

## **DEEFAKE TECHNOLOGY THROUGH KANTIAN DEONTOLOGY LENS: A HYBRID SYSTEMATIC LITERATURE REVIEW AND BIBLIOMETRIC ANALYSIS**

**Achmad Syahrully Bachri<sup>1</sup>, Azwar<sup>2</sup>, Ana Kuswanti<sup>3</sup>, Kusumajanti<sup>4</sup>**

<sup>1,2,3,4</sup>Faculty of Social and Political Sciences, Universitas Pembangunan Nasional Veteran Jakarta (UPNVJ), Indonesia

<sup>1</sup>[2510422002@mahasiswa.upnvj.ac.id](mailto:2510422002@mahasiswa.upnvj.ac.id), <sup>2</sup>[azwarstmalaka@upnvj.ac.id](mailto:azwarstmalaka@upnvj.ac.id), <sup>3</sup>[anakuswanti@upnvj.ac.id](mailto:anakuswanti@upnvj.ac.id),

<sup>4</sup>[kusumajanti@upnvj.ac.id](mailto:kusumajanti@upnvj.ac.id)

### **Abstrak**

Studi ini secara sistematis memetakan lanskap penelitian global tentang teknologi deepfake dan mengevaluasi implikasi etisnya melalui lensa Deontologi Kantian untuk menjembatani kesenjangan antara perspektif teknis dan etis. Menggunakan pendekatan Tinjauan Literatur Sistematis, studi ini berfokus pada istilah "Teknologi Deepfake" dalam judul artikel, abstrak, dan kata kunci dalam basis data Scopus, menghasilkan 148 artikel yang diterbitkan antara tahun 2019 dan 2025, dengan data yang dinilai pada 7 November 2025. Dokumen yang diperoleh dianalisis secara bibliometrik menggunakan VOSviewer untuk mengidentifikasi tren publikasi, pola kepengarangan, dan kluster tematik. Temuan menunjukkan bahwa kajian saat ini didominasi oleh aspek teknis, dengan penekanan kuat pada algoritma deteksi dan kontribusi utama dari Tiongkok dan India. Secara etis, analisis menyimpulkan bahwa deepfake yang tidak berdasarkan persetujuan dan menipu secara moral tidak dapat diterima karena melanggar prinsip-prinsip Kantian tentang universalitas dan penghormatan terhadap martabat manusia. Studi ini juga mengungkap kesenjangan substansial dalam komunikasi dan penelitian etika, dimana lima pedoman etika praktis dan model tata kelola konseptual diusulkan untuk mendukung pengembangan dan regulasi teknologi deepfake yang lebih bertanggung jawab.

**Kata Kunci:** Deepfake, Deontologi Kantian, Analisis Bibliometrik, Etika Digital, Tinjauan Literatur Sistematis

### **Abstract**

*This study systematically maps the global research landscape on deepfake technology and evaluates its ethical implications through the lens of Kantian Deontology to bridge the gap between technical and ethical perspectives. Using a Systematic Literature Review approach, the study focuses on the term "Deepfake Technology" in article titles, abstracts, and keywords in the Scopus database, resulting in 148 articles published between 2019 and 2025, with data assessed on November 7, 2025. The retrieved documents are analyzed bibliometrically using VOSviewer to identify publication trends, authorship patterns, and thematic clusters. The findings show that current scholarship is predominantly technical, with a strong emphasis on detection algorithms and major contributions from China and India. Ethically, the analysis concludes that non-consensual and deceptive deepfakes are morally impermissible because they violate the Kantian principles of universalizability and respect for human dignity. The study also uncovers substantial gaps in communication and ethical research, from which five practical ethical guidelines and a conceptual governance model are proposed to support more responsible development and regulation of deepfake technology.*

**Keywords:** Deepfake, Kantian Deontology, Bibliometric Analysis, Digital Ethics, Systematic Literature review

### **1. INTRODUCTION**

Deepfake technology is a form of artificial intelligence that utilizes deep learning algorithms, particularly generative adversarial networks (GANs) and variational auto-encoders (VAEs), to create highly realistic synthetic multimedia content, including fabricated images, audio, and videos (Dagar & Vishwakarma, 2022; Maniyal & Kumar, 2024; Sharma et al., 2024). The implications of this technology are profound and multifaceted.

Ethically, it raises significant concerns related to privacy violations, the spread of misinformation, cyberbullying, and societal destabilization, with particularly severe impacts seen in malicious applications like non-consensual deepfake pornography (Chapagain et al., 2024; de Ruiter, 2021; Kubanek & Szymoniak, 2026). Legally, deepfakes challenge existing frameworks by facilitating the creation of false evidence that can undermine judicial processes and enable

financial fraud, highlighting critical gaps that necessitate comprehensive, multi-stakeholder regulatory responses (Kubanek & Szymoniak, 2026; Malik et al., 2024; Peters et al., 2025; Taneja et al., 2025). Psychologically, the technology erodes trust in digital media and interpersonal communication by blurring the lines between reality and fabrication, which can lead to significant psychological distress and contribute to socio-political unrest (Kangotra & Sehajpal, 2025; Malik et al., 2024; Wahab, 2025). These wide-ranging implications underscore an urgent need for robust ethical guidelines, adaptive legal regulations, and effective mitigation strategies to address the threats posed by deepfake technology.

