evangelos Moustakas and others, 'Blurring lines between fiction and reality: Perspectives of experts on marketing effectiveness of virtual influencers', *2020 International Conference on Cyber Security and Protection of Digital Services (Cyber Security)* (IEEE 2020) <<http://dx.doi.org/10.1109/cybersecurity49315.2020.9138861>> accessed 21 June 2024.

²² Leon Festinger, 'A Theory of Social Comparison Processes' (1954) 7(2) *Human Relations* 117, <<http://dx.doi.org/10.1177/001872675400700202>> accessed 21 June 2024.

²³ Youngme Moon, 'Don't Blame the Computer: When Self-Disclosure Moderates the Self-Serving Bias' (2003) 13(1-2) *Journal of Consumer Psychology* 125, <http://dx.doi.org/10.1207/s15327663jcp13-1&2_11> accessed 21 June 2024.

²⁴ Brittan Heller and Avi Bar-Zeev, 'The Problems with Immersive Advertising: In AR/VR, Nobody Knows You Are an Ad' (2021) 1(1) *Journal of Online Trust and Safety* <<http://dx.doi.org/10.54501/jots.v1i1.21>> accessed 21 June 2024.

²⁵ Massimo Durante, 'The Online Construction of Personal Identity Through Trust and Privacy' (2011) 2(4) *Information* 594, <<http://dx.doi.org/10.3390/info2040594>> accessed 23 June 2024.



prevalence of fake presences may eradicate our sense of reality in the virtual realm. Virtual agents may be purposefully designed for or tricked into (e.g., by untruthful or low-quality online data) spreading misinformation and other unethical communications²⁶. The background stories of virtual influencers, the content they share, and, most importantly, their visual appearance can create false representations in society, like unrealistic perceptions of beauty standards²⁷. This can be problematic, i.e., consumers having difficulties distinguishing virtual from human influencers²⁸, as these consumers do not realize that they are comparing themselves to a non-human and may feel anxious about the way they look, to the point of inhibiting their ability to live well²⁹.

This risk gets worse when influencers are involved in marketing activities. As consumers are more likely to rely on recommendations from individuals that have views and beliefs similar to their own, making consumers falsely believe they are engaged in communications with humans, might suspend consumers' abilities to identify and critically evaluate persuasive marketing tactics.

Therefore, knowing the exact status of a virtual agent as computer or AI generated might be strongly beneficial for users, resulting in a positive impact on various markers of health, ranging from lower body dissatisfaction, unhealthy dieting behavior and disordered eating behaviors, as well as increased self-esteem, improvement of mood. Some virtual influencers are transparent about their virtual identity. However, this is not always the case. The reason is that disclosing it may negatively affect the effectiveness of the communication.

The source credibility model shows that influencers' perceived characteristics may impact their trustworthiness, expertise, and attractiveness and affect the desired results of their messages³⁰). Having low source credibility, influencers will lose the ability to engage consumers with sponsored posts. In this regard, knowing exactly the human or non-human nature of an influencer is pretty relevant. The outcomes of marketing research show that anthropomorphism increases brand liking and purchase intentions, while disclosing virtuality may lead people to feel uncomfortable or become more suspicious of persuasion attempts³¹.

Thus, disclosing virtuality in commercial communications may lead to lowered brand trust and attitudes, lower purchase intentions, and engagement³².

²⁶ Mekhail Mustak and others, 'Deepfakes: Deceptions, mitigations, and opportunities' (2023) 154 Journal of Business Research 113368, <<http://dx.doi.org/10.1016/j.jbusres.2022.113368>> accessed 21 June 2024.

²⁷ Rosalind Gill, *Perfect: Feeling Judged on Social Media* (Polity Press 2023).

²⁸ Claudia Franke, Andrea Groeppel-Klein and Katrin Müller, 'Consumers' Responses to Virtual Influencers as Advertising Endorsers: Novel and Effective or Uncanny and Deceiving?' [2023] Journal of Advertising 1, <<http://dx.doi.org/10.1080/00913367.2022.2154721>> accessed 21 June 2024.

²⁹ Fengyi Deng and Xia Jiang, 'Effects of human versus virtual human influencers on the appearance anxiety of social media users' (2023) 71 Journal of Retailing and Consumer Services 103233, <<http://dx.doi.org/10.1016/j.jretconser.2022.103233>> accessed 21 June 2024.

³⁰ Roobina Ohanian 'The impact of celebrity spokespersons' perceived image on consumers' intention to purchase' (1991) 31(1) Journal of Advertising Research 46.

³¹ Parker J Woodroof and others, 'What's done in the dark will be brought to the light: effects of influencer transparency on product efficacy and purchase intentions' (2020) 29(5) Journal of Product & Brand Management 675, <<http://dx.doi.org/10.1108/jpbm-05-2019-2362>> accessed 21 June 2024.

³² Sophie C Boerman, 'The effects of the standardized instagram disclosure for micro- and meso-influencers' (2020) 103 Computers in Human Behavior 199, 103 <<http://dx.doi.org/10.1016/j.chb.2019.09.015>> accessed 21 June 2024.

7. A review of the legal framework

On these grounds, the WEF raised the issue of “Authenticity of digital entities” and highlighted the importance to give “representational consideration to promote authenticity”³³.

The EPRS report on the Metaverse highlighted as well the “considerable scope for a wide range of illegal and harmful behaviours and practices in the metaverse environment. This makes it essential to consider how to attribute responsibility, *inter alia*, for fighting illegal and harmful practices and misleading advertising practices”³⁴.

There is a debate on the need to revise legislation on advertising in order to address its metaverse implications. Some experts believe that the regulatory framework governing advertising in the metaverse is the one referable to the discipline provided for advertising in video games, others argue that it’s sufficient to adapt the existing rules to the new virtual world.

Other experts argue that regulations should be crafted to limit the scope of emotion-responsive advertising to restrict virtual product placement within the metaverse and improve transparency³⁵.

Another issue concerns the distinction among the various techniques available to virtual agents to influence users beliefs and behaviours, such as persuasion, manipulation, deception, coercion and exploitation³⁶. Only the subsuntion of any influencing behaviours within the right modes of influence may bring to establish effective legal remedies³⁷.

What’s certain is that currently limits the possibility of using the behaviors and emotions of virtual avatars to promote virtual products in the metaverse is becoming a pressing challenge³⁸.

