

Criminalising Pornographic Deep Fakes: A Gender-Specific Inspection of Image-Based Sexual Abuse

Rüya Tuna Toparlak, MLaw

Academic Assistant, Faculty of Law University of Lucerne

E-mail: rueya.toparlak@unilu.ch

SciencesPo Law School
The 10th Graduate Conference
Law & Technology
16-17 June 2022, Paris

Criminalising Pornographic Deep Fakes: A Gender-Specific Inspection of Image-Based Sexual Abuse

Rüya Tuna Toparlak, MLaw

Abstract

Pornographic deep fakes account for the majority of deep fakes online. They overwhelmingly target women. This otherwise promising technology is getting easier to use and accessible to everyone. With the democratisation of this technology, it is becoming possible for everyone to create non-consensual pornographic deep fakes of regular people. In this sense pornographic deep fakes are exacerbating the broader problem of image-based sexual abuse. Although there are applicable laws to pornographic deep fakes, they have shortcomings in addressing the issue effectively. This paper inspects pornographic deep fakes under the applicable personality rights of Swiss Criminal Law and illustrates the need for their criminalisation under a broader context of image-based sexual abuse. The case is made to provide symbolic justice to victims, deter potential perpetrators, and mobilise a proper societal response.

Contents

- 1. Introduction 1
- 2. Setting the Scene: A closer look at pornographic deep fakes..... 2
 - 2.1. The terminology 3
 - 2.2. Harms and Effects of Image-Based Sexual Abuse on Victims 4
 - 2.2.1. Psychological, emotional, and financial damage 5
 - 2.2.2. Gendered element of victimisation in deep fake pornography and image-based sexual abuse .. 6
 - 2.3. Societal Attitudes Towards Non-Consensual Pornographic Deep Fakes 7
 - 2.4. Interconnection of Image-Based Sexual Abuse with Digital Violence and Other Types of Abuse .. 8
- 3. Regulatory Approach 10
 - 3.1. Applicable Law and Legal Challenges 11
 - 3.1.1. Applicable personality rights under Swiss civil law 11
 - 3.1.2 The need for a criminal response 20
 - 3.2. Regulatory Efforts at the European Union Level..... 23
- 4. Conclusion 25

1. Introduction

*So will I turn her virtue into pitch
And out of her own goodness make the net
That shall enmesh them all.*
- *W. Shakespeare*¹

When actress Kristen Bell found out there were pornographic deep fake videos of her circulating online, her first reaction was shock.² Her face was superimposed onto a pornographic actress's body, making it appear as though she was the one acting in the pornographic video. "This is my face, it belongs to me. I did not consent to this. That's why it's not okay even if it's labelled as fake." she said.³ Her story is one of many. Numerous celebrity women like Emma Watson, Gal Gadot, and Taylor Swift are being exploited by non-consensual pornographic deep fakes.⁴ Celebrity women were the first victims of this technology, due to the large amounts of their data like images and video being available to train the deep fake software.⁵ In recent years, this technology has become better, faster, and easily accessible to everyone.⁶ With the democratisation of deep fakes, it is now possible for everyone to use it with ease and create non-consensual pornographic deep fakes of regular people.⁷

Deep fakes are not completely unregulated, albeit non-consensual pornographic deep fakes bring certain specific legal challenges that should be identified. For this purpose, this paper will take a gender-based approach and examine non-consensual pornographic deep fakes in a broad context of image-based sexual abuse.⁸ A closer look should be provided at the harms and effects of

¹ *Iago in Othello Act II Scene iii.*

² Cleo Abram. "The Most Urgent Threat of Deepfakes Isn't Politics. It's Porn." Vox. Vox, June 8, 2020. <https://www.vox.com/2020/6/8/21284005/urgent-threat-deepfakes-politics-porn-kristen-bell>.

³ Ibid.

⁴ Danielle Keats Citron, "Sexual Privacy." *Yale Law Journal* 128 (2019): 1921, 1922.

⁵ Henry Ajder et al., "The State of Deepfakes: Landscape, Threats, and Impact." Deeptrace labs, 2019. <https://sensity.ai/reports/>

⁶ Mariëtte van Huijstee et al., "Tackling Deepfakes in European Policy." Panel for the Future of Science and Technology | European Parliament, 2021,

[https://www.europarl.europa.eu/stoa/en/document/EPRS_STU\(2021\)690039](https://www.europarl.europa.eu/stoa/en/document/EPRS_STU(2021)690039); Robert Chesney and Danielle Keats Citron, "Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security." *California Law Review* 107, no. 6 (2019): 1762.

⁷ Tobias Lantwin, "Deep Fakes - Düstere Zeiten für den Persönlichkeitsschutz?" *MMR* 2019, no. 9 (2019): 575.

⁸ See chapter 2.1. of this paper.

pornographic deep fakes on women and societal attitudes towards deep fakes should be inspected.⁹ The paper will identify legal challenges and explore the applicable personality rights under the Swiss Civil Code.¹⁰ Followingly, the need for the criminalisation of pornographic deep fakes under a clause that would address all abusive behaviour involving non-consensual taking and sharing of a person's sexual images will be illustrated. Thereby, the paper advocates for effective prosecution and proper societal evaluation and condemnation of this behaviour.

2. Setting the Scene: A closer look at pornographic deep fakes

There are many analogue and digital techniques that can be used to create videos and other content that misrepresent people and events. Deep fakes have been recently popularised due to the hyper realistic results they create. They are a sophisticated trend in digital impersonations.¹¹ The technology uses artificial intelligence and machine learning to swap faces and bodies in pictures and videos.¹² It can also be used to manipulate existing video and audio content. It creates realistic impersonations, making individuals appear in places they've never been in, doing things they never were doing, saying things they never said.¹³ This promising new technology can be a great tool for self-expression and has many beneficial uses.¹⁴ Yet non-consensual pornographic deep fakes of women constitute the vast majority of the deep fakes circulating online.¹⁵

An analysis conducted in 2019 by the cybersecurity company Deeptrace found that 96% of all deep fakes online are pornographic and disproportionately female.¹⁶ The top four websites dedicated to deep fake pornography received more than 134 Million views on videos targeting hundreds of women worldwide.¹⁷ Deep fake phenomenon emerged around 2018 when a Reddit user released "FakeApp" that enabled users to easily create deep fakes for free.¹⁸ Democratisation of this technology, created a massive surge in the non-consensual pornographic deep fakes.¹⁹ A now banned subreddit was dedicated to creating deep fakes by inserting faces of female celebrities into

⁹ See chapters 2.2. and 2.3. of this paper respectively.

¹⁰ See chapter 3.1. of this paper.

¹¹ Chesney and Citron, "Deep Fakes," 1758

¹² Ibid.

¹³ Ibid.

¹⁴ They can be used in positive ways like satirical, educational, parody purposes. Ibid. 1769.

¹⁵ Ajder et al., "The State of Deep Fakes," 1, 2.

¹⁶ Ibid.

¹⁷ Ibid.

¹⁸ Rebecca Delfino, "Pornographic Deepfakes — Revenge Porn's Next Tragic Act — the Case for Federal Criminalization." *Fordham Law Review* 88, no. 3 (2019): 893.

¹⁹ Chesney and Citron, "Deep Fakes," 1762, 1764; Lantwin, "Deep Fakes," 575; Edvinas Meskys et al., "Regulating Deep Fakes: Legal and Ethical Considerations." *Journal of Intellectual Property Law & Practice* 15, no. 1 (2020): 25.

pornographic videos.²⁰ Within a few months of its release, FakeApp had been downloaded more than 100.000 times, enabling regular people to create realistic deep fakes.²¹

The FakeApp project is currently dead, however, the technology is thriving. Similar software has since then surfaced and amassed big user bases.²² One such software is called “Deepfacelab” and according to its developers, more than 95% of deep fakes on the internet are created using this software.²³ Another popular software is called “Reface”, which was launched in January 2020 and has been downloaded 70 Million times in one year and given a 5.5 Million USD seed round.²⁴ These are all easy-to-use software and require no technical background. For a realistic result, one needs to identify the video to be used and feed as few as 500 images of the victim to be superimposed.²⁵ A highly realistic result may take up to 40.000 source images.²⁶ While highly realistic deep fakes require a high amount of source images and computing power, there are alternatives for “undressing women” that operate on a much easier level. One such software called “deepnude” was released in 2019. It allowed users to “strip” clothed women in 30 seconds and generate naked body parts.²⁷ Although creators took the official website down, the software continues to be independently distributed.²⁸

2.1. The terminology

Scholars have adopted different terms to comprehensively refer to the diverse range of abusive behaviour involving sexual images, like revenge porn, non-consensual pornography, digital sexual violence, and image-based sexual abuse.²⁹ Most colloquially used term “revenge porn” emerged to

²⁰ Chesney and Citron, “Deep Fakes,” 1772.

²¹ Adam Dodge and Erica Johnstone. “Using Fake Video Technology to Perpetuate Intimate Partner Abuse.” California Partnership to End Domestic Violence. <https://www.cpedv.org/publication/using-fake-video-technology-perpetuate-intimate-partner-abuse>.

²² “8 Best Deepfake Apps and Tools in 2022.” RankRed, January 10, 2022. <https://www.rankred.com/best-deepfake-apps-tools/>.

²³ “Iperov/Deepfacelab: Deepfacelab Is the Leading Software for Creating Deepfakes.” GitHub. <https://github.com/iperov/DeepFaceLab>. Accessed May 29, 2022.

²⁴ Natasha Lomas, “Reface Grabs \$5.5m Seed Led by a16z to Stoke Its Viral Face-Swap Video App.” TechCrunch, December 8, 2020. <https://techcrunch.com/2020/12/08/reface-grabs-5-5m-seed-led-by-a16z-to-stoke-its-viral-face-swap-video-app/>.

²⁵ Dodge and Johnston, “Fake Video Partner Abuse,” 5; For a comparative analysis of the number of source images, see Andrea Hauser, “This Is How You Create the Perfect Deepfake.” Scip, November 22, 2018. <https://www.scip.ch/en/?labs.20181122>.

²⁶ “How Many Iterations Does a Realistic Deepfake Image Need?” Deepfake Now, April 21, 2020. <https://deepfakenow.com/how-many-iterations-realistic-deepfake-image/>. Accessed May 29, 2022.

²⁷ Samantha Cole, “Creator of DeepNude, App That Undresses Photos of Women, Takes It Offline.” VICE, June 27, 2019. <https://www.vice.com/en/article/qv7agw/deepnude-app-that-undresses-photos-of-women-takes-it-offline>.

²⁸ Ajder et al., “The State of Deep Fakes,” 8.

²⁹ For a comparative analysis of the terminology, see Sophie Maddocks, “From Non-Consensual Pornography to Image-Based Sexual Abuse: Charting the Course of a Problem with Many Names.” *Australian Feminist Studies* 33, no. 97 (2018): 345–61.

describe the non-consensual sharing of nude/sexually suggestive images of a person, mostly an ex-partner.³⁰ The term is highly criticised by feminist and legal scholars alike.³¹ The critique convenes around the term “revenge porn” focusing on the revenge intent behind and the content of the images, downplaying the act as something similar to consensual pornography, and the victim-blaming connotations.³² Research shows that perpetrators share non-consensual sexual images with differing motivations, not just for revenge purposes.³³

Adopting a terminology that highlights the act of abuse and effects rather than the pornographic content of the medium also provides proper societal designation.³⁴ Thus, this paper will use the most comprehensive term image-based sexual abuse while referring to all abusive behaviour that would involve the non-consensual taking and sharing of sexual images.³⁵ Non-consensual pornographic deep fakes will be observed within image-based sexual abuse, as a new technology that is exacerbating this problem.³⁶ For the purposes of this paper, deep fake pornography will solely refer to non-consensual depictions, even when it is not explicitly stated. Consensual deep fake pornography shall be left out of the scope of this paper. The term digital violence shall be used to encompass many forms of abuse happening online, including image-based sexual abuse, cyberstalking, harassment, doxxing,³⁷ etc.

