Journal of Innovative and Creativity
Journal Homepage: https://joecy.org/index.php/joecy
e-mail: journaljoecy@gmail.com

Faculty of Education
University of Pahlawan Tuanku Tambusai
ISSN 2775-771X (Online)
ISSN 2962-570X (Printed)
6 (1) 2026, Page: 1440-1447

The Dynamics of Gender-Based Violence Through Generative Technology (Deepfake & Voice Cloning) in a Digital Communication Perspective

Ilmi Sila Ayu[1], Febi Ramadhani Rusdin[2], Muhammad Taslim[3]
[1]Affiliation Universitas Media Nusantara Citra[1,2]
Universitas Almarisah[3]

Corresponding Author: Ilmi Sila Ayu, email: ilmisilaayu@gmail.com

Published: January, 2026

ABSTRACT

The proliferation of generative Artificial Intelligence (AI) technologies has transformed the digital communication landscape, while simultaneously emerging new variants of online gender-based violence (OGBV). This study aims to analyze how deepfake and voice cloning technologies are utilized as instruments of symbolic terror, blurring the lines between reality and fabrication. Employing a qualitative approach with case study methods and critical discourse analysis, this research examines the phenomenon of AI misuse within the context of Computer-Mediated Communication (CMC). The findings indicate that AI-facilitated violence creates a "hyperreality" of victimization; wherein victims experience tangible psychological and social impacts even though the abusive material is digitally synthesized. From a communication perspective, this phenomenon degrades digital identity integrity and creates a chilling effect that silences women's participation in the public sphere. Indonesian regulations, including the ITE Law and the Sexual Violence Crime Law (UU TPKS), are considered to suffer from a cultural lag in responding to the rapid adoption of this technology.

.

Keywords: Deepfake, Online Gender-Based Violence, Digital Communication, Hyperreality, Media Regulation.

© The Author(s). 2021 Open Access This article is licensed under a Creative Commons Atibution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0

PENDAHULUAN

The evolution of communication technology has reached a stage where the boundary between the message creator (sender) and the reality of the message itself has become increasingly blurred. The presence of Generative Artificial Intelligence (Gen-AI) allows for the manipulation of audio and visuals with an alarming degree of precision, creating what is known as synthetic media (Chesney & Citron, 2019). Over the past decade, optimism regarding the democratization of digital content has slowly eroded due to the misuse of this technology for exploitative purposes, particularly targeting women.

The phenomenon of deepfake video manipulation based on machine learning to swap one person's face onto another person's body and voice cloning have become new weapons within the spectrum of Online Gender-Based Violence (OGBV). UN Women (2025) reports that AI-facilitated violence is now a global threat that exacerbates gender inequality in the digital space. The 2024 Annual Notes Report (CATAHU) by Komnas Perempuan (National Commission on Violence Against Women) also recorded a significant surge in cyber violence cases, with the modus operandi of digital content manipulation occupying an alarming rank (Komnas Perempuan, 2024).

Unlike conventional non-consensual intimate content distribution (revenge porn) which requires authentic original material, AI-based violence can affect anyone without requiring any prior history of intimate interaction. Samosir (2023) highlights that OGBV in this era distributes fear through messaging platforms like Telegram and social media X, where deepfake content is produced and disseminated en masse.

From the perspective of Communication Science, this is not merely a technological issue but a crisis of meaning and communication ethics. Jean Baudrillard (1981) once predicted the era of simulacra, where representation replaces reality. Today, deepfake porn is the most brutal manifestation of this simulacra: a technically produced lie that nonetheless carries socially destructive consequences for the victim's reputation. Cangara (2023) emphasizes the importance of digital communication ethics to prevent dehumanization in cyberspace, but the acceleration of technology often outpaces users' ethical adaptation.

This research is urgent given the limited communication literature in Indonesia that specifically dissects the impact of deepfake from the viewpoint of symbolic victimization and communication policy. This study will answer the questions: How does generative technology reconstruct forms of gender violence in the digital space, and how does the communication perspective view the failure of regulation in protecting the digital identity integrity of citizens?

