Deepfake Threats And EU Law: Navigating Disinformation, Cyber Violence, And The Risks Of Digital Manipulation[1]

MARIAM MAKHNIASHVILI[2]
RESEARCH FELLOW
THE ECONOMIC POLICY RESEARCH CENTER, GEORGIA

Abstract: *The purpose of this study is to examine the challenges posed by AI-generated deepfakes. They are one form of so-called 'synthetic media', which draw on advances in AI, using algorithms and deep learning to change elements of a photo, video or audio track, or to recreate a person's voice or face with life-like subtlety. creating the illusion of staged actions by someone else.[3] Earlier versions of generative AI required coding capabilities and technical proficiency; today, anyone with internet access is constrained only creativity is required.[4] The goal of this paper is to explore critical areas where EU legislative regulation is essential to protect fundamental rights.*
This research explores deepfake narratives, their impact on elections, social engineering, and harassment, and how the European Union (EU) is responding through policy measures and legal frameworks. Additionally, it examines the legal accountability of deepfake creators and platforms, particularly in cases of image-based sexual abuse and political deception. By analyzing the intersection of Directive (EU) 2024/1385 with freedom of expression, this study aims to evaluate the effectiveness of EU strategies in defending truth and combating disinformation. Since false information falls into three categories—disinformation, misinformation, and malinformation—deepfakes are classified as disinformation.[5] Additionally, this study will discuss the AI Act and Digital Services Act, assessing their roles in regulating deepfake-generated content, ensuring transparency, and addressing platform accountability within the EU's legal framework.
Key words: *image-based sexual abuse, Deepfake, cyber harassment, EU policy on AI accountability, Disinformation.*

INTRODUCTION

The photos, video, and audio recordings were once considered the "gold standard" for proof of reality, new artificial intelligence (AI) models that can create realistic (but fake) content has fundamentally changed past notions of trust.[6] Deepfakes are videos in which one face

[1] This research was supported by a Marie Curie Staff Exchange within the Horizon Europe Programme (grant acronym: CARSI, nr.: 101086415)
[2] Mariam Makhniashvili (mmakh21@freeuni.edu.ge) is a MSCA fellow at EPRC, Georgia. She studies law at Free University of Tbilisi. Her academic and practical experience includes legal research and political analysis, and she is a practicing lawyer specializing in corporate, administrative, and criminal law. ORCID: 0009-0000-5145-828X.
[3] This definition is based on insights from Sam Gregory, Heard About Deepfakes? Don't Panic. Prepare., World Economic Forum, November 23, 2018, Heard about deepfakes? Don't panic. Prepare | World Economic Forum and GREVIO's General Recommendation No. 1 on the Digital Dimension of Violence Against Women, Council of Europe, Strasbourg, 2021. 1680a49147
[4] UNESCO, Exposing Technology-Facilitated Gender-Based Violence in an Era of Generative AI, World Trends in Freedom of Expression and Media Development Series, 2nd edition, Paris, UNESCO, 2023, ISBN 978-92-3-100669-2 387483eng.pdf
[5] Tom Dobber, Nadia Metoui, Damian Trilling, Natali Helberger, and Claes de Vreese, "Do (Microtargeted) Deepfakes Have Real Effects on Political Attitudes?" The International Journal of Press/Politics, vol. 26, no. 1, 2021, 69–91. Chochia, A.; Nässi, T. (2021). Ethics and Emerging Technologies – Facial Recognition. IDP Revista de Internet Derecho y Política, 34, 1−12. DOI: 10.7238/idp.v0i34.387466.
[6] Beena Ammanath, in a world of deepfakes, we must build a case for trustworthy synthetic AI content, World Economic Forum, May 30, 2024. We must build a case for trustworthy AI synthetic content | World Economic Forum

has been (seamlessly) replaced by another face and sound is manipulated, so as to create the illusion that another person's actions are being staged.[7]

Deepfake technology represents a transformative advancement in artificial intelligence and digital manipulation. It refers to the use of deep learning algorithms to create compelling and often deceptive media content.[8] They include images, audio files and moving images, either individually or in combination.[9] These manipulated media are created using machine learning, and can make it seem as if a person says or does something, while in reality, they have never said or done anything of the sorts.[10] The existence of deepfakes may lead people to lose faith in visual material when processing information. [11]

Experiments have proven that visual messages influence society more strongly than verbal ones. [12] Further studies showed that ''image bites" are more effectively to shape voter opinions than "sound bites."[13] Prior later demonstrated that survey respondents exhibited higher knowledge retention when factual recall questions included both visual and verbal information.[14] Deepfakes have the ability to generate both voice and visuals within the same material, making them highly deceptive and dangerous. That's precisely why it is extremely dangerous and worth paying close attention to.

AI technologies which can generate 'deepfakes' can also be used in the production of both misinformation and disinformation. According to the Global Risks Report 2024, misinformation and disinformation were identified as the most significant global risk over the next two years, Synthetic content will manipulate individuals, damage economies and fracture societies in numerous ways over the next two years. However, in the long term, their ranking is expected to shift to fifth place among the most pressing global risks.[15]

1. DEEPFAKE NARRATIVES: AI AND THE MANIPULATION OF TRUTH

Various terms have emerged to describe information that polarizes, and deceives. The most popular terms are misinformation and disinformation, they are sometimes used interchangeably, but researchers agree they are separate and distinct.[16]

---

[7] GREVIO (Group of Experts on Action against Violence against Women and Domestic Violence), General Recommendation No. 1 on the digital dimension of violence against women, Strasbourg, Council of Europe, 2021. 1680a49147

[8] 2023 State Of Deepfakes: Realities, Threats, And Impact, Security Hero, 2023.2023 State Of Deepfakes: Realities, Threats, And Impact

[9] Dan Cavedon-Taylor, Deepfakes and Democracy: A Catch-22?, Journal of the American Philosophical Association, Cambridge University Press, April 23, 2025. Deepfakes and Democracy: A Catch-22? | Journal of the American Philosophical Association | Cambridge Core

[10] Tom Dobber, Nadia Metoui, Damian Trilling, Natali Helberger, Claes de Vreese, do (Microtargeted) Deepfakes Have Real Effects on Political Attitudes? The International Journal of Press/Politics, Sage Publications, July 25, 2020, p. 69–91. Do (Microtargeted) Deepfakes Have Real Effects on Political Attitudes? – Tom Dobber, Nadia Metoui, Damian Trilling, Natali Helberger, Claes de Vreese, 2021

[11] Teresa Weikmann, Hannah Greber, Alina Nikolaou, After Deception: How Falling for a Deepfake Affects the Way We See, Hear, and Experience Media, Europeana, University of Vienna, 2024.. Europeana.eu-1300-https___phaidra_univie_ac_at_o_2083412-a2c2effd53574d0f522252480124759d (4).pdf

[12] D. A. Graber, Seeing is Remembering: How Visuals Contribute to Learning from Television News, Journal of Communication, vol. 40, no. 3, 1990, pp. 134–156.

[13] Maria Elizabeth Grabe & Erik Page Bucy, Image Bite Politics: News and the Visual Framing of Elections, Oxford University Press, 2009.

[14] Markus Prior, Visual Political Knowledge: A Different Road to Competence? The Journal of Politics

[15] Global Risks Report 2024, World Economic Forum, January 10, 2024. WEF_The_Global_Risks_Report_2024.pdf

[16] Christina Nemr & William Gangware, Weapons of Mass Distraction: Foreign State-Sponsored Disinformation in the Digital Age. Weapons-of-Mass-Distraction-Foreign-State-Sponsored-Disinformation-in-the-Digital-Age.pdf

Print ISSN 0248-2851 • On-line/Linking ISSN 2743-4052 • ISBN 978-630-354-213-3

Misinformation is understood as the false information that is not intended to cause harm, unlike disinformation, which has a malicious intent.

