International Conference on Current Research Trends
ICCRT 2025 - Hybrid (Hafizabad, Pakistan & Online)
20-21 December 2025

AI-Generated Content and Violation of Human Rights: United Nations Initiatives and International Community Response

Jalil Ahmad [1*]

[1]Human Rights Institute, Southwest University of Political Science & Law, China

* Corresponding Email: jalilkhan4400@gmail.com

Published Online: 24 January 2026 by JCPP *DOI: https://doi.org/10.64060/ICPP.11*

ABSTRACT

While technologies offer significant social and economic benefits, they have also new forms of harm, including non-consensual deepfakes, synthetic child sexual abuse material, and large-scale disinformation. These practices pose direct threats to internationally protected human rights, particularly the rights to privacy, dignity, equality, freedom of expression, and participation in public life. The article demonstrates that current international and domestic responses struggle to address challenges related to attribution of responsibility, enforcement, and victim protection in cases involving synthetic media. It concludes that a more targeted, human-rights-based approach to AI-generated content is required, one that moves beyond general ethical principles and provides clear legal guidance, shared responsibility among key actors, and meaningful remedies for affected individuals. This article argues that although the United Nations has increasingly acknowledged the human rights risks associated with artificial intelligence, its existing governance frameworks remain largely abstract and insufficiently tailored to the specific harms caused by AI-generated content. The absence of precise legal standards and enforceable obligations has contributed to fragmented national responses, inconsistent platform accountability, and limited access to effective remedies for victims. Using a doctrinal and policy-analysis approach, the study examines relevant UN initiatives alongside comparative regulatory practices in the United States, European Union, United Kingdom, China, Russia, and Pakistan.

Keywords: AI-Generated Content; Violation of Human Rights; International Community; Human Rights

1. Introduction

AI systems that generate text, images, audio, and video have moved rapidly from experimental tools to influential actors within global information ecosystems. Generative AI now affects how people communicate, form opinions, engage in political processes, and construct social reality[1]. Alongside legitimate uses in creativity, education, and innovation, these technologies have made it significantly easier to fabricate realistic but false content, impersonate individuals, and scale abuse at unprecedented speed[2]. Practices such as non-consensual sexual deepfakes, synthetic child sexual abuse material, and automated disinformation campaigns have transformed the nature of digital harm[3]. These practices directly interfere with internationally protected rights, including privacy, dignity, equality, freedom of expression, and participation in public affairs[4]. Unlike earlier forms of online manipulation, AI-generated content often lacks clear human authorship, complicating traditional legal concepts of intent, responsibility, and liability. At the international level, the United Nations and its agencies have increasingly addressed the human rights implications of artificial intelligence through resolutions, thematic reports, and ethical guidelines[5][6]. Instruments developed by the Office of the High Commissioner for Human Rights and UNESCO reflect a growing consensus that AI systems must operate in ways that respect human dignity and fundamental freedoms[7]. However, these initiatives tend to treat AI as a broad category, offering high-level principles while giving limited attention to the specific and recurring harms produced by generative content.

This article argues that this gap has concrete consequences. The absence of clear, enforceable guidance on AI-generated content at the UN level has contributed to fragmented national responses, uneven

platform practices, weak accountability, and limited access to remedies for victims. By treating AI-generated content as a distinct human rights challenge rather than a secondary issue within general AI governance, this study aims to clarify the nature of the harm, assess the adequacy of existing international frameworks, and identify more effective regulatory approaches grounded in human rights law.

2. AI-Generated Content as a Source of Human Rights Violations

AI-generated content constitutes a direct and practical source of human rights harm rather than a speculative future risk[8]. Generative tools enable the rapid and inexpensive production of highly realistic images, videos, and audio, often without the knowledge or consent of the individuals depicted[9]. This fundamentally alters how violations occur: harm can be inflicted remotely, anonymously, and at scale, with limited traceability and accountability[10]. One of the most severe forms of harm arises in the context of sexual and gender-based abuse. Non-consensual intimate deepfakes disproportionately target women and girls, reinforcing existing patterns of gender inequality. Faces sourced from social media or public records are digitally inserted into explicit material or used to generate entirely fabricated sexual images[11]. Although the content is synthetic, victims experience real consequences, including reputational damage, loss of employment, psychological distress, and social exclusion. These practices violate the rights to privacy and dignity and undermine the principle of non-discrimination under international human rights law[12]. Children face particularly acute risks. Generative AI is increasingly used to produce artificial child sexual abuse material by manipulating real images or generating child-like representations. While such material may not involve direct physical contact with a child, it normalizes sexual exploitation, creates demand for abuse-related content, and may retraumatize identifiable victims whose likenesses are used[13]. These practices undermine children's rights to protection from sexual exploitation and challenge legal frameworks that were designed around the identification of direct perpetrators and victims.