2.2. Harms and Effects of Image-Based Sexual Abuse on Victims

When victims discover that they have been used in pornographic deep fakes, the damage may be profound.³⁸ Victims may experience a range of psychological and emotional harm as well as financial damage. The following chapters will illustrate these harms and effects and the gendered element of their severity.

³⁰ Ibid., 347.

³¹ Ibid., 348; Nicola Henry et al., *Image-Based Sexual Abuse: A Study on the Causes and Consequences of Non-Consensual Nude or Sexual Imagery* (London: Routledge, Taylor & Francis Group, 2021), 3-6.; Clare McGlynn, Erika Rackley, and Ruth Houghton. “Beyond ‘Revenge Porn’: The Continuum of Image-Based Sexual Abuse.” *Feminist Legal Studies* 25, no. 1 (2017): 30.

³² Henry et al., *Image-based Sexual Abuse*, 3-6.

³³ Some of these motivations are power, control, wanting to have fun, prove “masculinity” Henry et al., *Image-based Sexual Abuse*, 72-76.

³⁴ Maddocks, “Problem with Many Names,” 350.

³⁵ The term image-based sexual abuse has been developed by Clare McGlynn and Erika Rackley. It is adopted in the literature to address all abusive behaviour that would involve the non-consensual taking and sharing of sexual images, e.g., upskirting, downblousing. McGlynn, Rackley, and Houghton, “Beyond Revenge Porn,” 26.

³⁶ Deep fakes are transforming yesterday’s notion of revenge porn into unprecedented categories. Karolina Mania, “The Legal Implications and Remedies Concerning Revenge Porn and Fake Porn: A Common Law Perspective.” *Sexuality & Culture* 24, no. 6 (2020): 2089.

³⁷ Doxxing is publicly revealing identifying information about someone like their address and phone number, typically with malicious intent.

³⁸ Citron, “Sexual Privacy” 1926.

2.2.1. Psychological, emotional, and financial damage

A qualitative study conducted in 2017 with 18 image-based sexual abuse survivors revealed serious mental health effects for victims, including PTSD, suicidality, anxiety and depression.³⁹ The study highlighted the similarity of the effects on the victims with other types of sexual assault.⁴⁰ Other key themes in victims of image-based sexual abuse were identified as constrained liberty, isolation, and withdrawal from online life.⁴¹ This was the case when Rana Ayyub, an investigative journalist from India, was victimised by pornographic deep fakes.⁴² Her face was superimposed into a pornographic video and circulated widely. The stress sent her to the hospital with heart palpitations, high blood pressure, and anxiety. She withdrew from online life and started self-censoring.⁴³

The harmful culture of victim blaming that is pertinent in other forms of image-based sexual abuse is also present in non-consensual deep fake pornography.⁴⁴ In an interview, one victim noted the victim blaming and harassment she experienced after her digitally altered images were made public. She stated that the harassment was also racially motivated.⁴⁵ Another intersectionality happens with homophobia and transphobia.⁴⁶ Deep fakes may be used to share information about someone's sexuality, identity and perpetrate discrimination.⁴⁷ One such example was from Malaysia, where homosexuality is still illegal. A political aid was arrested following a video depicting him having homosexual intercourse with the Malaysian minister of economic affairs.⁴⁸ The Malaysian prime minister alleged that the video was a deep fake, but the political aid lost his job, got arrested and both men suffered damages, nonetheless.

³⁹ Samantha Bates, "Revenge Porn and Mental Health: A Qualitative Analysis of the Mental Health effects of Revenge Porn on Female Survivors." *Feminist Criminology* 12, no. 1 (2017): 30.

⁴⁰ *Ibid.*, 39.

⁴¹ Henry et al., *Image-based Sexual Abuse*, 11.

⁴² Rana Ayyub, "I Was the Victim of a Deepfake Porn Plot Intended to Silence Me." HuffPost UK. HuffPost UK, November 21, 2018. https://www.huffingtonpost.co.uk/entry/deepfake-porn_uk_5bf2c126e4b0f32bd58ba316.

⁴³ *Ibid.*

⁴⁴ Citron, "Sexual Privacy" 1876; Danielle Keats Citron and Mary Anne Franks, "Criminalizing Revenge Porn," *Wake Forest Law Review* 49, no. 2 (2014): 348, 367.; Olga Jurasz and Kim Barker, "Sexual Violence in the Digital Age: A Criminal Law Conundrum?" *German Law Journal* 22, no. 5 (2021): 795; Carolyn A. Uhl et al., "An Examination of Nonconsensual Pornography Websites." *Feminism & Psychology* 28, no. 1 (2018): 61, 62.

⁴⁵ The victim was harassed online with comments that said "she deserved this abuse because she was a brown woman" Henry et al., *Image-based Sexual Abuse*, 56.

⁴⁶ Citron, "Sexual Privacy" 1898, 1920, 1926; Henry et al., *Image-based Sexual Abuse*, 34, 35, 55.

⁴⁷ Henry et al., *Image-based Sexual Abuse*, 30.

⁴⁸ Jarny Blakkarly, "A Gay Sex Tape Is Threatening to End the Political Careers of Two Men in Malaysia." SBS News, June 17, 2019. <https://www.sbs.com.au/news/article/a-gay-sex-tape-is-threatening-to-end-the-political-careers-of-two-men-in-malaysia/ilgqdaqo5>; Tom Simonite, "The Web Is Drowning in Deepfakes and Almost All of Them Are Porn." WIRED UK. WIRED UK, October 13, 2019. <https://www.wired.co.uk/article/deepfakes-porn>

The harms of image-based sexual abuse certainly have an economic dimension. The stigma nude images carry, especially for women and girls, may cause consequences in the job market.⁴⁹ Job loss or problems of securing new employment were certainly observed as harms caused by other forms of image-based sexual abuse.⁵⁰ It is fair to predict that deep fakes will enhance this problem. Another aspect of financial loss might occur when pornographic deep fakes are used for extortion.⁵¹ A “revenge porn” website operator was convicted of six counts of extortion due to him charging 350\$ to victims to remove their pictures and information from his website.⁵² Pornographic deep fakes present a new possibility for extortion happening alongside image-based sexual abuse.⁵³ Through this illustration of psychological, emotional, and financial damages, this subtitle concludes the varying harms caused by pornographic deep fakes and that deep fakes are exacerbating the image-based sexual abuse phenomenon.⁵⁴

2.2.2. Gendered element of victimisation in deep fake pornography and image-based sexual abuse

As with other forms of image-based sexual abuse, young adult women are more likely to be victimised by non-consensual deep fake pornography.⁵⁵ Majority of pornographic deep fakes online are female.⁵⁶ This parallels broader statistics on image-based sexual abuse.⁵⁷ Statistics on “revenge porn” shared by the Cyber Civil Rights Initiative in 2014 highlight that 90% of victims are female and 68% of victims were between the ages of 18-30.⁵⁸ A recent content analysis scanned seven so-called revenge porn websites and found that 92% of victims were women.⁵⁹ Six of the seven websites allowed photographs of both men and women, while one was specifically for sharing non-consensual nude images of women.⁶⁰

Women are not just more likely to be victimised by image-based sexual abuse, but they also experience the effects in a more profound way.⁶¹ A study conducted with a similar number of

⁴⁹ Chesney and Citron, “Deep Fakes,” 1773; Huijstee et al., “Tackling Deepfakes in European Policy” 30.

⁵⁰ Citron and Franks, “Criminalizing Revenge Porn” 352.; Citron, “Sexual Privacy” 1891, 1921; According to a Microsoft study from 2010, Unsuitable photos, videos, and information about the applicant that comes up in the search engines constitute one of the leading reasons for rejection.

⁵¹ Chesney and Citron, “Deep Fakes,” 1772.

⁵² “Revenge Porn Website Operator Jailed.” The Guardian, April 4, 2015. <https://www.theguardian.com/us-news/2015/apr/04/revenge-porn-website-operator-jailed>

⁵³ Chesney and Citron, “Deep Fakes,” 1772.

⁵⁴ Jenny-Kerstin Bauer and Ans Hartmann, “Formen digitaler geschlechtsspezifischer Gewalt” in *Geschlechtsspezifische Gewalt in Zeiten der Digitalisierung*, ed. Nivedita Prasad (Germany: Transcript, 2019), 85.

⁵⁵ Chesney and Citron, “Deep Fakes,” 1773.

⁵⁶ Ajder et al., “The State of Deep Fakes,” 1.

⁵⁷ Citron, “Sexual Privacy” 1908.

⁵⁸ “End Revenge Porn.” Cyber Civil Rights Initiative. <https://cybercivilrights.org/>

⁵⁹ Uhl et al., “Nonconsensual Pornography Websites” 50.

⁶⁰ Ibid., 56-58.

⁶¹ Henry et al., *Image-based Sexual Abuse*, 48.

male/female victims found that female victims were much more likely to report feelings of unsafety, harm to their health, relational impacts and reputational concerns.⁶² Nude images carry a stigma, especially for women and girls and this stigma heightens the negative effects victims experience.⁶³ Non-consensual deep fake pornography provides new ways for perpetrators of image-based sexual abuse to victimise women. They no longer need “actual” nude images to threaten, humiliate, poke fun at, and control their victims.⁶⁴

2.3. Societal Attitudes Towards Non-Consensual Pornographic Deep Fakes

Above mentioned statistics illustrate the profound harms and effects of image-based sexual abuse and how pornographic deep fakes are likely to exacerbate the problem.⁶⁵ Yet, the majority of people are indifferent to this phenomenon.⁶⁶ The harms are downplayed, and digital sexual violence is regarded as less serious than offline crimes of violence.⁶⁷ The cases of sexual violence that get reported are already low⁶⁸ and reports of online sexual violence are often seen as more of a stretch of limited police sources.⁶⁹ In the cases of specifically non-consensual pornographic deep fakes, the reason for this “downplaying” has a couple of different dimensions.

Many perpetrators do not recognise the harms their creations are doing and regard their motivations as “more of a joke”.⁷⁰ Especially in such depictions where the video is clearly fake, perpetrators and viewers might falsely assume there is no real harm. Although pornographic deep fakes do not actually depict featured victim’s actual private parts, they still exercise dominion over people’s sexuality and exhibit it to others without consent.⁷¹ Regardless of their level of realism, all pornographic deep fakes reduce victims to sex objects and violate the depicted person’s rights.⁷² Victims feel the same level of humiliation, alienation, and fright. This is illustrated in journalist Rana Ayyub’s experience of finding out pornographic deep fake videos of herself.⁷³ She said “You

⁶² Ibid.

⁶³ Chesney and Citron, “Deep Fakes,” 1773.

⁶⁴ Perpetrators have differing motivations for their acts of image-based sexual violence. They might find it funny and sexy, do it to impress friends, to control and embarrass the victim. Henry et al., *Image-based Sexual Abuse*, 72, 76.

⁶⁵ see fn. 39, 40, 49, 54.

⁶⁶ Meskys et al., “Regulating Deep Fakes,” 27.

⁶⁷ Citron and Franks, “Criminalizing Revenge Porn” 350.

⁶⁸ In Switzerland, only 8% of female sexual violence victims reported the crime to the police. Christina Bornatici, “Junge Frauen in der Schweiz: Wo stehen sie heute?” Eidgenössische Kommission für Frauenfragen EKF, March 22, 2022. <https://www.ekf.admin.ch/ekf/de/home/dokumentation/fachzeitschrift--frauenfragen-/frauenfragen-2022.html>

⁶⁹ Jurasz and Barker, “Sexual Violence in the Digital Age,” 794.; Citron and Franks, “Criminalizing Revenge Porn” 366

⁷⁰ Henry et al., *Image-based Sexual Abuse*, 72, 77.