We follow others in defining online disinformation as intentional behavior that purposively misleads and online misinformation as unintentional behavior that inadvertently misleads. Deepfakes are disinformation because they originate with intentional acts, but they become misinformation, too, if circulated online by people who mistakenly believe them to be truthful representations. For the purposes of this study, this distinction is not germane because we do not seek to explain the factors shaping a decision to share a deepfake.[17]

The spread of both disinformation and misinformation can have a range of harmful consequences, such as threatening our democracies, polarizing debates, and putting the health, security and environment of EU citizens at risk.[18]

Disinformation is widespread and harmful, epistemically and practically. We are facing a global information crisis that the Secretary-General of World Health Organization (WHO) has declared an 'infodemic'[19]

2. ARE DEEPFAKES REALLY A THREAT TO ELECTIONS?

What is the biggest threat from deepfakes? If you had asked people a year ago, many would have said their ability to disrupt global elections and create a misinformation and disinformation apocalypse, but that's not what happened. [20] In general, targeted phishing attacks on election officials or political parties can compromise sensitive information or introduce malware, often using deepfake technology. Beyond that, regarding disinformation and misinformation, AI-generated deepfake videos or audio recordings can be used to spread false information about candidates or manipulate public perception. Additionally, organized campaigns spreading misinformation through social media or other platforms can influence public opinion, undermine trust in election integrity, and ultimately sway election results. In reality, expectations were not met in reality and did not have the impact on the 2024 elections that many had predicted.[21]

We do not dispute that artificial intelligence has the ability to generate deepfakes, which can subsequently be used to spread disinformation and misinformation. However, its impact on society is very low when it comes to elections. Unfortunately, many studies focus on the capabilities of artificial intelligence, while discussing real-world outcomes less frequently.[22] In 2018, it was believed that disinformation campaigns could be part of hybrid threats to internal security, including election processes, particularly when combined with cyberattacks.[23]

---

[17] Cristian Vaccari & Andrew Chadwick, Deepfakes and Disinformation: Exploring the Impact of Synthetic Political Video on Deception, Uncertainty, and Trust in News, social media + Society, Sage Publications, February 19, 2020. Deepfakes and Disinformation: Exploring the Impact of Synthetic Political Video on Deception, Uncertainty, and Trust in News

[18] Tackling Online Disinformation | Shaping Europe's Digital Future, European Commission, 2025. Tackling online disinformation | Shaping Europe's digital future

[19] Mona Simion, Knowledge and Disinformation, Episteme, vol. 21, no. 4, December 2024, pp. 1208–1219. Published online by Cambridge University Press, June 29, 2023. Knowledge and Disinformation | Episteme | Cambridge Core

[20] Nathan Hamiel, *Deepfakes Are Here to Stay and We Should Remain Vigilant*, *World Economic Forum*, January 10, 2025. Deepfakes are here to stay and we should remain vigilant | World Economic Forum

[21] Gretchen Bueermann & Daniel Dobrygowski, Elections, Cybersecurity, and AI: The Growing Threat of Deepfakes and Social Engineering, World Economic Forum, November 8, 2023. Here's what to know about elections, cybersecurity and AI | World Economic Forum

[22] N. Hamiel, Illusion of Influence: The AI-Generated Misinformation Apocalypse That Wasn't, Perilous Tech, July 18, 2024. Illusion of Influence: The AI-Generated Misinformation Apocalypse That Wasn't | Perilous Tech

[23] Tackling Online Disinformation: A European Approach, European Commission, COM/2018/236 final, April 26, 2018. EUR-Lex – 52018DC0236 – EN – EUR-Lex

Print ISSN 0248-2851 • On-line/Linking ISSN 2743-4052 • ISBN 978-630-354-001-6

Everyone anticipated that generative AI would have a significant impact on the 2024 elections. According to the World Economic Forum (2024), Over the next two years, elections—including those in India—were expected to face serious challenges due to misinformation and disinformation, potentially harming democratic processes.[24] Factually, the prediction did not come true, as Prime Minister Narendra Modi's party, which led in election campaigns, mobilized its voter base using violent rhetoric against Muslims. The harm caused stemmed directly from the hateful rhetoric itself, rather than from the AI tools used to spread it.[25]

Numerous facts demonstrate that its influence on elections is minimal. From Meta's report, we read that less than 1% of AI-generated content was fact-checked misinformation during the 2024 elections. It appears that Meta took sufficient measures to minimize the impact of Generative AI on elections.[26] Additionally, they collaborated with other companies within the same industry, such as Amazon, Adobe X, Microsoft, OpenAI, TikTok, and others, culminating in the signing of The Tech Accord to Combat Deceptive Use of AI in 2024 Elections. Participating to mitigate the impact of disinformation, misinformation, and deepfakes on elections.[27] Deepfakes will not have deleterious consequences for democratic procedures, chiefly elections, because there is no evidence of interference in elections. To make such a prediction for the future, deepfakes should already be having some impact on elections. However, there is no evidence that they have.[28] Therefore, for real, the risks of deepfakes and AI-generated misinformation in elections are overestimated. AI-driven misinformation does not significantly change people's minds on highly polarized topics.

3. DEFENDING TRUTH: EU STRATEGIES AGAINST DISINFORMATION

Although deepfakes no longer influence elections, there are areas where they can still be used to manipulate people, for instance, Moldova's Prime Minister, Dorin Recean, described how AI-generated deepfakes, spread by foreign forces, impacted the public. These images showing children in military uniforms standing next to the EU flag, pushing a narrative that EU integration would lead to war. Suggesting that if Moldova seeks EU integration, the same outcome as in Ukraine would follow. Moldova's case is a clear example of how disinformation can manipulate public emotions and fears, shaping perceptions through strategic misinformation.[29] That's why it requires the European Union's attention.

Large-scale disinformation campaigns are a major challenge for Europe and require a coordinated response from EU countries, EU institutions, online platforms, news media and EU citizens as well.[30] In here, distinguishing between ethics, policy and law in the regulatory framework

[24] The Global Risks Report 2024, World Economic Forum, January 10, 2024. WEF_The_Global_Risks_Report_2024.pdf
[25] Russell Brandom, India's Election Wasn't the Deepfake Doomsday Many Feared, Rest of World, May 30, 2024. https://restofworld.org/2024/exporter-india-deepfake-trolls/
[26] Nick Clegg, What We Saw on Our Platforms During 2024's Global Elections, Meta, December 3, 2024. https://about.fb.com/news/2024/12/2024-global-elections-meta-platforms/
[27] Munich Security Conference, AI Elections Accord: A Tech Accord to Combat Deceptive Use of AI in 2024 Elections, Munich Security Conference, February 16, 2024. AI Elections Accord – Munich Security Conference
[28] Dan Cavedon-Taylor, Deepfakes and Democracy: A Catch-22? Journal of the American Philosophical Association, vol. 10, no. 1, March 2024, pp. 24–45. Published online by Cambridge University Press. Deepfakes and Democracy: A Catch-22? | Journal of the American Philosophical Association | Cambridge Core
[29] Jesus Serrano, Protection vs. Censorship: How to Strike the Right Balance Curbing Online Disinformation, World Economic Forum, April 8, 2025. How to strike a balance curbing online disinformation | World Economic Forum
[30] Chochia, Archil; Sicat, Eden Grace Niñalga (2023). Ethics and Modern Technologies: Example of Navigating Children's Rights in an AI-Powered Learning Environment. In: Ramiro Troitiño, D.; Kerikmäe, T.; Hamuľák, O. (Ed.). Digital Development of the European Union. (129−141). Springer, Cham. DOI: 10.1007/978-3-031-27312-4_9.