2. Literature Review

2.1. Gender-Based Violence (OGBV) and AI

OGBV or Technology-Facilitated Gender-Based Violence (TFGBV) is defined as acts of violence facilitated, exacerbated, or amplified by the use of information technology. This phenomenon has undergone a significant evolution, where the modus operandi has shifted from the exploitation of authentic material (such as traditional revenge porn) to the fabrication of synthetic content. Danielle Citron (2022) in her work The Fight for Privacy asserts that violence in cyberspace does not merely violate physical boundaries, but fundamentally attacks a person's "sexual privacy," which is the core of self-autonomy and dignity.

The utilization of generative artificial intelligence, particularly the Generative Adversarial Networks (GANs) algorithm, has transformed OGBV into AI-facilitated violence. This modus allows perpetrators to create obscene material without needing direct interaction or access to the victim's intimate photos. Deepfake and voice cloning serve as new weapons that manipulate the victim's basic biometric data—face and voice to generate simulations that are virtually indistinguishable from reality. Therefore, this violence is algorithmic, attacking the victim's visual and auditory integrity massively and anonymously.

When AI is used to create pornographic deepfakes, perpetrators not only violate privacy but also effectively usurp the victim's right to their own image (right to publicity). In the context of Indonesian law, Mahardika (2025) highlights that deepfake often resides in a gray area due to the difficulty in proving the element of "originality" or "authenticity" in the pornographic material, a prerequisite often sought by law enforcement. The lack of clear legal limits for engineered content creates a large loophole that benefits the perpetrators.

Ultimately, KBGO-AI is positioned as a systemic amplification of existing misogyny. The UN Women (2025) report underscores that this technology exacerbates gender inequality in the digital space by creating an inherently unsafe environment for women. The threat of creating and distributing synthetic content functions as a tool for social control, limiting women's participation in public and political spaces, thereby strengthening patriarchal structures through modern technological instruments.

2.2. Hyperreality and Digital Identity

Jean Baudrillard's Theory of Hyperreality is relevant for dissecting the deepfake phenomenon. Baudrillard (1981) explained that society has entered the third order of simulacra, where the sign or representation (simulacra) no longer reflects reality (the second reality) but rather replaces it, becoming more "real" than the original. In the context of deepfake, the fake video or audio created technically is a perfect sign, but it carries real and destructive social and psychological functions.

The role of social media is crucial in accelerating this process of hyperreality. Nasrullah (2018) emphasizes that in a digital ecosystem dominated by the logic of virality and sensation, the speed of content dissemination often far surpasses the verification process of truth. Consequently, deepfake content spreads as truth before the victim has time to prove its falsehood. This temporal lag allows the falsified sign to take root in public consciousness, making it a collective 'truth,' despite being fabricated.

For the victim, even though the content is false (fake), the trauma, social stigma, and destruction of reputation received are hyper-real (very real). This phenomenon creates severe digital identity distortion. The victim's digital identity, built through years of authentic interactions and representation, is instantly replaced by a synthetic persona constructed by AI. Victims lose control over their own self-narrative, and the fake image begins to dictate how the public and social networks perceive them.

Therefore, KBGO-AI is not just a privacy violation, but also an attack on the individual's ontological security. The ability to trust one's own eyes and ears (seeing is believing) is destroyed, as the boundary between the original and the fake has vanished. The consequences of this erosion of trust ripple throughout the digital ecosystem, which in turn creates an environment fraught with suspicion, where any digital evidence can be denied (termed the liar's dividend), making victims even more vulnerable.

2.3. Communication Privacy Management

The Communication Privacy Management (CPM) Theory proposed by Petronio (2002) offers a framework for understanding an individual's right to control access to their personal information through rule-based boundaries (privacy boundaries) and shared ownership (collective privacy boundaries) when that information is disclosed. The central tenet of this theory is that privacy is a dialectical process: individuals constantly balance the need for disclosure and the need for confidentiality.