AI-generated content also affects freedom of expression and participation in public life. Deepfake videos and fabricated audio recordings are used to intimidate journalists, discredit political opponents, and target minorities[14]. Rather than facilitating open debate, these practices create hostile environments that silence individuals through fear and harassment. The resulting chilling effect is particularly pronounced for women journalists, human rights defenders, and activists, many of whom withdraw from public discourse to protect their safety[15]. Beyond individual harms, widespread synthetic media erodes the social conditions necessary for the enjoyment of human rights. When realistic images and audio can no longer be trusted, genuine documentation of abuse or wrongdoing may be dismissed as fake, while false content circulates as truth. This collapse of shared factual reference points weakens accountability mechanisms and undermines the right to information. Human rights protection depends not only on formal legal rules but also on public trust in evidence and institutions; AI-generated content increasingly disrupts both. A central concern across these harms is the lack of effective remedies. Victims often face uncertainty about who bears responsibility: the individual user, the platform, or the AI developer. Many legal systems have not yet adapted to address this diffusion of responsibility, leaving victims with slow takedown processes, limited access to compensation, and weak institutional support. This gap between harm and remedy exposes structural weaknesses in existing human rights protection frameworks when confronted with generative technologies[16].

3. Methodology

This study adopts a qualitative doctrinal and policy-analysis approach to examine how AI-generated content gives rise to human rights violations and how these risks are addressed within United Nations frameworks and selected national legal systems. Given that the central issues concern legal responsibility, regulatory design, and access to remedies, doctrinal analysis provides an appropriate method for evaluating the adequacy of existing norms. Primary sources include international human

rights treaties, resolutions and reports of the UN General Assembly and Human Rights Council, thematic reports of UN Special Rapporteurs, and policy instruments issued by bodies such as the Office of the High Commissioner for Human Rights and UNESCO. To assess how international guidance is translated into practice, the study examines legislation and official policy documents from the United States, European Union, United Kingdom, China, Russia, and Pakistan. Secondary sources consist of peer-reviewed academic literature and reports by civil society organizations documenting harms linked to AI-generated content, including non-consensual deepfakes, child sexual abuse material, and coordinated disinformation campaigns. The analysis proceeds in three stages: mapping forms of AI-generated content against affected rights; evaluating how UN instruments conceptualize these harms and allocate responsibility; and comparing national responses to identify areas of convergence, divergence, and regulatory gaps. The study does not include empirical interviews or technical testing. Its contribution lies in clarifying normative and institutional shortcomings and identifying where more precise human-rights-based responses to AI-generated content are required.

4. United Nations Initiatives on AI and Human Rights: Progress and Limitations

In recent years, the United Nations has increasingly recognized the human rights risks associated with artificial intelligence. Various UN bodies, including the Office of the High Commissioner for Human Rights (OHCHR), UNESCO, and the Human Rights Council, have produced reports, resolutions, and policy instruments aimed at promoting a human-rights-based approach to AI governance. These initiatives represent an important step in acknowledging that AI systems can interfere with fundamental rights and must therefore be subject to international human rights standards. The OHCHR has repeatedly warned that AI technologies may undermine rights such as privacy, equality, freedom of expression, and non-discrimination if deployed without adequate safeguards. Similarly, UNESCO's Recommendation on the Ethics of Artificial Intelligence emphasises principles such as human dignity, transparency, accountability, and fairness. These documents reflect a growing consensus that AI governance should not be driven solely by innovation and efficiency but must also prioritize the protection of human rights. However, despite this normative progress, UN initiatives remain largely high-level and abstract when addressing AI-generated content. Most instruments discuss AI risks in general terms and do not sufficiently distinguish between different forms of AI use. As a result, specific and recurring harms caused by generative technologies—such as non-consensual deepfakes, synthetic child sexual abuse material, and AI-driven disinformation—receive limited direct attention. This lack of specificity makes it difficult for states and platforms to translate international guidance into concrete legal obligations.

A key limitation of UN frameworks lies in their reliance on non-binding, ethics-based language. While ethical principles are valuable, they do not create enforceable duties or clear accountability mechanisms. Victims of AI-generated harm are rarely mentioned as rights-holders entitled to effective remedies. UN instruments provide little guidance on who should bear responsibility when harm results from the interaction of users, platforms, and AI developers, leaving significant gaps in protection and enforcement. Moreover, UN initiatives offer limited direction on access to justice and remedies. Issues such as rapid takedown of harmful synthetic content, compensation for victims, and cross-border enforcement are largely absent from existing guidance. This omission is particularly problematic given the transnational nature of AI-generated content and the speed at which harm can spread. Without clearer standards on remedies and due diligence obligations, victims are often left dependent on platform discretion rather than enforceable legal rights.

