

Rassegna

**di Diritto della Moda
e delle Arti**

dottrina
e giurisprudenza
commentata

Rivista semestrale
Fasc. 2/2022

dirittomodaearti.it

ISSN 2785-6259

Rassegna di diritto della moda e delle arti

Rivista semestrale

Direttore scientifico

Prof. Enrico Damiani

Comitato di direzione

Enrico Al Mureden (Alma Mater Studiorum Università di Bologna), Andrea Barengi (Università degli Studi del Molise), Laura di Bona (Università degli Studi Carlo Bo), Enrico del Prato (Università degli Studi di Roma La Sapienza), Romolo Donzelli (Università degli Studi di Macerata), Laura Marchegiani (Università degli Studi di Macerata), Sandro Nardi (Università degli Studi di Foggia), Barbara Pozzo (Università degli Studi dell' Insubria), Stefano Villamena (Università degli Studi di Macerata).

Comitato scientifico

Benedetta Agostinelli (Università Roma Tre), Luca Barchiesi (Università degli Studi di Macerata), Nathan Allix (Université Paris II Panthéon-Assas), Alessio Bartolacelli (Università degli Studi di Macerata), Giovanni Berti De Marinis (Università di Perugia), Charles-Edouard Bucher (Université de Nantes), Ermanno Calzolaio (Università degli Studi di Macerata), Ilaria Caggiano (Università degli Studi Suor Orsola Benincasa), Francesca Cangelli (Università degli Studi di Foggia), Marta Cenini (Università dell' Insubria), Tommaso Dalla Massara (Università degli Studi di Roma Tre), Manuel Espejo Lerdo De Tejada (Universidad de Sevilla), Chiara Feliziani (Università degli Studi di Macerata), Francesco Gambino (Università degli Studi di Macerata), Fabrizio Marinelli (Università degli Studi dell' Aquila), Manuel Garcia Mayo (Universidad de Sevilla), Lucilla Gatt (Università degli Studi Suor Orsola Benincasa), Juan Pablo Murga Fernández (Universidad de Sevilla), Guido Noto La Diega (University of Stirling), Andreas Rahmatian (University of Glasgow), Carmela Robustella (Università degli Studi di Foggia), Claudio Scognamiglio (Università degli Studi di Roma Tor Vergata), Antoine Touzain (Université Rouen-Normandie), Laura Vagni (Università degli Studi di Macerata), Camilo Zufelato (Universidade de São Paulo).

Comitato di redazione

Caporedattrici: Giorgia Vulpiani (Caporedattrice sezione moda – Università degli Studi di Macerata), Chiara Iorio (Caporedattrice sezione arte – Università degli Studi di Macerata).
Alessandro Berrettini (Università degli Studi di Camerino), Chantal Bompreszi (Università di Bologna), Sara Cavagnero (Northumbria University), Maxime Cormier (Université Paris II Panthéon-Assas), Martina D' Onofrio (Università degli Studi di Verona), Jeanne De Dinechin (Université Paris II Panthéon-Assas), Francesca Ferretti (Università degli Studi di Camerino), Clémentine Hebrard (Université Paris II Panthéon-Assas), Eva Helesbeux (Université Paris II Panthéon-Assas), Carlotta Ippoliti Martini (Alma Mater Studiorum Università di Bologna), Daria Onitui (University of Edinburgh); Cristina Evangelia Papadimitriu (Università degli Studi di Messina).

Comitato di referaggio

Alessandra Dignani (Università di Macerata), Matilde D' Ottavio (Università Politecnica delle Marche), Francesca Guaricci (Università di Macerata), Miriam Larocca (Università di Camerino), Emanuele Montagna (Università di Camerino), Tanya Tiberi (Università di Macerata).

Registrazione presso il Tribunale di Macerata al n. 652/2022. Direttore responsabile: Lisa De Leonardis.

Editore: Associazione per la promozione dello studio e della ricerca in materie giuridiche, con sede in Macerata (MC), via Carducci n. 67

INDICE

SAGGI

ENRICO DAMIANI, <i>Cripto-arte e non fungible tokens: i problemi del civilista</i>	352
ARIANNA ALPINI, <i>Dalla ‘ platform economy’ alla ‘ clout economy’ . La discussa natura giuridica degli NFTs</i>	365
MARTA CENINI, <i>La digitalizzazione della bellezza</i>	375
GEO MAGRI, <i>Falsa o erronea attribuzione di opera d’arte e rimedi contrattuali</i>	385
DARIA ONITIU, <i>Problematising the (deep)fakē Self in Fashion for AI Governance</i>	412
FEDERICO ONNIS CUGIA, <i>La natura dei passion investments tra attività di consumo e prodotti finanziari</i>	429
FRANCESCA ROTOLO, <i>Creatività vs. originalità: la qualificazione di «opera dell’ ingegno» nell’ era dell’ intelligenza artificiale</i>	466
GIACOMO FURLANETTO, <i>La responsabilità estesa del produttore come key factor per una moda sostenibile: riflessioni comparatistiche</i>	488
VALENTINA ZAMPAGLIONE, <i>Diritto, moda e nuove tecnologie: chi risponde dei danni ambientali prodotti all’ estero?</i>	502

NOTE

SANDRO NARDI, <i>Il diritto all’ immagine del bene culturale</i>	518
ANGELA MENDOLA, <i>L’ illecito sfruttamento online dell’ immagine e dello pseudonimo del testimonial: spunti di comparazione con il sistema tedesco</i>	534
INDICE DEGLI AUTORI.....	561

SAGGI

Problematising the (deep)fake Self in Fashion for AI Governance

Daria Onitii

Postdoctoral researcher - University of Oxford

Imagine you being the protagonist in a movie where you, feeling trapped in a mirror, rip your clothes off, intending to create a rupture between the self and your reflection. This scene illustrates the character's dialectic experience, entailing the management of appearance and self-perception. It is a metaphorical embodiment of the character experiencing an inherent conflict – how shapes and presence are mirrored- and between the seen and unseen, the objective and subjective reality. Imagine now, how the characters' actions in a more contemporary setting, are shaped by AI techniques reproducing a so-called deepfake entailing the individual's image, whereby algorithms perform the actions' metaphorical significance, including disruptions between the body, the clothing, and self-perception.

This paper argues for a closer inspection of the interactional implications of deepfakes in fashion, focusing on the inter-relationship between the individual and technology performing narratives in fashion, and with particular reference to the EU Commission's proposal for a Regulation on Artificial Intelligence.

SUMMARY: 1. Fake Authenticity. 2. Mirrored Fakery. – 3. A Rupture for AI in fashion and governance. 4. Prohibiting subjective reality. – 5. Towards Snowfakes. – 6. Exhausted identity and concluding thoughts.

1. Synthetic media technologies transcend the spheres of intended and unintended uses of technology.¹ Advancement of unsupervised deep learning techniques allow for the proliferation of deepfakes altering the transmission or creation of information in video, audio, or text for their viewers.² The paradox with deepfakes is that whilst the internals work based on an adversarial game within a constrained latent space,³ their representations illustrate contradictory narratives: a political fact including a sincere lie, an intended action of revenge porn including unintended intimidation, or an honest testimony including fabrication of facts. With the proliferation of deepfakes, the question arises how to articulate the proliferation of synthetic content including the degrees of permissible and impermissible “fake authenticity”.