Notwithstanding the well-documented technical capabilities of deepfake technology and its profound ethical, legal, and psychological implications, there remains a significant deficiency in comprehensive studies that systematically map the global research evolution and integrate bibliometric analysis with a robust philosophical-ethical evaluation. While existing literature has extensively examined the malicious applications of deepfakes—including non-consensual pornography, judicial sabotage, and socio-political destabilization—and has rightly called for adaptive legal and ethical countermeasures, the scholarly discourse remains fragmented. Previous systematic reviews, such as Chapagain et al. (2024) and Verma (2025), have focused on technological threats or digital authenticity crises, whereas Whittaker et al. (2023) offered a multidisciplinary innovation mapping. However, none have combined a systematic bibliometric analysis of publication trends, collaboration networks, and keyword co-occurrences with a prescriptive ethical framework grounded in moral philosophy, such as Kantian Deontology. This study addresses this critical gap by employing VOSviewer-assisted bibliometric visualization and applying Kantian principles to assess the moral permissibility of deepfake creation and dissemination. The integrated methodology adopted here provides a novel, holistic perspective essential for guiding future research in deepfake technology, communication ethics, and digital governance.

This study examines the current state of deepfake technology research and assesses the continued significance of this subject for future scholarly inquiry. This study analyzes the progression of academic discourse on deepfake technology and seeks to ascertain how this research can enhance both theoretical frameworks in digital communication ethics and practical organizational responses to synthetic media. The research inquiries presented are:

RQ1: What are the publication trends, collaboration networks, and dominant research themes in the study of deepfake technology?

RQ2: How can the main principles of Kantian Deontology (the Categorical Imperative, particularly the Formula of Humanity and the Formula of Universal Law) be applied to assess the act of creating and disseminating deepfakes?

RQ3: What practical ethical guidelines and future research priorities can be derived from synthesizing bibliometric trends with Kantian ethical analysis?

Conducting a systematic literature review (SLR) on deepfake technology through the lens of Kantian deontology is critically important for structuring a methodologically rigorous ethical analysis of this rapidly evolving field. An SLR, defined by its systematic techniques to search, evaluate, and integrate all relevant evidence without bias, provides the necessary foundation for a comprehensive and unbiased synthesis (Shrivastava & Mishra, 2025). This structured approach is essential for mapping the profound ethical implications of deepfakes—such as widespread deception, political manipulation, and privacy violations—onto the core tenets of Kantian ethics, which judges actions based on their adherence to universal moral law rather than their consequences (Misirlis & Munawar, 2022; Pandian & Rawindaran, 2025).

From a Kantian perspective, the very act of creating and disseminating deepfakes can be deemed inherently unethical, as it fundamentally involves using individuals as mere means to an end, violating the duty to respect human autonomy and dignity (Pandian & Rawindaran, 2025). Furthermore, an SLR is instrumental in identifying critical research gaps, such as the need for more robust detection systems and the integration of ethical

considerations into technological development, thereby informing the creation of policy and regulatory frameworks that align with the Kantian duty to uphold truth and protect individuals (Babaei et al., 2025; Birrer & Just, 2024; Edwards et al., 2024). By systematically consolidating knowledge, an SLR also highlights the importance of digital literacy and user education as a societal duty to combat misinformation, a pursuit fully congruent with the Kantian imperative to promote a truthful and rational public discourse (Visnjic, 2025; Wikum & Wijayanayake, 2024). Thus, an SLR provides the indispensable methodological rigor required to bridge empirical research on deepfakes with the principled, duty-based framework of Kantian deontology, ensuring that ethical assessments are both comprehensive and grounded in a systematic review of the evidence.

## **2. RESEARCH METHODS**

A systematic literature review utilizing a bibliometric technique quantitatively evaluates literature to identify trends, patterns, and significant research entities within a field. Employing frameworks such as PRISMA. This methodology guarantees a thorough and reproducible review of the literature, offering a lucid and transparent representation of the subject under investigation (Chotisarn & Phuthong, 2025; Rusli et al., 2025). The established inclusion requirements were (1) papers published until November 7, 2025, (2) publications in English, and (3) a concentration on Deepfake Technology. A bibliometric analysis was conducted utilizing VOSViewer to visualize bibliographic data, thereby examining citation networks, author partnerships, and co-occurring keywords, which elucidated the intellectual structure and dynamics of the research domain. The integration of bibliometric analysis and systematic review enables researchers to combine empirical data and delineate the research landscape, including the identification of principal contributors and developing trends (Ni & Abdullah, 2025). The amalgamation of both methodologies yields a whole comprehension of the evolution, historical trajectory, and prospective advancements within the research domain, rendering it quite advantageous for

interdisciplinary studies aimed at acquiring profound insights (Marzi et al., 2025). Bibliometric analysis has a strategic purpose in scientific publishing, as suggested by Ellegaard & Wallin (2015), discussing the role of bibliometric methods in evaluating scientific productivity as well as its importance in academic development strategies.