Deepfake and voice cloning technologies brutally violate these boundaries. The violations are non-consensual and non-reciprocal; the perpetrator does not need to be involved in a communication relationship with the victim. Research by Musyaffa and Effendi (2022) confirms that interactions on social media often open the door to exploitation. Biometric data voluntarily uploaded by users (e.g., profile pictures or voice samples on TikTok) are then "stolen" and misused by third parties without consent, transforming innocent public data into an instrument of private harm.

The violation of CPM by AI is a form of seizing control over self-representation. Citron (2022) emphasizes that privacy is fundamentally about control: the right to decide who sees, uses, and processes our personal data. In the case of KBGO-AI, perpetrators not only access information but also actively produce a new image entirely outside the victim's control. The victim's body and voice are reduced to data points that can be exploited, creating a destructive digital asset owned by an irresponsible third party.

The ultimate consequence of the collapse of privacy boundaries through KBGO-AI is the emergence of a chilling effect. The fear of potential identity and synthetic content forgery makes women tend to withdraw from the digital public sphere or practice self-censorship. They consciously limit the uploading of photos of their face, voice, or even opinions, to prevent their biometric material from being stolen and misused. This results in a drastic shrinkage of their public sphere, damaging their capacity to participate fully in social and political dialogue in cyberspace.
.

3. Research Methods

This research adopts a qualitative approach with a primary focus on case studies and critical discourse analysis. The selection of the qualitative approach is based on the need to explore the depth and complexity of the phenomenon of AI-mediated Online Gender-Based Violence (KBGO-AI), which is subjective, contextual, and rich in symbolic meaning. Case studies are used to intensively dissect specific instances of deepfake and voice cloning content distribution that occurred in the Indonesian digital space throughout the period 2023–2024. The object of study is not just the content itself, but also the patterns of dissemination, the narratives constructed around the content, and the media/regulatory response to the hyperreality crisis of victimization it creates.

Primary and secondary data collection techniques are carried out through two main stages. The first stage is Digital Observation, which involves monitoring and collecting textual and visual data from social media platforms frequently used as a means of KBGO-AI distribution, such as X (Twitter) and Telegram. This observation focuses on identifying cybercrime patterns, perpetrator anonymity mechanisms, and how the synthetic content is constructed and narrated to maximize the psychological and social impact on the victims. The second stage is Document Study or textual analysis, where researchers critically examine annual reports from digital rights and human rights advocacy institutions (e.g., Komnas Perempuan and SAFEnet) and scrutinize the formal texts of relevant regulations (the ITE Law and the UU TPKS).

The data collected, from both digital observation and document study, is then analyzed using the interactive model of Miles and Huberman, which emphasizes the process of data reduction, data presentation, and conclusion drawing. Data reduction is performed by sorting information relevant to the focus of KBGO-AI, separating authentic data from unverified claims, and grouping findings based on the mechanism of violence (algorithmic) and the psychosocial impact (chilling effect). Data presentation is done in the form of descriptive narrative and analytical argumentation strengthened by evidence from the findings (such as statistical data on violence trends from Komnas Perempuan).

The analytical tool used in this research is Critical Discourse Analysis (CDA), specifically to dismantle the power relations operating through generative technology and the language of regulation. CDA allows the researcher not only to describe what happens but also to analyze how AI-based violence is institutionalized and justified in public discourse, or conversely, how regulation fails to represent the victimization experience of deepfake victims. This approach specifically aims to expose the patriarchal power relations that utilize technological sophistication to control and silence women's participation in cyberspace.