When we look at commercial deepfakes, such as the RefaceAI Application, which allowed end-users face swapping with fashion brands' branded content, we see that the goal of advancements in deep learning

¹ As argued by Nina I Brown, ‘[a]s deepfake technology matures and improves, it can potentially be abused in myriad ways’, see NI. BROWN, *Deepfakes and the Weaponization of Disinformation*, in *Virginia Journal of Law and Technology*, 20120, p. 9.

² Y. MIRSKY and W. LEE, *The Creation and Detection of Deepfakes: A Survey*, in *ACM Computing Surveys*, 2021, p. 1-2.

³ K. DE VRIES, *You never fake alone. Creative AI in action*, in *Information, Communication & Society*, 2020, p. 2114.

enables the *approximation* of realistic content.⁴ In particular, this technology works with Generative Adversarial Networks (GANs) which is a method of unsupervised learning, involving two neural networks pitted against each other to generate results ‘that are convincing enough that the second neural network believes the [output] are examples from the real world...’.⁵ Katja de Vries provides an illuminating discussion of the GAN’s adversarial process, which she describes as looking like a ‘sadistic game’ between the generator learning to produce an output that could overcome the hurdle of, let’s say a “face of a famous celebrity”, without knowing the parameters of how that celebrity could look like, and the discriminator who needs to expose the output as real or fake.⁶

Many popular application areas of GANs in fashion focus on fashion design, but there are other emerging contexts,⁷ where we see the proliferation of deepfakes in the fashion domain. For example, the fashion brand “Zalando” developed a campaign entailing the model Cara Delevingne to create alternative voice fonts and shots for personalised advertising back in 2018.⁸ Moreover, the Fashion Innovation Agency, Superspersonal and Hanger developed a virtual try-on application that allows consumers to ‘believably visualise themselves’ within the brand’s ad campaigns on their mobile phones.⁹

What policymakers are grappling with the most when confronted with commercial deepfakes is how to regulate the “fake” in the reproduction of reality.¹⁰ This is because there are many practical dangers to the proliferation of synthetic media technology, whereby deepfakes can accelerate fake news and micro-

⁴ This was a collaboration between the RefaceAI application and the luxury fashion brand ‘Gucci’ back in 2020; B. ROBERTS-ISLAM, *Why Fashion Needs More Imagination When It Comes To Using Artificial Intelligence*, in *Forbes* of 21 September 2020 www.forbes.com/sites/brookeroberstislam/2020/09/21/why-fashion-needs-more-imagination-when-it-comes-to-using-artificial-intelligence/?sh=503929233f63; C. MALLEY, *We Can All Be The Next Face of Gucci — Thanks to Deepfakes*, in *HYPERBEAST* of 4 September 2020 <https://hypebeast.com/2020/9/reface-ai-deepfakes-artificial-intelligence-fashion-interview>.

⁵ L. LUCE, *Artificial Intelligence for Fashion How AI is Revolutionizing the Fashion Industry*, Berkeley California (US), 2018, p. 134; see also, S. SYLVESTER, *Don’t Let Them Fake You Out: How Artificially Mastered Videos Are Becoming the Newest Threat in the Disinformation War and What Social Media Platforms Should Do About It*, in *Federal communications law journal*, 2021, p. 373.

⁶ K. DE VRIES, *You never fake alone. Creative AI in action*, cit., p. 2114- 2115.

⁷ L. LUCE, *Artificial Intelligence for Fashion How AI is Revolutionizing the Fashion Industry*, cit., p. 14; K. SOHN, C. EUYOUNG SUNG, G. KOO and O. KWON, *Artificial intelligence in the fashion industry: consumer responses to generative adversarial network (GAN) technology*, in *International Journal of Retail & Distribution Management*, 2020, p. 62-63; C. STOKEL-WALKER, *AI can change a fashion model’s pose and clothing*, in *New Scientist*, 2021, p. 18.

⁸ K. CHITRAKORN, *How deepfakes could change fashion advertising*, in *Vogue Business* of 11 January 2021 www.voguebusiness.com/companies/how-deepfakes-could-change-fashion-advertising-influencer-marketing.

⁹ FIA, SUPERPERSONAL and HANGER, *Using “Deep-Fake” Virtual Try-On To Bring LFW Attendees Into Fashion Presentations*, in *Fashion Innovation Agency* www.fialondon.com/projects/hanger-x-superspersonal/; see also, J. BURTON, *The Changing Face Of Fashion - How Virtual Fashion Is Going To Make The Unimaginable Real: e chat to the head of London College of Fashion’s Innovation Agency, Matthew Drinkwater, about fashion of the future. All-digital clothes, here we come*, in *Huffpost* of 3 September 2020 www.huffingtonpost.co.uk/entry/the-changing-face-of-fashion-how-virtual-fashion-is-going-to-make-the-unimaginable-real_uk_5eeca00dc5b6c1f6518b2087.

¹⁰ M. WESTERLUND, *The Emergence of Deepfake Technology: A Review*, in *Technology Innovation Management Review*, 2019, p. 42; J. KIETZMANN, LW LEE, IP MCCARTHY and T. ENKERLIN, *Deepfakes: Trick or treat?*, in *Business Horizons*, 2020, p.136; compare with A. YAMOKA-ENKERLIN, *Disrupting Disinformation: Deepfakes And The Law*, in *New York University Journal of Legislation and Public Policy*, 2020, p. 728.

targeting techniques, non-consensual pornography, identity theft, and reputational damage, to name a few.¹¹ Nevertheless, a recent report that was requested by EU Panel for the Future of Science and Technology (STOA) equally emphasises the positive uses of synthetic content, such as educational uses or examples of deepfakes uses for satire.¹² However, how we should promote the positive uses of deepfakes, such as enabling more diverse representations of body shapes regarding fashion ad campaigns, whilst oppressing any harmful uses, such as potential bias uses of this technology, remains an unanswered question.

The aspects of unreality in synthetic content is not readily discernible to the human eye, nor do an observer's beliefs develop around objective facts.¹³ Indeed, the role of AI techniques- from the use of algorithms to direct disinformation campaigns during the U.S election in 2016 and based on the Facebook user's "fashion taste" for avant-garde brands, to the fashion retailers' avatar-creation for targeted ads¹⁴- all these illustrate that the proliferation of evolving technology including deepfakes could transcend many contexts, without clear boundary lines of sensitivity and risk of harm, beyond truth and falsehood. Accordingly, many risks are based on the acceleration and accessibility of deepfake technology, whereby there are no benchmarks to assess the output's reliability, truthfulness, and objectivity in individual circumstances. What follows is that many are concerned that deepfakes will challenge our 'visual experience' with regard to 'any kind of information, whether true or false' (emphasis added).¹⁵

My thesis is that our concern with deepfakes is our ability to maintain the authenticity of our dialectic experience of truth and falsehood including the approximations of replicated reality in synthetic content. This is because the GAN's "cat and mouse" game to distinguish between fake and real for authenticity paradoxically moulds into peoples' perceptual beliefs of a narrative that is fake and authentic at the same time. In doing so, I believe that actually a distinction between objective and fabricated facts for regulating harmful uses of technology is an illusionary one with regard to deepfakes in fashion. This is because an individual's immersive experience – as a reproduced identity in a deepfake- is neither true nor false but

¹¹ See also, D. LU, *Dubbing with deepfakes*, in *New Scientist*, 2019, p. 8.