⁷¹ Citron, “Sexual Privacy” 1921.

⁷² Ibid., 1914.

⁷³ Ayyub, “Victim of a Deepfake Porn”

can call yourself a journalist, you can call yourself a feminist but, in that moment, I just couldn't see through the humiliation.”⁷⁴

In a gender specific inspection, scholars suggested that ritualised acts of sexual humiliation against women are used for men's collective amusement in the service of proving one's position in male social hierarchy.⁷⁵ It is suggested that the emergence of this technology with depicting celebrities played a part in the initial indifference attitude.⁷⁶ The rise of deep fake celebrity pornography was categorised as disenfranchised men, whose efforts to exert power over women have failed, seeking out means through which to consolidate “male” identity and status.⁷⁷ These were/are the women that most men will never be able to touch or possess.⁷⁸

Gendered messages minimising the impact of non-consensual pornographic deep fakes are powerful and often internalised by victims.⁷⁹ All these reasons mentioned here, contribute to most of the cases going unreported.⁸⁰ Luckily in recent years, societal attitude towards non-consensual pornographic deep fakes has started to shift and some surveys illustrated that majority of respondents condemn these videos, even though their popularity on pornographic websites keeps increasing.⁸¹ Criminal law reform, in particular, is certainly needed to help in subverting the normalisation of non-consensual pornographic deep fakes.⁸²

2.4. Interconnection of Image-Based Sexual Abuse with Digital Violence and Other Types of Abuse

It is not uncommon for image-based sexual abuse to occur alongside other types of digital violence. One such behaviour is “doxxing”, which heightens the effects of image-based sexual abuse.⁸³ A study conducted with 1244 victims found that more than 50% of survivors' full names and links to social media profiles accompanied the naked photos and that 20% of survivors' email addresses and phone numbers were posted alongside their photos.⁸⁴ A more recent study with 6109

⁷⁴ Ibid.

⁷⁵ Henry et al., *Image-based Sexual Abuse*, 96.

⁷⁶ Meskys et al., “Regulating Deep Fakes,” 27.

⁷⁷ Henry et al., *Image-based Sexual Abuse*, 98.

⁷⁸ Ibid.

⁷⁹ Citron and Franks, “Criminalizing Revenge Porn” 353.

⁸⁰ Prasad, “Digitalisierung geschlechtsspezifischer Gewalt” in *Geschlechtsspezifische Gewalt*, ed. Prasad 85.

⁸¹ Matthew B Kugler and Carly Pace. “Deepfake Privacy: Attitudes and Regulation.” *Northwestern University Law Review* 116, no. 3 (2021): 641; Matt Burgess, “Deepfake Porn Is Now Mainstream. and Major Sites Are Cashing In.” WIRED UK, August 27, 2020. <https://www.wired.co.uk/article/deepfake-porn-websites-videos-law>

⁸² Jurasz and Barker, “Sexual Violence in the Digital Age,” 785; Delfino, “Pornographic Deepfakes” 902.

⁸³ Jurasz and Barker, “Sexual Violence in the Digital Age,” 789.

⁸⁴ Citron and Franks, “Criminalizing Revenge Porn” 350, 351.

respondents found that for 2/3 of victims, their images were shared alongside identifying personal information.⁸⁵ This makes the “online” crimes committed against victims feel much more real and they could physically be harmed at any moment.⁸⁶ Digital violence perpetuates offline harassment and stalking.⁸⁷ It reduces the victims into a passive and silent state where they cannot defend themselves.⁸⁸ Threats of sexual violence “literally, albeit not physically, penetrates the body” and non-consensual pornographic deep fakes can transform rape threats into a terrifying reality.⁸⁹

Forms of sexual violence are interconnected and non-consensual deep fake pornography is no exception. Digital sexual violence does not function separately from analogue violence, it is supplementary to violent relationship dynamics.⁹⁰ In one qualitative study with 75 participants, around one-third of the victims experienced image-based sexual abuse alongside or in the context of other forms of domestic violence.⁹¹ Similarly, a quantitative study with 957 respondents found that 59.4% of victims had experienced poly-victimisation from which, 8.3% experienced it alongside physical abuse and 28.2% ongoing harassment.⁹² As follows, it is easy to see image-based sexual abuse as an extension of patriarchal power. These examples show the importance of contextualising the acts of image-based sexual abuse, including non-consensual pornographic deep fakes alongside other forms of sexual abuse.

These statistics illustrate the continuous character of sexual violence.⁹³ In this sense, image-based sexual abuse and non-consensual pornographic deep fakes pass into other forms of sexual violence and share the basic common characters.⁹⁴ This continuum framework rejects the hierarchy between physical and digital acts of sexual violence.⁹⁵ This framework is helpful to challenge societal outlooks and call for comprehensive regulation, however, it is not fully compatible with criminal law. “Fair labelling” of a crime belongs to the legal certainty and security principles of a democratic society. For both offenders and victims, as well as for the entire legal community, it is not only important whether an offender is convicted at all but also based on which elements of the crime

⁸⁵ Henry et al., *Image-based Sexual Abuse*, 25.

⁸⁶ Citron and Franks, “Criminalizing Revenge Porn”, 353, 354.

⁸⁷ Ibid.

⁸⁸ Sanne Kalf, “What Does a Feminist Approach to Deepfake Pornography Look like?” Masters of Media | University of Amsterdam, October 24, 2019. <http://mastersofmedia.hum.uva.nl/blog/2019/10/24/what-does-a-feminist-approach-to-deepfake-pornography-look-like/>

⁸⁹ Chesney and Citron, “Deep Fakes” 1773; Citron, “Sexual Privacy” 1925.

⁹⁰ Prasad, “Digitalisierung geschlechtsspezifischer Gewalt” in *Geschlechtsspezifische Gewalt*, ed. Prasad 21.

⁹¹ Henry et al., *Image-based Sexual Abuse*, 80.

⁹² Ibid., 21.

⁹³ See fn. 90, 91, 92.

⁹⁴ For more on the continuum of sexual violence, see McGlynn, Rackley, and Houghton, “Beyond Revenge Porn,” 25-46; Jurasz and Barker, “Sexual Violence in the Digital Age,” 784-799.

⁹⁵ McGlynn, Rackley, and Houghton, “Beyond Revenge Porn,” 35.

they are convicted.⁹⁶ The designation of the crime and its systematic classification in the structure of the criminal code provides information about how the committed wrong is evaluated by society.⁹⁷ Therefore, a criminal provision that targets image-based sexual abuse apropos other sexual offences would be optimal.

3. Regulatory Approach

There are already existing laws applicable to pornographic deep fakes. However, pornographic deep fakes present unique challenges for the application of the existing law.⁹⁸ Firstly, the absence of consistent definitions for image-based sexual abuse and deep fakes as well as significant discrepancies in legislative content present challenges to identifying which rules are applicable to which deep fakes.⁹⁹ The information and communication environment of today means that once a video is online, it is close to impossible to get it off of the internet, which presents another legal issue.¹⁰⁰ Scholars also mention that the existing law puts the responsibility to hold the perpetrators liable on the victims.¹⁰¹ It is questionable whether this approach is appropriate in the case of pornographic deep fakes as the victims are hardly in a position to take appropriate measures after the material and psychological harms they suffered.¹⁰² Indeed both quantitative and qualitative studies confirm this assumption.¹⁰³ These challenges result in a low level of protection of the rights of the victims.¹⁰⁴ This shortcoming is being discussed both at the European Union, member state, and Swiss levels.¹⁰⁵

⁹⁶ Nora Scheidegger, Agota Lavoyer, and Tamara Stalder. “Reformbedarf im Schweizerischen Sexualstrafrecht.” *sui generis*, 2020, 71.

⁹⁷ Ibid.

⁹⁸ One European Union level study states that, although there are applicable rules that protect victims, the legal route remains challenging. Huijstee et al., “Tackling Deepfakes in European Policy” 39.

⁹⁹ Mania, “Revenge Porn and Fake Porn” 2088.

¹⁰⁰ Chesney and Citron, “Deep Fakes” 1764; Tyrone Kirchengast, “Deepfakes and Image Manipulation: Criminalisation and Control.” *Information & Communications Technology Law* 29, no. 3 (2020): 313.

¹⁰¹ Huijstee et al., “Tackling Deepfakes in European Policy” 50.

¹⁰² Ibid.

¹⁰³ In an Austrian 2018 study, 79% of female digital violence victims reported that they did not ask for support. “Gewalt Im Netz Gegen Frauen Und Mädchen.” Weisser Ring, March 8, 2021, 74. <https://www.weisser-ring.at/gewalt-im-netz-gegen-frauen-und-maedchen/>; In Switzerland, only 8% of female sexual violence victims reported the crime to the police. Bornatici, “Junge Frauen in der Schweiz” 97.

¹⁰⁴ Mania, “Revenge Porn and Fake Porn” 2088.

¹⁰⁵ Aengus Collins, and Touradj Ebrahimi. “Spotlight on risk: Risk governance and the rise of deepfakes” International Risk Governance Center EPFL, May 12, 2021. <https://www.epfl.ch/research/domains/irgc/specific-risk-domains/projects-cybersecurity/forging-authenticity-governing-deepfake-risks/>; Huijstee et al., “Tackling Deepfakes in European Policy”.

3.1. Applicable Law and Legal Challenges

Depending on the case facts, certain criminal provisions¹⁰⁶, data protection¹⁰⁷, and intellectual property¹⁰⁸ rules may also be applicable to non-consensual pornographic deep fakes. However, on the Pan-European level, non-consensual pornographic deep fakes are mostly discussed within the right to privacy and personality rights (*Persönlichkeitsrechte*).¹⁰⁹ The right to privacy has an important social function because there's a close connection between our ability to control who has access to us and our ability to create and maintain different sorts of social relationships with different people.¹¹⁰ Therefore, the following sections will inspect applicable personality rights under Swiss civil law and go only as far as illustrating the need for a specific criminal provision targeting image-based sexual abuse.

3.1.1. Applicable personality rights under Swiss civil law

Under the Swiss Civil Code (“ZGB”) personality as a whole is protected in Art. 28 ZGB. The clause refrains from listing individual personality rights and thus, remains applicable through changing times and technological advancements.¹¹¹ Although Art. 28 ZGB is applicable for any right that is inseparably linked to the person as a bearer, the doctrine and jurisprudence have formed categories of personality rights.¹¹² As follows, the main distinction can be made between the physical, psychological, and social spheres of protection (*die physische, psychische und soziale Schutzbereiche*).¹¹³

Under the physical sphere of personality rights, freedom of movement and sexual freedom and integrity are categorised as central rights.¹¹⁴ Here sexual integrity focuses on self-determination over one's own body and the sense of shame, which can also be violated through salacious behaviour

¹⁰⁶ see chapter 3.1.2. of this paper.

¹⁰⁷ “Right to be forgotten” under data protection rules could provide an additional remedy. Lantwin, “Deep Fakes” 576. Huijstee et al., “Deepfakes in European Policy” 38.

¹⁰⁸ In the cases of deep fakes that are produced by taking a pornographic video and montaging another face in it, the owners of the original video might raise intellectual property claims. However, this is not a fitting solution as the core of the violation is not on a property level. Citron and Franks, “Criminalizing Revenge Porn” 360.

¹⁰⁹ Huijstee et al., “Tackling Deepfakes in European Policy” 40, 64.

¹¹⁰ Nora Scheidegger, “Balancing Sexual Autonomy, Responsibility, and the Right to Privacy: Principles for Criminalizing Sex by Deception.” *German Law Journal* 22, no. 5 (2021): 780.