Print ISSN 0248-2851 • On-line/Linking ISSN 2743-4052 • ISBN 978-630-354-213-3

of AI is important, but so it is to understand the alignment among principles, policies and legal acts.[31] Because it ensures coherence, facilitates innovation while increasing safety and trust. Such considerations are also seen to address societal changes and guides responsive legislative development.[32]

To provide an effective contribution to addressing the issue of disinformation, it is crucial to have the support of a multidisciplinary community including fact-checkers, academic researchers and other relevant stakeholders. The European Digital Media Observatory (EDMO) is a project that supports the independent community working to combat disinformation. The main mission is to combat online disinformation by fostering collaboration between fact-checkers, researchers, media organizations, and policymakers across Europe.[33]

The adoption of the Digital Services Act can be considered a milestone of the European constitutional strategy. In order to face the challenges raised by platform power[34], together with the Digital Markets Act.[35] It includes a range of obligations for platforms.

European Commission has brought together major online platforms, emerging and specialized platforms, players in the advertising industry, fact-checkers, research and civil society organizations to deliver a strengthened Code of Practice on Disinformation,[36] which sets out principles and commitments for online platforms and the advertising sector to counter the spread of disinformation online in the EU, which its signatories agreed to implement. It is the world's first self-regulatory instrument to fight disinformation.[37] The Digital Services Act (DSA) follows a similar narrative. First, the DSA is predicated on the concept that "what is illegal in the real world must also be illegal on platforms." The DSA mandates online platforms to support law enforcement, addressing the rapid pace and complexity of digital interactions that challenge traditional regulatory frameworks.[38] People trust business more than they trust media, with brands playing a fundamental role in the media environment as they directly shape narratives and indirectly support quality content and trustworthy information through their marketing investments.[39]

Beyond existing efforts, combating disinformation requires proactive measures. Local and regional authorities can launch awareness campaigns to promote ethics and tolerance in social media, while governments can fund educational initiatives or collaborate with civil society to address fake news. Larger programs can help citizens distinguish truth from disinformation, ensuring a more

[31] Zengin, Aysegul; Chochia, Archil; Martínez-Ramil, Pablo (2025). Copyright in Text and Data Mining: Legal Challenges and the AIA's Role in the EU. International and Comparative Law Review, 25 (1).

[32] Şaban İbrahim Göksal, Maria Claudia Solarte Vasquez & Archil Chochia, The EU AI Act's Alignment within the European Union's Regulatory Framework on Artificial Intelligence, International and Comparative Law Review, vol. 24, no. 2, December 2024, p 25–53. The-EU-AI-Acts-Alignment-within-the-European-Unions-Regulatory-Framework-on-Artificial-Intelligence (1).pdf

[33] European Digital Media Observatory (EDMO), European Commission, 2025. European Digital Media Observatory (EDMO) | Shaping Europe's digital future

[34] Giovanni De Gregorio, Digital Constitutionalism and Freedom of Expression, Digital Constitutionalism in Europe, Cambridge University Press, May 6, 2022, chapter 5, pp. 157–215. Digital Constitutionalism and Freedom of Expression (Chapter 5) – Digital Constitutionalism in Europe

[35] *Proposal for a Regulation on Contestable and Fair Markets in the Digital Sector (Digital Markets Act)*, *European Commission*, December 15, 2020. EUR-Lex – 52020PC0842 – EN – EUR-Lex

[36] Mona Simion, Knowledge and Disinformation, Episteme, vol. 21, no. 4, December 2024, pp. 1208–1219. Published online by Cambridge University Press, June 29, 2023. Knowledge and Disinformation | Episteme | Cambridge Core

[37] Guidance to Strengthen the Code of Practice on Disinformation, European Commission, May 26, 2021. https://ec.europa.eu/commission/presscorner/api/files/document/print/en/qanda_21_2586/QANDA_21_2586_EN.pdf

[38] Joel P. Trachtman, Platform Responsibility: An International Legal Synthesis, Defeating Disinformation, Cambridge University Press, vol. 14, January 30, 2025. Platform Responsibility (Chapter 14) – Defeating Disinformation

[39] Principles for the Future of Responsible Media in the Era of AI, World Economic Forum, January 2024. WEF_Principles_for_the_Future_of_Responsible_Media_in_the_Era_of_AI_2024.pdf

Print ISSN 0248-2851 • On-line/Linking ISSN 2743-4052 • ISBN 978-630-354-001-6

informed and resilient society.[40] It is noteworthy that Fact-checkers who often called upon to mitigate the spread of deepfakes, play crucial role in today's information environment. They are able to keep up with technological changes in a disinformation environment quickly. [41] Fact-checking as a profession has significantly transformed over the past decades. While journalists traditionally carried out fact-checking as part of their work, today, independent organizations have taken on this role.[42] They are able to keep up with technological changes in a disinformation environment quickly. [43]

4. DEEPFAKES AND THEIR GENUINE INFLUENCE ON SOCIETY

Generative AI presents several real risks. Deepfakes cause harm in two main categories: harassment and social engineering.

4.1. Deepfakes as a Tool for Social Engineering

The first set of risks involves social engineering attacks. Deepfakes can be used to exploit both people and technology. For example, an attacker can use AI to clone a family member's voice and use it to try and scam a relative out of money. This type of scam is on the rise, prompting organizations such as the US Federal Trade Commission to issue a consumer alert.[44]

Scammers used AI-based software to impersonating a UK-based energy firm leader and successfully demanded a fraudulent transfer of $243,000 in 2019. Experts in cybercrime have described this incident as a rare case of AI being leveraged for hacking.[45] In early 2020, a bank manager in Hong Kong was tricked by scammers using AI voice-cloning technology and made large money transfers believing the request was real.[46] scammers used AI voice-cloning technology to trick at least eight elderly people in Canada, leading to a total loss of $200,000.[47]

4.2 Deepfakes and Harassment: A Growing Threat

Deepfakes can be used to target and harass individuals. The most extreme form of harassment involves the creation of non-consensual pornography, which can be used outright or as content for blackmail and sextortion scams.[48] This form of violence against women is referred to as 'image-based sexual abuse' ("IBSA"), "This term refers to all forms of the non-consensual creating,

[40] Hate Speech and Fake News: The Impact on Working Conditions of Local and Regional Elected Representatives, Congress of Local and Regional Authorities of the Council of Europe, 43rd Session, CG (2022)43-11final, October 25, 2022. https://rm.coe.int/0900001680a8340b

[41] Teresa Weikmann & Sophie Lecheler, cutting through the Hype: Understanding the Implications of Deepfakes for the Fact-Checking Actor-Network, Digital Journalism, vol. 12, no. 2, March 2023. Cutting through the Hype: Understanding the Implications of Deepfakes for the Fact-Checking Actor-Ne

[42] Wilson Lowrey, The Emergence and Development of News Fact-Checking Sites: Institutional Logics and Population Ecology, Journalism Studies, vol. 18, no. 3, March 2017, pp. 376–394.