7.1 Transparency obligations under the Artificial Intelligence Act

Specific obligations to disclose AI generated contents, including audio-visual contents resembling existing persons, arise from the recently adopted AI Act³⁹ with the purpose to minimize the “new risks of misinformation and manipulation at scale, fraud, impersonation and consumer deception” and restore “the integrity and trust in the information ecosystem” (recital 133).

Drawing on the experience of a plethora of AI ethics charters and guidelines⁴⁰, setting transparency, articulated as the duty to make an object or entity knowable⁴¹, as a

³³ WEF, *Metaverse Identity: Defining the Self in a Blended Reality*, 2024, 32.

³⁴ EPRS, *Metaverse Opportunities, risks and policy implications*, 2024.

³⁵ See Louis Rosenberg, ‘Regulation of the Metaverse: A Roadmap’ [2022] 6th International Conference on Virtual and Augmented Reality Simulations (ICVARs 2022) .

³⁶ Robert Noggle, ‘The Ethics of Manipulation’, *The Stanford Encyclopedia of Philosophy* (Metaphysics Research Lab, Stanford University 2022)

³⁷ Iason Gabriel et al., *The Ethics of Advanced AI Assistants*, Google DeepMind, 2024, 81.

³⁸ DLA Piper, *Metaverse: business opportunities and legal challenges*, 24.

³⁹ European Parliament legislative resolution of 13 March 2024 on the proposal for a regulation of the European Parliament and of the Council on laying down harmonised rules on Artificial Intelligence (Artificial Intelligence Act) and amending certain Union Legislative Acts (COM(2021)0206 – C9-0146/2021 – 2021/0106(COD)).

⁴⁰ Anna Jobin, Marcello Lenca and Effy Vayena, ‘The global landscape of AI ethics guidelines’ (2019) 1(9) *Nature Machine Intelligence* 389, <<http://dx.doi.org/10.1038/s42256-019-0088-2>> accessed 21 June 2024.

⁴¹ Paul Hayes, ‘An ethical intuitionist account of transparency of algorithms and its gradations’ (2020) 13(3) *Business Research* 849, <<http://dx.doi.org/10.1007/s40685-020-00138-6>> accessed 21 June 2024.

pillar principle for any AI deployment⁴², the new rules establish the general principle according to which users should be made aware when they are interacting with AI. In particular, Article 50 establishes a two-tier system of disclosure duties. The first, set out in para. 1, concerns “providers” as defined in Art. 3(3), placing on the EU market or putting into service AI systems or general-purpose AI models. The latter, contained in para 4, regards deployers as defined in Art. 3(4) of such AI systems using the output produced by it in the Union.

Our analysis is potentially concerned by both of these categories, because brands could develop their own conversational agents or avail of models for conversational agents already placed on the market by providers and personalize or customize them for their professional purposes. Therefore, a brand may theoretically be both a provider or a deployer.

Under para 1, a provider shall ensure that “AI systems intended to interact directly with natural persons are designed and developed in such a way that the natural persons concerned are informed that they are interacting with an AI system”⁴³.

The notice is not an absolute requirement. Natural persons should be notified that they are interacting with an AI system “unless this is obvious from the point of view of a natural person who is reasonably well-informed, observant and circumspect, taking into account the circumstances and the context of use.”

This approach is however problematic because it leaves any evaluation to the providers and adds a layer of subjectivity and uncertainty. The obviousness may depend on the level of anthropomorphism of the fictional character. Following the taxonomy mentioned before, we can speculate that the disclosure should be needed only for hyper-realistic characters.

Also for this reason, the exception does not seem consistent with the rationale followed by the same EU legislator when establishing the duty to disclose the advertising nature of ad-contents. In that case, the notice that a content has been sponsored must be given even if it could be inferred by the context, for example because the sponsor company name is part of the message.

The European Committee of the Regions pointed out this aspect in its Opinion over the AI Act proposal⁴⁴, inviting to remove the exception on the grounds that “natural persons should always be duly informed whenever they encounter AI systems and this should not be subject to interpretation of a given situation. Their rights should be guaranteed at all times in interactions with AI systems”. Even if the final wording of the article has been modified after the amendments approved by the Parliament⁴⁵, the exception has been maintained.

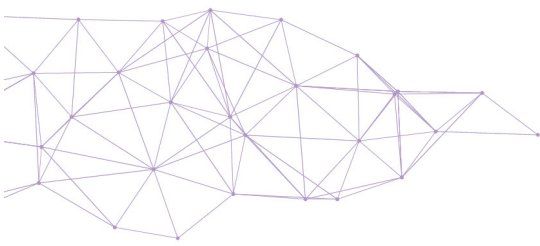
As clarified by recital 133 and para 2, the duty of disclosure should be addressed by providers of systems generating synthetic contents by design. This means that marking or labelling techniques should be implemented at the level of the system or at the level of the

⁴² Ugo Pagallo and Massimo Durante, ‘The Good, the Bad, and the Invisible with Its Opportunity Costs: Introduction to the ‘J’ Special Issue on “the Impact of Artificial Intelligence on Law” (2022) 5(1) J 139, <<http://dx.doi.org/10.3390/j5010011>> accessed 21 June 2024.

⁴³ This obligation shall not apply to AI systems authorised by law to detect, prevent, investigate, and prosecute criminal offences, unless those systems are available for the public to report a criminal offence.

⁴⁴ Opinion of the European Committee of the Regions – European approach to artificial intelligence – Artificial Intelligence Act (2022/C 97/12).

⁴⁵ Article 50 establishes that “Providers shall ensure that AI systems intended to interact with natural persons are designed and developed in such a way that the natural person concerned are informed that they are interacting with an AI system, unless this is obvious from the point of view of a natural person who is reasonable well-informed, observant and circumspect, taking into account the circumstances and the context of use”.



model, thereby ensuring “that the outputs are marked in a machine-readable format and detectable as artificially generated or manipulated”. This should facilitate the fulfilment of the disclosure obligation by the downstream provider of the AI system.

Examples of marking techniques are watermarks, metadata identifications, cryptographic methods for proving provenance and authenticity of content, logging methods, fingerprints. However, nothing prevents to adopt other “appropriate” techniques. The “appropriateness” should be measured in terms of reliability, interoperability, effectiveness, and robustness, taking “into account the specificities and the limitations of the different types of content and the relevant technological and market developments in the field, as reflected in the generally acknowledged state-of-the-art”.

Further to the technical solutions employed by the providers of the system, the EU legislator introduces specific duties upon the deployer of the technology. Such duties are autonomous from those placed upon the providers, even if potentially deployers could avail themselves of the marking or labelling techniques already implemented by providers.