5. Community Norms and AI-Generated Content

Community norms play a central role in shaping how rights are respected or violated in everyday life. They define informal expectations about consent, privacy, truth, and acceptable behavior, both online and offline. AI-generated content places significant strain on these norms by enabling practices that were previously difficult, costly, or socially discouraged[17]. When realistic synthetic images, videos,

and audio can be produced and shared with minimal effort, behaviors that would once have been seen as clearly abusive or deceptive risk becoming normalized, even in the absence of explicit legal approval. One of the most visible shifts concerns norms around consent and personal dignity. In many societies, there is a widely shared understanding that using another person's image for sexual, humiliating, or deceptive purposes without permission is unacceptable. AI-generated content disrupts this understanding by introducing the claim that harm is reduced when the material is "not real." This framing weakens the social meaning of consent and allows perpetrators to distance themselves from responsibility. As synthetic intimate images and deepfake memes circulate more widely, victims are often expected to tolerate or ignore the abuse, while objections are dismissed as overreactions(Flynn et al., 2025). Over time, this erodes respect for bodily autonomy and personal dignity, particularly for women and girls.

AI-generated content also challenges norms related to truth and trust. Visual and audio media have traditionally carried strong evidentiary weight in public discourse. The spread of convincing synthetic media undermines this assumption, making it harder for communities to agree on what is real. False content can be shared as authentic, while genuine documentation of abuse or wrongdoing can be rejected as fabricated[19][20]. This dynamic weakens the social foundations of accountability and public debate, creating an environment in which deception becomes easier and denial more plausible. The resulting uncertainty does not affect all actors equally; those with power and resources are better able to contest evidence, while marginalized individuals face higher barriers to being believed. Norms governing public participation and speech are also affected. AI-generated harassment, impersonation, and coordinated disinformation campaigns are increasingly used to intimidate journalists, activists, and political opponents. When such practices become common, they reshape expectations about who can safely speak in public and at what cost. Many individuals, particularly women, minorities, and human rights defenders, respond by withdrawing from online spaces or limiting their participation to avoid further harm[21][22]. This informal exclusion does not rely on formal censorship, yet it has a comparable effect on freedom of expression and democratic engagement.

Children and young people are especially vulnerable to these normative shifts. For many, digital platforms are primary spaces for social interaction, learning, and identity formation. Regular exposure to AI-generated sexualized imagery, humiliation, or dehumanizing content can alter perceptions of consent, privacy, and respect. When synthetic manipulation of images is treated as entertainment or humor, young users may internalize the idea that other people's bodies and identities are freely available for modification and circulation. These early normative lessons can have lasting consequences for how rights and boundaries are understood later in life[23]. Online platforms and digital communities play a decisive role in reinforcing or resisting these changes. Platform rules and moderation practices signal what kinds of behavior are tolerated. Where AI-generated abuse remains visible, spreads rapidly, or is inconsistently moderated, users receive an implicit message that such conduct falls within acceptable limits. Conversely, clear labelling of synthetic content, rapid removal of non-consensual material, and visible support for victims can help sustain norms that prioritize dignity and consent[24]. At present, platform responses remain uneven, particularly across languages and regions, contributing to fragmented normative environments. Cultural and social context further shapes how AI-generated content affects community norms. In societies where honor, reputation, and family standing are deeply embedded social values, a single synthetic image can cause severe social consequences, including exclusion or violence against the victim. In other contexts, stronger emphasis on expressive freedom may lead to the initial dismissal of synthetic abuse as harmless or humorous[25]. These differences highlight that AI-generated content does not operate in a social vacuum; it interacts with existing inequalities and power relations, often amplifying them.

6. Key Challenges for AI-Generated Content Regulations

Regulating AI-generated content presents a set of challenges that differ in both scale and nature from earlier forms of digital governance. Existing legal frameworks were largely designed for human-created content, identifiable authorship, and slower modes of dissemination. Generative AI disrupts these assumptions by enabling rapid, large-scale production of realistic content with limited traceability[26]. As a result, even where laws addressing online harm exist, they often struggle to capture the specific risks posed by synthetic media. A first challenge concerns legal definition. Many jurisdictions lack clear and consistent definitions of AI-generated content, deepfakes, or synthetic media. Some legal approaches focus narrowly on impersonation of real individuals, while others rely on broader concepts such as "false" or "manipulated" information. This lack of precision creates uncertainty for courts, regulators, and platforms. Narrow definitions leave harmful practices unregulated, while overly broad rules risk restricting legitimate expression, including satire, artistic work, and political criticism[27]. The absence of clear legal categories also complicates international cooperation, as states often regulate similar harms under incompatible terms.