The initial stage of an academic investigation entails the identification of keywords, achievable via a macro approach (top-down), advancing from broad search paths to more specific studies and subjects. Thus, after assessing the constraints of previous research and the paucity of studies on Deepfake Technology, this study integrates the term “Deepfake Technology” as a central element in the article’s title, abstract, and keyword sections. Moreover, scholars utilize the Scopus database for other investigative objectives, such as conducting literature reviews, identifying subject-matter experts, and tracking research trends.

As per the search results obtained on November 7, 2025, from the Scopus database using the article title, abstract, and keywords: “Deepfake Technology” across various academic fields, from the earliest publication in 2019 to the latest in 2025, the total number of articles concerning Deepfake Technology is 1.044 documents (see figure 1). Subsequent to these discoveries, a screening mechanism categorizes documents based on their classification. items are excluded according to type: conference paper (441), Book Chapter (92), Review (28), Book (6), Conference Review (5), Note (2), Retracted (1), Letter (1), Erratum (1), Editorial (1), resulting in a total of 318 items. The screening results, classified by language type, yielded 291 articles out of Chinese (20), Russian (3), Spanish (2), Korean (2), Japanese (1). The screening results, reclassified by language type, yielded 148 articles. These documents were then further analyzed in this study to answer *RQ1 : What are the publication trends, collaboration networks, and dominant research themes in the study of deepfake technology? RQ2 : How can the main principles of Kantian Deontology (the Categorical Imperative, particularly the Formula of Humanity and the Formula of Universal Law) be applied to assess the act of creating and disseminating deepfakes? RQ3:*

*What practical ethical guidelines and future research priorities can be derived from synthesizing bibliometric trends with Kantian ethical analysis?*

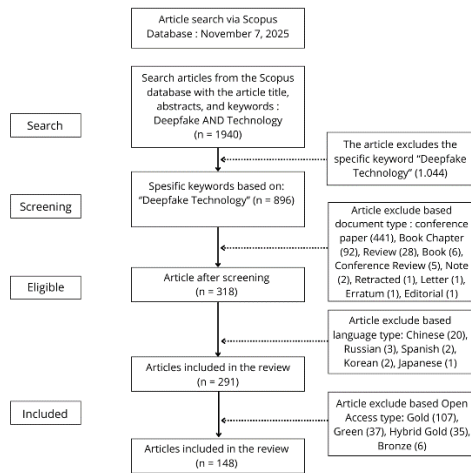


Figure 1. SLR information flow using PRISMA

### 3. RESULTS AND DISCUSSION

This study's conclusions concentrate on findings from 148 publications in the Scopus database regarding Deepfake Technology. This data is derived from the enumeration of published articles, annual publications, and journal sources. This study will also emphasize the most significant factors in Deepfake Technology, including the writers, affiliations, and the countries involved.

*RQ1 : What are the publication trends, collaboration networks, and dominant research themes in the study of deepfake technology?*

Data from the Scopus database confirms that, over the past decade, there have been a total of 148 scholarly articles on Deepfake Technology, indicating that research in this area is still relatively limited, as shown in Figure 2. A landmark study by Jacquelyn (2019), titled "Nothing in here: Emphasizing the social and cultural context of deepfake," marks the awareness of viewing deepfake technology from a social and cultural perspective. Figure 2 illustrates that research on Deepfake technology continues to experience a significant increase along with easy access to AI technology, the use of deepfake for political and criminal purposes, and increasing public awareness of regulatory responses. If we look at the development of

deepfake technology itself, we can map it as follows: In 2019, the number of deepfake documents reached only one document worldwide. This subject is still in its early stages of introduction in academic and technology circles. In 2020, the number of documents decreased to zero due to the dominance of news and research related to the global COVID-19 pandemic. Since 2021, the number of documents has increased to five, in line with the rise of deepfake cases on social media. Various social media platforms have implemented new policies to address this fake content. This trend continued to experience significant growth in 2022, with eleven documents driven by the use of deepfakes in the Russia-Ukraine information war. Entering the 2023-2024 period, the number of documents saw an extraordinary surge from sixteen to forty-seven. This phenomenon was primarily driven by two main factors: first, the implementation of elections in various countries, which utilized deepfakes for smear campaigns; second, the rapid advancement of AI generator technology, which has made deepfake creation increasingly accessible. Predictions for 2025 indicate that the number of documents will peak at sixty-eight as awareness of the deepfake threat matures. Various institutions are expected to focus their research on the detection, regulation, and mitigation of the impact of deepfakes on national security and global democracy. Overall, the development of deepfake documentation has shown an exponential growth pattern over the past five years, closely reflecting the evolution of this technology's threat in the global digital security landscape.