MATERIALS AND METHODS

3.1 Mekanisme Teror: Dari Fisik ke Algoritmik

The results of the investigation show a fundamental shift in the pattern of OGBV, moving from violence based on authentic material (such as genuine photos or videos) toward violence entirely facilitated by artificial intelligence algorithms, particularly through Generative Adversarial Networks (GANs). This shift transforms OGBV into AI-facilitated violence, which enables perpetrators to

perform visual and audio manipulation with near-perfect precision and without requiring any history of physical or intimate contact with the victim. The object of violence shifts from the victim's physical body to their basic biometric data face, voice, and public image which is then synthesized into damaging content.

This algorithmic terror mechanism begins with the "theft" of biometric data, which is often obtained legally and voluntarily by the victim through uploads on social media such as profile pictures, short videos on TikTok, or voice recordings on podcasts. This public data is then fed into an AI model for training, resulting in deepfakes that place the victim's face onto non-consensual sexual content, or voice cloning that makes the victim appear to utter obscene or threatening statements. From a communication perspective, this process is perfect identity hijacking, where the victim's visual and auditory integrity is seized, and their self-representation in the public sphere is controlled by a manipulative synthetic persona.

The uniqueness of KBGO-AI lies in the ease of mass production and the anonymity it offers, which significantly differentiates it from conventional OGBV. Perpetrators can produce hundreds or even thousands of deepfake content pieces at relatively low cost and high speed, enabling a "carpet bombing" attack on the victim's image. The anonymous nature of distribution platforms such as Telegram and the dark web further exacerbates the difficulty in tracing and law enforcement. According to the Council of Europe (2022), this technology expands the spectrum of violence because it allows anonymous perpetrators to attack victims remotely with permanent and indelible reputational damage.

The consequence of this mechanism is the creation of a "toxic digital asset." Once deepfake or voice cloning content is distributed in the digital sphere, the material becomes permanent and can reappear at any time, even after takedown efforts on the initial platform. The victim's inability to completely erase this fake digital footprint creates a cycle of continuous victimization. The limitations of Indonesian legal mechanisms in reaching global platforms and the difficulty in proving AI-generated content lead to massive vulnerability, as victims must constantly live under the threat of a digital shadow that they never created.

3.2 Psychosocial Impact: The Chilling Effect

The psychosocial impact of KBGO-AI extends far beyond reputational damage, reaching into the realm of acute mental trauma. Case studies and reports from SAFEnet victims (2023) consistently show that victims experience extreme anxiety, panic attacks, depression, and profound social isolation. The main source of trauma is not merely the content itself, but the feeling of absolute powerlessness to fight the false reality created by AI. They must face a skeptical public that finds it difficult to distinguish between the authentic and the fake, which exacerbates the psychological suffering. Wahyuni and Saputra (2023) confirm that OGBV victims experience a drastic decline in mental resilience, often leading to decisions to withdraw from social and professional interaction.

This phenomenon also sharply triggers what is called the liar's dividend, a concept popularized by Chesney and Citron (2019). This dividend is the paradoxical benefit gained by perpetrators of crime from the existence of deepfake technology. On one hand, fake content is accepted as authentic, harming the victim; on the other hand, perpetrators of genuine crimes can deny real video or audio evidence of their offenses by claiming it as a deepfake. For KBGO-AI victims, this creates a hostile communication environment, where their attempts to prove the content's falsehood are often dismissed as manipulation, forcing them to be trapped in a web of lies created by technology.

The deepfake threat directly results in a significant chilling effect on women's participation in the digital public sphere. Especially for women who work as journalists, human rights activists,

politicians, or public figures, this threat becomes an effective tool of terror to silence criticism or free expression. The fear of their biometric content potentially being stolen and misused leads them to practice self-censorship or withdraw completely from platforms that require uploading authentic photos or voices. This damages the principles of deliberative democracy, as critical women's voices are eliminated from social dialogue, which ultimately strengthens existing dominant narratives.