¹² M. VAN HUIJSTEE, P. VAN BOHEEMEN, D. DAS, L. NIERLING, J. JAHNEL, M. KARABOGA, M. FATUN, L. KOOL and J. GERRITSEN, *Tackling deepfakes in European policy*, in *Panel for the Future of Science and Technology (STOA)* of 30 July 2021, p. 26; see also, B. VAN DER SLOOT and Y. WAGENSVELD, *Deefakes: regulatory challenges for the synthetic society*, in *Computer Law & Security Review*, 2022, p. 3-4

¹³ D. FALLIS, *The Epistemic Threat of Deepfakes*, in *Philosophy & Technology*, 2021, p. 625; see also, KR. HARRIS, *Real Fakes: The Epistemology of Online Misinformation*, in *Philosophy & Technology*, 2022, p. 11-12.

¹⁴ M. FERRIER, *Christopher Wylie: 'The fashion industry was crucial to the election of Donald Trump'*, in *The Guardian* of 29 November 2018 www.theguardian.com/fashion/2018/nov/29/christopher-wylie-the-fashion-industry-was-crucial-to-the-election-of-donald-trump; J. PEREIRA, *Deepfakes and Fashion Advertising*, in *Medium* of 1 March 2021 <https://medium.com/futurists-club-by-science-of-the-time/deepfakes-and-fashion-advertising-bc99d308357e>.

¹⁵ DG. JOHNSON and N. DIAKOPOULOS, *What to do about deepfakes*, in *Communications of the ACM*, 2021, p. 33; M. VAN HUIJSTEE, P. VAN BOHEEMEN, D. DAS, L. NIERLING, J. JAHNEL, M. KARABOGA, M. FATUN, L. KOOL and J. GERRITSEN, *Tackling deepfakes in European policy*, cit., p. III.

can involve narratives that are harmful to self-perception. Not having the space to interrogate these contradictory beliefs leads to the practical consequence of deepfakes in the commercial context we should be worried about, and that goes beyond the need to detect and expose “false content”.

2. How do re-assess the implications of deepfake for the individual and society? It is important to begin with a fundamental misconception of the implications of deepfakes for AI governance, to elaborate on my thesis above. In particular, deepfakes in fashion requires us to move away from a conception of controlling a “mirrored fakery” of the self and consider the interactional implications of AI techniques for the inter-relationship between clothing, body and self-perception.

Let us focus on two examples to investigate some aspects of transparency regarding deepfakes in fashion. Consider an individual interacting with a virtual try-on application and engaging with the algorithms’ realistic pattern, capturing the individual’ s face, small mannerisms, and combining the person’ s features with an approximate representative body including the brand’ s ads.¹⁶ This “recombinant version of the self” sustains the end user’ s experience shaping and being shaped by deepfakes. In particular, it raises interesting questions concerning what form of information disclosure establishes the link between the ‘ authentic or truthful’ self and the ‘ artificially generated or manipulated’ content.¹⁷

Conversely, consider an individual interacting with a virtual try-on application and being shown a “disturbing” approximation of the self, including algorithms exaggerating the person’ s mannerisms, and features in relation to the brand’ s ads on beauty advice and cosmetic surgery.¹⁸ In this example, we see how “fashion” (i.e. the material components attached to the reflected body), being culturally informed by the technologies’ negotiation of the mirrored self, still plays a driving action in the individual’ s perception of appearance. Algorithmic infrastructures may evoke knowledge production and attribution beyond the end-users awareness.¹⁹ However, deepfakes add another dimension to the way technology mediates the substance of fashion based on the meanings attached to the individual’ s virtual presence.

¹⁶ See also, FIA, SUPERPERSONAL and HANGER, *Using "Deep-Fake" Virtual Try-On To Bring LFW Attendees Into Fashion Presentations*, cit.; K. BARON, *Digital Doubles: The Deepfake Tech Nourishing New Wave Retail*, in *Forbes* of 29 July 2019 www.forbes.com/sites/katiebaron/2019/07/29/digital-doubles-the-deepfake-tech-nourishing-new-wave-retail/?sh=d7c15124cc7b.

¹⁷ Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL LAYING DOWN HARMONISED RULES ON ARTIFICIAL INTELLIGENCE (ARTIFICIAL INTELLIGENCE ACT) AND AMENDING CERTAIN UNION LEGISLATIVE ACTS. COM/2021/206 final. 21 April 2021, art 52 (3).

¹⁸ See also, A. HOUSTON, *Ad of the Day: Dove deepfakes highlight toxic beauty advice on social media*, in *The Drum* of 27 April 2022 www.thedrum.com/news/2022/04/27/ad-the-day-dove-deepfakes-highlight-toxic-beauty-advice-social-media.

¹⁹ See for instance, M. HILDEBRANDT and B-J. KOOPS, *The Challenges of Ambient Law and Legal Protection in the Profiling Era*, in *Modern Law Review*, 2010, p. 428; D.MC QUILLAN, *Data Science as Machinic Neoplatonism*, in *Philosophy & Technology*, 2018, p. 262.

Whether labelling requirements of the deepfake need to highlight these socially and culturally concealed values, such as beauty standards on cosmetic surgery, that is another compelling question that needs to be examined. Further, how do we re-assess the contours of the social and cultural notions of fashion informing impermissible uses of technology, is an additional element that should inform discourse on the socio-legal implications of deepfakes.

Both examples show how neural networks resemble the individuals' performative role of "fashion" to manage and perceive appearance, but those algorithmic approximations direct the individuals' interactional presence within a social and cultural construct, based on neural nets' enablers learning from the training data. A more compelling concern is the deepfake's dynamics of the re-combinations of fashion on the individual and whether a rupture between the person and the mirrored self gives rise to the implications of synthetic technologies' manipulating individual behaviour. Kati Chitrakorn describes that deepfakes in fashion allow fashion brands to tailor content using more representative notions of different body shapes or skin tones, but those "mirrored identities" can create the individual to develop a sense of 'psychological ownership' of the 'extensions of themselves' and being nudged to 'buy more products, at higher prices, and even to willingly promote those products among their friends'.²⁰

The real question here is not *only* about promulgating information disclosure requirements exposing "recombinant identities", but on a fundamental level, we need to ask ourselves whether an end-user can effectively control fragments of self that are simultaneously mirrored whilst containing fabricated narratives on fashion, such as notions on the "representative body shape"? Hence, can we demand a sense of control of the mirrored fakery of the self, which only relies on our *subjective* experience?

This question brings us to an inherent conflict, which is our desire to control the truthfulness of an artificial embodiment of the self, whilst acknowledging that the metaphorical significance of deepfakes to replicate "fashion" is only a tiny aspect of our management and perception of appearance. It is a metaphorical embodiment of the individual experiencing an inherent conflict – how shapes and presence are mirrored- and between the seen and unseen, the objective and subjective reality which needs to be protected from the outset.