[43] Teresa Weikmann & Sophie Lecheler, cutting through the Hype: Understanding the Implications of Deepfakes for the Fact-Checking Actor-Network, Digital Journalism, vol. 12, no. 10, March 2023. Cutting through the Hype: Understanding the Implications of Deepfakes for the Fact-Checking Actor-Ne

[44] Nathan Hamiel, Deepfakes Proved a Different Threat Than Expected: Here's How to Defend Against Them, World Economic Forum, January 10, 2025. Deepfakes are here to stay and we should remain vigilant | World Economic Forum

[45] Catherine Stupp, Fraudsters Used AI to Mimic CEO's Voice in Unusual Cybercrime Case: Scams Using Artificial Intelligence Are a New Challenge for Companies, The Wall Street Journal. Fraudsters Used AI to Mimic CEO's Voice in Unusual Cybercrime Case – WSJ

[46] Thomas Brewster, Fraudsters Cloned Company Director's Voice in $35 Million Heist, Police Find, Forbes, October 14, 2021. Fraudsters Cloned Company Director's Voice In $35 Million Heist, Police Find

[47] Ryan Cooke, How Scammers Likely Used Artificial Intelligence to Con Newfoundland Seniors Out of $200K, CBC News, March 22, 2023. How scammers likely used artificial intelligence to con Newfoundland seniors out of $200K | CBC News

[48] Extortion, also called "sextortion," refers to the act of using the threat of publishing sexual content (images, videos, deepfakes, sexual rumors) to menace, coerce, or blackmail someone—either for more sexual content, for money, or both. *Source: GREVIO (Group of Experts on Action against Violence against Women and Domestic Violence), General Recommendation No. 1 on the digital dimension of violence against women, Strasbourg, Council of Europe, 2021.*

Print ISSN 0248-2851 • On-line/Linking ISSN 2743-4052 • ISBN 978-630-354-213-3

taking or sharing of intimate images or videos, including altered or manipulated media, and threats to distribute such material".[49] Significantly, IBSA goes beyond distribution, to encompass the non-consensual creation of intimate images or videos, particularly using artificial intelligence to alter material to make it sexual or pornographic.[50] Many Jurisdictions are recognizing deepfake pornographic materials as the new IBSA, intimate material can be quickly and cheaply generated using increasingly common and user-friendly technology.[51] A staggering 98% of deepfake videos circulating online are pornographic in nature.[52]

Image-based sexual abuse can have devastating consequences, profoundly impacting victims' lives. It often results in a drastic psychological and emotional break, where individuals feel an intrusive violation so severe that they mentally separate their life into two distinct phases: before and after the abuse.[53] One of the most pressing dangers online today is the financial sextortion of young people.[54] IBSA violates fundamental rights (In European scholarship, IBSA is considered a serious violation of sexual autonomy)[55] and undermines human dignity, eroding victims' self-worth and respect.[56]

5. EU POLICY AND ACTION ON IMAGE-BASED SEXUAL ABUSE: THE LEGAL DEVELOPMENT PATH

This type of harassment can have devastating consequences, making the adoption of legal measures essential. The EU still has a long road ahead in effectively combating violence against women, though its efforts can be traced back to a European Parliament Resolution in 1986 on violence against women.[57]

After the 2019 European elections, gender equality became a key focus in European politics. This was strongly supported by Ursula von der Leyen, the first female President of the European Commission.[58] Soon after, the Commission introduced the Gender Equality Strategy.[59] As part of its Strategy, the EU finalized its accession to the Council of Europe Convention on Preventing and

[49]Carlotta Rigotti & Clare McGlynn, Towards an EU Criminal Law on Violence Against Women: The Ambitions and Limitations of the Commission's Proposal to Criminalize Image-Based Sexual Abuse, New Journal of European Criminal Law, vol. 13, no. 4, 2022. Towards an EU criminal law on violence against women: The ambitions and limitations of the Commission's proposal to criminalise image-based sexual abuse – Carlotta Rigotti, Clare McGlynn, 2022

[50] Danielle K. Citron & Robert Chesney, Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security, California Law Review, vol. 107, no. 6, December 2019. Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security

[51] Carlotta Rigotti and Clare McGlynn, *Towards an EU Criminal Law on Violence Against Women: The Ambitions and Limitations of the Commission's Proposal to Criminalise Image-Based Sexual Abuse*, *New Journal of European Criminal Law* 13, no. 4, 2022.

[52] 2023 State of Deepfakes: Realities, Threats, and Impact, AI Advisory Boards, April 9, 2024. 2023 State Of Deepfakes: Realities, Threats, And Impact

[53] Clare McGlynn, Kelly Johnson, Erika Rackley, Nicola Henry, Nicola Gavey, Anastasia Powell & Asher Flynn, "It's Torture for the Soul": The Harms of Image-Based Sexual Abuse, Social & Legal Studies, vol. 30, no. 4, 2021, p 541.

[54] The Intervention Journey: A Roadmap to Effective Digital Safety Measures, World Economic Forum, March 2025.WEF_The_Intervention_Journey_A_Roadmap_to_Effective_Digital_Safety_Measures_2025.pdf

[55] Anja Schmidt, The Abuse of Sexual Images between Liberal Criminal Law and the Protection of Sexual Autonomy, Criminalizing Intimate Image Abuse: A Comparative Perspective, Oxford University Press, February 2024.

[56] Clare McGlynn & Erika Rackley, *Image-Based Sexual Abuse*, *Oxford Journal of Legal Studies*, vol. 37, no. 3, Autumn 2017. Image-Based Sexual Abuse | Oxford Journal of Legal Studies | Oxford Academic

[57] Council Conclusions on the Eradication of Violence Against Women in the European Union, Council of the European Union, March 8, 2010. COUNCIL OF

[58] Petra Debusscher, The EU Gender Equality Strategy 2020-2025: The Beginning of a New Season?, European Trade Union Institute, 2023. 05__The EU Gender Equality Strategy 2020-2025 the beginning of a new season.pdf

[59] Gender Equality Strategy 2020-2025, European Commission, 2025. Gender equality strategy – European Commission

Print ISSN 0248-2851 • On-line/Linking ISSN 2743-4052 • ISBN 978-630-354-001-6

Combating Violence Against Women and Domestic Violence, known as the "Istanbul Convention".[60] It should be noted that, several EU countries opposed the ratification of the Istanbul Convention. In 2019, the European Parliament requested a legal opinion from the ECJ on the EU's accession to the Istanbul Convention. In 2021, the ECJ ruled that Articles 78(2), 82(2), 84, and 336 TFEU allow the Council to adopt the Convention by qualified majority, without requiring full Member State agreement.[61] In 2021, the European Parliament urged the EU and its Member States to take measures against cyber violence targeting women, recognizing it as an extension of the sexual violence and abuse they face in offline spaces.[62] While negotiations were on-going regarding the EU's accession to the Istanbul Convention, the European Commission published a proposal addressing specific forms of cyber violence, including non-consensual sharing of manipulated material, cyber-stalking, cyber harassment, and cyber incitement to violence or hatred. The proposal mandated enhanced coordination and cooperation among Member States, acknowledging that the regulation of cyber violence remains fragmented, leaving victims vulnerable and unprotected.[63] In this regard, the Commission stressed that cyber violence is as prevalent and significant as physical forms of gender-based violence.

In May 2024, the European Union adopted landmark legislation, the first binding law, aiming to harmonize legislative and policy responses to violence against women and domestic violence across all its Member States.[64] However, across the EU there is little legislative action in this field, with recent Irish legislation being a notable exception. The comprehensive framework of the Istanbul Convention offers additional measures for prevention, prosecution, and coordination, helping to strengthen protections against digitally manipulated sexual abuse.[65] According to the Group of Experts on Action against Violence against Women and Domestic Violence (GREVIO), this provision applies to image-based sexual abuse (IBSA), including deepfake-generated sexual images.[66]

While not a complete solution, this provision marks an initial step for the EU and its Member States to strengthen protections against IBSA. Following several rounds of inter-institutional negotiations between the European Parliament and the Council, a political agreement was achieved in February 2024. In summary, while the EU institutions agreed to remove rape from the list of crimes included, Other significant proposals were retained including the criminalization of cyberflashing. Generally, in terms of the legal basis, the Commission relies on Article 83.1 TFEU which provides competence to establish minimum rules and define criminal offences "in the areas

[60] Ionel Zamfir, EU Accession to the Istanbul Convention, European Parliamentary Research Service, February 2023. EU accession to the Istanbul Convention