¹¹¹ Regina E. Aebi-Müller, “ZGB 28,” in *CHK - Handkommentar zum Schweizer Privatrecht*, eds. Peter Breidschmidt and Alexandra Jungo (Schulthess Juristische Medien: Switzerland, 2016), 104, rn. 2.

¹¹² Andrea Büchler, “ZGB 28,” in *ZGB Kommentar Schweizerisches Zivilgesetzbuch*, eds. Jolanta Kren Kostkiewicz, et al. (Orell Füssli: Switzerland, 2021) 150, rn. 2; Bianka S. Dörr, “ZGB 28” in *ZGB Kurzkomentar*, eds. Andrea Büchler and Dominique Jakob (Helbing Lichtenhahn: Switzerland) 93, rn 3.

¹¹³ Büchler, “ZGB 28,” in *Kommentar Schweizerisches Zivilgesetzbuch*, eds. Kostkiewicz, et al. 150, rn. 2.

¹¹⁴ *Ibid.*

and verbal attacks.¹¹⁵ Whereas the psychological integrity encases the emotional life.¹¹⁶ Violations occurring here can be exemplified as psychological terror and putting a person in a panic state.¹¹⁷ The social aspect of personality rights encapsulates a number of rights and most centrally includes the right to protection of privacy (*Schutz der Privatsphäre*).¹¹⁸ Although violations through non-consensual deep fakes might concern all spheres of personality herein, protection of privacy and related concrete rights hold particular significance.

a. The right to one's image, voice, and expression

The protection of social personality and protection of privacy under Art. 28 ZGB is further concretised as the right to one's image, voice, and expression.¹¹⁹ The right to one's image covers both the taking/making of the image and its publication.¹²⁰ In principle, no one may be depicted, i.e., photographed, drawn, or filmed, without their consent.¹²¹ The first challenge of pornographic deep fakes arises here. One could argue that pornographic deep fakes do not raise the same level of privacy concerns as other forms of image-based sexual abuse, because they do not “actually” depict a person who exists.¹²² This would not be a convincing argument as the Federal Court of Switzerland has previously ruled in favour of the woman whose personality rights were violated through a non-consensual portraiture painting, depicting her partially nude.¹²³ It was stated that despite the changing moral concepts, the depiction of a woman with a naked upper body represents a severe intrusion into her personality.¹²⁴

There are many types of deep fakes with varying levels of sophistication. It might be questionable if the right to one's image would be relevant for a low-quality deep fake in which a private person's manipulated face is imposed onto a pornographic video.¹²⁵ In a case concerning violation of the right to one's image through photomontage, Federal Court of Switzerland has drawn attention to the identifiability of the complainant.¹²⁶ In this case, the head was taken from a photograph of the complainant and the naked body was taken from an unknown third party, yet the complainant was

¹¹⁵ Aebi-Müller, “ZGB 28” in *Handkommentar zum Schweizer Privatrecht*, eds. Breidschmidt and Jungo, 106, rn. 11.

¹¹⁶ Dörr, “ZGB 28” in *ZGB Kurzkommentar*, eds. Büchler and Jakob, 94 rn 6

¹¹⁷ Büchler, “ZGB 28,” in *Kommentar Schweizerisches Zivilgesetzbuch*, eds. Kostkiewicz, et al., 150, rn. 4.

¹¹⁸ *Ibid.*, rn. 5.

¹¹⁹ Dörr, “ZGB 28” in *ZGB Kurzkommentar*, eds. Büchler and Jakob, 94, rn. 7.

¹²⁰ Büchler, “ZGB 28,” in *Kommentar Schweizerisches Zivilgesetzbuch*, eds. Kostkiewicz, et al., 151, 9.

¹²¹ *Ibid.*

¹²² Delfino, “Pornographic Deepfakes” 897.

¹²³ BGE 5P.40/2003 Decision Date: 27.05.2013.

¹²⁴ BGE 5P.40/2003, 2. Decision Date: 27.05.2013.

¹²⁵ Federal Court of Switzerland has previously ruled that a personality right violation through photomontage has occurred in the form of an honour violation. BGE 5A_376/2013.

¹²⁶ BGE 5A_553/2012, 2. Decision Date: 14.04.2014

still recognisable. Here the question was raised whether the right to one's image also includes the added foreign body parts in the case of a photomontage.¹²⁷

The Court decided that the face is crucial for the identification of a person and found that the entire appearance is ultimately attributed to the person in question.¹²⁸ Followingly, in the context of pornographic deep fakes, it could be assumed that the private person's face which is "technically constructed" by the perpetrator and montaged onto another body would still be considered a part of the right to one's image. The Federal Court of Switzerland also disregarded the fact that photomontage was obvious to notice and reiterated the identifiability of the face.¹²⁹ In this respect, even the low-quality pornographic deep fakes where it is obvious that the video is fake, or ones that are labelled as deep fakes would constitute a personality right violation for the person whose face is recognisable.

Therein another question would be whether the pornographic actress who consented to the original video could raise a claim under personality rights. Here the issue would not lie with her consenting to the original performance as consent needs to be specific regarding both the image and the purpose of use.¹³⁰ However, in the light of the discussion above, it could be disputed that her body is immediately recognisable. Here doctrine mentions that in cases of personality violations by the media, the injured party must be recognisable by third parties.¹³¹ It is not enough if the injured party is the sole person that can recognise his/herself in the disputed content.¹³² Nonetheless, depending on the concrete case facts redress could be sought in other areas of personality rights, e.g., violations of honour. Even though challenges mentioned here have presented solutions and certain precedent, they still cause victims of pornographic deep fakes continuing agony and stress.

b. Violations of honour

Another concretisation of social personality that would be relevant for pornographic deep fakes is the protection of honour.¹³³ Art. 28 ZGB concerns the human-moral validity of a person and the respect that a person may expect because they adhere to the prevailing moral standards.¹³⁴ In

¹²⁷ BGE 5A_553/2012, 3.1. Decision Date: 14.04.2014

¹²⁸ BGE 5A_553/2012, 3.1. Decision Date: 14.04.2014

¹²⁹ BGE 5A_553/2012, 3.1. Decision Date: 14.04.2014

¹³⁰ Aebi-Müller, "ZGB 28" in *Handkommentar zum Schweizer Privatrecht*, eds. Breidschmidt and Jungo, 112, rn. 25.

¹³¹ Caroline Kirchschräger, "ZGB 28, ZGB28a" in *Haftpflichtkommentar – Kommentar zu den schweizerischen Haftpflichtbestimmungen*, eds. Willi Fischer and Thierry Luterbacher (Dike: Switzerland, 2016) 1147, rn. 20.

¹³² Ibid.

¹³³ For protection of honour under social personality see Kirchschräger, "ZGB 28, ZGB28a" in *Haftpflichtkommentar* eds. Fischer and Luterbacher, 1145, 1146, rn. 17.

¹³⁴ Aebi-Müller, "ZGB 28" in *Handkommentar zum Schweizer Privatrecht*, eds. Breidschmidt and Jungo, 109, rn. 18.

contrast to violations of honour addressed under criminal law, Art. 28 ZGB also protects the social standing of a person regarding essential areas of life such as profession, politics, army, sport, etc.¹³⁵ Federal Court of Switzerland has previously considered the photomontage of a well-known male homosexual singer's face onto a female singer's body under violations of honour.¹³⁶ This violation occurred after he lost the Eurovision song contest to said female singer, thus the court also referred to the effect the violation had on his standing regarding his professional singing life.¹³⁷

This paper has illustrated that pornographic deep fakes cause immense humiliation to victims.¹³⁸ Therefore, it might be assumed that pornographic deep fakes (and deep fakes in general) could also be considered under violations of honour. This might be attractive especially because untrue statements are considered defamatory.¹³⁹ Moreover, untrue, and honour-infringing statements are always considered to be a violation by the Federal Supreme Court.¹⁴⁰ Nonetheless, through a gender-based lens it should be mentioned that applying honour-based clauses to pornographic deep fakes, labels pornographic content as “damaging” to a woman's reputation, further perpetuating gender inequality. Here, the main concern with pornographic deep fakes is the violation of a person's sexual autonomy and self-determination.¹⁴¹ Thus, they should be mainly addressed as such.

c. The unlawfulness of personality rights violations

Under the Swiss Civil Code, not every impairment of personality is considered a violation. For the protection of Art. 28 ZGB to be relevant, the impairment of personality needs to be of noticeable disturbance and certain intensity and cause a serious threat to the personality right.¹⁴² Perpetrator of a poorly constructed pornographic deep fake could argue that it was meant as just a “joke” and it meant no harm.¹⁴³ The intensity of the impairment is judged upon objective grounds, rather than the personal feelings of the injured party.¹⁴⁴ In cases of violations of personality rights by the media, it is considered whether the social reputation of a person appears to be impaired from the point of view of the average consumer of that media.¹⁴⁵

¹³⁵ For more on the criminal law to Art 28 ZGB relationship see chapter 3.1.1.f. of this paper; *Ibid.*; Kirchschräger, “ZGB 28, ZGB28a” in *Haftpflichtkommentar* eds. Fischer and Luterbacher, 1146, rn. 17.

¹³⁶ BGE 5A_376/2013, 4.4. Decision Date: 29.10.2013.

¹³⁷ BGE 5A_376/2013, 4.3, 4.4. Decision Date: 29.10.2013.

¹³⁸ See chapter 2.2. of this paper.

¹³⁹ Aebi-Müller, “ZGB 28” in *Handkommentar zum Schweizer Privatrecht*, eds. Breidschmidt and Jungo, 153, rn. 14.

¹⁴⁰ *Ibid.*

¹⁴¹ Citron, “Sexual Privacy” 1898.

¹⁴² Büchler, “ZGB 28,” in *Kommentar Schweizerisches Zivilgesetzbuch*, eds. Kostkiewicz, et al., 153, rn. 14.

¹⁴³ See fn. 53, 54, 55.

¹⁴⁴ Kirchschräger, “ZGB 28, ZGB28a” in *Haftpflichtkommentar* eds. Fischer and Luterbacher, 1147, rn. 21.

¹⁴⁵ Kirchschräger, “ZGB 28, ZGB28a” in *Haftpflichtkommentar* eds. Fischer and Luterbacher, 1147, rn. 22.

Majority of the Swiss public which was surveyed by Amnesty International found that sending intimate photos of oneself to other people without consent is not acceptable under any circumstances.¹⁴⁶ Considering this strong attitude against one type of image-based sexual abuse, a similar reaction may be assumed in cases of non-consensual pornographic deep fakes. Thus, it is feasible to say that almost all non-consensual pornographic deep fakes would surpass this threshold and be regarded as a violation of personality rights. Nonetheless, the injured person is to prove the violation of personality right as per Art. 8 ZGB and proving the intensity of their suffering puts the victims under continuous scrutiny.¹⁴⁷

All violations are considered unlawful unless there are grounds for justification.¹⁴⁸ In the case of non-consensual pornographic deep fakes, overriding public and private interests may be examined as possible grounds for justification.¹⁴⁹ Here the public's need for information is considered a weighty public interest.¹⁵⁰ Even for persons who are in the public eye, in which public interest in information is given greater weight, barely any cases are conceivable that would justify the personality right violation through a pornographic deep fake on the grounds of overriding public interest.¹⁵¹ In overriding private interests, perpetrators may claim for example freedom of expression, freedom of arts, and satire.¹⁵²

In a case concerning the non-consensual partially nude portraiture painting of a woman, Federal Courts have dismissed the claim for overriding freedom of arts.¹⁵³ Satire, humour, and caricature necessarily contain a value judgement and by definition, alienate or exaggerate facts. In this respect, their unlawfulness is to be judged upon a different standard and only to be assumed if the depiction “exceeds the limits inherent in its nature to an intolerable extent”.¹⁵⁴ It is stated in the doctrine that

¹⁴⁶ “Repräsentative Studie: Mehrheit Für ‘Nur-Ja-Heisst-Ja.’” Amnesty International Schweiz, April 12, 2022. <https://www.amnesty.ch/de/themen/frauenrechte/sexualisierte-gewalt/dok/2022/mehrheit-fuer-nur-ja-heisst-ja>.