Para 4 establishes that “deployers of an AI system that generates or manipulates image, audio, or video content constituting a deep fake, shall disclose that the content has been artificially generated or manipulated”. To this purpose, art. 3(60) defines ‘deep fake’ as “AI-generated or manipulated image, audio or video content that resembles existing persons, objects, places or other entities or events and would falsely appear to a person to be authentic or truthful”.

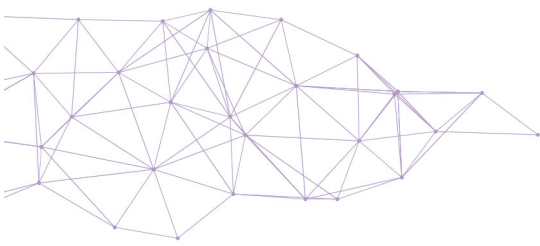
The proper balance between the safeguard of integrity and the freedom of art and creativity requires that the negative impact on the latter should be minimized. Therefore, when the content “forms part of an evidently artistic, creative, satirical, fictional analogous work or programme”, the disclosure should be done in “an appropriate manner that does not hamper the display or enjoyment of the work”.

While fictional characters should be normally considered as an artistic or creative work, I do not expect that this provision should be applied if they are used for a primary commercial purpose.

Surprisingly, the waiver from disclosure when the artificial nature is obvious does not apply here. Therefore, we should conclude that the duty to mark artificially generated contents is different in scope ranging from providers and deployers even if para 5 sets disclosure requirements that are the same for both (information should “be provided to the natural persons concerned in a clear and distinguishable manner at the latest at the time of the first interaction or exposure.”).

7.2 Self-regulation standards

Art. 50(7) of the AI Act encourages “the drawing up of codes of practice at Union level to facilitate the effective implementation of the obligations regarding the detection and labelling of artificially generated or manipulated content”. This provision should leave space to the advertising self-regulation authorities to set appropriate standards as it has already been done for influencer marketing, particularly with the European Advertising Standard Authority (EASA) Best Practice Recommendation on Influencer Marketing and the similar provisions contained in local codes and national guidelines. As a result, we expect that deployers shall adopt disclaimers in the form of hashtags followed by the notice, building on the past experiences concerning the disclosure of the commercial nature of ad-contents.



In the meantime, platforms are already taking matters into their own hands and developing their own standards. For example, TikTok updated its platform guidelines to require that synthetic or manipulated media that shows realistic scenes be clearly disclosed. This can be done using a sticker or caption, such as “synthetic”, “fake”, “not real”, or “altered”. The guidelines require disclosure to be directly in the videos, not just in the virtual influencer’s bio⁴⁶.

Meta has a manipulated media policy since 2020, covering videos that are created or altered by AI to make a person appear to say something they didn’t say. This approach has been judged too since later on people have developed other kinds of realistic AI-generated content like audio and photos, and this technology is quickly evolving and the need to address manipulation that shows a person doing something they didn’t do arised. In February, Meta announced that further to the self-disclosure process it has been working with industry partners on common technical standards for identifying AI content, including video and audio based on a “Made with AI” label on AI-generated video, audio and images⁴⁷.

It will be based on detection of industry-shared signals of AI images or people self-disclosing that they’re uploading AI-generated content. We will see how accurate this AI detection software shall be.

It already added “Imagined with AI” to photorealistic images created using the Meta AI feature.

Meta plans to start labeling organic AI-generated content in May 2024, and stop removing content solely on the basis of its manipulated video policy in July.

In March 2024, YouTube announced a way for creators to self-label when their videos contain AI-generated or synthetic material⁴⁸. The checkbox appears in the uploading and posting process, and creators are required to disclose “altered or synthetic” content that seems realistic. That includes things like making a real person say or do something they didn’t; altering footage of real events and places; or showing a “realistic-looking scene” that didn’t actually happen.

On the other hand, disclosures won’t be required for things like beauty filters, special effects like background blur, and “clearly unrealistic content” like animation. For most videos, this added transparency will appear in the expanded description. Only for videos that touch on more sensitive topics – like health, news, elections, or finance – a more prominent label on the video shall be included and penalties, including content removal or suspension from the YouTube Partner Program shall be introduced for creators for repeatedly not disclosing videos that are meaningfully altered or synthetically generated. But for now, the YouTube feature just relies on the honor system – creators have to be honest about what’s appearing in their videos.

It announces also that a label could be applied in the near future to videos in cases where creators haven’t disclosed.

⁴⁶ ‘TikTok Creator Academy: Empowering Creators to Grow and Succeed on TikTok | TikTok For Creator’ (TikTok - Make Your Day) <www.tiktok.com/creators/creator-portal/en-us/community-guidelines-and-safety/ai-generated-content-label/> accessed 22 June 2024. .

⁴⁷ ‘Our Approach to Labeling AI-Generated Content and Manipulated Media | Meta’ (Meta) <<https://about.fb.com/news/2024/04/metas-approach-to-labeling-ai-generated-content-and-manipulated-media/>> accessed 22 June 2024.

⁴⁸ ‘New Disclosures and Labels for Generative AI Content on YouTube - YouTube Community’ (Google Help) <<https://support.google.com/youtube/thread/264550152/new-disclosures-and-labels-for-generative-ai-content-on-youtube?hl=en>> accessed 22 June 2024.

7.3 Other existing duties of disclosure concerning virtuality

The main gap of the transparency obligations set out in the AI Act is that their object and scope is limited to AI Systems or content generated by AI Systems or models.

This means that fictional characters which do not amount to AI systems and are just computer generated are not covered by the above mentioned duties.

This may be problematic because also CGI may bring to similar consumers' deception issue. Therefore, the AI Act should not wave the EU institutions to set a similar obligation also for virtual influencers which donot technically amount to an AI system, but still may perfectly resemble a human and pose specific risks of deception.

Actually, only few legislations provide guidance on what brands should do to avoid misleading consumers about the real nature of synthetic agents and most of them refer only to the specific case of virtual influencers.

The first and, at the time of writing, only jurisdiction establishing that virtual influencers "must additionally disclose consumers that they are not interacting with a real human being" is India. Following the Consumer Protection Act of 2019, the Advertising Standards Council of India (ASCI) in 2021 became the first national regulator to require an "upfront and prominent" disclosure of this kind⁴⁹.

India has been followed by France. The Influencers Act, which came into effect on 1 June 2023, supplements the pre-existing regulations on advertising establishing that content with altered or artificially intelligent images must be accompanied by statements such as "virtual images" in order to limit the psychological impact on the public.