What follows is that our focus needs to be on the interactional implications of deepfakes, rather than the deepfake's factual representation of the self. This is the misconception on tackling commercial deepfakes for AI governance in that trustworthiness depends on our beliefs or the statistical significance attached

²⁰ K. CHITRAKORN, *How deepfakes could change fashion advertising*, cit.; reference to, CP. KIRK, *How Customers Come to Think of a Product as an Extension of Themselves*, in *Harvard Business Review* of 17 September 2018 <https://hbr.org/2018/09/how-customers-come-to-think-of-a-product-as-an-extension-of-themselves>.

to truthfulness and fabrication.²¹ Whilst I acknowledge the value of this approach with regard to the *intentional* use of deepfakes to accelerate political disinformation, reputational damage, or identity theft, I do feel that with deepfakes in personalisation and advertising we need to develop further guidelines for regulation. Two aspects are difficult to prove, as well as detect algorithmically, which are the *vulnerability* of the subject and the inauthenticity of the content. The first point is an issue with regard to commercial deepfakes in fashion, advertising and personalisation, whereas the second concern hints at solutions of technical impossibility to comprehensively find the ‘silver bullet’ detecting deepfakes.²²

Hence, I see the interactional implications of deepfakes in promoting contradictory narratives *about* an individual’s subjective reality, such as a fashion brand promoting products suiting the end-users face and mannerisms adapted to social and cultural narrative. We need to pay closer attention to the deepfake’s disruptions between the body, clothing, and self-perception for AI governance. In other words, we need to address how can we regulate deepfakes in fashion for end-users to control, focusing on (i) vulnerability regarding the interactive experience and, (ii) the inauthenticity of the dialectic content, in order to create a rupture between the self and your reflection. Both aspects will be examined by discussing the EU Commission’s proposal for a Regulation on Artificial Intelligence (AI Act proposal) including the relevant provisions on synthetic content.²³ In addition, I shall consider the Council of the European Union’s compromise text from the 6th of December 2022 (Council General approach), which illustrates the most recent version at the time of writing.²⁴

3. The AI Act proposal creates an inherent tension in assessing the interactional implications of deepfakes, precluding a nuanced approach to the risks of this technology in the fashion domain. This raises the question of how we should problematise the role of deepfakes –as a technology shaping the

²¹ Compare with K. DE VRIES, *You never fake alone. Creative AI in action*, cit., p.2119.

²² A. ELIAZAT, *European and UK Deepfake Regulation Proposals Are Surprisingly Limited*, in *Adolfo Eliazat* of 6 April 2022 <https://adolfoeliazat.com/2022/04/06/european-and-uk-deepfake-regulation-proposals-are-surprisingly-limited/>; I. SAMPLE, *What are deepfakes – and how can you spot them?*, *The Guardian* in 13 January 2020 www.theguardian.com/technology/2020/jan/13/what-are-deepfakes-and-how-can-you-spot-them; compare with J. HSU, *Deepfake detector could protect world leaders*, in *New Scientist*, 2022, p. 10; indeed, there are research challenges regarding the detection of fake audio as illustrated in Z. ALMUTAIRI and H. ELGIBREEN, *A Review of Modern Audio Deepfake Detection Methods: Challenges and Future Directions*, in *Algorithms*, 2022, p. 18; consider also the results of the so-called Deepfake Detection Challenge (DFDC) dataset that was published by Facebook AI and which highlights that ‘[d]eepfake detection is extremely difficult and still an unsolved problem’, see B. DOLHANSKY, J. BITTON, B. PFLAUM, J. LU, R. HOWES, M. WANG and C. CANTON FERRER, *The DeepFake Detection Challenge (DFDC) Dataset*, in ArXiv of 28 October 2020 <https://arxiv.org/pdf/2006.07397.pdf>.

²³ Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL LAYING DOWN HARMONISED RULES ON ARTIFICIAL INTELLIGENCE (ARTIFICIAL INTELLIGENCE ACT) AND AMENDING CERTAIN UNION LEGISLATIVE ACTS. COM/2021/206 final. 21 April 2021.

²⁴ Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts- General Approach. 2021/0106(COD)-14954/22, 25 November 2022.

inter-relationship between identity, body, and appearance- - from a governance perspective to ensure notions of trustworthy AI.²⁵

The AI Act proposal illustrates a top-down approach to AI governance in that it provides a restrictive view to the functional requirements of harmful deepfakes as well as an ambiguous view on the way technologies' substantive aspects need to be addressed. An important aspect of the AI Act proposal is its risk-based approach, distinguishing between requirements for high-risk systems, specific systems requiring transparency obligations, prohibited practices and all other AI systems that are of minimal risk.²⁶ Whilst a risk-based approach allows to capture evolving threats and unintended uses of technology, this is a methodology that evolved with open-ended standards focusing on the definitional aspects of prohibited practices and transparency obligations in the AI Act proposal.²⁷ A bottom-up approach concerning a risk-based methodology should clearly outline how deepfakes amplify vulnerabilities and which concrete measures are necessary to ensure meaningful control regarding the individual's approximations of replicated reality in synthetic content. For instance, a risk-classification methodology needs to include those voices who are impacted and endured by the way technology shapes "fashion"- from civil society to the fashion designer working with AI- to complement a contextual approach to AI governance.²⁸ The AI Act falls short of this nuance based on a loose classification between limited risk systems and prohibited practices.

As highlighted above, the regulation of deepfakes in fashion requires an understanding of how to balance between the uses of technology, whilst looking at the metaphorical embodiment of deepfakes that shaping the end-user's interactional experience. This requires us to debunk two important elements of the AI Act proposal including its risk-based approach to address issues of end-user vulnerability and inauthenticity dialectic experience regarding deepfakes. I am going to examine first the role of unacceptable uses of technology in Article 5 (1) (a) – (b) prohibiting a narrow view of subjective reality applicable to synthetic content technology.²⁹ Then, I will discuss the way the AI Act proposal promulgates a patchwork of "snowflakes" for the regulation of commercial deepfakes (in fashion).

²⁵ On the notion of trustworthy AI, please see High-Level Expert Group on Artificial Intelligence, *Ethics Guidelines for Trustworthy AI*, 8 April 2019, p. 2.

²⁶ Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL LAYING DOWN HARMONISED RULES ON ARTIFICIAL INTELLIGENCE (ARTIFICIAL INTELLIGENCE ACT) AND AMENDING CERTAIN UNION LEGISLATIVE ACTS, cit., Explanatory Memorandum; see also, N. ERIKSSON, *Europe draws up regulations to control AI risks*, in *World today*, 2021, p. 5.

²⁷ Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL LAYING DOWN HARMONISED RULES ON ARTIFICIAL INTELLIGENCE (ARTIFICIAL INTELLIGENCE ACT) AND AMENDING CERTAIN UNION LEGISLATIVE ACTS, cit.

²⁸ L. EDWARDS, *Regulating AI in Europe: four problems and four solutions*, in *Ada Lovelace Institute* of 31 March 2022, p.11.

²⁹ Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL LAYING DOWN HARMONISED RULES ON ARTIFICIAL INTELLIGENCE (ARTIFICIAL INTELLIGENCE ACT) AND

4. The AI Act proposal highlights that some actions entailing the misuse of technology for ‘manipulative, exploitative and social control practices’ should be banned as ‘they contract Union values...fundamental rights, including the right to non-discrimination, data protection and privacy and the rights of the child’.³⁰ Having said that, whilst emergent risks can only be witnessed incrementally, the ‘cascading impacts’ of deepfake technology ‘is not limited to a single type or category of risk, but rather to a combination of cascading impacts at different levels’.³¹ As rightly observed by Natalie Smuha, Emma Ahmed-Rengers, Adam Harkens et al, ‘[f]uture uses of AI systems can be hard to predict, and it seems premature to permanently fix the list of prohibited AI practices’.³² An important question is whether we can effectively identify the manner some manipulative AI techniques practices create a spill over effect altering an individual’s subjective reality with regard to a sensory stimuli.