¹⁴⁷ The Federal Court has previously ruled that if, according to general life experience, the injury can cause intense harm, it is sufficient to prove this injury; the intensity of the harm then no longer needs to be proven. BGE 5A_376/2013, 8.1. Decision Date: 29.10.2013.; In another case concerning the non-consensual use of an actor's old movie performance in an advertisement, Swiss courts have decided that although there was a violation of personality rights, the severity was not enough to justify granting him damages. BGE 129 III 715, 4 Decision Date: 25.08.2003.

¹⁴⁸ Büchler, “ZGB 28,” in *Kommentar Schweizerisches Zivilgesetzbuch*, eds. Kostkiewicz, et al., 153, rn. 15.

¹⁴⁹ These are consent, overriding private and public interest, and legal grounds. Kirchschräger, “ZGB 28, ZGB28a” in *Haftpflichtkommentar* eds. Fischer and Luterbacher, 1149, 1150. In non-consensual pornographic deep fakes, overriding private and public interests are the only relevant justification grounds.

¹⁵⁰ Kirchschräger, “ZGB 28, ZGB28a” in *Haftpflichtkommentar* eds. Fischer and Luterbacher, 1149, 1155, rn. 27, 55.

¹⁵¹ Lantwin, “Deep Fakes” 578.

¹⁵² Kirchschräger, “ZGB 28, ZGB28a” in *Haftpflichtkommentar* eds. Fischer and Luterbacher, 1149, rn. 27.

¹⁵³ BGE 5P.40/2003 Decision Date: 27.05.2003.

¹⁵⁴ Aebi-Müller, “ZGB 28” in *Handkommentar zum Schweizer Privatrecht*, eds. Breidschmidt and Jungo, 109, rn. 20.

the publication of manipulated photographs is a rather serious violation of personality because the image is perceived by the average viewer as a particularly credible source of information.¹⁵⁵

The Federal Court has previously negated the satire defence in a case depicting a well-known homosexual male singer in women's clothing through photomontage.¹⁵⁶ In this respect, reasons for justification for photomontages are considered limited.¹⁵⁷ In the case of pornographic deep fakes, their hyper-realism and sensitive material bring the violation to a new level of intensity. Regardless of if they depict private people or people in the public eye¹⁵⁸, hardly any constellations are conceivable in which the dissemination and consumption of pornographic deep fakes would outweigh the protection of personality rights.¹⁵⁹

d. Possible defendants of the claim

When a personality right is violated, action regarding Art. 28 ZGB could be directed against anyone.¹⁶⁰ The person committing the violation and anyone who participates in the violation of personality rights is liable.¹⁶¹ In a traditional media sense, this allows for action to be taken against many people who participate in the violation like the journalist, chief editor along with the media outlet.¹⁶² Pornographic deep fakes however are strictly the concern of new media.¹⁶³ As identifying the perpetrators, i.e., the person who has created and uploaded the video, is one of the biggest challenges of deep fakes,¹⁶⁴ raising a possible claim against websites where these videos are hosted might be considered. Contrary to other jurisdictions that exempt internet service providers from liability for the content they host to a certain extent,¹⁶⁵ Switzerland does not regulate specific

¹⁵⁵ Ibid.

¹⁵⁶ BGE 5A_376/2013, 5.1 Decision Date: 29.10.2013.

¹⁵⁷ Aebi-Müller, "ZGB 28" in *Handkommentar zum Schweizer Privatrecht*, eds. Breidschmidt and Jungo, 112, rn. 25. Federal Court of Switzerland has previously accepted satire as a justification on a photomontage of a big pharma CEO's head onto a naked body. The photomontage was done for a political purpose and the critique was directed to his position as a high earning CEO, rather than him personally. BGE 5A_553/2012, 3.2, 3.3, 3.7. Decision Date: 14.04.2014.

¹⁵⁸ More on the privacy expectations of people on the public eye see Aebi-Müller, "ZGB 28" in *Handkommentar zum Schweizer Privatrecht*, eds. Breidschmidt and Jungo, 116.

¹⁵⁹ Lantwin "Deep Fakes" 578.

¹⁶⁰ Büchler, "ZGB 28," in *Kommentar Schweizerisches Zivilgesetzbuch*, eds. Kostkiewicz, et al., 152-153, rn 13.

¹⁶¹ Dörr, "ZGB 28" in *ZGB Kurzkommentar*, eds. Büchler and Jakob, 95, rn 13.

¹⁶² The mere placing of a link on a website to a newspaper is not considered sufficient as an act of participation. Büchler, "ZGB 28," in *Kommentar Schweizerisches Zivilgesetzbuch*, eds. Kostkiewicz, et al., 152-153, rn 13.; Aebi-Müller, "ZGB 28" in *Handkommentar zum Schweizer Privatrecht*, eds. Breidschmidt and Jungo, 116, rn. 37.

¹⁶³ See chapter 2 of this paper.

¹⁶⁴ Lantwin "Deep Fakes" 578.; Huijstee et al., "Deepfakes in European Policy" VII; Delfino, "Pornographic Deepfakes," 901.

¹⁶⁵ For more information on the exemption Section 230 grants in the U.S. see Chesney and Citron, "Deep Fakes" 1795-1800.; For more information on the exemption e-Commerce Directive grants in the EU see Huijstee et al., "Deepfakes in European Policy" 41-46.

provisions for internet service providers in this context.¹⁶⁶ Thus, content on the internet is to a large extent governed by the same rules that apply in offline situations.¹⁶⁷ However, the issues of identifying the perpetrators, problems arising out of jurisdiction, and enforcing the rules for service providers situated abroad are recognised on a Swiss level as well.¹⁶⁸

e. Types of claims to be raised in a civil lawsuit

In cases of personality rights violations, types of claims the injured party may bring are listed as general in Art. 28a ZGB. The first paragraph of Art. 28a ZGB foresees that the following claims may be raised; the prohibition of a threatened violation; the remedy of an existing violation (removal); and the declaration of a violation as unlawful if it continues to have a disruptive effect. In the context of pornographic deep fakes, the claim to remedy an existing violation (removal) has the most relevance, as victims are rarely aware of the deep fake's existence before it is made public.¹⁶⁹ The prerequisites for removal under Art. 28a par. 1 ZGB are firstly that the violation is ongoing at the time of the judgment, secondly that it is possible to remedy the violation and thirdly that the order to remove the violation is proportionate.¹⁷⁰

Removal examples of such injunctions in a traditional media context are the prohibition of the distribution of a book that violates personality rights, which is already on the market, the removal of defamatory posters that have already been put up, or the surrender of documents that violate personality rights.¹⁷¹ In a new media context the Federal Court of Switzerland has previously ruled for the removal of a blog post from the website on which it was hosted.¹⁷² The third party providing the hosting services was held equally liable for the removal of the violating post.¹⁷³ Thus, victim of a pornographic deep fake could request the removal of the violating content, provided it is possible to remove it. Here, the above-mentioned challenges of identifying the perpetrators and enforcing the rules remain, as these videos may be hosted in remote jurisdictions.¹⁷⁴

¹⁶⁶ Platform liability is left out of the scope of this paper. For more information on this situation in Switzerland see Bundesamt für Kommunikation. "Zivilrechtliche Verantwortlichkeit Der Internet-Provider." BAKOM, April 11, 2019. <https://www.bakom.admin.ch/bakom/de/home/digital-und-internet/internet/zivilrechtlicher-verantwortlichkeit-der-internet-provider.html>.

¹⁶⁷ Lukas Morscher and Lukas Staub "e-Commerce 2021, Switzerland." Lenz & Staehelin. Lexology, 2021. 100. https://www.lenzstaehelin.com/fileadmin/user_upload/GTDT_E-Commerce_Switzerland_2021.pdf.

¹⁶⁸ Bundesrat. "Kein Handlungsbedarf Bei Zivilrechtlicher Verantwortlichkeit Von Providern: Essential Results of the Federal Council Report" Eidgenössisches Justiz- und Polizei Department, December 11, 2015. 8. <https://www.ejpd.admin.ch/ejpd/de/home/aktuell/news/2015/2015-12-110.html>.

¹⁶⁹ Delfino "Pornographic Deepfakes" 935.

¹⁷⁰ Büchler, "ZGB 28," in *Kommentar Schweizerisches Zivilgesetzbuch*, eds. Kostkiewicz, et al., 156, rn. 3.

¹⁷¹ Ibid.

¹⁷² The blog that was found to be violating the personality right of the claimant, was written on a website where anyone could write a blog post. BGE 5A_792/2011 Decision Date: 14.01.2013.

¹⁷³ BGE 5A_792/2011 Decision Date: 14.01.2013; see chapter 3.1.1.d. of this paper.

¹⁷⁴ See chapter 3.1.1.d. of this paper; Chesney and Citron, "Deep Fakes" 1774.

Art. 28a par. 3 ZGB foresees that damages and compensation, as well as the surrender of a profit, may also be claimed. Compensation for financial loss and emotional suffering is another important claim for victims of pornographic deep fakes, which is raised in connection to the general norms of the liability law as per Art. 41 ff. in Swiss Code of Obligations (“OR”).¹⁷⁵ If the victim of a pornographic deep fake suffered pecuniary loss, they would be entitled to damages.¹⁷⁶ The requirements here are based on OR 41 and are as follows; damage, wrongfulness, adequate causal connection and fault.¹⁷⁷ It is difficult to prove fault, especially in connection with personality violations by the media,¹⁷⁸ which would be enhanced by the other challenges victims of pornographic deep fakes experience while raising civil claims.¹⁷⁹

Under this article, compensation for emotional suffering can be claimed in connection with the fault-based liability regulated under Art. 49 OR.¹⁸⁰ The suffering must be of a severity that justifies financial compensation, and it must not have already been redressed in some other way.¹⁸¹ Furthermore, the claim for compensation must be quantified and the injured party must explain the circumstances that caused their emotional pain.¹⁸² Here, the judge has considerable discretion not only with regard to the qualification of the suffering that gives rise to compensation but also with regard to the compensation amount.¹⁸³ Proving the intensity of their emotional suffering justifies a compensation, along with proving fault are obstacles for victims of pornographic deep fakes.¹⁸⁴

Earlier this paper concluded that in many cases pornographic deep fakes occur alongside other types of abuse.¹⁸⁵ In this context Art. 28b ZGB becomes relevant for personality rights violations through violence, threats, or stalking. The provision provides several measures that concretise and

¹⁷⁵ Kirchschräger, “ZGB 28, ZGB28a” in *Haftpflichtkommentar* eds. Fischer and Luterbacher, 1150, rn. 33

¹⁷⁶ For the financial harm that victims of pornographic deep fakes experience, see chapter 2.2.1. of this paper. The person whose personality has been unlawfully infringed has suffered pecuniary loss, he or she is entitled to damages. Büchler, “ZGB 28,” in *Kommentar Schweizerisches Zivilgesetzbuch*, eds. Kostkiewicz, et al., 157, rn. 10.

¹⁷⁷ *Ibid.*

¹⁷⁸ *Ibid.*, For the context of new media, the Federal Council of Switzerland takes the view that the criterion of the proximity of the provider to the content, i.e., the provider’s editorial influence, is decisive when determining liability, see Bundesrat, “Federal Council Report,” 3. Thus, proving fault for the internet service provider in which the pornographic deep fake is hosted is assumed to be even more difficult, as the website’s editorial influence would most likely be very minimal.

¹⁷⁹ See chapter 3.1.1.d. of this paper.