In the U.S., the Federal Trade Commission released an updated version of the Endorsement Guides⁵⁰, which makes clear that brands may be held liable for virtual influencers' unfair commercial practices as it happens for human endorsers. This means that virtual influencers should avoid making statements implying their humanity or a personal experience with the product.

Notwithstanding that, the Guides do not provide for any duty of disclosing virtual identity and some scholars have already recommended to fill this gap⁵¹. It is also noteworthy to say that the state of California recently introduced a ban from using avatars in political communication, but this is valid only "within 60 days of an election"⁵².

7.4 Other prospective and forthcoming regulations

Other specific *ad hoc* legislative interventions may be expected in the next future.

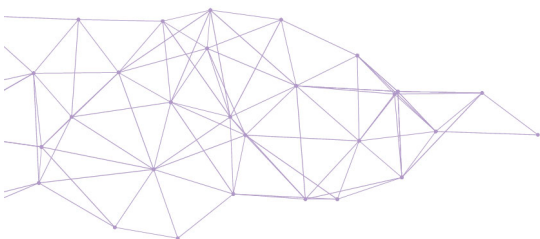
On 23 April, the Italian Council of Ministers passed a bill (AI Bill) aimed at introducing national AI provisions, supplementing the AI Act. The AI Bill consists of five parts. The first, devoted to establish "Principles and purposes", set a general principal of transparency. The

⁴⁹ ASCI (2021). The Code for self-regulation of advertising content in India. See Article 1.4 of the Guidelines for influencer advertising in digital media providing that "A virtual influencer must additionally disclose to consumers that they are not interacting with a real human being. This disclosure must be upfront and prominent". On this topic, see Priyanka Patnaik, 'Regulations for Social Media Influencers and Celebrity Endorsement' (2021) 3(1) Indian Journal of Law and Legal Research 1.

⁵⁰ 16 CFR Part 255: Guides Concerning the Use of Endorsements and Testimonials in Advertising.

⁵¹ Jim Masteralexis, Steve McKelvey and Keevan Statz, '#IAMAROBOT: Is It Time for the Federal Trade Commission to Rethink Its Approach to Virtual Influencers in Sports, Entertainment, and the Broader Market?' (2021) 2021(12) Harvard Journal of Sports & Entertainment Law 353, 376.

⁵² A.B. 730, 2019 Leg., Reg. Sess. (Cal. 2019).



bill introduces as well provisions on identifying textual, photographic, audiovisual, and radio content generated or altered by AI systems. Providers of audiovisual and radio broadcasting services can generate or alter content using AI tools. And they can present facts and information that are not real as actual data. But they have to get the consent of the relevant right holders and identify – with the acronym AI – the content by inserting identification elements or signs that are visible and recognizable by users.

The identification must be present at the beginning and end of broadcasts and content. It's not required for creative, satirical, artistic, or fictitious content unless it's detrimental to the rights and freedoms of third parties. In addition, the Italian Communication Authority (AGCOM) will have to promote forms of co-regulation and self-regulation using a code of conduct with both audiovisual and radio media service providers and video-sharing platform providers.

The bill also states that the unlawful dissemination of content generated or manipulated by AI to mislead as to its genuineness (eg deepfakes) is punishable by imprisonment from six months to three years and from one to five years if it causes unjust damage.

It is also worth noting 'The Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence' released by the Biden administration on 30 October 2023. It makes reference to the need "to protect Americans from AI-enabled fraud and deception by establishing standards and best practices for detecting AI-generated content and authenticating official content". The Department of Commerce will develop guidance for content authentication and watermarking to clearly label AI-generated content.

8. Existing legal framework establishing transparency obligations in advertising

Other obligations to disclose virtual identity do not exist at the EU level. Therefore, without *ad hoc* legislative interventions, courts should interpret extensively transparency obligations already established by EU law, even if they do not strictly concern fictional characters.

In particular, a general duty to inform users that they are interacting with a fictional person should arise from several different EU pieces of legislation, from the Unfair Commercial Practice Directive, to the Digital Services Act. I will analyse the relevant provision in the next paragraphs.

8.1 Omission of material information under the Unfair Commercial Practices Directive

Any BtoC practice that materially distorts or is likely to distort the economic behaviour of an average consumer normally amounts to a misleading practice, as regulated by the Directive 2005/29/EC (Art. 6) concerning unfair commercial practices (UCPD). Misleading practices could be either by actions or by omissions. Articles 7(1) and (2) establish a positive obligation on traders to provide all the 'material information' that the average consumer needs to make an informed purchasing decision.

The UCPD does not define 'material information'. However, by way of interpretation, it is possible to argue that the virtual nature of an endorser should be considered as such. Relevant to this purpose is Article 7.4 which mandates to disclose "whether the third party offering the products is a trader or not" even if limited to the specific case of an 'invitation

to purchase' on online marketplaces. Furthermore, in the *Wathelet* case⁵³, the Court stressed that "it is essential that consumers are aware of the identity of the seller". The same principle should be extended to an influencer as a person 'acting in the name of or on behalf of a trader'. Indeed, the EU Commission clarified that for the purposes of the UCPD, an influencer may be qualified as a 'trader'⁵⁴. Consequently, the obligation to be clear about the identity concerns directly all persons that carry out promotional activities towards consumers on behalf of a trader.

8.2 The duty to disclose the commercial intent of a commercial practice

Other arguments in favour of a disclosure duty may be derived by analogy with other information requirements established directly by EU law⁵⁵.

First, Article 6(a) of the e-Commerce Directive⁵⁶, Articles 9, 10 and 28(b) of the Audiovisual Media Services Directive (AMSD)⁵⁷, similarly to Article 7(2) UCPD, establish that failing to identify the commercial intent of a practice is regarded as a misleading omission. Coherently, the driving force behind all influencer marketing regulation adopted by advertising self-regulation authorities is the principle that influencers must disclose when they have a material connection with brands they promote through clear and understandable disclaimers such as #ad or #sponsored⁵⁸.

This principle has been recently established by the Italian Authority for Communications Guarantees (AGCOM) in its new guidelines on influencer marketing⁵⁹. The guidelines require influencers to comply with certain rules of the Consolidated Act on audio-visual media services (TUSMA)⁶⁰, which implements in Italy the AMSD. This means that influencers should comply with its article 9, stating that "audiovisual commercial communications shall be readily recognisable as such". Unfortunately, the guidelines do not mention virtual influencers and the duty of disclosure does not concern specifically their status.