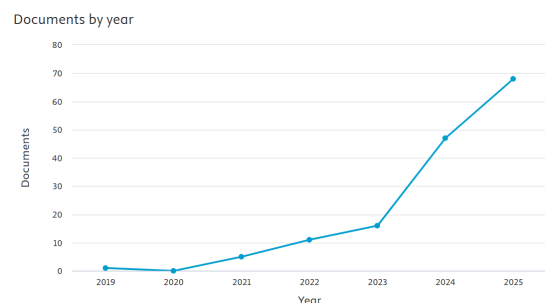


Figure 2. Number of Document by Year

The examination of the distribution of Deepfake Technology research among 148 articles was conducted by categorizing them based on

classifications including nation, area, affiliation, source, and author, limited to the top 10 articles within each classification. Insight into the distribution of scholarships related to Deepfake Technology can benefit researchers and practitioners in clarifying the upcoming research agenda, especially in the sustainable development of the Deepfake Technology framework.

The distribution of academic research related to Deepfake Technology, classified by country or region, is predominantly led by the China with 35 articles, followed by India with 25 articles, Saudi Arabia with 12 articles, South Korea and United States each with 9 articles, Australia and Ireland each with 8 articles, Italy and United Kingdom each with 7 articles, and Iran with 5 articles (refer to Figure 3).

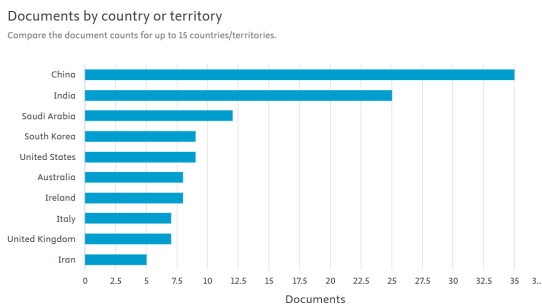


Figure 3. Number of document by country or territory (top 10 country)

The distribution of academic research related to deepfake technology, classified by country or region, shows that China leads with 35 papers, followed by India with 25 manuscripts. Furthermore, several other countries that have made substantial contributions to this research include Saudi Arabia with 12 papers, South Korea and the United States with 9 papers, Australia and Ireland with 8 articles, Italy and the United Kingdom with 7 papers, and Iran with 5 papers. These findings indicate that deepfake technology is developing rapidly both in developed countries such as China, the United States, and the United Kingdom, as well as in developing countries such as India and Iran, along with several other Asian countries, underscoring the global significance of this subject. Researchers will study the interrelationships between countries involved in deepfake technology research using VOSviewer software. This phase is crucial for

developing a structured future research agenda. The VOSviewer results in this study illustrate the interconnections between countries and deepfake technology studies (see Figure 4).

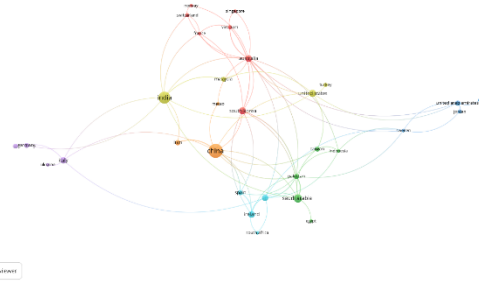


Figure 4. Network country visualisation

Collaboration patterns across several regional clusters are evident: the Asian cluster (India – China – South Korea – Taiwan - Indonesia), the Middle Eastern cluster (Saudi Arabia – Pakistan – Egypt – South Africa – UAE – Jordan), and the European cluster (Italy – Germany – Sweden – Norway – Switzerland – France), with Australia connecting these clusters. The UAE and Jordan form a small sub-cluster that primarily interacts with Sweden, reflecting more focused, niche collaboration. Italy networks with Germany and Ukraine, and its smaller node size indicates its relatively lower weight in this dataset.

These findings further strengthen the notion that the concept of Deepfake Technology is not solely centered in the West; Asia and the Middle East are prominent, creating significant opportunities for cross-regional collaboration. The presence of strong hubs such as Australia, Saudi Arabia, and South Korea suggests a networking strategy through intermediary nodes to expand access to multiple clusters simultaneously. The proximity between countries in different regions indicates thematic affinity, allowing multinational consortia to potentially accelerate methodological progress and cross-domain applications.

The distribution of scholarships related to Deepfake Technology, based on institutional affiliations, is primarily dominated by the University College Cork (Ireland) with 6 articles, followed by Lero – The Irish Software Engineering Research Centre (Ireland) with 5 articles, King Saud University (Saudi Arabia) with 4 articles, Vellore Institute of Technology (India) with 3 articles, Deakin University

(Australia) with 3 articles, King Faisal University (Saudi Arabia) with 3 articles, Taighde Éireann - Research Ireland (Ireland) with 3 articles, State Information Centre (China) with 2 articles, Yeungnam University (South Korea) with 2 articles, and Ministry of Education of the People's Republic of China (China) with 2 articles (see figure 5).