¹⁸⁰ Büchler, “ZGB 28,” in *Kommentar Schweizerisches Zivilgesetzbuch*, eds. Kostkiewicz, et al., 157, rn. 11.

¹⁸¹ *Ibid.*

¹⁸² *Ibid.*

¹⁸³ Amongst other things, the degree of fault is considered in the determination of compensation amount. Kirchschräger, “ZGB 28, ZGB28a” in *Haftpflichtkommentar* eds. Fischer and Luterbacher, 1151, rn. 37.

¹⁸⁴ see fn. 166, 167, 178.

¹⁸⁵ see chapter 2.4. of this paper.

supplement the defensive measures against personality violations regulated in Art. 28a ZGB.¹⁸⁶ These measures in Art. 28b ZGB are listed as follows; prohibition of approaching the victim (par. 1 fig. 1); prohibition of being in certain locations (par. 1 fig. 2); prohibition of contacting the victim (par. 1 fig. 3); and banning the perpetrator from the shared home (par. 2).

In the context of Art. 28b ZGB, violence is the direct impairment of the physical, psychological, sexual, or social integrity of a person.¹⁸⁷ A threat is understood to be the threat of violence against the claimant or persons close to them.¹⁸⁸ Stalking means the compulsive pursuit and harassment of the claimant, and it must occur over a longer period of time and cause objective fear.¹⁸⁹ In cases where violations through pornographic deep fakes occur in the context of or alongside the violence, threats, or stalking within the meaning of Art. 28b ZGB, these measures could also be raised.

f. From Art. 28 ZGB to criminal law

The protection of personality under private law is supplemented by protection under criminal law.¹⁹⁰ There is no general criminal clause corresponding to Art. 28 ZGB however, certain individual areas of personality are also protected by provisions of criminal law e.g., honour.¹⁹¹ As illustrated above, Art. 28 ZGB is relevant when it comes to non-consensual pornographic deep fakes. However, it has a significant enforcement deficit.¹⁹² Once these videos are uploaded online, it becomes almost impossible to deal with using methods of conventional civil law.¹⁹³ It is very hard to identify and locate the perpetrators, they may be using software that makes their IP address untraceable.¹⁹⁴ Deep fakes are also easily copied, shared, and stored in remote jurisdictions.¹⁹⁵ Considering all of these challenges, supplementing the protection for personality rights in pornographic deep fakes with a criminal provision addressing the broader context of image-based sexual abuse is a valid option.¹⁹⁶ Criminalising non-consensual pornographic deep fakes by itself, would not completely solve all challenges. However, having the resources of the criminal justice system would help.

¹⁸⁶ Büchler, “ZGB 28,” in *Kommentar Schweizerisches Zivilgesetzbuch*, eds. Kostkiewicz, et al., 159, rn 1.

¹⁸⁷ Büchler, “ZGB 28,” in *Kommentar Schweizerisches Zivilgesetzbuch*, eds. Kostkiewicz, et al., 159, rn. 2.

¹⁸⁸ *Ibid.*

¹⁸⁹ *Ibid.*

¹⁹⁰ Aebi-Müller, “ZGB 28” in *Handkommentar zum Schweizer Privatrecht*, eds. Breidschmidt and Jungo, 105, 106, rn. 9.

¹⁹¹ *Ibid.*

¹⁹² see chapter 3.1.1.d. of this paper

¹⁹³ Lantwin, “Deep Fakes” 576.

¹⁹⁴ Delfino, “Pornographic Deepfakes” 899.; Huijstee et al., “Deepfakes in European Policy” 50.

¹⁹⁵ Chesney and Citron, “Deep Fakes” 1774.

¹⁹⁶ Huijstee et al., “Deepfakes in European Policy” 61.

Criminal law provides victims with better legal remedies and minimises further victimisation.¹⁹⁷ Additionally, the deterrence function of criminal law would allow for the effective discouragement of potential creators of non-consensual pornographic deep fakes.¹⁹⁸ By contrast, the deterrence and remedy civil law can offer stands relatively modest.¹⁹⁹ Furthermore, practical concerns and the continuous scrutiny victims suffer, often render civil law solutions more theoretical than real.²⁰⁰ Due to these shortcomings, the demand to rely on milder alternatives to criminal law in cases of clear violations of the right to sexual self-determination, seems troublesome.²⁰¹ Especially considering pornographic deep fakes and image-based sexual abuse pass on or occur in the context of other types of sexual abuse, they should be addressed in a way that reflects this.²⁰²

In addition to the violation of sexual integrity and autonomy of the victim, the societal effects and harms of pornographic deep fakes should not be overlooked. These videos perpetuate gender inequality and hostility towards women and minorities.²⁰³ Criminal response to all forms of image-based sexual abuse is valuable to inspire and influence proper societal response.²⁰⁴ As follows, the European Parliament also recognises extending the current criminal legal framework at member state levels as a valid policy option.²⁰⁵ Some states have introduced criminal provisions targeting image-based sexual abuse.²⁰⁶ Similar efforts are being made in Switzerland as well. In light of this, the following chapters are intended to provide only an overview of these reform efforts in Switzerland.

3.1.2 The need for a criminal response

There are several already existing clauses in the Swiss Criminal Code (“StGB”) that might be indirectly relevant for non-consensual deep fakes, depending on the specific case facts. This chapter will provide a brief overview of these. The third section of StGB (Art. 173-179novies) regulates

¹⁹⁷ Delfino, “Pornographic Deepfakes” 902.

¹⁹⁸ Mania, “Revenge Porn and Fake Porn” 2090.

¹⁹⁹ Citron and Franks, “Criminalizing Revenge Porn” 357.

²⁰⁰ Ibid.

²⁰¹ Nora Scheidegger, “Revision des Sexualstrafrechts: die Verankerung des Konsensprinzips im StGB” in *Recht und Geschlecht*, ed. Juristinnen Schweiz (Switzerland: Dike, 2022), 208.

²⁰² See chapter 2.4. of this paper.

²⁰³ Huijstee et al., “Deepfakes in European Policy,” 29; Chesney and Citron, “Deep Fakes” 1776.

²⁰⁴ Jurasz and Barker, “Sexual Violence in the Digital Age,” 785.

²⁰⁵ Huijstee et al., “Deepfakes in European Policy,” 61.

²⁰⁶ In Germany, Art. 201a of StGB was regulated in 2014 to address the violation of the most personal sphere of life and personality rights through image recordings. This clause is sectioned under violation of personal life and privacy (15. Abschnitt, *Verletzung des persönlichen Lebens- und Geheimbereichs*) under the criminal code. In 2020, Art. 184k of StGB was regulated to address the violation of the intimate area through image recordings. This clause is sectioned under criminal offences against sexual self-determination (13. Abschnitt, *Straftaten gegen die sexuelle Selbstbestimmung*), properly contextualising image-based sexual abuse within where it belongs, thus making pornographic deep fakes punishable.

the criminal acts against one's honour and private sphere (*Strafbare Handlungen gegen die Ehre und den Geheim- oder Privatbereich*). The protection of honour under criminal law is limited to the moral sphere, namely to the reputation and the feeling of the person concerned to be an honourable person.²⁰⁷ Herein, the concept of honour under criminal law is narrower than that under civil law.²⁰⁸ Under criminal acts against one's honour, defamation, and slander (*Ehrverletzungen, üble Nachrede*) in Art. 173 StGB, and libel (*Verleumdung*) in Art. 174 StGB might have relevance for pornographic deep fakes.

Defamation and slander under Art. 173 are relevant for whoever accuses another of dishonourable conduct or other facts likely to damage their reputation. Libel under Art. 174 StGB is a qualified form of defamation and is fulfilled when the perpetrator accuses the injured party of dishonourable facts or behaviour, which do not correspond to reality.²⁰⁹ In contrast to defamation, the objective elements of Art. 174 StGB, therefore, require in every case that the defamatory statement is untrue.²¹⁰ Art. 176 StGB foresees that verbal defamation and libel are deemed equivalent to statements made in images and pictures. In addition to defamation, slander, and libel, Art. 177 StGB states that anyone who attacks someone's honour in another way including through an image shall be punished. Thus, these laws might apply to certain pornographic deep fakes, provided the concrete case fulfils the objective requirements.

Depending on the type of data gathered for the deep fake, Art. 179novies StGB might be relevant as the unauthorised acquisition of sensitive personal data that is not freely accessible. In addition to the crimes against honour and private sphere, laws on extortion and blackmail regulated in Art. 156 StGB might be relevant in some cases of pornographic deep fakes.²¹¹ The case facts might contain a version of cybercrimes in Art. 143 StGB, like hacking into the victim's computer or social media accounts. Additionally, criminal laws on pornography pursuant to Art. 197/4 and Art. 197/5 StGB might be significant if the deep fake involves the content of an underage person, animals, and/or violence.

However, similar to violations of honour under civil law, applying crimes against honour to pornographic deep fakes has the problem of overlooking the core problem with these videos.²¹²

²⁰⁷ Similarly, to the violation of one's honour under civil law, criminal offences are to be interpreted in a constitutional manner with a view to freedom of opinion, information and, media. Andreas Donatsch "StGB 173-179novies" in *StGB: Kommentar Schweizerisches Strafgesetzbuch*, eds. Andreas Donatsch, et al. (Orell Füssli: Switzerland, 2022) 451, rn. 2.

²⁰⁸ Ibid.; see chapter 3.1.1.b. of this paper.

²⁰⁹ Ibid. 461, rn 1.

²¹⁰ Ibid.

²¹¹ McGlynn, Rackley, and Houghton, "Beyond Revenge Porn," 34.

²¹² See chapter 3.1.1.b. of this paper.

The core issue of pornographic deep fakes is the non-consensual violation of a person's sexual autonomy and self-determination, not the damaging or libellous character of these videos. This is further illustrated in the interconnection these videos have to other types of sexual abuse.²¹³ Hence, they should be defined and addressed under the offences against sexual integrity section of the Criminal Code.²¹⁴ Moreover the laws overviewed in this chapter are applicable to only certain cases depending on the situation. This unclear application of law further muddles the challenges of pornographic deep fakes.²¹⁵ In turn, the legal system stays inaccessible to victims. A risk governance study conducted in Switzerland advocated for the prohibition of non-consensual pornographic deep fakes for similar reasons.²¹⁶ Even if the enforcement of such a prohibition would be difficult, the law plays an important role in signalling societal boundaries.²¹⁷

a. The efforts of addressing image-based sexual abuse under the reform of the Swiss Criminal Code on Sexual Offences

Discussions on a reform of the Swiss Criminal Code on Sexual Offences (*Sexualstrafrecht*) were instigated after Istanbul Convention's entry in the force in Switzerland in 2018.²¹⁸ On 01.02.2021, Commission for Legal Affairs of the Council of States under the Swiss Parliament sent proposals for a revision of the Swiss Criminal Code on Sexual Offences to consultation of political parties, academic institutions, and civil society actors.²¹⁹ On 08.08.2021, the report on the outcome of this consultation procedure was published.²²⁰

The revision aims to also address image-based sexual abuse under a few different clauses. The sexual harassment clause of Art. 198 StGB, which previously excluded all types of non-physical contact,²²¹ was amended to include "harassment by means of images". This was supported by almost all participants in the consultation process.²²² A considerable portion of participants welcomed the regulation on criminalising the "forwarding" of pornographic selfies under the rules

²¹³ See chapter 2.4. of this paper.

²¹⁴ The recent efforts to address image-based sexual abuse under the sexual offences section of StGB see chapter 3.1.2.a. of this paper.

²¹⁵ Delfino, "Pornographic Deepfakes," 901. Main challenges are identifying the perpetrators, pornographic deep fakes being hosted in remote jurisdictions, unclear platform liability. See chapter 3.1.1. of this paper.

²¹⁶ Collins and Ebrahimi, "Spotlight on Risk," 3.

²¹⁷ Ibid.