A second challenge lies in responsibility and attribution. Harmful AI-generated content typically involves multiple actors, including users who generate the content, platforms that host and amplify it, and developers that design and release generative models. Existing liability frameworks rarely clarify how responsibility should be distributed among these actors(Ganai & Naikoo, 2025). Victims therefore face uncertainty about whom to hold accountable and which legal pathways are available. This diffusion of responsibility weakens the right to an effective remedy and allows harmful practices to persist without meaningful consequences. Detection and enforcement present a third major challenge. As generative models become more advanced, distinguishing synthetic content from authentic material becomes increasingly difficult. Technical detection tools exist, but they are unevenly available, often controlled by private companies, and subject to continuous technological competition. Many regulators, courts, and law-enforcement agencies lack the technical expertise or resources required to assess synthetic media reliably[29]. This gap between legal standards and enforcement capacity means that regulatory obligations frequently remain ineffective in practice.

The cross-border nature of digital platforms further complicates regulation. AI-generated content can be created in one jurisdiction, hosted in another, and accessed globally within seconds. National laws on privacy, free expression, and platform liability vary widely, allowing harmful content to circulate through jurisdictions with weaker regulatory oversight[30]. Existing international cooperation mechanisms are slow and poorly adapted to the speed at which synthetic content spreads, leaving victims without timely protection and limiting the reach of domestic enforcement actions. Another significant challenge involves balancing the protection of human rights. Measures aimed at restricting harmful AI-generated content must protect individuals from abuse, exploitation, and disinformation without becoming tools for censorship or political control. In some contexts, broadly framed rules on "fake news" or "misuse of technology" have been used to silence journalists, critics, and human rights defenders[31]. This risk is particularly acute where judicial independence is weak or regulatory powers are concentrated in executive bodies. Effective regulation therefore requires safeguards that prevent abuse while addressing genuine harm. Regulatory dependence on private platforms and technology companies also poses structural concerns. In practice, platforms often act as primary decision-makers regarding content removal, labelling, and visibility. Their policies and enforcement practices are shaped by commercial priorities and vary across regions and languages. Communities in the Global South frequently receive less effective moderation and slower responses, reinforcing existing inequalities in protection. This form of private governance raises questions about transparency, accountability, and democratic oversight in the protection of human rights.

7. Notable Practices from the International Community

Different countries are trying to deal with AI-generated content in very different ways. None of them has a perfect model, but each offers useful ideas for how to protect people's rights while still allowing

innovation. This section highlights some of the more promising approaches from the USA, EU, UK, China, Russia and Pakistan, especially where they focus on consent, transparency, platform duties and victim protection.

7.1 United States

In the United States, regulation of AI-generated content is mostly built from the bottom up. Many rules come from individual states rather than a single national law. Several states now have specific deepfake laws, for example banning deceptive deepfakes in election campaigns close to voting day, or treating non-consensual sexual deepfakes as a form of image-based sexual abuse[32]. These laws are narrow but practical: they focus on clear harms (election manipulation, sexual exploitation) and give prosecutors concrete tools. A major federal step is the *take it down act 2025.* This law requires online platforms to remove non-consensual intimate "visual deceptions" – including AI-generated deepfake nudes – within a short time after a victim reports them[33]. It creates a standard process: victims can send a notice, platforms must act quickly, and failure to remove content can trigger legal consequences. While it does not solve all problems, it is a strong example of a victim-centered removal regime for AI-generated sexual abuse. The US debate also shows another "best practice": thinking beyond punishment. Some federal proposals, such as the *deep fakes accountability act,* would require clear labels or watermarks on AI-generated audio, video and images, so that viewers know when content is synthetic. Although these proposals are not yet fully in force, they reflect an important idea: prevention and transparency (for example, labelling and detection tools) should go together with criminal and civil liability[34].

7.2 European Union

The European Union has taken one of the most systematic approaches. The *EU AI Act* (Regulation 2024/1689) is a broad, risk-based law that covers many uses of AI, not just deepfakes. For AI systems that generate or manipulate images, audio or video, the Act creates transparency duties: when content is AI-generated or heavily altered, users should be told clearly, unless this is already obvious[35]. This is meant to protect fundamental rights such as privacy, dignity and freedom of information. The AI Act works together with the *Digital Services Act (DSA),* which sets strict obligations for very large online platforms. Platforms must assess and reduce "systemic risks", including disinformation, gender-based violence and election manipulation linked to deepfakes, and they must be more transparent about their content moderation decisions and algorithms[36]. This combination, duties on AI providers plus duties on platforms, is a key feature of the EU model. The EU is also updating its *criminal and victim-protection laws*. A new directive on violence against women and domestic violence recognizes non-consensual intimate images, including those generated with AI, as a serious form of abuse that requires protection and remedies. Reforms to child sexual abuse law also aim to cover AI-generated material involving minors[37]. Taken together, the EU offers a strong example of how to integrate AI-generated content into a wider human-rights and platform-governance framework.