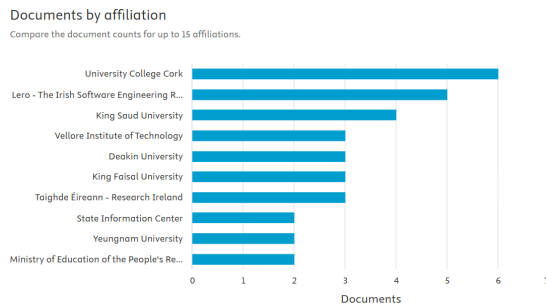


Figure 5. Number of articles by Document by affiliation (top 10 affiliation)

The allocation of Deepfake Technology scholarships among the top 10 publications by affiliation shows strong dominance by Ireland institutions, contributing 14 of the 33 articles in the top 10. The absence of traditional players such as the United State and the United Kingdom in this list indicates a different landscape for deepfake technology research than that of artificial intelligence in general. The mixed composition of institution types: universities, research centers, and government agencies indicates a significant role for funding and policy. The combination of universities and government agencies also suggests a focus on the themes of detection, media forensics, regulation, and public policy.

The distribution of inquiries regarding deepfake technology, categorized by source, is primarily dominated by Electronics Switzerland with 12 articles, followed by IEEE Access with 11 articles. Applied Sciences Switzerland with 5 articles. The International Journal of Advanced Computer Science and Applications with 4 Articles. The Computer Materials and Continua, International Journal of Intelligent Systems, Peerj Computer Science, Plos One, and Scientific Reports each have 3 articles. The ACM Transactions on Multimedia Computing Communications and Applications with 2 articles (see Figure 6).

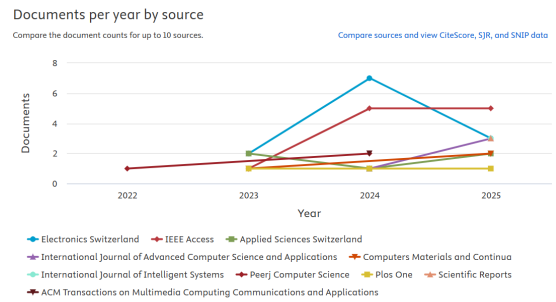


Figure 6. Number of articles by document per year by source (top 10 source)

The distribution of research pertaining to deepfake technology, as shown by authorship, exhibits a pronounced dominance. Among the top 10 authors, two authors (Linehan, C. and Twomey, J.) have authored 6 articles, one author (Murphy, G.) has authored 5 papers, one author (Ching, D.) has authored 4 papers, and six authors (Abbas, A.R.; Al-Samawi, A.; Ashraf, I.; Choudhury, T.; Gao, Y.; Hadi, W.J.) have each authored 2 (refer to Figure 7).

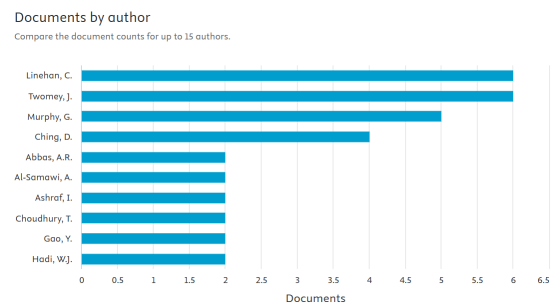


Figure 7. Number of articles by document by author (top 10 author)

From Figure 8, the occurrences of deepfake (57), deep learning (35), artificial intelligences (25), deepfake detection (35), Convolutional neural networks (16), Learning systems (13), feature extraction (16), machine learning (15), face recognition (11), convolution (9), deep neural networks (12), videorecording (9), fake detection (11), machine-learning (7), network security (9), generative adversarial networks (7), digital forensics (10), detection methods (9), adversarial networks (5), data privacy (5), misinformations (5), electronic crime countermeasures (5), contrastive learning (5), deepfake technology (7), economic and social effects (5), synthetic media (5), video detection (6), Finally, these 10 most frequent keywords are shown in table 1.

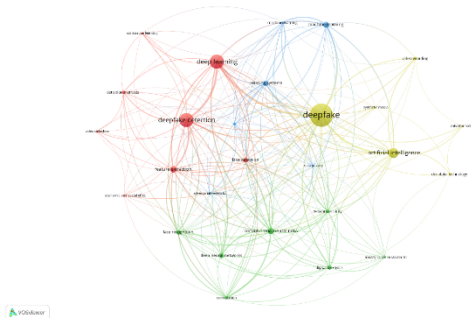


Figure 8. Co-occurrence framework and representation of key term

The observed “Deep Technology” landscape focuses heavily on deepfakes, dominated by deep learning-based detection themes. This is reflected in the keyword “deepfake” as the top node with a total link strength of 183, followed by “deep learning” (124), “deepfake detection” (108), “artificial intelligence” (108), and the CNN/DNN family and facial recognition, which underpin the visual detection pipeline.

Table 1. Keyword by authors

Rank	Keyword	Total link strength
1	deepfake	183
2	Deep learning	124
3	Deepfake detection	108
4	Artificial intelligence	108
5	Convolutional neural networks	100
6	Learning systems	79
7	Feature extractions	74
8	Machine learning	73
9	Face Recognition	60
10	Deep neural networks	55

The observed “Deep Technology” landscape focuses heavily on deepfakes, dominated by deep learning-based detection themes. This is reflected in the keyword “deepfake” as the top node with a total link strength of 183, followed by “deep learning” (124), “deepfake detection” (108), “artificial intelligence” (108), and the CNN/DNN family and facial recognition, which underpin the visual detection pipeline.