²¹⁸ Scheidegger, "Revision des Sexualstrafrechts" in *Recht und Geschlecht*, ed. Juristinnen Schweiz, 193, 194.

²¹⁹ "Kommission eröffnet Vernehmlassung zu einer Revision des Sexualstrafrechts." Die Bundesversammlung - Das Schweizer Parlament, February 1, 2021. <https://www.parlament.ch/press-releases/Pages/mm-rk-s-2021-02-01.aspx>.

²²⁰ "Bundesgesetz Zu Einer Revision des Sexualstrafrechts; Bericht über das Ergebnis des Vernehmlassungsverfahrens." Die Bundesversammlung - Das Schweizer Parlament, August 8, 2021. <https://www.parlament.ch/de/organe/kommissionen/sachbereichskommissionen/kommissionen-rk/berichte-vernehmlassungen-rk/vernehmlassung-rk-s-18-043>.

²²¹ Scheidegger, "Revision des Sexualstrafrechts" in *Recht und Geschlecht*, ed. Juristinnen Schweiz, 207.

²²² "Ergebnis des Vernehmlassungsverfahrens" p. 35.

on pornography in Art. 197/8bis StGB.²²³ Some participants specifically addressed “revenge porn” and advocated for its criminalisation.²²⁴ Discussions on this consultation report reflect modern rape reformers calling for the criminalisation of other non-physical forms of coercion as well.²²⁵

After above mentioned consultation process, by 11 votes to 1 the Commission decided on proposing a new offence targeting “revenge porn” under Art. 197a (draft)StGB.²²⁶ The new offence is titled “unauthorized forwarding of non-public sexual content” (*unbefugtes Weiterleiten von nicht öffentlichen sexuellen Inhalten*). Art. 197a (draft)StGB will make the forwarding of non-public sexual content, namely writings, sound or image recordings, images, objects, or performances, to a third person without the consent of the person identifiable therein, punishable up to one year or to a monetary penalty. In case the content is publicly shared, the offence will be punishable up to three years or a monetary penalty.

On 17.02.2022, the finalised draft of the revision of Criminal Code on Sexual Offences, and the report was submitted to the Council of States and Federal Council of Swiss Parliament. The draft is planned to be discussed in the coming summer session in Parliament.²²⁷ It is encouraging to see the reforms, and regulatory efforts to recognise image-based sexual abuse. The ultima ratio character of criminal law is certainly important and society's excessive punitiveness, its lust for punishment, and the use of criminal law as a cure-all are valid criticisms on some instances.²²⁸ However, this criticism cannot be applied to the required reform of the Swiss Criminal Code on Sexual Offences. The reform merely aims to systematically orient criminal law on the right to sexual self-determination and is certainly needed for criminalisation and effective prosecution of all non-consensual sexual contact fully and appropriately.²²⁹

3.2. Regulatory Efforts at the European Union Level

At the European Union level, many different regulations are being adopted to ensure safer technological development and digital space that respects users' fundamental rights. As follows,

²²³ Ibid., 31.

²²⁴ Ibid., 40.

²²⁵ Scheidegger, Lavoyer, and Stalder “Reformbedarf,” 773.

²²⁶ “Nein Heisst Nein?: Entwurf Für Ein Modernes Sexualstrafrecht Verabschiedet.” Die Bundesversammlung - Das Schweizer Parlament, February 18, 2022. <https://www.parlament.ch/press-releases/Pages/mm-rk-s-2022-02-18-2.aspx>.

²²⁷ Ibid.; “Strafrahmenharmonisierung Und Anpassung des Nebenstrafrechts an Das Neue Sanktionenrecht.” Die Bundesversammlung - Das Schweizer Parlament. <https://www.parlament.ch/de/ratsbetrieb/suche-curia-vista/geschaefte?AffairId=20180043>.

²²⁸ Scheidegger, “Revision des Sexualstrafrechts” in *Recht und Geschlecht*, ed. Juristinnen Schweiz, 207.

²²⁹ Ibid.

risks of deep fake technology are recognised, and several regulatory efforts are being made.²³⁰ Although they are not directly applicable to Switzerland, pornographic deep fakes are a problem that transcends jurisdiction and international efforts are valuable to combat them. In this context, the proposals for Artificial Intelligence Act (“AI Act”) and Digital Services Act (“DSA”) have relevance.²³¹ Under this chapter, this paper aims to provide a brief overview of these proposals in their relation to deep fakes.

AI Act adopts a risk-based approach to artificial intelligence and defines “high-risk” AI systems that pose significant risks to the health and safety or fundamental rights of persons. These high-risk systems are then subject to further requirements. Deep fakes are not regarded within high-risk systems. However, the AI systems used by law enforcement to detect deep fakes are regarded as high-risk.²³² Deep fakes themselves are regulated under Title IV, due to the specific risks of manipulation they pose.²³³ They are defined under Art. 52/3 AI Act as 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.

Art. 52/3 AI Act poses transparency obligations on the users of deep fakes; accordingly, it shall be disclosed that the content has been artificially generated or manipulated. The legal use by law enforcement, freedom of expression, and freedom of the arts and sciences are listed as legitimate exceptions to this disclosure obligation. AI Act does not differentiate between different types of deep fakes and regulates a general transparency obligation with rather wide exceptions. Transparency in the context of pornographic deep fakes is not enough as harms and effects are felt by the victims even when the video is clearly fake.²³⁴

Blanket regulations are not suitable for deep fakes as the harms, effects, and benefits occur on different areas of life and law.²³⁵ The wide range of applications means some deep fakes are going to be high-risk, while others are completely harmless. This fact has also been addressed by a study conducted under the European Parliamentary Research Service.²³⁶ Herein, it was recommended to

²³⁰ Huijstee et al., “Deepfakes in European Policy,” 37.

²³¹ Proposal for a Regulation of the European Parliament and Of the Council Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) dated 21.04.2021.; Proposal for a Regulation of the European Parliament and of the Council on a Single Market for Digital Services and amending Directive 200/31/EC dated 15.12.2020.

²³² AI Act preamble 38; AI Act Annex III, 6.

²³³ AI Act, Explanatory Memorandum 5.2.4.

²³⁴ See the chapter 2.2. of this paper.

²³⁵ Chesney and Citron, “Deep Fakes” 1771-1785; Kirchengast, “Deepfakes and Image Manipulation,” 313.

²³⁶ Huijstee et al., “Deepfakes in European Policy,” 59.

clarify which AI practices should be prohibited, to ban certain AI applications like pornographic deep fakes, and to foresee legal obligations for deep fake technology providers.²³⁷

For pornographic deep fakes, DSA also has significance as it aims to ensure that online platforms are held accountable regarding illegal and harmful content their services may host.²³⁸ DSA conceptualises illegal content broadly and states that unlawful and non-consensual sharing of private images fall under this definition.²³⁹ The act foresees transparency and due diligence obligations for all digital service providers.²⁴⁰ Further obligations are conceptualised for very large platforms, proportionate to their capability, adapted to issues and illegal content they most likely host.²⁴¹ The proposed measures contribute to better responding to illegal content including non-consensual sharing of users' private images and data, which non-consensual pornographic deep fakes would fall under.²⁴² Herein, notice and action and information obligations have specific relevance as it is going to provide a swift take-down response.²⁴³

4. Conclusion

There are many types of deep fakes. While there are some beneficial uses for this technology, pornographic deep fakes account for the overwhelming majority. It is apparent that blanket regulations are not effective in dealing with deep fakes; different rules depending on the types of deep fakes are needed. Consistent legal definitions should be adopted on what would constitute a non-consensual pornographic deep fake. Here, the level of sophistication and clear identifiability of the victim should not matter, as the harms and effects are felt the same way. All non-consensual pornographic deep fakes violate the sexual integrity and autonomy of victims.

Although pornographic deep fakes started with the depiction of female celebrities, democratisation of this technology means that anyone can be a victim. Perpetrators create these videos with different motivations, amongst which are taking revenge upon an ex-partner, shaming and silencing victims, or simply having fun. These videos have a gendered element. Victims of pornographic deep fakes are overwhelmingly women. Victims experience severe psychological, emotional, and

²³⁷ Huijstee et al., "Deepfakes in European Policy," VII, 61, 80, 81.

²³⁸ "Digital Services Act: Commission Welcomes Political Agreement on Rules Ensuring a Safe and Accountable Online Environment." European Commission, April 23, 2022. https://ec.europa.eu/commission/presscorner/detail/en/ip_22_2545.

²³⁹ DSA preamble 12; Art. 2(g).

²⁴⁰ DSA preamble 22.

²⁴¹ "Impact Assessment of the Digital Services Act." European Commission, December 15, 2020. part 1, 73, rn 298 <https://digital-strategy.ec.europa.eu/en/library/impact-assessment-digital-services-act>.

²⁴² Ibid., part 1., 64, rn. 249.

²⁴³ DSA Art. 8.

financial damage. The severity of the effects also has a gendered dimension, as women feel the harm in a more pronounced way.

Pornographic deep fakes should be contextualised and addressed within a broader sense of image-based sexual abuse as all abusive behaviour involving the non-consensual taking and sharing of sexual images. In this sense pornographic deep fakes are exacerbating image-based sexual abuse, as perpetrators no longer need actual nude images. This paper also illustrated that pornographic deep fakes frequently occur alongside or in the context of other types of abuse. Image-based sexual abuse is interconnected with and passes on to other forms of sexual abuse. Therefore, it should be contextualised and addressed alongside other types of sexual abuse under criminal law.

There are applicable civil laws under personality rights to pornographic deep fakes. However, they have a significant enforcement deficit. Identifying the perpetrators, enforcing the rules in remote jurisdictions where the videos are hosted, and holding the platforms accountable is challenging. Once these videos are uploaded online, it is very hard to undo the damage. Although it would not solve all the challenges, criminalising pornographic deep fakes would be helpful to provide symbolic justice to victims, deter potential perpetrators, and mobilise a proper societal response. Herein, further research is needed from criminal law scholars to examine how conventional standards of fault should be conceptualised to provide appropriate criminalisation and effective prosecution in the cases of image-based sexual abuse.