Having said that, the AI Act proposal’s definition of prohibited practices only creates some soft lines among some contextual uses of technology, whilst leaving out the complexity of synthetic content technology in the commercial context. An important omission in Article 5’s definition of prohibited practices is to account for the different nuances of risks regarding commercial deepfakes.³³ In this respect, Article 5 (1) (a)- (b) of the AI Act proposal lists several contexts whereby a deepfake *could* ‘reasonably likely cause a person...physical or psychological harm’ when that technology is employing ‘subliminal techniques beyond a person’s consciousness [and] *in order to materially* distorting a person’s behaviour’ or when the tool is exploiting an individual’s specific vulnerabilities relating to a disability, or age

AMENDING CERTAIN UNION LEGISLATIVE ACTS, cit., art 5 (1) (a), art 5 (1) (b); compare with, Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts- General Approach, cit., amendment 1.3, art. 5 (1) (a), art 5 (1) (b).

³⁰ Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL LAYING DOWN HARMONISED RULES ON ARTIFICIAL INTELLIGENCE (ARTIFICIAL INTELLIGENCE ACT) AND AMENDING CERTAIN UNION LEGISLATIVE ACTS, cit., Recital 15; Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts- General Approach, cit., Recital 15.

³¹ M. VAN HUIJSTEE, P. VAN BOHEEMEN, D. DAS, L. NIERLING, J. JAHNEL, M. KARABOGA, M. FATUN, L. KOOL and J. GERRITSEN, *Tackling deepfakes in European policy*, cit., p. IV.

³² NA. SMUHA, E. AHMED-RENGERS, A. HARKENS, W. LI, J. MACLAREN, R. PISELLI and K. YEUNG, *How the EU Can Achieve Legally Trustworthy AI: A Response to the European Commission’s Proposal for an Artificial Intelligence Act*, in SSRN of 31 August 2021 https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3899991, p. 20; see also, G. SHARKOV, C. TODOROVA and P. VARBANOV, *Strategies, Policies, and Standards in the EU Towards a Roadmap for Robust and Trustworthy AI Certification*, in *Information & Security*, 2021, p.15.

³³ Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL LAYING DOWN HARMONISED RULES ON ARTIFICIAL INTELLIGENCE (ARTIFICIAL INTELLIGENCE ACT) AND AMENDING CERTAIN UNION LEGISLATIVE ACTS. cit., art 5.

(emphasis added).³⁴ The Council General approach alters the provision's requirement of 'intent', highlighting that the practices may include subliminal techniques or exploit vulnerabilities 'with the objective to *or the effect*' to materially alter individual behaviour and causing tangible harm (emphasis added).³⁵ Moreover, the Council General approach to the AI Act proposal further adds to the list of vulnerabilities with regard to technology, including practices that exploit 'a specific group of persons ... due to their social or economic situation'.³⁶ Whilst this reasoning enables us to establish some boundary work in how deepfakes can undermine an individual's autonomy of choice and produce systematic risks of bias,³⁷ these provisions do not give a conclusive answer of *when* the AI techniques *supporting* deepfakes should be banned in practice whilst interacting with the end-user(s).

In particular, one may not easily locate the extent the 'manipulation of reality' relates to the end user's ability to 'resist' any subliminal components in synthetic technologies,³⁸ as well as considering that the AI Act proposal does not define degrees of physical or psychological harm. The AI Act proposal focusing on the 'audio, image, [or] video stimuli' directing the consumers' conscious awareness does not highlight the degrees of manipulation that should be subject to an outright ban.³⁹ Indeed, the Council General

³⁴ Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL LAYING DOWN HARMONISED RULES ON ARTIFICIAL INTELLIGENCE (ARTIFICIAL INTELLIGENCE ACT) AND AMENDING CERTAIN UNION LEGISLATIVE ACTS, cit., art 5 (1) (a), art 5 (1) (b).

³⁵ Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts- General Approach, cit., art 5 (1) (a), art 5 (1) (b); consider other legislative frameworks, such as the Digital Services Act (DSA), which regulates the online platforms' use of subliminal techniques and dark patterns that may 'either on purpose or in effect' materially distort user choice, decision-making and autonomy; 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 277/1, 17 October 2022, Recital 67, art 25 (1), art 25 (3); it is important to note that the Digital Services Act considers dark patterns which are not captured by other legislative frameworks including the General Data Protection Regulation, and the Unfair Commercial Practices Directive, European Data Protection Board, *Guidelines 3/2022 on Dark patterns in social media platform interfaces: How to recognise and avoid them*, adopted 14 March 2022 https://edpb.europa.eu/system/files/2022-03/edpb_03-2022_guidelines_on_dark_patterns_in_social_media_platform_interfaces_en.pdf; European Commission, *Commission Notice – Guidance on the interpretation and application of Directive 2005/29/EC of the European Parliament and of the Council concerning unfair business-to-consumer commercial practices in the internal market*. OJ C 526 1, 29 December 2021; on the connection between subliminal techniques and dark patterns in the AI Act proposal, see F. LUPIANEZ- VILLANUEVA, A. BOLUDA, F. BOGLIACINO, G. LIVA, L. LECHARDOV and T. RODRIGUEZ DE LAS HERAS BALLELL, *Behavioural study on unfair commercial practices in the digital environment: dark patterns and manipulative personalisation*, in *Directorate-General for Justice and Consumers* of 16 May 2022 <https://op.europa.eu/en/publication-detail/-/publication/606365bc-d58b-11ec-a95f-01aa75ed71a1/language-en/format-PDF/source-257599418>, p. 83.

³⁶ Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts- General Approach, cit., art 5 (1) (b).

³⁷ Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts- General Approach, cit., Recital 16.

³⁸ Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts- General Approach, cit., Recital 16.

³⁹ Compare with Risto Uuk stating that '[a] stimuli would only be considered subliminal if it was presented for less than 50 milliseconds'; R. UUK, *Manipulation and the AI Act*, in *Future of Life Institute* of 18 January 2022) https://futureoflife.org/wp-content/uploads/2022/01/FLI-Manipulation_AI_Act.pdf; reference to, MR. IONESCU, *Subliminal perception of complex visual stimuli*, in *Romanian Journal of Ophthalmology*, 2016, p. 226.

approach further adds that ‘it is not necessary for the provider or the user to have the intention to cause the physical or psychological harm, as long as such harm results from the manipulative or exploitative AI-enabled practices’.⁴⁰ However, this statement contradicts with the argument that a provider and user’s intention to manipulate shall not be presumed ‘if the distortion results from ‘factors *external* to the AI system which are outside of [their] control’ (emphasis added).⁴¹ Hence, whilst the AI Act proposal’s definition of subliminal techniques seems broad on first sight,⁴² its reasoning focusing on intent and degrees of harm does not give a holistic picture on the way deepfake technology may manipulate consumers beyond conscious perception. In addition, the provisions, providing an ambiguous interpretation of the users and providers’ degree of control, does clarify which design choices exemplify patterns for manipulation and which practices correlate with an individual’s perceptual experience only. One important challenge is that deepfakes are explicitly calibrated to reflect, as well as distort an individual’s beliefs.⁴³ By way of illustration, suppose a virtual avatar showing the consumer ads resembling an audiences’ demographic and language.⁴⁴ One important factor is that the virtual avatar may drive user impression management to the degree of what *ought* to be a reflection of the self. Nevertheless, this technology needs to alter an individual’s *subjective experience*, considering that an AI system may materially distort behaviour and (likely) causing physical and psychological harm. In other words, much will depend on how we can verify an individual’s ‘sense of affinity’⁴⁵ with the virtual avatar to establish a necessary degree of harm.

However, this would make the detection of ‘subliminal techniques’ with regard to deepfakes difficult in practice. This is because deepfakes already embody visible and invisible features used for subliminal messaging, such as resembling the consumer’s face and using a calming voice for personalised messaging, for example. Detangling the visibility from the invisibility in distinguishing between the end user’s

⁴⁰ Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts- General Approach, cit., Recital 16.