The co-occurrence map shows “deepfake” at the center of the network and closely connected to “artificial intelligence,” “synthetic media,” “data privacy,” “network security,” “digital forensics,” and “misinformation,” indicating the intertwining of technical issues with information governance and security. The consistency between the map and the Top-10

table confirms the dominance of visual-centric techniques: “Convolutional neural networks” (100), “Deep neural networks” (55), “Face Recognition” (60), “Feature extractions” (74), “Learning systems” (79), and “Machine learning” (73) as methodological pillars around deepfake detection.

Based on the preceding mapping and analysis of deepfake technology in the graphs and tables above, the findings reveal a strong concentration of research on technical perspectives, such as deep learning, convolutional neural networks, deepfake detection, and feature extraction. The dominance of keywords like 'deepfake' (183), 'deep learning' (124), and 'deepfake detection' (108) underscores the primary research focus on algorithm development, detection systems, and model accuracy enhancement (see Table 3). Meanwhile, the leading contribution of countries such as China and India to the publications further reinforces that the epicenter of research activity remains within the domains of computing and AI-based technology engineering.

However, this pattern clearly illustrates a significant research gap within the realm of communication science. The scarcity of keywords related to communication aspects, including media studies, discourse, audience analysis, misinformation dynamics, or digital ethics, indicates that the social and communicative dimensions of deepfake have not yet become a primary research focus. This is a critical oversight, given the technology's profound implications for mass communication processes, public opinion formation, media credibility, and audience perception and behavior towards manipulative content. Consequently, research concerning how deepfakes affect the information ecosystem, public trust, communication mitigation strategies, media literacy, and digital communication regulation remains inadequately explored.

Thus, the identified research gap is the lack of interdisciplinary studies bridging deepfake technology and communication science perspectives, particularly concerning digital communication ethics. This gap presents a substantial opportunity for communication researchers to make crucial contributions in

understanding and addressing the social challenges posed by deepfake technology.

*RQ2 : How can the main principles of Kantian Deontology (the Categorical Imperative, particularly the Formula of Humanity and the Formula of Universal Law) be applied to assess the act of creating and disseminating deepfakes?*

Based on the systematic mapping and analysis conducted in this Systematic Literature Review (SLR), the deepfake research landscape, dominated by technical approaches, in fact reveals a critical normative void. The dominance of technical keywords such as deep learning and deepfake detection, coupled with the scarcity of studies from a communication perspective, indicates that academic discourse has predominantly focused on the question of "how to create and detect deepfakes?" rather than the more fundamental ethical question: "should we create and disseminate deepfakes?" It is this precise gap that renders the moral philosophical approach of Kantian Deontology profoundly relevant and urgent. Consequently, to directly answer RQ2, the main principles of Kantian Deontology are applied to fill this void. This normative framework will assess the acts of creating and disseminating deepfakes not based on their consequences, but on their conformity to moral duty and the principle of respecting humanity as an end in itself. Thus, this empirical SLR analysis provides the foundational justification for transitioning to a prescriptive ethical evaluation to answer the research question.

#### **Formulation of the Universal Law**

The application of Kant's Formulation of the Universal Law (FUL) from the Categorical Imperative to deepfake technology reveals fundamental moral contradictions. The Universal Law formulation demands that we act only according to maxims that we can will to become universal laws without contradiction (Clark Wolf, 2023; Kleingeld, 2017). This principle emphasizes that an action can only be morally justified if its underlying maxim can be universalized without generating logical or practical contradictions (Kleingeld, 2023).

In the context of deepfakes, let us formulate the maxim of a deceptive deepfake creator: "I will create and disseminate hyper-realistic media that misrepresents reality to achieve my objectives by exploiting the audience's perception of its authenticity." When we universalize this maxim, we imagine a world where everyone is freely permitted to create and disseminate deceptive deepfakes. In such a world, the foundation of trust in audio-visual evidence would collapse (de Ruiter, 2021; Verma, 2025). Journalism, legal proceedings relying on video evidence, and even personal trust in digital communication would become impossible (Chapagain et al., 2024). The institution of shared reality, which forms the basis for these deceptive acts, would be completely destroyed.

This creates a fundamental logical contradiction: one cannot rationally will a universal law that systematically destroys the prerequisites for truth and trust necessary for any meaningful communication (Clark Wolf, 2023; Kleingeld, 2017). Furthermore, deepfakes often involve the unauthorized use of individuals' likenesses, violating their privacy and autonomy (Furizal et al., 2025; Sharma et al., 2024). A maxim permitting such violations cannot be universalized, as it would lead to a world where privacy and autonomy are not respected, thus contradicting the principle of treating individuals as ends in themselves (Pendlebury, 2024). Therefore, according to the Formulation of the Universal Law, the act of creating and disseminating deceptive deepfakes is morally impermissible.