7.3 United Kingdom

The United Kingdom focuses heavily on online safety and image-based abuse. The *Online Safety Act 2023* creates a duty of care for platforms to tackle illegal and harmful content and makes the sharing and threatening to share intimate images, including deepfakes, a criminal offence. Regulators can demand risk assessments, transparency and better reporting tools, especially where children are involved[38]. Building on this, the government has announced new offences that will specifically criminalize the creation of sexually explicit deepfakes, not only their sharing. This is an important shift: it recognizes that harm begins at the point of fabrication, especially when victims are blackmailed or humiliated even before the content goes viral. Enforcement is handled by Ofcom, which has already started to use its powers, for example by fining pornography websites that fail to put strong age-verification and other safety checks in place under the new law[39]. While the Online Safety Act is

controversial and raises free-speech questions, it offers a concrete model of how to combine criminal offences, platform duties and an independent regulator with real sanctions.

7.4 China

China was among the first states to adopt detailed rules directly aimed at synthetic media. The *Administrative Provisions on Deep Synthesis Internet Information Services* require providers of "deep synthesis" tools (including deepfakes) to ensure that AI-generated content is clearly labelled, often through visible marks or watermarks, and not used to create fake news or other illegal information[40]. Providers must also obtain consent before using a person's biometric data (e.g. face or voice) in deep synthesis services. In 2023, China also introduced the *Interim Measures for the Management of Generative AI Services,* which cover public-facing generative AI models. These measures require service providers to monitor content, respond quickly to illegal use, protect users' personal information and add marks to AI-generated images and videos in line with the deep synthesis rules. The rules sit on top of broader laws on cyber security, data security and personal information[41]. As a result, China's framework illustrates a strong "provider-responsibility" model: AI and deepfake service providers are treated as content producers, with duties to label, filter and correct harmful content. From a human-rights perspective, this shows how far proactive obligations can go, but it also raises serious concerns about over-broad content control and the chilling of legitimate expression.

7.5 Russia

Russia does not yet have a single deepfake law, but it has started to regulate AI-generated content through its existing information-control framework[42]. In 2022, *Federal Law No. 32-FZ* changed the Criminal Code to punish the spread of "false information" about the armed forces, which includes manipulated videos and deepfakes used for disinformation about military operations[43]. This links deepfake regulation very closely to national security and information warfare. Public debate in Russia has moved towards more specific rules. Opinion polls and legal discussions show broad support among lawyers and citizens for the idea that deepfakes should be regulated at the legislative level, and lawmakers are considering rules that would require labelling AI-generated content and clarifying liability. The main "best practice" element here is the early recognition that deepfakes are not only a private or civil issue but also a serious challenge for information integrity and public order. At the same time, there is a real risk that such laws can be used to target political speech rather than to protect victims of harassment or sexual abuse.

7.6 Pakistan

Pakistan's main tool for dealing with online harms is the *Prevention of Electronic Crimes Act (PECA) 2016* updated by later rules and amendments[44]. PECA was written before the deepfake era, but some of its provisions can be used against AI-generated abuse[45]. Section 21 criminalizes image-based abuse, including capturing, editing, sharing or threatening to share sexually explicit images or videos, and digital rights groups now interpret this to cover deepfake sexual content as well[46]. Section 37 gives the Pakistan Telecommunication Authority (PTA) power to remove or block unlawful online content, and has been used to block pornography and other harmful material[47]. In 2025, new amendments introduced Section 26-A, which criminalizes the intentional dissemination of false and fake information online. This can be applied to certain deepfake disinformation campaigns, especially when they target public figures or state institutions[48]. The Federal Investigation Agency's Cyber Crime Wing investigates cases under PECA and has already registered FIRs against people who allegedly uploaded deepfake "immoral" videos of political leaders. Pakistan also shows how legal measures can be supported by civil society and international partners**.** Reports by UNFPA and local NGOs document technology-facilitated gender-based violence and deepfake abuse, and call for better victim support, digital evidence handling and specialized training for police and courts. These efforts, together with awareness campaigns about Section 21 PECA, are slowly building a practical response to AI-generated abuse even before a dedicated deepfake law exists[49]. At the same time, scholars and

rights groups' note that PECA's broad language is sometimes used to restrict journalism and free expression, and that the law still does not clearly mention synthetic media or AI-generated content. The lesson from Pakistan is that existing cybercrime and image-based abuse laws can be used as a starting point, but clearer and more balanced rules are needed to deal with AI-generated content in a way that protects both victims and fundamental freedoms.