⁴¹ Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts- General Approach, cit., Recital 16; Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL LAYING DOWN HARMONISED RULES ON ARTIFICIAL INTELLIGENCE (ARTIFICIAL INTELLIGENCE ACT) AND AMENDING CERTAIN UNION LEGISLATIVE ACTS. COM/2021/206 final. 21 April 2021, cit., Recital 16.

⁴² The General Approach of the Council also includes broader design choices in immersive environments, Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts- General Approach, cit., Recital 16.

⁴³ T. LEIGH DOWDESWELL and N. GOLTZ, *The clash of empires: regulating technological threats to civil society*, in *Information & Communications Technology Law*, 2020, p. 204.

⁴⁴ See D. LU, *Deepfakes are being used to dub adverts into different languages*, in *New Scientist* of 22 October 2019 www.newscientist.com/article/2220628-deepfakes-are-being-used-to-dub-adverts-into-different-languages/.

⁴⁵ This term is taken from Masahiro Mori’s description of the ‘uncanny valley phenomenon’; see M. MORI, *The Uncanny Valley: The Original Essay by Masahiro Mori* “The Uncanny Valley” by Masahiro Mori is an influential essay in robotics. This is the first English translation authorized by Mori, in *IEEE Spectrum* of 12 June 2002) <https://spectrum.ieee.org/the-uncanny-valley>.

subjective experience from tangible harm that is ‘likely to occur’ can not be based on an individual’s subjective reality only. Rather, it would be interesting to investigate how does a virtual avatar alter an individual’s subjective experience, based on the incorporation of fashion narratives, such as adopting a certain mode communicating with end-users.

What follows is that Article 5 (1) (a) of the AI Act proposal would tap into instances regulating an individual’s perception of subjective reality, whereby the extent and likelihood of physical or psychological injury would be almost impossible to prove.⁴⁶ Referring back to our example regarding the virtual avatar; how would an end-user when interacting with a deepfake conveying a “calming voice” know that the avatar’s actions are *not* down to his or her perceptual experience that caused a prolonged shopping addiction?

Article 5 (1) (b) of the AI Act proposal focuses on the intent to drive the end -user’s assumed characteristics, whereby the Council General approach similarly modified it to correspond to other regulatory frameworks including the Unfair Commercial Practices Directive.⁴⁷ Article 5 (1) (b) of the Council General approach stipulates that an ‘AI system that exploits any of the vulnerabilities of a specific group of persons due to their age, disability, a specific social or economic situation.’⁴⁸ Nevertheless, it arguably extends to the implications of AI systems, as well as commercial deepfake technology that produces systematic risks including concerns of non-discrimination and bias. For instance, algorithmic personalisation can cause price discrimination, causing individuals to pay more based on their demographic characteristics, as well as inadvertently and disproportionately targeting individuals based on their ethnic background.⁴⁹

⁴⁶ Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL LAYING DOWN HARMONISED RULES ON ARTIFICIAL INTELLIGENCE (ARTIFICIAL INTELLIGENCE ACT) AND AMENDING CERTAIN UNION LEGISLATIVE ACTS cit., art 5 (1) (a).

⁴⁷ Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL LAYING DOWN HARMONISED RULES ON ARTIFICIAL INTELLIGENCE (ARTIFICIAL INTELLIGENCE ACT) AND AMENDING CERTAIN UNION LEGISLATIVE ACTS, cit., art 5 (1) (b); there has been criticism that this provision significantly differs from other regulatory frameworks including the Unfair Commercial Practices Directive which may include ‘commercial practices that are also unintentionally directed towards them’ as stipulated in I. GEORGIEVA, T. TIMAN and M. HOEKSTRA, *Regulatory divergences in the draft AI act*, in *Panel for the Future of Science and Technology (STOA)* of March 2022 [https://www.europarl.europa.eu/RegData/etudes/STUD/2022/729507/EPRS_STU\(2022\)729507_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/STUD/2022/729507/EPRS_STU(2022)729507_EN.pdf), page IV; reference to Directive 2005/29/EC of the European Parliament and of the Council of 11 May 2005 concerning unfair business-to-consumer commercial practices in the internal market and amending Council Directive 84/450/EEC, Directives 97/7/EC, 98/27/EC and 2002/65/EC of the European Parliament and of the Council and Regulation (EC) No 2006/2004 of the European Parliament and of the Council (‘Unfair Commercial Practices Directive’). OJ L 149/22, 11 May 2005, art 5 (3); see also, VL. RAPOSO, *Ex machina: preliminary critical assessment of the European Draft Act on artificial intelligence*, in *International Journal of Law and Information Technology*, 2022, p. 93; compare with, Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts- General Approach, cit., recital 16.

⁴⁸ Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts- General Approach, cit., art 5 (1) (b).

⁴⁹ See research by J. LARSON, S. MATTU and J. ANGWIN, *Unintended Consequences of Geographic Targeting*, in *Technology*

However, Article 5 (1) (b) of the AI Act proposal including the recent compromise text fall short to explicitly define the dimension of AI systems to perpetuate discriminatory outcomes.⁵⁰ As argued by Iliana Georgieva, Tjerk Timan and Marissa Hoekstra, the AI Act proposal ‘portrays a significant gap in the protection of persons who might be subject to AI manipulation on the basis of other protected characteristics under EU equality law, such as ethnicity, religion, race, sex...’.⁵¹ Indeed, an important aspect is that the AI Act proposal needs to ensure consistency with fundamental rights’ and that does not preclude us to consider the right to non-discrimination with regard to the interpretation of prohibited practices.⁵² By way of illustration, one important point is that deepfake technology for revenge porn imply a ‘gendered dimension’, being predominantly directed to women.⁵³ Similarly, Jacuelyn Burkell and Chandell Gosse observe that deepfake technology, whilst ‘not problematic in and of itself’ is embedded within social and cultural attitudes that can solidify harmful outcomes, such as putting women at an increased risk of objectification and intimidation.⁵⁴

What follows is that Article 5 (1) (b) in the Council General approach text does allow us to develop a progressive interpretative framework regarding the systematic risks of commercial deepfakes, but we need more guidance in how to define unacceptable risks, considering evolving forms of algorithmic bias and unfair treatment. This is because “clothing” – from the physical appearance in the workplace informing social conventions on gender identity to the socio-cultural norms shaping variables of attractiveness and correlating with age - is a sensory experience in itself informing human interaction in

Science., 2015, <https://techscience.org/a/2015090103/>.

⁵⁰ Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL LAYING DOWN HARMONISED RULES ON ARTIFICIAL INTELLIGENCE (ARTIFICIAL INTELLIGENCE ACT) AND AMENDING CERTAIN UNION LEGISLATIVE ACTS, cit., art 5 (1) (b); Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts- General Approach, cit., art 5 (1) (b); I. GEORGIEVA, T. TIMAN and M. HOEKSTRA, *Regulatory divergences in the draft AI act*, cit., p. IV; European Digital Rights (EDRi), Access Now, Panoptikon Foundation, epicenter.works, AlgorithmWatch, European Disability Forum (EDF), Bits of Freedom, Fair Trials, PICUM, and ANEC, EU Artificial Intelligence Act for Fundamental Rights, in European Digital Rights (EDRi) of 30 November 2021 <https://edri.org/wp-content/uploads/2021/12/Political-statement-on-AI-Act.pdf>, p. 2.

⁵¹ I. GEORGIEVA, T. TIMAN and M. HOEKSTRA, *Regulatory divergences in the draft AI act*, cit., p. IV.