Second, the UCPD prohibits as misleading by default "falsely claiming or creating the impression that the trader is not acting for purposes related to his trade, business, craft or profession or falsely representing oneself as a consumer". The purpose of this information requirement is to make sure that consumers always understand the very nature of the communication and know with whom they are interacting

⁵³ EUCJ, 9 November 2016, *Sabrina Wathelet v Garage Bietheres & Fils SPRL*, C-149/15, para 37.

⁵⁴ Guidance on the interpretation and application of Directive 2005/29/EC concerning unfair business-to-consumer commercial practices in the internal market, 21 December 2024.

⁵⁵ Article 7(5) UCPD clarifies that 'information requirements established by EU law in relation to commercial communication, including advertising', shall be regarded as material information by "default".

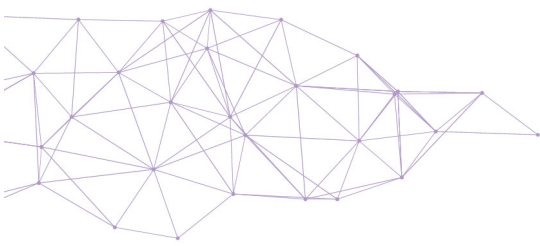
⁵⁶ Directive 2000/31/EC on certain legal aspects of information society services, in particular electronic commerce, in the Internal Market.

⁵⁷ 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.

⁵⁸ Jacopo Ciani and Massimo Tavella, 'La riconoscibilità della natura pubblicitaria della comunicazione alla prova del digital: native advertising tra obbligo di disclosure e difficoltà di controllo' (2017) 2(1) *Informatica e diritto* 485.

⁵⁹ AGCOM, Linee-guida volte a garantire il rispetto delle disposizioni del testo unico da parte degli influencer e istituzione di un apposito tavolo tecnico, Allegato A delibera n. 7/24/CONS, available at <https://www.agcom.it/documents/10179/32926720/Allegato+16-1-2024/2e637eaf-dec5-4ded-ab19-bcf99904163d?version=1.0>.

⁶⁰ Legislative Decree No. 208 of 8 November 2021.



online. The same transparency interest exists when consumers are facing virtual influencers.

8.3 Transparency obligations under the Consumer Rights Directive and the Digital Services Act

EU law is generally averse to any form of hidden marketing.

Article 8(5) of the Consumer Rights Directive⁶¹, in the case of telemarketing, held the trader to “disclose the identity and, where applicable, the identity of the person on whose behalf he makes that call, and the commercial purpose of the call”⁶².

In the same direction, social media platforms have recently seen their transparency obligations reinforced by the Digital Services Act^{63 64}. Recital 68 states that “providers of online platforms should...be required to ensure that the recipients of the service have certain individualised information necessary for them to understand when and on whose behalf the advertisement is presented”. Article 26 establishes that for any advertisement presented, the recipients should be “able to identify, in a clear, concise and unambiguous manner and in real time, the following:... (b) the natural or legal person on whose behalf the advertisement is presented; (c) the natural or legal person who paid for the advertisement if that person is different from the natural or legal person referred to in point (b)”⁶⁵.

Based on these principles, it is not surprising that the European Consumers Organisation⁶⁶ recently supported the introduction of two disclosure obligations regarding “edited” or “altered” content (e.g. when a picture has been photoshopped), and “virtual picture” or content for virtually created images (via AI for instance).

9. Legislation and practices concerning edited and retouched bodies and images

Other hard and soft law legislation concern specifically the practice of editing or retouching bodies and images of natural persons. This is relevant to our purposes, to the extent to which the discrimen between fictional or edited images may be difficult to draw but the practice of photo retouching raises similar issues of altering consumers’ perception.

Some countries, like France⁶⁷ and Norway⁶⁸, implemented legal provisions that require informing consumers whenever an advertising image or video has been digitally retouched.

In Norway, the Regulation do not specify any requirements of the scope of the alterations before the labeling requirement is triggered. As a result of that, the

⁶¹ Directive 2011/83/EU on consumer rights.

⁶² Joasia Luzak, ‘Passive Consumers vs. The New Online Disclosure Rules of the Consumer Rights Directive’ [2015] SSRN Electronic Journal <<http://dx.doi.org/10.2139/ssrn.2553877>> accessed 21 June 2024.

⁶³ Regulation (EU) 2022/2065 on a Single Market for Digital Services.

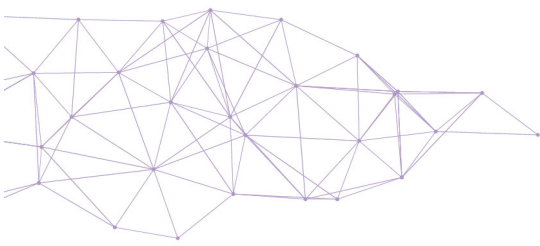
⁶⁴ Caroline Cauffman and Catalina Goanta, ‘A New Order: The Digital Services Act and Consumer Protection’ [2021] European Journal of Risk Regulation 1, <<http://dx.doi.org/10.1017/err.2021.8>> accessed 21 June 2024.

⁶⁵ Providers of online platforms are also requested to “provide recipients of the service with a functionality to declare whether the content they provide is or contains commercial communications”.

⁶⁶ From Influence to responsibility. Time to regulate influencer marketing, Position paper No BEUC-X-2023-093, 18 July 2023.

⁶⁷ Décret n° 2017-738 du 4 mai 2017 relatif aux photographies à usage commercial de mannequins dont l'apparence corporelle a été modifiée.

⁶⁸ Marketing Control Act.



advertisement must be labeled if one removes one single pimple, tattoo, wound, etc. Furthermore, one is required to label the picture or film regardless if the alteration contains beautification or not. General post-editing of pictures or films, for example, brightness, contrast, saturation, black and white editing, among other things, does not have to be labelled. However, general post-editing may in some cases, especially if the post-editing appears as speculative in the shape, size, or skin of the body, be labeled. Whether post-editing prompts the advertisement to be labeled must be assessed from case to case. Relevant considerations are the extent of the editing and whether it is likely to entail body-image pressure. The purpose of the label requirement is to counteract body-image pressure and applies to alterations in the relevant advertising person's body, size, or skin. Changing other matters in the picture or film does not contribute to body-image pressure. Retouching or manipulation of other matters than a person in the photo or film does not have to be labeled, such as buildings, nature, etc. The Regulation set out requirements regarding the design, size, and placement of the label. The label must be placed in the upper left corner of the advertising, must contrast with the background and cover approximately 7 % of the image surface. In case of film, it must be displayed for the entire duration, even if the retouched or manipulated person is not visible during the whole film. The use of the label will also fulfill the requirement to identify advertising. In practice, this means that posts do not have to be additionally labeled as advertising.