[61] *Opinion of the Court (Grand Chamber) on the Istanbul Convention*, *Court of Justice of the European Union*, October 6, 2021. EUR-Lex – 62019CV0001(02) – EN – EUR-Lex

[62] European Parliament, *Resolution on Combating Gender-Based Violence: Cyberviolence*, P9_TA(2021)0489, 14 December 2021, Texts adopted – Combating gender-based violence: cyberviolence – Tuesday, 14 December 2021

[63] *Proposal for a Directive on Combating Violence Against Women and Domestic Violence*, *European Commission*, March 8, 2022. EUR-Lex – 52022PC0105 – EN – EUR-Lex

[64] Directive (EU) 2024/1385 on Combating Violence Against Women and Domestic Violence, European Parliament & Council of the European Union. Directive – EU – 2024/1385 – EN – EUR-Lex

[65] Carlotta Rigotti and Clare McGlynn, *Towards an EU Criminal Law on Violence Against Women: The Ambitions and Limitations of the Commission's Proposal to Criminalise Image-Based Sexual Abuse*, *New Journal of European Criminal Law* 13, no. 4 (2022), Towards an EU criminal law on violence against women: The ambitions and limitations of the Commission's proposal to criminalise image-based sexual abuse – Carlotta Rigotti, Clare McGlynn, 2022

[66]Group of Experts on Action Against Violence Against Women and Domestic Violence (GREVIO), General Recommendation No. 1 on the Digital Dimension of Violence Against Women, 2024, https://rm.coe.int/grevio-rec-no-on-digital-violence-against-women/1680a49147

Print ISSN 0248-2851 • On-line/Linking ISSN 2743-4052 • ISBN 978-630-354-213-3

of particularly serious crime with a cross-border dimension resulting from the nature or impact of such offences". [67]

Additionally, the final agreement included the Council's amendment stipulating that cyber stalking and harassment, along with the non-consensual sharing of intimate images online, should only be considered criminal offenses across the EU when such actions are likely to cause serious psychological harm or instill fear for the victims' safety. Under Article 5 of Directive (EU) 2024/1385, Member States must ensure that making sexually explicit images, videos, or similar material publicly accessible via information and communication technologies (ICT) without the consent of the person depicted is punishable as a criminal offense—if the act is likely to cause serious harm to the victim.[68] The phrase " cause serious harm to that person" presents a challenge since victims may not immediately perceive the harm they've suffered. If serious harm becomes the primary criterion for punishment, offenders could escape accountability when the damage is not clearly acknowledged or proven. Therefore, I believe that the creation of such material using deepfake technology should be punishable, regardless of whether harm is present or not.

According to Article 5(1)(b) of Directive (EU) 2024/1385, requires Member States to criminalize creation and distribution of AI-generated intimate deepfakes. This provision acknowledges the rapid rise of artificial intelligence and other technologies used to manipulate or alter images to falsely depict individuals as engaging in sexually explicit activities without their consent. It further criminalizes making such manipulated content publicly accessible through information and communication technologies (ICTs). According to Article 5(1)(c) of Directive (EU) 2024/1385, the scope of criminalized conduct is extended to include threats to distribute intimate or manipulated material covered under Article 5(1)(a) and 5(1)(b).[69]

Under Article 5(1)(b) of Directive (EU) 2024/1385, the definition of altered and deepfake imagery is narrow, covering only depictions of active engagement in sexually explicit activities. This excludes nude images, even if they are sexual in nature, leaving a gap in legal protection for victims. Regarding Article 5(1)(c) of Directive (EU) 2024/1385, it criminalizes threats related to the distribution of intimate or manipulated material but only when used to coerce, pressure, or manipulate a person into acting, agreeing to something, or refraining from a certain action. However, the provision does not explicitly cover threats made without coercive intent.[70]

The wording "making accessible to the public" in Article 5 of Directive (EU) 2024/1385 presents a significant legal issue. While deepfake-generated intimate content may not always reach a wide audience, its targeted dissemination—such as within a victim's family, workplace, or social circles—can still cause severe harm, including psychological trauma, reputational damage, and even fatal consequences. One of the most significant gaps in Article 5 of Directive (EU) 2024/1385 is that it only criminalizes the distribution or the threat of distribution of intimate or manipulated material. The mere creation of deepfake sexual images is not explicitly punishable, leaving a major loophole.

Article 7(b) addresses content that has been created or digitally altered to make it seem as though an individual is participating in sexual activities, without their consent. This provision is intended to regulate deepfakes or fake porn. Article 7(b) has a narrow scope, applying only to manipulated material where a person appears to be engaged in sexual activities. This definition could

[67] *Consolidated Versions of the Treaty on European Union and the Treaty on the Functioning of the European Union*, *EUR-Lex*, October 26, 2012. eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:12012E/TXT

[68] European Parliament and Council, *Directive (EU) 2024/1385 on Combating Violence Against Women and Domestic Violence*, 14 May 2024, Directive – EU – 2024/1385 – EN – EUR-Lex

[69] European Union, *Directive (EU) 2024/1385 on Combating Violence Against Women and Domestic Violence*, 14 May 2024, Directive – EU – 2024/1385 – EN – EUR-Lex

[70] Carlotta Rigotti, Clare McGlynn, and Franziska Benning, *Image-Based Sexual Abuse and EU Law: A Critical Analysis*, *German Law Journal* 25, no. 9 (2024)

Print ISSN 0248-2851 • On-line/Linking ISSN 2743-4052 • ISBN 978-630-354-001-6

lead to ambiguity, as it may exclude non-sexual nudity, such as images of people changing clothes or using restrooms, even if the content is considered sexualized. Due to these limitations, the draft risks excluding certain harmful content. A more clear and comprehensive approach would be to criminalize all intimate images, including nude and sexual content, as done in Article 7(a). Article 7(c) rightly criminalizes the intentional threat to distribute intimate images or videos, including deepfakes, recognizing the severe harm such threats inflict. By addressing both manipulated material and deepfake-related threats, this provision signals a strong commitment to tackling abuse and will necessitate legal changes across Member States. It should also cover blackmail where the perpetrator demands money to prevent distribution, a more common form of extortion where adults are the victims[71]

Overall, the regulation of IBSA and, in a broader context, online and technology-facilitated violence against women, was considered a significant milestone with the potential to bolster women's fundamental rights and safety in cyberspace. However, it was widely acknowledged that the current wording of the Directive was narrow in scope and did not adequately reflect the experiences of victim.[72]

5.1. Legal Sanctions for Deepfake-Related Violence

Article 10(4) of the EU Directive on combating violence against women and domestic violence sets a minimum punishment of at least one year of imprisonment for offenses covered in Article 7, which includes crimes related to deepfake abuse. [73] This penalty has been criticized by the EESC which considers that the sanction should be equivalent to the minimum for cyber-stalking which is a maximum of two years' imprisonment. Determining the severity of violence against women offenses is challenging, as victims' experiences can differ significantly. Simply raising criminal penalties may not be the most effective solution. Instead, criminalization plays a crucial role in recognizing image-based sexual abuse as a serious crime, but its primary goal should be to shift societal attitudes and prevent abuse, rather than solely increasing prison sentences.[74]

6. LEGAL SOLUTIONS: SUING THE CREATORS OF DEEP FAKES OR SUING THE PLATFORMS

First and foremost, it should be noted that victims often avoid filing lawsuits because the 'Streisand Effect' influences on their decision. The Streisand Effect tends to loom over the choice to sue, especially when the deepfake is humiliating or damages their reputation. Legal action inevitably draws public attention, and unless the victim is allowed to file under a pseudonym, pursuing a lawsuit may intensify the harm they experience.[75]