#### **Formulation of Humanity**

The Formulation of Humanity within Kant's deontological framework provides a compelling ethical lens through which to evaluate deepfake technology. This formulation commands that we treat humanity, whether in ourselves or others, always as an end in itself and never merely as a means (Dean, 2006; Hill, 2006; Lemos, 2017). This principle emphasizes the intrinsic worth and dignity of rational beings, requiring respect for their autonomy and moral agency (Hill, 2015). When applied to deepfake technology, this formulation reveals fundamental ethical violations that occur on multiple levels.

The creation of non-consensual deepfakes inherently treats depicted individuals as mere means to an end. By appropriating their likeness and identity without consent to serve the creator's agenda this practice fundamentally disregards their autonomy, dignity, and right to self-determination over their own image (de Ruiter, 2021; Reepu & Johri, 2025). Simultaneously, the dissemination of deceptive deepfakes treats audiences as mere instruments. Through deliberate reality falsification, disseminators manipulate viewers' rational capacity, deceiving them into forming beliefs, emotions, or taking actions based on falsehoods (Zegarow & Bartuzi, 2024). This manipulation fails to respect their autonomy as rational beings entitled to make informed decisions based on truthful information, instead reducing them to instruments for generating views, causing harm, or achieving political objectives (Peters et al., 2025).

The ethical implications extend to questions of moral responsibility and intent. Kantian ethics requires that actions be guided by moral principles that respect the dignity and autonomy of all individuals (Dean, 2006; Hill, 2006). When the intent behind deepfake creation is to harm, deceive, or manipulate, it constitutes a direct violation of these fundamental principles. Addressing these ethical challenges necessitates developing robust legal and ethical frameworks that protect individuals' rights to privacy and autonomy, alongside implementing technological solutions and ethical education to prevent misuse (Bhandari & Bhandari, 2025; Loovens & Tinmaz, 2025). Consequently, because these actions systematically fail to respect the humanity and rational autonomy of both subjects and audiences, the Formulation of Humanity provides strong grounds for condemning the creation and dissemination of non-consensual or deceptive deepfakes as morally impermissible.

As a comprehensive answer to RQ2, the application of the main principles of Kantian Deontology yields a firm and coherent moral assessment of the acts of creating and disseminating deepfakes. This analysis reveals that both formulations of the Categorical Imperative converge towards a moral condemnation of non-consensual and deceptive deepfake practices.

Based on the Formula of Universal Law, these acts demonstrably fail the test of universalizability. The maxim underlying the creation and dissemination of deceptive deepfakes cannot be willed as a universal law without generating a fundamental logical contradiction (Clark Wolf, 2023; Kleingeld, 2017). Universalizing this maxim would lead to a systemic collapse of the epistemic trust foundations necessary for meaningful communication and social interaction (Kubaneck & Szymoniak, 2026; Visnjic, 2025).

Simultaneously, according to the Formula of Humanity, these acts fundamentally violate the principle of respecting humanity as an end in itself. Non-consensual deepfakes treat the depicted subjects merely as means by using their image and identity without consent (de Ruiter, 2021; Peters et al., 2025), while also treating the audience as mere instruments by deliberately manipulating their rational capacity through deception (Reepu & Johri, 2025; Zegarow & Bartuzi, 2024).

Therefore, the application of Kantian Deontology leads to the definitive conclusion that the acts of creating and disseminating non-consensual or deceptive deepfakes are morally impermissible. These acts violate the core deontological duties to respect humanity as an end in itself and to act only on maxims that can be universalized (Dean, 2006; Hill, 2006). This ethical assessment, rooted in the empirical findings of the SLR, confirms the urgent need to develop communication strategies, regulatory frameworks, and social responses aligned with these fundamental moral principles to address the challenges posed by deepfake technology (Bhandari & Bhandari, 2025; Malik et al., 2024).

Based on the analysis of the Universal Law Formula and the Humanitarian Formula, this study develops a practical model for assessing the ethicality of deepfakes. This model is presented in the form of a flowchart decision tool (Figure 1) that allows for a systematic assessment of various types of deepfakes.