8. Conclusion

This study shows that AI-generated content is not just a "tech problem" or a "social media issue." It has become a direct pathway to human rights harm. When synthetic images, videos, and audio are used for non-consensual pornographic deepfakes, child sexual abuse material, targeted harassment, hate campaigns, and political manipulation, the damage is real: people lose dignity, privacy, safety, and equal access to public life. A second key finding is that AI-generated content also weakens community norms that normally protect society. Consent starts to look optional, truth becomes unclear, and online cruelty is treated like entertainment. This creates a dangerous cycle: once harmful synthetic content becomes common, victims are blamed or ignored, and abusers feel encouraged. Over time, the result is a colder and less trustworthy information environment where women, children, minorities, journalists, and human rights defenders carry the heaviest burden. Regulation must be guided by human rights principles, especially dignity, privacy, equality, and child protection, rather than leaving decisions mainly to private companies or treating the issue as a side topic of "general AI governance".

References

[1] J. Afzal, "Development of Legal Framework of Digital Laws," in *Implementation of Digital Law as a Legal Tool in the Current Digital Era*, Singapore: Springer Nature Singapore, 2024, pp. 139–154. doi: 10.1007/978-981-97-7106-6_7.

[2] A. O. Mohamed, "The Effect of Simulating Virtual Scenes using Artificial Intelligence Techniques in Producing Various Media Materials," *J. Ecohumanism*, vol. 3, no. 8, Nov. 2024, doi: 10.62754/joe.v3i8.4771.

[3] J. Afzal, "Best Practice of Digital Laws and Digital Justice," in *Implementation of Digital Law as a Legal Tool in the Current Digital Era*, J. Afzal, Ed., Singapore: Springer Nature Singapore, 2024, pp. 95–120. doi: 10.1007/978-981-97-7106-6_5.

[4] J. Afzal, *Contemporary Theory of Digitalization (CToD)*. 2026. doi: 10.2139/ssrn.6018858.

[5] J. Afzal, *Implementation of digital law as a legal tool in the current digital Era*. Springer, 2024.

[6] J. Afzal, *Exploring the New Horizons of International Law Concerning Globalization of Economy*. Springer Nature.

[7] F. R. Moreno, "Generative AI and deepfakes: a human rights approach to tackling harmful content," *Int. Rev. Law Comput. Technol.*, Sept. 2024, Accessed: Jan. 20, 2026. [Online]. Available: https://www.tandfonline.com/doi/abs/10.1080/13600869.2024.2324540

[8] K. de Vries, "Let the Robot Speak! AI-Generated Speech and Freedom of Expression," in *YSEC Yearbook of Socio-Economic Constitutions 2021: Triangulating Freedom of Speech*, S. Hindelang and A. Moberg, Eds., Cham: Springer International Publishing, 2022, pp. 93–115. doi: 10.1007/16495_2021_38.

[9] D. V. Voinea, "AI AND COPYRIGHT - WHO OWNS AI GENERATED CONTENT?," July 2023, doi: 10.5281/ZENODO.15252004.

[10] Y. Wang, "Synthetic Realities in the Digital Age: Navigating the Opportunities and Challenges of AI-Generated Content," Aug. 18, 2023. doi: 10.36227/techrxiv.23968311.v1.

[11] S. Firmino Pinto, "AI Friend? Risks, Implications, and Recommendations on Generative AI for Children," 2024.

[12] S. Ali *et al.*, "Children as creators, thinkers and citizens in an AI-driven future," *Comput. Educ. Artif. Intell.*, vol. 2, p. 100040, Jan. 2021, doi: 10.1016/j.caeai.2021.100040.

[13] Z. Li, W. Zhang, H. Zhang, R. Gao, and X. Fang, "Global Digital Compact: A Mechanism for the Governance of Online Discriminatory and Misleading Content Generation," *Int. J. Human–Computer Interact.*, Jan. 2025, Accessed: Jan. 21, 2026. [Online]. Available: https://www.tandfonline.com/doi/abs/10.1080/10447318.2024.2314350

[14] J. Ahmad and A. Haider, "Firewall Technology Testing in Pakistan: The Fine Line Between National Security and Freedom of Expression," *J. Eng. Sci. Technol. Trends*, vol. 2, no. 1, 2025, Accessed: Jan. 21, 2026. [Online]. Available: https://journals.scopua.com/index.php/JESTT/article/view/11

[15] K. Goth, "AI-Generated Content and the Pollution of the Information Sphere: A Freedom of Expression Analysis under Article 10 ECHR," Master Thesis, 2025. Accessed: Jan. 21, 2026. [Online]. Available: https://studenttheses.uu.nl/handle/20.500.12932/49552