Furthermore, the erosion of trust caused by KBGO-AI undermines the individual's ontological security. The belief in common sense, the ability to trust what one sees or hears, is destroyed when the boundary between simulation and reality is lost. Victims lose a fundamental sense of security in their digital identity. When every image and voice can be engineered, the integrity of self-representation on the internet collapses, creating a condition of chronic suspicion that spreads throughout the digital communication ecosystem, making every citizen vulnerable to these AI-based identity manipulation attacks.

3.3 Regulatory Void and the Challenges of the ITE Law

Although Indonesia possesses a legal framework for cyber violence through Law No. 1 of 2024 (Second Revision of the ITE Law) and is specifically regulated by Law No. 12 of 2022 concerning the Crime of Sexual Violence (UU TPKS), the response to AI-facilitated OGBV (KBGO-AI) still faces what is termed a cultural lag. Article 14 of the UU TPKS explicitly prohibits the recording, distribution, or possession of non-consensual sexual content. However, this regulation was designed in an era where the contested content was authentic material obtained illegally, not content entirely fabricated by AI.

The main challenge in law enforcement lies in the difficulty of proving the element of "authenticity" or "originality" of the material. Susilo (2022) notes that law enforcement often focuses on proving the legal fact that the material is a photo or video that "truly" depicts the victim. When the material is proven to be digitally "fake" or "engineered," cases often become weak or are directed towards defamation articles, which do not encompass the dimension of sexual violence. The argument that engineered material is not a "legal fact" disregards the communication principle that the social consequences of such simulacra are real, resulting in impunity for deepfake perpetrators.

Therefore, a shift in the legal paradigm is necessary, moving from a content-based approach to a harm-based approach. The harm-based approach views the tangible damage suffered by the victim trauma, job loss, social stigma as the basis for prosecution, regardless of whether the content of the violence is authentic or synthetic. The reinterpretation of Article 14 of the UU TPKS is crucial; the clause must be expanded to cover acts of digital engineering intended to produce and disseminate non-consensual sexual imagery, even if the subject matter is a digital synthesis.

In addition to evidentiary challenges, regulation also faces issues of jurisdiction and platform accountability. Much of the KBGO-AI content is distributed through global encrypted messaging applications like Telegram, which are beyond the direct reach of Indonesian law. This necessitates stronger clauses for platform accountability, compelling service providers to adopt synthetic content detection technologies (such as watermarking) and expedite the takedown process for violating content. Arawinda (2021) emphasizes the importance of adopting strict digital ethics at the corporate level to close the regulatory loopholes exploited by KBGO-AI perpetrators.

CONCLUSION

AI-facilitated Online Gender-Based Violence (KBGO-AI) is the latest manifestation of misogyny in the digital space. The findings of this research confirm that generative technology, particularly deepfake and voice cloning, has surpassed the boundaries of conventional tools of violence. They function as a medium of communication that creates a hyperreality of victimization. Although the

distributed obscene content is a digital fabrication, the impacts of psychological trauma, social stigma, and reputational ruin experienced by the victims are real (hyper-real). This crisis fundamentally demonstrates the failure of the digital space to guarantee the ontological security of its users, as the integrity of digital identity can now be damaged merely by a few lines of algorithmic code.

The deepfake phenomenon must be viewed as a communication crisis, not merely a technological problem. Theoretically, generative technology forcibly collapses the privacy boundaries explained in CPM theory (Petronio, 2002), where basic biometric data can be stolen to construct a false identity. Furthermore, the existence of convincing synthetic content triggers the liar's dividend (Chesney & Citron, 2019), which erodes public trust in digital truth, making it difficult for victims to prove the material's falsehood. The most severe consequence is the chilling effect silencing the participation of female journalists, activists, and public figures which ultimately damages the principles of deliberative democracy and freedom of expression in the cyber realm.