---

[71] Carlotta Rigotti and Clare McGlynn, *Towards an EU Criminal Law on Violence Against Women: The Ambitions and Limitations of the Commission's Proposal to Criminalise Image-Based Sexual Abuse*, *New Journal of European Criminal Law* 13, no. 4, 2022, Towards an EU criminal law on violence against women: The ambitions and limitations of the Commission's proposal to criminalise image-based sexual abuse – Carlotta Rigotti, Clare McGlynn, 2022

[72] Carlotta Rigotti, Clare McGlynn, and Franziska Benning, *Image-Based Sexual Abuse and EU Law: A Critical Analysis*, *German Law Journal* 25, no. 9, 2024, Image-Based Sexual Abuse and EU Law: A Critical Analysis | German Law Journal | Cambridge Core

[73] European Parliament and Council, Directive (EU) 2024/1385 on Combating Violence Against Women and Domestic Violence, 14 May 2024. Directive – EU – 2024/1385 – EN – EUR-Lex

[74] Carlotta Rigotti and Clare McGlynn, *Towards an EU Criminal Law on Violence Against Women: The Ambitions and Limitations of the Commission's Proposal to Criminalise Image-Based Sexual Abuse*, *New Journal of European Criminal Law* 13, no. 4, 2022. Towards an EU criminal law on violence against women: The ambitions and limitations of the Commission's proposal to criminalise image-based sexual abuse – Carlotta Rigotti, Clare McGlynn, 2022

[75] Mike Masnick, *Since When Is It Illegal to Just Mention a Trademark Online?*, *Techdirt*, January 5, 2005. Since When Is It Illegal To Just Mention A Trademark Online? | Techdirt.

Print ISSN 0248-2851 • On-line/Linking ISSN 2743-4052 • ISBN 978-630-354-213-3

If a victim decides to sue the creator of a deep fake, multiple legal frameworks will come into play, including intellectual property and tort law.[76] If the original photo was taken by the victim, she will hold the copyright to that image[77]. The most severe cases of victims being targeted by fake news can fall under defamation law. However, suing every internet troll is wildly impractical. In most cases, online platforms such as Facebook or Twitter are the most viable defendants in a defamation lawsuit related to fake news.[78] Victims may also sue in tort for intentional infliction of emotional distress. This requires proof of "extreme and outrageous conduct."[79] It is often stated that conduct qualifies for the tort only if it is "so outrageous in character, and so extreme in degree, as to go beyond all possible bounds of decency, and to be regarded as atrocious, and utterly intolerable in a civilized community."[80]

It is challenging to hold individual deepfake creators accountable, so creators are not the only parties who might bear responsibility. Since content platform play a key role in enabling the distribution of deep fakes, and harmful deepfake creators can be difficult to track down, the most efficient and effective way to mitigate harm may be to impose liability on the platforms themselves. Online platforms are largely protected from civil liability for user-generated content under the Communications Decency Act (CDA), Section 230. This law grants immunity, preventing platforms from being treated as publishers or speakers of harmful content posted by users. Even if platforms encourage certain content, courts have interpreted Section 230 broadly, allowing them to avoid liability.[81]

In 2018, lawmakers modified the statute by enacting the Allow States and Victims to Fight Online Sex Trafficking Act ("FOSTA") to address websites' facilitation of sex trafficking. The idea is that online platforms should face the same liability for enabling illegal sex-trafficking, as offline outlets do.

7. REGULATING AI-GENERATED PORNOGRAPHIC DEEPFAKES UNDER THE DIGITAL SERVICES ACT AND AI ACT

In 2022, the EU introduced the Digital Services Act (DSA) as a fundamental element of its digital policy, aiming to protect fundamental rights while fostering a secure online environment.[82] DSA regulates online platforms, imposing size-based obligations. Under Article 2(h), "illegal content" is defined as any information that does not comply with EU or national law. Recital 12 specifically highlights the non-consensual sharing of private images, aligning with existing national legislation on certain forms of image-based sexual abuse (IBSA). This legal connection is expected to be reinforced by Article 5 of the Directive.

[76] Robert Chesney & Danielle Keats Citron, Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security, California Law Review, vol. 107, no. 6, 2019, pp. 1753–1767.
[77] Megan Farokhmanesh, Is It Legal to Swap Someone's Face into Porn Without Consent?, The Verge, January 30, 2018. Is it legal to swap someone's face into porn without consent? | The Verge
[78] Eric Goldman, Combatting Deepfakes Through the Right of Publicity, Lawfare, March 28, 2019. Perma | Combatting Deep Fakes through the Right of Publicity – Lawfare
[79] Robert Chesney & Danielle Keats Citron, Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security, California Law Review, vol. 107, no. 6, 2019, pp. 1753–1819.
[80] Benjamin C. Zipursky, Snyder v. Phelps, Outrageousness, and the Open Texture of Tort Law, DePaul Law Review, vol. 60, no. 2, 2011, p 473
[81] Robert Chesney & Danielle Keats Citron, Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security, California Law Review, vol. 107, no. 6, 2019, pp. 1753–1819.
[82] Aina Turillazzi, Federico Casolari, Mariarosaria Taddeo, and Luciano Floridi, *The Digital Services Act: An Analysis of Its Ethical, Legal, and Social Implications*, SSRN, 12 January 2022.

Print ISSN 0248-2851 • On-line/Linking ISSN 2743-4052 • ISBN 978-630-354-001-6

Additionally, Recital 87 refers to "illegal pornographic content", though its definition remains unclear "due to variations in pornography laws across Member States and the lack of EU competence in this area", the recital provides an explicit example: "content representing non-consensual sharing of intimate or manipulated material."[83] In Section 5, additional obligations are imposed for Very Large Online Platforms (VLOPs) and two as Very Large Online Search Engines (VLOSEs).[84] Platforms such as Facebook, Instagram, TikTok, fall into this category, though messaging apps like WhatsApp and Telegram are currently excluded, despite calls from various stakeholders for their inclusion. Under Article 34, VLOPs and VLOSEs must assess risks related to illegal content and gender-based violence. Article 35 requires prompt removal of non-consensual material, distinguishing deepfakes from authentic images, and collaborating with other platforms to prevent content from spreading further. While the DSA applies to all online platforms serving EU users, enforcement is decentralized, relying on Member States, which may lead to inconsistencies in addressing IBSA. The effectiveness of this framework depends on regulators' engagement and their willingness to hold platforms accountable.[85]

In 2021, the European Commission published its proposal for the AI Act. Three years later, the European Parliament and the Council approved the final text, but it will take additional years before this legislation becomes enforceable at the national level.[86] Additionally, the AI Act applies to developers, distributors, and users of AI systems within the EU market, ensuring accountability for AI-generated content. The European Commission is expected to adopt implementation measures for labeling and detection of deepfake materials, but enforcement remains decentralized, relying on Member States.[87]

8. THE INTERSECTION OF DIRECTIVE (EU) 2024/1385 AND FREEDOM OF EXPRESSION

The Council included a concluding paragraph in Article 5 of Directive (EU) 2024/1385, clarifying that its criminalization mandate must be applied in full respect of the rights, freedoms, and principles established in Article 6 TEU.[88] It is unclear what has prompted this provision, as EU law is already supposed to be fully harmonized, and this principle is taken into account when adopting directives. It also underscores that this criminalization is implemented without prejudice to fundamental principles related to freedom of expression. That's why I will dedicate a separate section to freedom of expression.

[83] Carlotta Rigotti, Clare McGlynn, and Franziska Benning, *Image-Based Sexual Abuse and EU Law: A Critical Analysis*, *German Law Journal* 25, no. 9 (2024), Image-Based Sexual Abuse and EU Law: A Critical Analysis | German Law Journal | Cambridge Core

[84] Access Now, *VLOPs or Flops: Is Big Tech Dodging Accountability in the EU?*, 4 May 2023, VLOPs or flops: Is Big Tech dodging accountability in the EU? – Access Now

[85] Carlotta Rigotti, Clare McGlynn, and Franziska Benning, *Image-Based Sexual Abuse and EU Law: A Critical Analysis*, *German Law Journal* 25, no. 9 (2024), Image-Based Sexual Abuse and EU Law: A Critical Analysis | German Law Journal | Cambridge Core

[86] European Union, *Artificial Intelligence Act*, 2024.