[16] N. Ahmad, A. W. Ali, and M. H. B. Yussof, "The Challenges of Human Rights in The Era of Artificial Intelligence," *UUM J. Leg. Stud. UUMJLS*, vol. 16, no. 1, pp. 150–169, Jan. 2025.
[17] R. Q. Idrees, S. Kiyani, and A. Shahid*, "The Role of International Law in Protecting Human Rights Globally," *Law Res. J.*, vol. 3, no. 2, pp. 71–85, May 2025.
[18] A. Flynn, A. Powell, A. Eaton, and A. J. Scott, "Sexualized Deepfake Abuse: Perpetrator and Victim Perspectives on the Motivations and Forms of Non-Consensually Created and Shared Sexualized Deepfake Imagery".
[19] T. Yu, Y. Tian, Y. Chen, Y. Huang, Y. Pan, and W. Jang, "How Do Ethical Factors Affect User Trust and Adoption Intentions of AI-Generated Content Tools? Evidence from a Risk-Trust Perspective," *Systems*, vol. 13, no. 6, June 2025, doi: 10.3390/systems13060461.
[20] J. Afzal, C. Yongmei, A. Fatima, and A. Noor, "Review of various Aspects of Digital Violence," *J. Eng. Sci. Technol. Trends*, vol. 1, no. 2, 2024.
[21] S. Nasiri and A. Hashemzadeh, "The Evolution of Disinformation from Fake News Propaganda to AI-driven Narratives as Deepfake," *J. Cyberspace Stud.*, vol. 9, no. 1, Jan. 2025, doi: 10.22059/jcss.2025.387249.1119.
[22] J. Afzal, *Exploring the New Horizons of International Law Concerning Globalization of Economy*. Springer Nature.
[23] A. Machado, A. Pesqueira, J. Rodrigues dos Santos, A. Sacavem, M. Sousa, and F. Teixeira, "ESG and Digital Transformation in Organizations," 2025, pp. 1–32. doi: 10.1007/978-3-031-86079-9_1.
[24] C. Easttom, "Malicious Use of Artificial Intelligence," in *2025 IEEE 15th Annual Computing and Communication Workshop and Conference (CCWC)*, Jan. 2025, pp. 00499–00507. doi: 10.1109/CCWC62904.2025.10903787.
[25] M. Krkić, "Cultural perspectives on AI usage and regulation in deepfake creation: how culture shapes AI practices," *Int. Commun. Chin. Cult.*, vol. 12, no. 2, pp. 225–237, June 2025, doi: 10.1007/s40636-025-00330-5.
[26] H. Qudah, "Tracing the development of environmental, social and governance (ESG) performs: a systematic review and bibliometric analysis," *Future Bus. J.*, vol. 12, Jan. 2026, doi: 10.1186/s43093-025-00705-5.
[27] K. Ding, "A Tiered Framework for Copyright Ownership in AI-Generated Content," *INNO-PRESS J. Emerg. Appl. AI*, vol. 1, no. 9, Dec. 2025, doi: 10.65563/jeaai.v1i9.52.
[28] P. A. Ganai and I. A. Naikoo, "The Ethical Paradox of AI-Generated Texts: Investigating the Moral Responsibility in Generative Models".
[29] G. Pandy, A. Murugan, and V. Pugazhenthi, "Generative AI: Transforming the Landscape of Creativity and Automation," *Int. J. Comput. Appl.*, Jan. 2025, doi: 10.5120/ijca2025924392.
[30] A. Gaidartzi and I. Stamatoudi, "Authorship and Ownership Issues Raised by AI-Generated Works: A Comparative Analysis," *Laws*, vol. 14, no. 4, Aug. 2025, doi: 10.3390/laws14040057.
[31] M. Anderljung, J. Hazell, and M. von Knebel, "Protecting society from AI misuse: when are restrictions on capabilities warranted?," *AI Soc.*, vol. 40, no. 5, pp. 3841–3857, June 2025, doi: 10.1007/s00146-024-02130-8.
[32] A. D. Andriani, "The Future of Digital Content: AI-Generated Texts, Images, Videos, and Real-Time Production," in *Impacts of AI-Generated Content on Brand Reputation*, IGI Global Scientific Publishing, 2025, pp. 143–170. doi: 10.4018/979-8-3373-4327-3.ch006.
[33] T. [R-T. Sen. Cruz, "All Info - S.146 - 119th Congress (2025-2026): TAKE IT DOWN Act." Accessed: Jan. 21, 2026. [Online]. Available: https://www.congress.gov/bill/119th-congress/senate-bill/146/all-info
[34] V. Ugwuoke and M. R. Sanfilippo, "The Current Landscape of Deepfake Legislation in the United States: Analysis of State-Level Responses," *J. Inf. Policy*, June 2025, doi: 10.5325/jinfopoli.15.2025.0004.
[35] D. Hartmann, J. R. L. de Pereira, C. Streitbörger, and B. Berendt, "Addressing the regulatory gap: moving towards an EU AI audit ecosystem beyond the AI Act by including civil society," *AI Ethics*, vol. 5, no. 4, pp. 3617–3638, Aug. 2025, doi: 10.1007/s43681-024-00595-3.
[36] L. Nannini, E. Bonel, D. Bassi, and M. J. Maggini, "Beyond phase-in: assessing impacts on disinformation of the EU Digital Services Act," *AI Ethics*, vol. 5, no. 2, pp. 1241–1269, Apr. 2025, doi: 10.1007/s43681-024-00467-w.
[37] M. Palade-Ropotan, "MEANS OF PROTECTING VICTIMS OF CRIME AT EUROPEAN LEVEL," *Int. J. Leg. Soc. Order*, vol. 5, no. 1, pp. 1–15, 2025.
[38] P. Coe, "Tackling online false information in the United Kingdom: The Online Safety Act 2023 and its disconnection from free speech law and theory*," *J. Media Law*, vol. 15, no. 2, pp. 213–242, July 2023, doi: 10.1080/17577632.2024.2316360.
[39] S. Law, "Effective enforcement of the Online Safety Act and Digital Services Act: unpacking the compliance and enforcement regimes of the UK and EU's online safety legislation," *J. Media Law*, vol. 16, no. 2, pp. 263–300, July 2024, doi: 10.1080/17577632.2025.2459441.