List of References

- Abram, Cleo. "The Most Urgent Threat of Deepfakes Isn't Politics. It's Porn." Vox. Vox, June 8, 2020. <https://www.vox.com/2020/6/8/21284005/urgent-threat-deepfakes-politics-porn-kristen-bell>. Accessed May 29, 2022
- Ajder, Henry, Giorgio Patrini, Francesco Cavalli, and Laurence Cullen. "Reports." The State of Deepfakes: Landscape, Threats, and Impact. Deeptrace labs, 2019. <https://sensity.ai/reports/>. Accessed May 29, 2022
- Ayyub, Rana. "I Was the Victim of a Deepfake Porn Plot Intended to Silence Me." HuffPost UK. HuffPost UK, November 21, 2018. https://www.huffingtonpost.co.uk/entry/deepfake-porn_uk_5bf2c126e4b0f32bd58ba316. Accessed May 29, 2022
- Bates, Samantha. "Revenge Porn and Mental Health." *Feminist Criminology* 12, no. 1 (2016): 22-42.
- Blakkarly, Jarny. "A Gay Sex Tape Is Threatening to End the Political Careers of Two Men in Malaysia." SBS News, June 17, 2019. <https://www.sbs.com.au/news/article/a-gay-sex-tape-is-threatening-to-end-the-political-careers-of-two-men-in-malaysia/ilgqdaqo5>. Accessed May 29, 2022
- Bornatici, Christina. "Frauenfragen 2022." Eidgenössische Kommission für Frauenfragen EKF, March 22, 2022. <https://www.ekf.admin.ch/ekf/de/home/dokumentation/fachzeitschrift--frauenfragen--frauenfragen-2022.html>. Accessed May 29, 2022
- Breidschmidt, Peter and Alexandra Jungo, eds., *Handkommentar zum Schweizer Privatrecht*. Switzerland: Schulthess Juristische Medien, 2016.
- "Bundesgesetz Zu Einer Revision des Sexualstrafrechts Bericht über das Ergebnis des Vernehmlassungsverfahrens." Die Bundesversammlung - Das Schweizer Parlament, August 8, 2021. <https://www.parlament.ch/de/organe/kommissionen/sachbereichskommissionen/kommissionen-rk/berichte-vernehmlassungen-rk/vernehmlassung-rk-s-18-043>. Accessed May 29, 2022
- Bundesrat. "Kein Handlungsbedarf Bei Zivilrechtlicher Verantwortlichkeit Von Providern." Eidgenössisches Justiz- und Polizei Department, December 11, 2015. <https://www.ejpd.admin.ch/ejpd/de/home/aktuell/news/2015/2015-12-110.html>. Accessed May 29, 2022
- Burgess, Matt. "Deepfake Porn Is Now Mainstream. and Major Sites Are Cashing In." WIRED UK, August 27, 2020. <https://www.wired.co.uk/article/deepfake-porn-websites-videos-law>. Accessed May 29, 2022
- Büchler, Andrea and Dominique Jakob, eds., *ZGB Kurzkomentar*. Switzerland: Helbing Lichtenhahn, 2018.
- Chesney, Robert, and Danielle Keats Citron. "Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security." *California Law Review* 107, no. 6 (2019): 1753–1820.
- Citron, Danielle Keats. "Sexual Privacy." *Yale Law Journal* 128 (2019): 1872–1960.
- Citron, Danielle Keats and Mary Anne Franks, "Criminalizing Revenge Porn," *Wake Forest Law Review* 49, no. 2 (2014): 345-391.
- Cole, Samantha. "Creator of DeepNude, App That Undresses Photos of Women, Takes It Offline." VICE, June 27, 2019. <https://www.vice.com/en/article/qv7agw/deepnude-app-that-undresses-photos-of-women-takes-it-offline>. Accessed May 29, 2022
- Collins, Angus, and Touradj Ebrahimi. "Forged Authenticity: Governing Deepfake Risks." International Risk Governance Center EPFL, May 12, 2021.

- <https://www.epfl.ch/research/domains/irgc/specific-risk-domains/projects-cybersecurity/forging-authenticity-governing-deepfake-risks/>. Accessed May 29, 2022
- Delfino, Rebecca. “Pornographic Deepfakes — Revenge Porn’s Next Tragic Act – the Case for Federal Criminalization.” *Fordham Law Review* 88, no. 3 (2019): 887-938.
- “Digital Services Act: Commission Welcomes Political Agreement on Rules Ensuring a Safe and Accountable Online Environment.” European Commission, April 23, 2022. https://ec.europa.eu/commission/presscorner/detail/en/ip_22_2545. Accessed May 29, 2022
- Dodge, Adam, and Erica Johnstone. “Using Fake Video Technology to Perpetuate Intimate Partner Abuse.” California Partnership to End Domestic Violence. Accessed May 29, 2022. <https://www.cpedv.org/publication/using-fake-video-technology-perpetuate-intimate-partner-abuse>. Accessed May 29, 2022
- Donatsch, Andreas, Stefan Heimgartner, Bernhard Isenring, Hans Maurer, Marcel Riesen-Kupper, and Ulrich Weder, eds., *StGB: Kommentar Schweizerisches Strafrechtbuch*. Switzerland: Orell Füssli, 2022.
- “End Revenge Porn.” Cyber Civil Rights Initiative. Accessed May 29, 2022. <https://cybercivilrights.org/>. Accessed May 29, 2022
- Fischer, Willi and Thierry Luterbacher, eds., *Haftpflichtkommentar – Kommentar zu den schweizerischen Haftpflichtbestimmungen*. Switzerland: Dike, 2016.
- “Gewalt Im Netz Gegen Frauen Und Mädchen.” Weisser Ring, March 8, 2021. <https://www.weisser-ring.at/gewalt-im-netz-gegen-frauen-und-maedchen/>. Accessed May 29, 2022
- Hauser, Andrea. “This Is How You Create The Perfect Deepfake.” Amount of Images, Lighting and Angles, November 22, 2018. <https://www.scip.ch/en/?labs.20181122>. Accessed May 29, 2022
- Henry, Nicola, Clare McGlynn, Asher Flynn, Kelly Johnson, Anastasia Powell, and Adrian J. Scott. *Image-Based Sexual Abuse: A Study on the Causes and Consequences of Non-Consensual Nude or Sexual Imagery*. London: Routledge, Taylor & Francis Group, 2021.
- “How Many Iterations Does a Realistic Deepfake Image Need?” Deepfake Now, April 21, 2020. <https://deepfakenow.com/how-many-iterations-realistic-deepfake-image/> Accessed May 29, 2022
- van Huijstee, Mariette, Pieter van Boheemen, Djurre Das, Jutta Jahnel, Murat Karaboga, and Martin Fatun. Tackling Deepfakes in European Policy. Panel for the Future of Science and Technology (STOA) | European Parliament, 2021. [https://www.europarl.europa.eu/stoa/en/document/EPRS_STU\(2021\)690039](https://www.europarl.europa.eu/stoa/en/document/EPRS_STU(2021)690039). Accessed May 29, 2022
- “Impact Assessment of the Digital Services Act.” European Commission, December 15, 2020. <https://digital-strategy.ec.europa.eu/en/library/impact-assessment-digital-services-act>. Accessed May 29, 2022
- “Iperov/Deepfacelab: Deepfacelab Is the Leading Software for Creating Deepfakes.” GitHub. Accessed May 29, 2022. <https://github.com/iperov/DeepFaceLab>. Accessed May 29, 2022
- Jurasz, Olga, and Kim Barker. “Sexual Violence in the Digital Age: A Criminal Law Conundrum?” *German Law Journal* 22, no. 5 (2021): 784-799.
- Juristinnen, Schweiz, ed. *Recht und Geschlecht*. Switzerland: Dike, 2022
- Kalf, Sanne. “What Does a Feminist Approach to Deepfake Pornography Look like?: Masters of Media.” Masters of Media | University of Amsterdam, October 24, 2019. <http://mastersofmedia.hum.uva.nl/blog/2019/10/24/what-does-a-feminist-approach-to-deepfake-pornography-look-like/>. Accessed May 29, 2022
- Kirchengast, Tyrone. “Deepfakes and Image Manipulation: Criminalisation and Control.” *Information & Communications Technology Law* 29, no. 3 (2020): 308-323.

- “Kommission eröffnet Vernehmlassung zu einer Revision des Sexualstrafrechts.” Die Bundesversammlung - Das Schweizer Parlament, February 1, 2021. <https://www.parlament.ch/press-releases/Pages/mm-rk-s-2021-02-01.aspx>. Accessed May 29, 2022
- Kostkiewicz, Jolanta Kren, Stephan Wolf, Marc Amstutz, Roland Fankhauser, eds. *ZGB Kommentar Schweizerisches Zivilgesetzbuch*. Switzerland: Orell Füssli, 2021.
- Kugler, Matthew B., and Carly Pace. “Deepfake Privacy: Attitudes and Regulation.” *Northwestern University Law Review* 116, no. 3 (2021): 611-680.
- Lantwin, Tobias. “Deep Fakes - Düstere Zeiten Für Den Persönlichkeitsschutz? .” *MMR* 2019, no. 9 (2019): 574-578.
- Lomas, Natasha. “Reface Grabs \$5.5m Seed Led by a16z to Stoke Its Viral Face-Swap Video App.” TechCrunch. TechCrunch, December 8, 2020. <https://techcrunch.com/2020/12/08/reface-grabs-5-5m-seed-led-by-a16z-to-stoke-its-viral-face-swap-video-app/>. Accessed May 29, 2022
- Maddocks, Sophie. “From Non-Consensual Pornography to Image-Based Sexual Abuse: Charting the Course of a Problem with Many Names.” *Australian Feminist Studies* 33, no. 97 (2018): 345–361.
- Mania, Karolina. “The Legal Implications and Remedies Concerning Revenge Porn and Fake Porn: A Common Law Perspective.” *Sexuality & Culture* 24, no. 6 (2020): 2080-2097.
- McGlynn, Clare, Erika Rackley, and Ruth Houghton. “Beyond ‘Revenge Porn’: The Continuum of Image-Based Sexual Abuse.” *Feminist Legal Studies* 25, no. 1 (2017): 25-46.
- Meskys, Edvinas, Aidas Liaudanskas, Julija Kalpokiene, and Paulius Jurcys. “Regulating Deep Fakes: Legal and Ethical Considerations.” *Journal of Intellectual Property Law & Practice* 15, no. 1 (2020): 24-31.
- Morscher, Lukas, and Lukas Staub. “e-Commerce 2021, Switzerland.” Lenz & Staehelin. Lexology, May 2021. https://www.lenzstaehelin.com/fileadmin/user_upload/GTDT_E-Commerce_Switzerland_2021.pdf. Accessed May 29, 2022
- Prasad, Nivedita. *Geschlechtsspezifische Gewalt in Zeiten der Digitalisierung*. Germany: Transcript, 2019.
- “Repräsentative Studie: Mehrheit Für ‘Nur-Ja-Heisst-Ja.’” Amnesty International Schweiz, April 12, 2022. <https://www.amnesty.ch/de/themen/frauenrechte/sexualisierte-gewalt/dok/2022/mehrheit-fuer-nur-ja-heisst-ja#>. Accessed May 29, 2022
- “Revenge Porn Website Operator Jailed.” The Guardian, April 4, 2015. <https://www.theguardian.com/us-news/2015/apr/04/revenge-porn-website-operator-jailed>. Accessed May 29, 2022
- Scheidegger, Nora, Agota Lavoyer, and Tamara Stalder. “Reformbedarf Im Schweizerischen Sexualstrafrecht.” *sui generis*, (2020): 58-75.
- Scheidegger, Nora. “Balancing Sexual Autonomy, Responsibility, and the Right to Privacy: Principles for Criminalizing Sex by Deception.” *German Law Journal* 22, no. 5 (2021): 769-783.
- Simonite, Tom. “The Web Is Drowning in Deepfakes and Almost All of Them Are Porn.” WIRED UK. WIRED UK, October 13, 2019. <https://www.wired.co.uk/article/deepfakes-porn>. Accessed May 29, 2022
- “Strafrahmenharmonisierung und Anpassung des Nebenstrafrechts an das neue Sanktionenrecht.” Die Bundesversammlung - Das Schweizer Parlament. Accessed May 30, 2022. <https://www.parlament.ch/de/ratsbetrieb/suche-curia-vista/geschaefte?AffairId=20180043>. Accessed May 29, 2022
- Uhl, Carolyn A, Katlin J Rhyner, Cheryl A Terrance, and Noël R Lugo. “An Examination of Nonconsensual Pornography Websites.” *Feminism & Psychology* 28, no. 1 (2018): 50-68.
- “Nein Heisst Nein’: Entwurf Für Ein Modernes Sexualstrafrecht Verabschiedet.” Die Bundesversammlung - Das Schweizer Parlament, February 18, 2022. <https://www.parlament.ch/press-releases/Pages/mm-rk-s-2022-02-18-2.aspx>. Accessed May 29, 2022

- “Zivilrechtliche Verantwortlichkeit Der Internet-Provider.” BAKOM, April 11, 2019.
<https://www.bakom.admin.ch/bakom/de/home/digital-und-internet/internet/zivilrechtlicher-verantwortlichkeit-der-internet-provider.html>. Accessed May 29, 2022.
- “8 Best Deepfake Apps and Tools in 2022.” RankRed, January 10, 2022.
<https://www.rankred.com/best-deepfake-apps-tools/>. Accessed May 29, 2022.