Currently, Indonesia faces a significant cultural lag in its regulatory response to KBGO-AI. Although the ITE Law and the UU TPKS exist, law enforcement tends to fixate on proving the "authenticity" of the material, which is inherently difficult to prove in cases of synthetic content. The argument that engineered material is not a "legal fact" ignores the real damage suffered by the victim. Therefore, a shift in the legal paradigm is needed, moving from a content-based approach to a harm-based approach. The reinterpretation of Article 14 of the UU TPKS must be bold enough to cover acts of digital engineering intended to inflict suffering and degrade sexual dignity, without getting trapped in the debate over whether the material is real or fake.

To mitigate this threat, an integrated response from cross-sectors is required. First, the public must be equipped with critical digital literacy that focuses on the detection of synthetic content. Second, social media platforms have an ethical and technological responsibility to adopt strict detection and watermarking systems for every piece of generative content uploaded. Third, the government and legislature need to accelerate the formation of regulations that explicitly govern the liability of AI developers and define deepfake as a criminal act regardless of non-commercial intent. With these steps, the digital identity integrity of citizens can be restored, and cyberspace can become a safer and more inclusive arena.

DAFTAR PUSTAKA

Arawinda, S. H. (2021). Perlindungan Hukum terhadap Perempuan Korban Kekerasan Berbasis Gender Online di Indonesia. Jurnal Yustika: Media Hukum dan Keadilan, 24(2), 76-90.

Baudrillard, J. (1981). Simulacra and Simulation. Ann Arbor: University of Michigan Press.

Cangara, H. (2023). Etika Komunikasi: Menjadi Manusia yang Santun Berkomunikasi dalam Era Digital. Jakarta: Prenada Media.

Chesney, R., & Citron, D. K. (2019). Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security. California Law Review, 107, 1753-1820.

Citron, D. K. (2022). The Fight for Privacy: Protecting Dignity, Identity, and Love in the Digital Age. New York: W.W. Norton & Company.

Council of Europe. (2022). Combating rising hate against women in Europe. Strasbourg: Gender Equality Commission.

Komnas Perempuan. (2024). Catatan Tahunan (CATAHU) 2024: Kekerasan Terhadap Perempuan di Ruang Siber. Jakarta: Komisi Nasional Anti Kekerasan Terhadap Perempuan.

Mahardika, M. I. (2025). Tinjauan Yuridis Terhadap Pelaku Deepfake Porn Sebagai Kekerasan Gender Berbasis Online Menurut UU Pornografi. Lex Privatum, 14(5), 1-15.

Musyaffa, R. A., & Effendi, S. (2022). Kekerasan Berbasis Gender Online dalam Interaksi di Media Sosial. Komunikologi: Jurnal Ilmiah Ilmu Komunikasi, 19(2), 120-135.

Nasrullah, R. (2018). Riset Khalayak Digital: Perspektif Ekologi Media dan Komunikasi Budaya. Bandung: Simbiosa Rekatama Media.

Petronio, S. (2002). Boundaries of Privacy: Dialectics of Disclosure. New York: State University of New York Press.
SAFEnet. (2023). Laporan Situasi Hak-Hak Digital Indonesia: Represi Digital di Tahun Politik. Jakarta: Southeast Asia Freedom of Expression Network.

Samosir, C. P. (2023). Kekerasan Berbasis Gender Online (KBGO) dan Cara Mendukung Korban: Analisis Konten Film Like dan Share. Jurnal PIKMA : Publikasi Ilmu Komunikasi Media Dan Cinema, 6(1), 230-245.

Susilo, M. E. (2022). Tinjauan Yuridis Terhadap Deepfake Pornography Sebagai Bentuk Kekerasan Berbasis Gender Online. Jurnal Hukum & Pembangunan, 52(1), 1-15.

UN Women. (2025). AI-powered online abuse: How AI is amplifying violence against women and what can stop it. New York: United Nations Entity for Gender Equality and the Empowerment of Women.

Wahyuni, L., & Saputra, N. E. (2023). Resiliensi dan Ketahanan Psikologi: Studi Kasus Kekerasan Berbasis Gender Online (KBGO). Jurnal Ketahanan Nasional, 29(2), 1-15.