⁵² See also Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL LAYING DOWN HARMONISED RULES ON ARTIFICIAL INTELLIGENCE (ARTIFICIAL INTELLIGENCE ACT) AND AMENDING CERTAIN UNION LEGISLATIVE ACTS, cit., Explanatory Memorandum; indeed, there has been a wave of criticism that the AI Act adopts a product-safety approach which does not illuminate on the way providers and users need to consider the implications of AI systems on fundamental rights; see Data & Society and European Center for Not-for-Profit Law, Mandating Human Rights Impacts Assessments in the AI Act, in Data & Society of 22 November 2021 <https://datasociety.net/wp-content/uploads/2021/11/HUDIERA-5-Pager-FinalR1.pdf>; compare with European Data Protection Supervisor, *Opinion 20/2022 on the Recommendation for a Council Decision authorising the opening of negotiations on behalf of the European Union for a Council of Europe convention on artificial intelligence, human rights, democracy and the rule of law*, 13 October 2022, p.2.

⁵³ B. CHESNEY and D. CITRON, *Deep Fakes: A Looming Challenge for Privacy, Democracy and National Security*, in *California Law Review.*, 2019, p. 1773.

⁵⁴ J. BURKELL and C. GOSSE, *Nothing new here: Emphasizing the social and cultural context of deepfakes*, in *First Monday.*, 2019, p. 1.

real- time.⁵⁵ Whether deepfakes- a technology that is deemed to be a ‘ catalyst for greater gender inequality’⁵⁶- will solidify and create prejudices and stereotyping regarding “clothing” and “appearance” is a topic that needs to be examined in further research. Nevertheless, we need to pay closer attention to the way deepfakes may set out the parameters of “clothing”, such as by recommending clothing using the audiences’ demographic characteristics, to clarify the terms of Article 5 (1) (b) including the systematic risks of emerging technology for manipulation, bias, and exclusion.⁵⁷

5. Turning to the specific transparency obligations in Article 52 (3) of the AI Act proposal,⁵⁸ it is important to note that we need to adopt a nuanced approach, which allows us to distinguish between the individual communicating and the deepfake “replicating” fashion narratives. Imagine a virtual avatar doing a fashion ad campaign with your favourite celebrity, promoting the new collection in sixteen different languages, and considering various demographics.⁵⁹ The emergence of computer-generated imaginary models – from the fashion brands’ use of the virtual influencer “Lil Miquela” posting about lifestyle choices or the model “Shudu” that was designed from the looks’ of a Barbie doll⁶⁰ – are illustrative to the way algorithms’ may produce fashion narratives alongside a digitally mediated reality. This is because both the “Lil Miquela” and the “Shudu” models are digitally created but the way technology portrayed the “human aspect” was based on the models’ process of self-representation such as by interacting with *various* end-users on social medial channels.⁶¹

For instance, “Shudu’ s” designer admitted that end-users mistakenly believed that the avatar was a human influencer.⁶² The transparency obligations in Article 52 (3) of the AI Act proposal intend to

⁵⁵ ME. ROACH-HIGGINS and JB. EICHER, *Dress and Identity*, in *Clothing and Textiles Research Journal*, 1992, p. 1; Add. ADOMAITIS, R. RASKIN and D. SAIKI, *Appearance Discrimination: Lookism and the Cost to the American Woman*, in *The Seneca Falls Dialogues Journal*, 2017, p. 75.

⁵⁶ M. VAN HUIJSTEE, P. VAN BOHEEMEN, D. DAS, L. NIERLING, J. JAHNEL, M. KARABOGA, M. FATUN, L.KOOL and J. GERRITSEN, *Tackling deepfakes in European policy*, cit., p. 24.

⁵⁷ Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts- General Approach, cit., art 5 (1) (b).

⁵⁸ Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL LAYING DOWN HARMONISED RULES ON ARTIFICIAL INTELLIGENCE (ARTIFICIAL INTELLIGENCE ACT) AND AMENDING CERTAIN UNION LEGISLATIVE ACTS, cit., art 52 (3).

⁵⁹ See A.MAHDAMI, *Why Bella Hadid and Lil Miquela’s kiss is a terrifying glimpse of the future*, in *The Guardian* of 21 May 2019 <https://www.theguardian.com/commentisfree/2019/may/21/bella-hadid-lil-miquela-terrifying-glimpse-calvin-klein>; K. CHITRAKORN, *How deepfakes could change fashion advertising*, cit.

⁶⁰ A. DU PARCQ and B. LONDON, *The man behind Shudu Gram & the world’s first ‘digital supermodels’ reveals the secrets behind his stratospheric success*, in *Glamour* of 13 September 2018 www.glamourmagazine.co.uk/article/shudu-gram-virtual-supermodels.

⁶¹ John Griffiths provides an illuminating outlook on how virtual, deepfake influencers interact with an audience, see J. GRIFFITHS, *Deepfake Influencers: The Future Of Fashion Advertising?*, in *Foundation* of 13 May 2022 <https://foundationagency.co.uk/blog/deepfake-influencers-the-future-of-fashion-advertising/>.

⁶² A. NEWBOLD, *Newbold, ‘The Numerous Questions Around The Rise Of CGI Models And Influencers*, in *Vogue* of 18 August 2018

precisely avoid this dilemma by specifying user obligations that allow end-users to distinguish the “fake” from the “real”.⁶³ Article 52 (3) of the AI Act proposal lists specific transparency obligations applicable to commercial deepfakes and excluding users who are acting in a personal capacity from the provision’s scope.⁶⁴ The labelling requirements stipulate that ‘[u]sers of an AI system that generates or manipulates image, audio or video content that appreciably resembles existing persons, objects, places or other entities or events and would falsely appear to a person to be authentic or truthful (‘deep fake’), shall disclose that the content has been artificially generated or manipulated’.⁶⁵

The provision’s aim is seemingly clear, relating to the protection of the human subjects versus the artificial content. However, the AI Act proposal’s distinction between the provider and user concerning these transparency obligations, raises interesting questions of enforcement.⁶⁶ Crucially, the provision’s wording leaves the open question on who decides on the risk, considering the increasing accessibility and proliferation of software enabling the creation of synthetic content.⁶⁷ Article 52 (3) of the AI Act proposal only provides labelling requirements for users who ‘shall disclose that the [synthetic] content has been artificially generated or manipulated’, leaving out the extent providers can contribute to these transparency requirements.⁶⁸ Indeed, the Strengthened Code of Practice on Disinformation puts the spotlight on platform operators developing a ‘cross-service understanding of manipulative behaviours’ of fake news or accounts including ‘malicious deepfakes’,⁶⁹ whereas the Digital Services Act provides a

www.vogue.co.uk/article/cgi-virtual-reality-model-debate.

⁶³ There are exemptions to this with regard to the detection and investigation of criminal offences or the use for artistic purposes and considering the freedom of expression, Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL LAYING DOWN HARMONISED RULES ON ARTIFICIAL INTELLIGENCE (ARTIFICIAL INTELLIGENCE ACT) AND AMENDING CERTAIN UNION LEGISLATIVE ACTS, cit., art 52 (3), art 3 (4); see also, I. VAROSANEC, *On the path to the future: mapping the notion of transparency in the EU regulatory framework for AI*, in *International review of law, computers & technology*, (2022), p. 104.

⁶⁴ Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL LAYING DOWN HARMONISED RULES ON ARTIFICIAL INTELLIGENCE (ARTIFICIAL INTELLIGENCE ACT) AND AMENDING CERTAIN UNION LEGISLATIVE ACTS, cit., art 3 (4); M.VEALE and F. BORGESIU, *Demystifying the Draft EU Artificial Intelligence Act*, in *Computer Law Review International*, (2021), p. 108.