In the UK, a proposal for a Digitally Altered Body Images Bill requiring advertisers, broadcasters and publishers to display a "fundamentally changed" logo in cases where an image of a human body or body part has been digitally altered in its proportions, lies in the House of Commons since 2022.

Notwithstanding that, the Advertising Standards Authority (ASA) adopted guidelines on the use of pre- and post-production techniques in cosmetic ads and on the use of production techniques in beauty ads. They make clear that such techniques should not misleadingly exaggerate the effect that a product is capable of achieving. Based on these provisions, the ASA has upheld complaints against a number of influencers for using post-production techniques such as beauty filters to affect features directly related to the product's performance.

In Spain, the Self-Regulation Code developed by Autocontrol in collaboration with the perfume and cosmetics sector similarly states that digital techniques should not alter the models' images to the extent that their body silhouettes or characteristics appear unrealistic or deceptive regarding the actual efficacy attributed to the product, nor attribute qualities or functions to the product that it does not possess.

The French Advertising Regulatory Authority (ARPP) has also expressed a similar stance, creating a code of recommendations in 2019.

Unlike neighboring countries, Italy has not included a specific provision concerning the use of editing tools in advertising. However, the Italian advertising self-regulation authority showed to share the same principle when it ordered a company to desist from an advertising campaign for an anti-wrinkle filler serum⁶⁹, using a face-rendering in order to make a before-after comparison and boasting the effectiveness of the advertised product in correcting and reducing wrinkles in just 5 minutes.

The authority considered the rendering misleading because the advertiser failed to prove that the captures were real images of the face. On these grounds, the use of images in advertising seems to burden the advertiser with the duty to give evidence that they have not artificially generated them.

⁶⁹ Decision no. 31/2023.

10. Conclusions

Based on this critical evaluation and review of the current and forthcoming normative framework, the article remarked at least two significant gaps.

First, existing legislation requests to be transparent about using virtual characters in their communications through disclaimers. This kind of obligation is placed by the AI Act for virtual agents which technically amount to an AI system and should be complemented by a similar duty of disclosure for those who are not AI-generated. CGI may still perfectly resemble a human and pose similar risks of impersonation or deception.

Luckily enough, this gap may be filled quite easily applying extensively duties of transparency already existing in EU law against hidden or misleading advertising. In addition to existing legislation, it might be helpful to build explicitly ethical standards into AI systems architecture, so it has no choice but to adhere to them. The dialogue management algorithm, for example, should be backed with various explicit principles, like “Don’t try to fool your user that you are a human”. A principle of this kind might be helpful in establishing trust with users. The second challenge is whether duties of disclosure are enough.

Meta spoused its Oversight Board’s view according to which a “less restrictive” approach to manipulated media like labels with context should be the better way to address the issue rather than unnecessarily restricting freedom of expression by removing manipulated media. My opinion is that this is not always the case and some more accurate distinction should be introduced.

For instance, whereby commercial communications contain a testimonial or endorsement of a product or service by a virtual influencer, a mere disclosure of virtuality does not prevent consumers’ deception. Indeed, in this case, the communication would not be, by definition, genuine, verifiable, and relevant.

It is relevant in this connection that the EU Omnibus Directive prohibits fake reviews and endorsements (such as ‘likes’ on social media) of products and requires platforms to verify their authenticity and take reasonable and proportionate steps to ensure that these reviews are genuine and reflect the experience of real consumers. The Directive also establishes that traders giving access to such reviews should clearly state how the reviews are obtained and checked, and how they ensure that these come from consumers who have used or purchased the product⁷⁰.

The same approach guided advertising self-regulation authorities⁷¹ to require that marketers must hold documentary evidence that a testimonial or endorsement used in a marketing communication is genuine, i.e. that the quote is from a real person, and it reflects what this said.

To the same point leads the legislation concerning the use of editing or retouching softwares, which disallows any alteration of the formal aspect and features of bodies and products.

On these grounds, we should carefully assess if testimonials or endorsements by virtual influencers should be banned by default because of their lack of authenticity.

⁷⁰ Mateja Đurović and Tim Kniepkamp, ‘Good advice is expensive – bad advice even more: the regulation of online reviews’ (2022) 14(1) Law, Innovation and Technology 128, <<http://dx.doi.org/10.1080/17579961.2022.2047523>> accessed 21 June 2024.

⁷¹ As the ICC under Article 13 Advertising and Marketing Communications Code or the UK Advertising Standard Authority – ASA under rule 3.45 of the UK Code of Non-broadcast Advertising and Direct & Promotional Marketing – CAP Code.



Bibliography

Aicardi, C. et al. (2020) 'Ethical and social aspects of Neurorobotics', *Science and Engineering Ethics*, 26(5), pp. 2533–2546. doi:10.1007/s11948-020-00248-8.

Advertising Standards Council of India (ASCI) (2021). The Code for self-regulation of advertising content in India.

Arsenyan, J. and Mirowska, A. (2021) 'Almost human? A comparative case study on the social media presence of virtual influencers', *International Journal of Human-Computer Studies*, 155, p. 102694. doi:10.1016/j.ijhcs.2021.102694.

Autorità per le Garanzie nelle Comunicazioni (AGCOM) Linee-guida volte a garantire il rispetto delle disposizioni del testo unico da parte degli influencer e istituzione di un apposito tavolo tecnico, Allegato A delibera n. 7/24/CONS, available at <https://www.agcom.it/documents/10179/32926720/Allegato+16-1-2024/2e637eaf-dec5-4ded-ab19-bcf99904163d?version=1.0>.

Assembly Bill 730, 2019 Leg., Reg. Sess. (Cal. 2019).

Benefits of biological ai: Soul machines (no date) Soul Machines_042324. Available at: <http://www.soulmachines.com/biological-ai> (Accessed: 15 July 2024).

BEUC Position paper From Influence to responsibility. Time to regulate influencer marketing, BEUC-X-2023-093, 18 July 2023.