[87] Carlotta Rigotti, Clare McGlynn, and Franziska Benning, *Image-Based Sexual Abuse and EU Law: A Critical Analysis*, *German Law Journal* 25, no. 9 (2024), Image-Based Sexual Abuse and EU Law: A Critical Analysis | German Law Journal | Cambridge Core

[88] European Union, *Directive (EU) 2024/1385 on Combating Violence Against Women and Domestic Violence*, 14 May 2024, Directive – EU – 2024/1385 – EN – EUR-Lex

Print ISSN 0248-2851 • On-line/Linking ISSN 2743-4052 • ISBN 978-630-354-213-3

Freedom of expression is one of the cornerstones of democratic society.[89] Digital technologies fundamentally change our understanding of what liberty of speech is. the digital age subtly alters our understanding of liberty of expression.[90]

Would suppressing deepfakes violate freedom of expression norms? Two simple answers are rejected: that deepfakes do not deserve protection under freedom of expression legislation because they are fake by definition; and that deepfakes can be targeted if but only if they are misleadingly presented as authentic.[91]

Based on Article 50.4 of AI act, the only exception is for deepfakes created as part of creative, satirical, artistic, or fictional work, thereby falling under freedom of expression. However, this exemption does not address the harm caused by sexually explicit deepfakes, where labeling alone does not mitigate violations of autonomy and sexual integrity.

As Koltay details, the EU is also developing multiple other measures. For present purposes, this can be illustrated by the Strengthened Code of Practice on Disinformation, the Digital Services Act (DSA). The measures together reflect positive and negative dimensions of democratic free speech, and they do things consistent with the positive dimensions of free expression in a way that is harder to imagine in the USA. The measures include indirect horizontal requirements on private actors, such as DSA obligations on platforms. The DSA requires very large platforms to assess and mitigate systemic risks while protecting fundamental rights, but enforcement challenges remain. Balancing risk regulation and free expression is complex, especially for US-based platforms with narrower rights frameworks.

9. CONCLUSION

Ultimately, it is clear that deepfakes represent one of the most serious challenges in today's digital reality. They serve as a powerful tool for spreading disinformation, which, according to the World Economic Forum, ranks among the top 10 threats facing humanity. The European Union's regulatory response is crucial in addressing this issue, ensuring accountability and transparency in AI-generated content.

Furthermore, collective action is essential in tackling disinformation effectively. A key demonstration of this commitment is the European Commission's initiative, which has brought together major online platforms, emerging and specialized digital services, advertising industry players, fact-checkers, researchers, and civil society organizations to establish a strengthened Code of Practice on Disinformation. This collaborative approach aims to enhance online integrity, content moderation, and platform responsibility in preventing the spread of misleading AI-generated narratives.

Image-based sexual abuse (IBSA) involves the non-consensual creation, distribution, or threats related to intimate materials. It is a serious gendered harm that has become increasingly common, especially with advancements in digital technology.

The European Union (EU) has recognized the urgency of addressing gender-based violence, including its online dimensions. A recent directive specifically targets IBSA by criminalizing the unauthorized sharing of intimate or manipulated content and cyberflashing. Additionally, the Digital

[89] Giovanni De Gregorio, Digital Constitutionalism and Freedom of Expression, Digital Constitutionalism in Europe: Reframing Rights and Powers in the Algorithmic Society, Cambridge University Press, 2022, Digital Constitutionalism and Freedom of Expression (Chapter 5) – Digital Constitutionalism in Europe

[90] Jack M. Balkin, Digital Speech and Democratic Culture: A Theory of Freedom of Expression for the Information Society, New York University Law Review 79, no. 1 2004, Digital Speech and Democratic Culture: a Theory of Freedom of Expression for the Information Society by Jack M. Balkin:: SSRN

[91] Alex Barber, *Freedom of Expression Meets Deepfakes*, *Synthese* 202, no. 40, 2023, Freedom of expression meets deepfakes | Synthese

Print ISSN 0248-2851 • On-line/Linking ISSN 2743-4052 • ISBN 978-630-354-001-6

Services Act (DSA) and AI Act strengthen regulations by holding online platforms, search engines, and AI developers accountable, ensuring transparency and safeguards in digital environments.

Despite certain **gaps in the Istanbul Convention and other legislative acts**, the **European Union's efforts** in addressing gender-based violence remain commendable. The EU has **taken significant steps** to regulate **image-based sexual abuse (IBSA) and online harms**, but there are **still areas that require improvement and refinement**.

Hopefully, **future revisions** will **fill existing gaps** and **enhance protections**, ensuring a **stronger and more comprehensive framework** against digital and gender-based violence. Continuous **legal evolution** and **multilateral cooperation** are essential in **strengthening these regulations**.

BIBLIOGRAPHY

1. AMMANATH, Beena. "In a World of Deepfakes, We Must Build a Case for Trustworthy Synthetic AI Content." *World Economic Forum*, May 30, 2024.
2. BALKIN, Jack M., "Digital Speech and Democratic Culture: A Theory of Freedom of Expression for the Information Society," *New York University Law Review* 79, no. 1 2004
3. BARBER, Alex, "Freedom of Expression Meets Deepfakes," *Synthese* 202, no. 40, 2023).
4. BREWSTER, Thomas, "Fraudsters Cloned Company Director's Voice in $35 Million Heist, Police Find," *Forbes*, October 14, 2021.
5. BUEERMANN, Gretchen; DOBRYGOWSKI, Daniel. "Elections, Cybersecurity, and AI: The Growing Threat of Deepfakes and Social Engineering." *World Economic Forum*, November 8, 2023.
6. CAVEDON-TAYLOR, Dan. "Deepfakes and Democracy: A Catch-22?" *Journal of the American Philosophical Association*, Cambridge University Press, April 23, 2025.
7. CHESNEY, Robert; CITRON, Danielle Keats, "Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security," *California Law Review* 107, no. 6 (2019), pp. 1753–1819.
8. CHOCHIA, A.; NÄSSI, T. (2021). Ethics and Emerging Technologies – Facial Recognition. IDP Revista de Internet Derecho y Política, 34, 1−12. DOI: 10.7238/idp.v0i34.387466.
9. CHOCHIA, Archil; SICAT, Eden Grace Niñalga (2023). Ethics and Modern Technologies: Example of Navigating Children's Rights in an AI-Powered Learning Environment. In: Ramiro Troitiño, D.; Kerikmäe, T.; Hamuľák, O. (Ed.). Digital Development of the European Union. (129−141). Springer, Cham. DOI: 10.1007/978-3-031-27312-4_9.
10. COOKE, Ryan, "How Scammers Likely Used Artificial Intelligence to Con Newfoundland Seniors Out of $200K," *CBC News*, March 22, 2023.
11. COUNCIL OF THE EUROPEAN UNION, *Council Conclusions on the Eradication of Violence Against Women in the European Union*, March 8, 2010.
12. COUNCIL OF EUROPE, *General Recommendation No. 1 on the Digital Dimension of Violence Against Women*, GREVIO, Strasbourg, 2021.
13. COURT OF JUSTICE OF THE EUROPEAN UNION, *Opinion of the Court (Grand Chamber) on the Istanbul Convention*, October 6, 2021.
14. DE GREGORIO, Giovanni, "Digital Constitutionalism and Freedom of Expression," in *Digital Constitutionalism in Europe: Reframing Rights and Powers in the Algorithmic Society*, Cambridge University Press, 2022.
15. DOBBER, Tom; METOUI, Nadia; TRILLING, Damian; HELBERGER, Natali; DE VREESE, Claes. "Do (Microtargeted) Deepfakes Have Real Effects on Political Attitudes?" *The International Journal of Press/Politics*, vol. 26, no. 1, 2021, pp. 69–91.
16. EUROPEAN COMMISSION, *Proposal for a Directive on Combating Violence Against Women and Domestic Violence*, 8 March 2022.
17. EUROPEAN COMMISSION, *Gender Equality Strategy 2020-2025*, 2025.
18. European Commission. *Proposal for a Regulation on Contestable and Fair Markets in the Digital Sector (Digital Markets Act)*, December 15, 2020.