⁶⁵ Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL LAYING DOWN HARMONISED RULES ON ARTIFICIAL INTELLIGENCE (ARTIFICIAL INTELLIGENCE ACT) AND AMENDING CERTAIN UNION LEGISLATIVE ACTS, cit., art 52 (3); Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts- General Approach, cit., art 52 (3).

⁶⁶ M.VEALE and F. BORGESIU, *Demystifying the Draft EU Artificial Intelligence Act*, cit., p. 108.

⁶⁷ See also commentary by Angelica Fernandez who speaks about problems of enforcement; A. FERNANDEZ, *Regulating Deep Fakes in the Proposed AI Act*, in *Media Laws: Law and Policy of the Media in a Comparative Perspective* of 23 March 2022 www.medialaws.eu/regulating-deep-fakes-in-the-proposed-ai-act/.

⁶⁸ Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL LAYING DOWN HARMONISED RULES ON ARTIFICIAL INTELLIGENCE (ARTIFICIAL INTELLIGENCE ACT) AND AMENDING CERTAIN UNION LEGISLATIVE ACTS, cit., art 52 (3); M. VAN HUIJSTEE, P. VAN BOHEEMEN, D. DAS, L. NIERLING, J. JAHNEL, M. KARABOGA, M. FATUN, L.KOOL and J. GERRITSEN, *Tackling deepfakes in European policy*, cit., p. 38.

⁶⁹ EU Commission, *The Strengthened Code of Practice on Disinformation 2022*, 16 June 2022 <https://digital-strategy.ec.europa.eu/en/library/2022-strengthened-code-practice-disinformation>, p. 15- 16.

much needed incentives for large online platforms to exercise mitigating measures, detect and remove ‘illegal content’.⁷⁰ Whether this holistic approach will provide a comprehensive remedy to the systematic risks’ of algorithms including synthetic content technology regarding disinformation, is an aspect that will turn on the way EU regulators and large platforms will define the contours of “illegal content”.

The diversity of “fashion” allows us to think about the contours of meaningful transparency regarding fashion narratives. In this respect, mere disclosure will not assist the consumer, as identified by Michael Veale and Frederick Borgesius.⁷¹ Regulating commercial content – such as the marketing of a virtual influencer- requires a balanced approach considering the intended and unintended uses of technology, and how algorithms shape fictional authenticity.

Jonathan Michael Square’ s thought-provoking perspective illuminates the way virtual content may portray “fashion narratives”. In particular, he argues that the virtual avatar “Lil Miquela”:

‘... is the product of the machinations of a design team that created a racially ambiguous woman of color whose appearance exploits wider society’ s preference for lighter skin. She also benefits from a non-White identity in this current era, in which being a person of color can afford a degree of authenticity and cachet among some internet circles. At the same time, she conforms to phenotypic preferences of mainstream society (i.e., young, thin, light-skinned, with bone-straight hair, “fine features,” normative speech patterns, and fashionable dress).’⁷²

Synthetic content – from the model “Lil Miquela” to the deepfake utilising algorithms for personalised advertising- benefits from a ‘degree of authenticity’ irrespective of the “fictional content”. This is based on the manner deepfakes may incorporate “fashion narratives”, such as entailing a virtual avatar encompassing the facial or body features of a “young women” whilst wearing and speaking about “clothing” and “style”. The way a deepfake may project these narratives is authentic to the audience whilst being fictional in its appearance. This may create important tensions for regulation, as the mere

⁷⁰ 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), cit., Recital 12, Recital 87, Recital 106, art 3 (h), art 34, art 35; The EU Commission has also developed specific transparency rules with regard to political advertising, see ‘European Democracy: Commission sets out new laws on political advertising, electoral rights and party funding’ (European Commission – Press release, 25 November 2021) < https://ec.europa.eu/commission/presscorner/detail/en/ip_21_6118 > accessed 13 January 2023; L. EDWARDS, *How to regulate misinformation*, in *Royal Society* of 25 January 2022 <https://royalsociety.org/blog/2022/01/how-to-regulate-misinformation/>.

⁷¹ M.VEALE and F.BORGESIOUS, *Demystifying the Draft EU Artificial Intelligence Act*, cit., p.108; see also, N. HELBERGER and N. DIAKOPOULOS, *The European AI Act and How it Matters for Research into AI in Media and Journalism*, in *Digital Journalism.*, 2022, p. 5.

⁷² J.M. SQUARE, *From Lil Miquela to Shudu Digital Slavery and the Twenty-First-Century Racialized Performance of Identity Politics*, in A. Kollnitz and M. Pecorari (eds), *Fashion, Performance & Performativity: The Complex Spaces of Fashion*, Bloomsbury Visual Arts, 2022, p. 135-136.

disclosure of “fake content” may not account for these nuances, and how fashion narratives shape the expression of perception and management of appearance.

6. Looking forward, it is important to build on the way deepfakes in fashion may perform actions and disruptions on an individual’s self-perception and appearance management for AI governance. This essay is not about providing a conclusive answer to the emergent threats of commercial deepfakes. Quite to the contrary, my analysis suggests that deepfake technology does not exist independently from an individual’s interactive experience, nor does human perception act only upon “fake content”. What this means is that we indeed need a nuanced approach to deepfakes, considering the way GANs may shape personalisation and advertising in fashion, and to exercise careful deliberation about the deep fakes’ metaphorical significance to shape narratives of the body, clothing, and self-perception. We need practical tools- from an interpretative framework on prohibited practices and labelling requirements in the AI Act proposal⁷³ – to enable to carve out end-user vulnerability and dialectic experience regarding the fictional representations of the self. Because deepfakes in fashion proclaim contradictory “truths” and are immersed in the individual’s performative role in fashion, I argue that we need to regulate commercial deepfakes based on the way algorithms exhaust fashion narratives, considering the individual’s perception and management of fashion.

I discussed that the AI Act proposal portrays the regulation of subjective reality, whereby technology can alter an ‘individual’s conscious experience’ to the degree that a person is likely to suffer physical or psychological harm.⁷⁴ However, we must note that deepfakes not only appeal to individual unconscious beliefs, such as emotions, but it is the way technology mirrors self-representation we should be concerned about. For instance, the way deepfakes in fashion can portray narratives about gender that are embedded with an individual’s appearance could give rise to a new form of “subjective neutrality” which in turn, can undermine an individual’s autonomy, as well as create new forms of bias. Exposing the role of fashion narratives by considering the interactive implications of deepfakes is one way for us to

⁷³ Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL LAYING DOWN HARMONISED RULES ON ARTIFICIAL INTELLIGENCE (ARTIFICIAL INTELLIGENCE ACT) AND AMENDING CERTAIN UNION LEGISLATIVE ACTS, cit., art 52 (3); Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts- General Approach, cit., art 5 (1) (a)- (b).

⁷⁴ Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts- General Approach, cit., art 5 (1) (a), art 5 (1) (b).

complement the specific transparency obligations in Article 52 (3),⁷⁵ as well as to preserve an individual's dialectic experience on the meanings attached to body, clothing, and perception.

⁷⁵ Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL LAYING DOWN HARMONISED RULES ON ARTIFICIAL INTELLIGENCE (ARTIFICIAL INTELLIGENCE ACT) AND AMENDING CERTAIN UNION LEGISLATIVE ACTS, cit., art 52 (3).