[40] M. Sheehan, “China’s AI Regulations and How They Get Made,” *Ho Ri Zo NS*, 2023.
[41] S. Migliorini, “China’s Interim Measures on generative AI: Origin, content and significance,” *Comput. Law Secur. Rev.*, vol. 53, p. 105985, July 2024, doi: 10.1016/j.clsr.2024.105985.
[42] E. Trikoz, E. Gulyaeva, and K. Belyaev, “Russian experience of using digital technologies and legal risks of AI,” *E3S Web Conf.*, vol. 224, p. 03005, 2020, doi: 10.1051/e3sconf/202022403005.
[43] D. Moskwa, “Russia’s battle for remembrance. Memory laws in Vladimir Putin’s Russia exemplified by the Russo-Ukrainian war,” *Rocz. Inst. Eur. Środ.-Wschod.*, vol. 22, no. 1, pp. 105–121, Nov. 2024, doi: 10.36874/RIESW.2024.1.6.
[44] C. Yongmei and J. Afzal, “Impact of enactment of ‘the prevention of electronic crimes act, 2016’as legal support in Pakistan,” *Acad. Educ. Soc. Sci. Rev.*, vol. 3, no. 2, pp. 203–212, 2023.
[45] I. U. Haq and S. M. Zarkoon, “Cyber Stalking: A Critical Analysis of Prevention of Electronic Crimes Act-2016 and Its Effectiveness in Combating Cyber Crimes, A Perspective from Pakistan,” *Pak. Multidiscip. J. Arts Sci.*, pp. 43–62, 2023, doi: 10.5281/zenodo.10450177.
[46] S. Naseer and C. Ashraf, “Gender-Based Violence in Pakistan’s Digital Spaces,” *Fem. Leg. Stud.*, vol. 30, no. 1, pp. 29–50, Apr. 2022, doi: 10.1007/s10691-021-09473-3.
[47] S. Khan, P. M. Tehrani, and M. Iftikhar, “Impact of PECA-2016 Provisions on Freedom of Speech: A Case of Pakistan,” *J. Manag. Info*, vol. 6, no. 2, pp. 7–11, 2019, doi: 10.31580/jmi.v6i2.566.
[48] M. A. Abbas and M. R. A. Tullah, “Criminalizing Disinformation in Pakistan: A Constitutional Analysis of Section 26-A of PECA and Freedom of Expression,” *Pak. J. LAW Anal. WISDOM*, vol. 4, no. 12, pp. 67–77, Dec. 2025.
[49] T. Nazakat and F. E. Malik, “Empowering Justice through AI: Addressing Technology-Facilitated Gender-Based Violence with Advanced Solutions,” *J. Law Soc. Stud.*, vol. 7, no. 1, pp. 26–42, Mar. 2025, doi: 10.52279/jlss.07.01.2642.

Declaration

Conflict of Study: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.
Funding Statement: This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.
Availability of Data and Material: Data will be made available by the corresponding author on reasonable request.