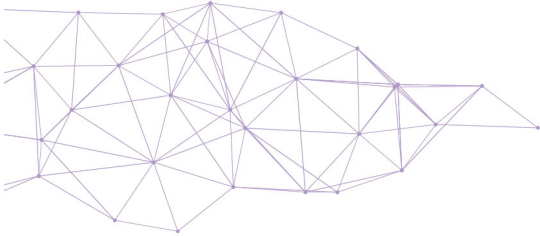
Boerman, S.C. (2020) 'The effects of the standardized Instagram disclosure for Micro- and Meso-influencers', *Computers in Human Behavior*, 103, pp. 199–207. doi:10.1016/j.chb.2019.09.015.

Brown, A. (2015) 'Picture [Im]Perfect: Photoshop Redefining Beauty in Cosmetic Advertisements, Giving False Advertising a Run for the Money', *Tex. Rev. Ent. & Sports L*, 87.

Cauffman, C. and Goanta, C. (2021) 'A new order: The digital services act and consumer protection', *European Journal of Risk Regulation*, 12(4), pp. 758–774. doi:10.1017/err.2021.8.

Ciani, J. (2024) 'The legal challenges of the digital afterlife industry and the new AI driven memorial chatbots', in *AI Approaches to the Complexity of Legal Systems*. Springer (forthcoming).

Ciani, J. and Tavella, M. (2017) 'La riconoscibilità della natura pubblicitaria della comunicazione alla prova del digital: native advertising tra obbligo di disclosure e difficoltà di controllo', *Informatica e diritto*, 2(1), p. 485.



Danaher, J. (2019) 'Welcoming robots into the moral circle: A defence of ethical behaviourism', *Science and Engineering Ethics*, 26(4), pp. 2023–2049. doi:10.1007/s11948-019-00119-x.

Décret n° 2017-738 du 4 mai 2017 relatif aux photographies à usage commercial de mannequins dont l'apparence corporelle a été modifiée, *French JORF* n°0106 5.5.2017

Decreto legislativo 8 novembre 2021, n. 208, Attuazione della direttiva (UE) 2018/1808 del Parlamento europeo e del Consiglio, del 14 novembre 2018, recante modifica della direttiva 2010/13/UE, relativa al coordinamento di determinate disposizioni legislative, regolamentari e amministrative degli Stati membri, concernente il testo unico per la fornitura di servizi di media audiovisivi in considerazione dell'evoluzione delle realtà del mercato, *Italian GU* n.293 of 10.12.2021.

Deng, F. and Jiang, X. (2023) 'Effects of human versus virtual human influencers on the appearance anxiety of social media users', *Journal of Retailing and Consumer Services*, 71, p. 103233. doi:10.1016/j.jretconser.2022.103233.

DLA PIPER, 'Metaverse: business opportunities and legal challenges', 24

Durante, M. (2011) 'The online construction of Personal Identity Through Trust and Privacy', *Information*, 2(4), pp. 594–620. doi:10.3390/info2040594.

Đurović, M. and Kniepkamp, T. (2022) 'Good advice is expensive – bad advice even more: The regulation of online reviews', *Law, Innovation and Technology*, 14(1), pp. 128–156. doi:10.1080/17579961.2022.2047523.

EPRS (2024) *Metaverse Opportunities, risks and policy implication*. rep.

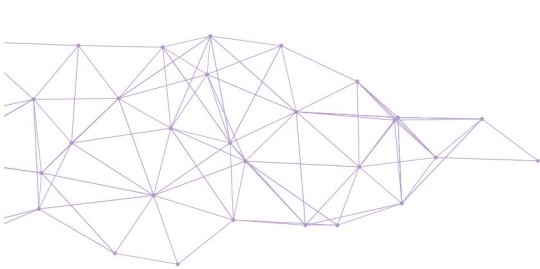
European Commission Guidance on the interpretation and application of Directive 2005/29/EC concerning unfair business-to-consumer commercial practices in the internal market, C/2021/9320.

European Committee of the Regions Opinion European approach to artificial intelligence – Artificial Intelligence Act (2022/C 97/12),

European Parliament and Council 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, OJ L 95, 15.4.2010.

European Parliament and Council Directive 2000/31/EC on certain legal aspects of information society services, in particular electronic commerce, in the Internal Market, OJ L 178, 17.7.2000.

European Parliament and Council Directive 2011/83/EU on consumer rights, OJ L 304, 22.11.2011.



European Parliament and Council Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market For Digital Services and amending Directive 2000/31/EC (Digital Services Act), OJ L 277, 27.10.2022.

European Parliament legislative resolution on the proposal for a regulation of the European Parliament and of the Council on laying down harmonised rules on Artificial Intelligence (Artificial Intelligence Act) and amending certain Union Legislative Acts (COM(2021)0206 – C9-0146/2021 – 2021/0106(COD)), 13 March 2024.

Farid, H. (2009) Seeing Is Not Believing. publication. IEEE Spectrum, 44.

Federal Trade Commission Endorsement Guides, 16 CFR Part 255: Guides Concerning the Use of Endorsements and Testimonials in Advertising.

Festinger, L. (1954) 'A theory of social comparison processes', Human Relations, 7(2), pp. 117–140. doi:10.1177/001872675400700202.

Franke, C., Groeppel-Klein, A. and Müller, K. (2023) 'Consumers' responses to virtual influencers as advertising endorsers: Novel and effective or uncanny and deceiving?', Journal of Advertising, 52(4), pp. 523–539. doi:10.1080/00913367.2022.2154721.

Gabriel, I., Manzini, A. (2024) 'The Ethics of Advanced AI Assistants', Google DeepMind, 81.

Gill, R. (2023) Perfect: Feeling judged on social media. Cambridge, UK: Polity.

Hayes, P. (2020) 'An ethical intuitionist account of transparency of algorithms and its gradations', Business Research, 13(3), pp. 849–874. doi:10.1007/s40685-020-00138-6.

Heller, B. and Bar-Zeev, A. (2021) 'The problems with immersive advertising: In Ar/VR, nobody knows you are an ad', Journal of Online Trust and Safety, 1(1). doi:10.54501/jots.v1i1.21.

Hildt, E., Laas, K. and Sziron, M. (2020) 'Editorial: Shaping ethical futures in brain-based and Artificial Intelligence Research', Science and Engineering Ethics, 26(5), pp. 2371–2379. doi:10.1007/s11948-020-00235-z.

Hudson, S. et al. (2019) 'With or without you? interaction and immersion in a virtual reality experience', Journal of Business Research, 100, pp. 459–468. doi:10.1016/j.jbusres.2018.10.062.

Istituto dell'Autodisciplina Pubblicitaria, Giuri, Decision no. 31/2023.