Print ISSN 0248-2851 • On-line/Linking ISSN 2743-4052 • ISBN 978-630-354-213-3

19. European Commission. *Tackling Online Disinformation: A European Approach*. COM/2018/236 final, April 26, 2018.
20. European Digital Media Observatory (EDMO), *European Commission*, 2025.
21. EUROPEAN PARLIAMENTARY RESEARCH SERVICE, *EU Accession to the Istanbul Convention*, February 2023.
22. EUROPEAN PARLIAMENT AND COUNCIL, *Directive (EU) 2024/1385 on Combating Violence Against Women and Domestic Violence*, 14 May 2024.
23. EUROPEAN PARLIAMENT, *Resolution on Combating Gender-Based Violence: Cyberviolence*, P9_TA(2021)0489, December 14, 2021.
24. EUROPEAN PARLIAMENT AND COUNCIL, *Directive (EU) 2024/1385 on Combating Violence Against Women and Domestic Violence*, 14 May 2024.
25. EUROPEAN TRADE UNION INSTITUTE, *The EU Gender Equality Strategy 2020-2025: The Beginning of a New Season?*, 2023.
26. EUROPEAN UNION, *Artificial Intelligence Act*, 2024.
27. EUROPEAN UNION, *Consolidated Versions of the Treaty on European Union and the Treaty on the Functioning of the European Union*, 26 October 2012.
28. FAROKHMANESH, Megan, "Is It Legal to Swap Someone's Face into Porn Without Consent?," *The Verge*, January 30, 2018.
29. GOLDMAN, Eric, "Combatting Deepfakes Through the Right of Publicity," *Lawfare*, March 28, 2019.
30. GÖKSAL, Şaban İbrahim; SOLARTE VASQUEZ, Maria Claudia; CHOCHIA, Archil, "The EU AI Act's Alignment within the European Union's Regulatory Framework on Artificial Intelligence," *International and Comparative Law Review* 24, no. 2 (December 2024), pp. 25–53.
31. GRABE, Maria Elizabeth; BUCY, Erik. *Image Bite Politics: News and the Visual Framing of Elections*. Oxford University Press, 2009.
32. GRABER, D. A. "Seeing is Remembering: How Visuals Contribute to Learning from Television News." *Journal of Communication*, vol. 40, no. 3, 1990, pp. 134–156.
33. Gregory, Sam. "Heard About Deepfakes? Don't Panic. Prepare." *World Economic Forum*, November 23, 2018.
34. Group of experts on action against women and domestic violence, Grevio, *General Recommendation No. 1 on the Digital Dimension of Violence Against Women*, 2024.
35. HAMIEL, Nathan. "Deepfakes Are Here to Stay and We Should Remain Vigilant." *World Economic Forum*, January 10, 2025.
36. MASNICK, Mike, "Since When Is It Illegal to Just Mention a Trademark Online?," *Techdirt*, January 5, 2005.
37. MCGLYNN, Clare, and RACKLEY, Erika, "Image-Based Sexual Abuse," *Oxford Journal of Legal Studies* 37, no. 3 (Autumn 2017).
38. MCGLYNN, Clare; JOHNSON, Kelly; RACKLEY, Erika; HENRY, Nicola; GAVEY, Nicola; POWELL, Anastasia; FLYNN, Asher, "It's Torture for the Soul": The Harms of Image-Based Sexual Abuse," *Social & Legal Studies* 30, no. 4 (2021), p. 541
39. Munich Security Conference. *AI Elections Accord: A Tech Accord to Combat Deceptive Use of AI in 2024 Elections*, February 16, 2024.
40. PRIOR, Markus. "Visual Political Knowledge: A Different Road to Competence?" *The Journal of Politics*.
41. RIGOTTI, Carlotta; MCGLYNN, Clare; BENNING, Franziska, "Image-Based Sexual Abuse and EU Law: A Critical Analysis," *German Law Journal* 25, no. 9 (2024).
42. RIGOTTI, Carlotta, and MCGLYNN, Clare, "Towards an EU Criminal Law on Violence Against Women: The Ambitions and Limitations of the Commission's Proposal to Criminalise Image-Based Sexual Abuse," *New Journal of European Criminal Law* 13, no. 4 (2022)
43. Security Hero. *2023 State Of Deepfakes: Realities, Threats, And Impact*, 2023.
44. SERRANO, Jesus. "Protection vs. Censorship: How to Strike the Right Balance Curbing Online Disinformation." *World Economic Forum*, April 8, 2025.
45. MONA, Simion. "Knowledge and Disinformation," *Episteme* 21, no. 4 (December 2024), pp. 1208–1219.

Print ISSN 0248-2851 • On-line/Linking ISSN 2743-4052 • ISBN 978-630-354-001-6

46. SCHMIDT, Anja, "The Abuse of Sexual Images between Liberal Criminal Law and the Protection of Sexual Autonomy," in *Criminalizing Intimate Image Abuse: A Comparative Perspective*, Oxford University Press, February 2024.
47. STUPP, Catherine, "Fraudsters Used AI to Mimic CEO's Voice in Unusual Cybercrime Case: Scams Using Artificial Intelligence Are a New Challenge for Companies," *The Wall Street Journal*.
48. Turillazzi, Aina, Casolari, Federico, Taddeo, Mariarosaria, and Floridi, Luciano, "The Digital Services Act: An Analysis of Its Ethical, Legal, and Social Implications," *SSRN*, 12 January 2022.
49. UNESCO. *Exposing Technology-Facilitated Gender-Based Violence in an Era of Generative AI*. World Trends in Freedom of Expression and Media Development Series, 2nd edition. Paris: UNESCO, 2023.
50. VACCARI, Cristian, and Andrew CHADWICK. "Deepfakes and Disinformation: Exploring the Impact of Synthetic Political Video on Deception, Uncertainty, and Trust in News." *Social Media + Society*, Sage Publications, February 19, 2020.
51. WEIKMANN, Teresa; GREBER, Hannah; NIKOLAOU, Alina. "After Deception: How Falling for a Deepfake Affects the Way We See, Hear, and Experience Media." *Europeana*, University of Vienna, 2024.
52. World Economic Forum. *The Global Risks Report 2024*. January 10, 2024.
53. World Economic Forum. *The Intervention Journey: A Roadmap to Effective Digital Safety Measures*, March 2025. AI ADVISORY BOARDS, *2023 State of Deepfakes: Realities, Threats, and Impact*, April 9, 2024.
54. ZIPURSKY, Benjamin C., "Snyder v. Phelps, Outrageousness, and the Open Texture of Tort Law," *DePaul Law Review* 60, no. 2 (2011), p. 473.
55. ZENGIN, Aysegul; CHOCHIA, Archil; MARTÍNEZ-RAMIL, Pablo (2025). Copyright in Text and Data Mining: Legal Challenges and the AIA's Role in the EU. International and Comparative Law Review, 25 (1).

Print ISSN 0248-2851 • On-line/Linking ISSN 2743-4052 • ISBN 978-630-354-213-3