Jobin, A., Ienca, M. and Vayena, E. (2019) 'The global landscape of AI ethics guidelines', Nature Machine Intelligence, 1(9), pp. 389–399. doi:10.1038/s42256-019-0088-2.

Kee, E. and Farid, H. (2011) 'A perceptual metric for photo retouching', Proceedings of the National Academy of Sciences, 108(50), pp. 19907–19912. doi:10.1073/pnas.1110747108.

Knott, A., Sagar, M. and Takac, M. (2021) 'The ethics of interaction with Neurorobotic Agents: A case study with Babyx', *AI and Ethics*, 2(1), pp. 115-128. doi:10.1007/s43681-021-00076-x.

Koebler, J. (2024) Deepfaked celebrity ads promoting Medicare scams run rampant on YouTube, 404 Media. Available at: http://www.404media.co/joe-rogan-taylor-swift-andrew-tate-ai-deepfake-youtube-medicare-ads/?mc_cid=f250e2b063&mp (Accessed: 15 July 2024).

Luzak, J. (2015) 'Passive consumers vs. the new online disclosure rules of the Consumer Rights Directive', *SSRN Electronic Journal [Preprint]*. doi:10.2139/ssrn.2553877.

Masteralexis, J., McKelvey, S. and Statz, K. (2021) '#IAMAROBOT: Is It Time for the Federal Trade Commission to Rethink Its Approach to Virtual Influencers in Sports, Entertainment, and the Broader Market?', *Harvard Journal of Sports & Entertainment Law*, 12, pp. 353-376.

Mende, M. et al. (2019) 'Service robots rising: How humanoid robots influence service experiences and elicit compensatory consumer responses', *Journal of Marketing Research*, 56(4), pp. 535-556. doi:10.1177/0022243718822827.

Moon, Y. (2003) 'Don't blame the computer: When self-disclosure moderates the self-serving bias', *Journal of Consumer Psychology*, 13(1-2), pp. 125-137. doi:10.1207/s15327663jcp13-1&2_11.

Mouritzen, S.L., Penttinen, V. and Pedersen, S. (2023) 'Virtual Influencer Marketing: The good, the bad and the Unreal', *European Journal of Marketing*, 58(2), pp. 410-440. doi:10.1108/ejm-12-2022-0915.

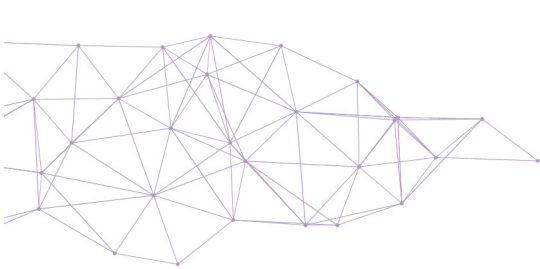
Moustakas, E. et al. (2020) 'Blurring lines between fiction and Reality: Perspectives of experts on marketing effectiveness of virtual influencers', 2020 International Conference on Cyber Security and Protection of Digital Services (Cyber Security) [Preprint]. doi:10.1109/cybersecurity49315.2020.9138861.

Mustak, M. et al. (2023) 'Deepfakes: Deceptions, mitigations, and opportunities', *Journal of Business Research*, 154, p. 113368. doi:10.1016/j.jbusres.2022.113368.

New disclosures and labels for Generative AI content on YouTube - YouTube community (no date) Google. Available at: <https://support.google.com/youtube/thread/264550152/new-disclosures-and-labels-for-generative-ai-content-on-youtube?hl=en>. (Accessed: 15 July 2024).

Noggle, R. (2022) 'The Ethics of Manipulation', *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University.

Our approach to labeling AI-generated content and Manipulated Media (2024) Meta. Available at: <https://about.fb.com/news/2024/04/metas-approach-to-labeling-ai-generated-content-and-manipulated-media> (Accessed: 15 July 2024).



Pagallo, U. and Durante, M. (2022) 'The good, the bad, and the invisible with its opportunity costs: Introduction to the "J" special issue on "the impact of Artificial Intelligence on law"', *J*, 5(1), pp. 139–149. doi:10.3390/j5010011.

Patnaik, P. (2021) 'Regulations for Social Media Influencers and Celebrity Endorsement', *Indian Journal of Law and Legal Research*, 3(1).

Rosenberg, L. (2022) 'Regulation of the metaverse: A roadmap', 2022 the 6th International Conference on Virtual and Augmented Reality Simulations [Preprint]. doi:10.1145/3546607.3546611.

Sabrina Wathelet v Garage Bietheres & Fils SPRL (Case C-149/15) [2016] ECLI:EU:C:2016:840.

Savageaux, L. (2022) 'Virtual Influencers: Harmless Advertising or Dystopian Deception?', De Pauw The Prindle Institute for Ethics [Preprint].

Sun, W., Gao, W. and Geng, R. (2021) 'The impact of the interactivity of internet celebrity anchors on consumers' purchase intention', *Frontiers in Psychology*, 12. doi:10.3389/fpsyg.2021.757059.

Tiggemann, M. and Anderberg, I. (2019) 'Social media is not real: The effect of "instagram vs reality" images on women's social comparison and Body Image', *New Media & Society*, 22(12), pp. 2183–2199. doi:10.1177/1461444819888720.

TikTok creator academy: Empowering creators to grow and succeed on TikTok: TikTok for creator (no date) TikTok. Available at: <http://www.tiktok.com/creators/creator-portal/en-us/community-guidelines-and-safety/ai-generated-content-label> (Accessed: 15 July 2024).

Tulving, E. (2002) 'Episodic memory: From mind to brain', *Annual Review of Psychology*, 53(1), pp. 1–25. doi:10.1146/annurev.psych.53.100901.135114.

Value creation in the metaverse. The real business of the virtual world. (2022) rep. McKinsey, 6.

What is the metaverse—and what does it mean for business? (2022) McKinsey & Company. Available at: <http://www.mckinsey.com/business-functions/mckinsey-digital/our-insights/what-is-the-Metaverse-and-what-does-it-mean-for-business> (Accessed: 15 July 2024).

Woodroof, P.J. et al. (2020) 'What's done in the dark will be brought to the light: Effects of influencer transparency on product efficacy and purchase intentions', *Journal of Product & Brand Management*, 29(5), pp. 675–688. doi:10.1108/jpbm-05-2019-2362.

World Economic Forum (2024) Metaverse Identity: Defining the Self in a Blended Reality. rep., p. 32.



MetaverseUA
Chair



Universitat d'Alacant
Universidad de Alicante