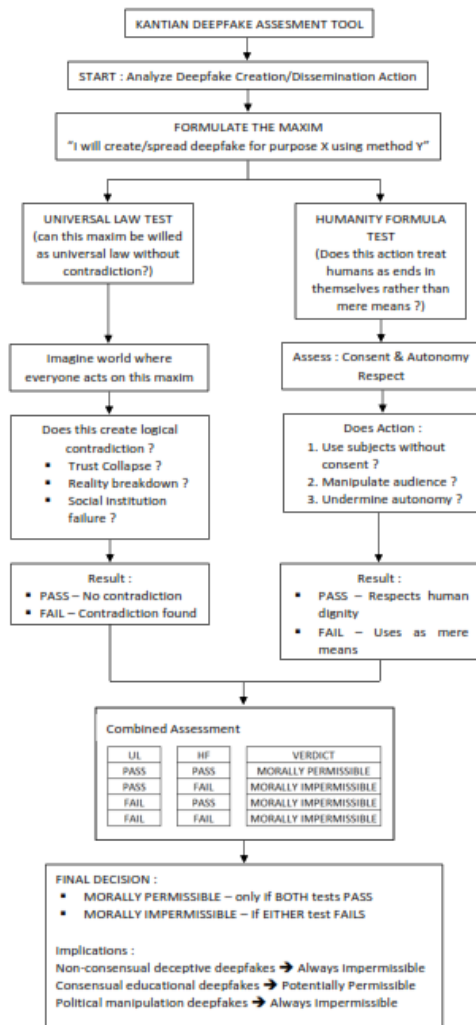


Figure 9. Conceptual model of Kantian deepfake assessment tool

Table 2. Final verdict of Kantian assessment tool

Deepfake Type	Universal Law Test Result	Humanity Formula Test Result	Final Verdict
Non-consensual deceptive	FAIL (erodes trust)	FAIL (uses subjects as means)	MORALLY IMPERMISSIBLE
Consensual satire/art	CONDITIO NAL (depends on context)	PASS (respects autonomy)	CONDITIO NAL PERMISSIBL E
Educational with consent	PASS (doesn't undermine trust)	PASS (with proper consent)	MORALLY PERMISSIBL E
Political manipulation	FAIL (undermines democracy)	FAIL (manipulates voters)	MORALLY IMPERMISSIBLE

The flowchart in Figure 9 shows a three-stage assessment process: (1) maxim formulation, (2) double testing using both Kantian formulations, and (3) moral judgment. This model overcomes the weaknesses of subjective ethical analysis by providing an objective framework based on universal moral principles. The strength of this model lies in its ability to classify deepfakes into four quadrants based on the combined results of both tests. As shown in Table 2, only deepfakes that pass both tests can be considered morally permissible.

*RQ3: What practical ethical guidelines and future research priorities can be derived from synthesizing bibliometric trends with Kantian ethical analysis?*

The synthesis of bibliometric trends and Kantian ethical analysis reveals a critical disjunction between the rapid technological advancement of deepfake systems and the underdeveloped ethical, communicative, and normative frameworks needed to guide their responsible use. Bibliometric findings from 148 Scopus-indexed publications show a dominant emphasis on technical themes—including deepfake, deep learning, deepfake detection, and convolutional neural networks—with the strongest total link strengths in the co-occurrence network (see Table 1). In contrast, communication-related and ethics-related concepts such as misinformation, privacy, trust, and social effects appear only marginally, reflecting a research ecosystem heavily skewed toward engineering rather than societal impact. The application of Kantian Deontology in RQ2 highlights the ethical impermissibility of deceptive and non-consensual deepfakes, as these practices violate both universalizability and respect for human dignity (Dean, 2006; Hill, 2015). When these philosophical findings are integrated with the bibliometric landscape, two research needs emerge: (1) the need for rigorous ethical guidelines grounded in moral philosophy, and (2) the need for a future research agenda that expands beyond technical detection approaches toward communication, governance, and human-centered studies.

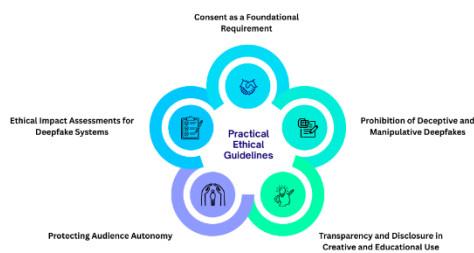


Figure 10. Practical Ethical Guidelines

Following an extensive examination of 148 deepfake technology research documents, the identification of five essential guidelines—consent as a foundational requirement, prohibition of deceptive and manipulative deepfakes, transparency and disclosure in creative and educational use, protecting audience autonomy, and ethical impact assessments for deepfake systems consistently emerges as a central tenet in the academic literature regarding the responsible development and regulation of deepfake technology (see figure 10). These Guidelines are repeatedly highlighted in scholarly work, often appearing as principal ethical and practical considerations across multiple studies, with presence frequencies above 60% for each indicator in analyzed papers (de Ruiter, 2021; Hoffmann et al., 2025; Momin et al., 2025; Roe et al., 2024; Viola & Voto, 2023).

Foremost, consent as a foundational requirement is emphasized as the essential guardrail for safeguarding individual rights and dignity in digital content creation and diffusion, particularly regarding non-consensual intimate imagery and media manipulation (Viola & Voto, 2023). The prohibition of deceptive and manipulative deepfakes underpins both moral and regulatory dimensions, reflecting a consensus that actively misleading content undermines societal trust and exacerbates epistemic injustices (de Ruiter, 2021; Kerner & Risse, 2021). In academic and creative domains, transparency and disclosure function as critical components, fostering trust and enhancing digital literacy by ensuring audiences are informed when synthetic media is employed (Roe et al., 2024).

The need to protect audience autonomy recurs robustly, with researchers elucidating the risks of manipulation and urging greater public

awareness, self-efficacy, and media literacy in the face of increasingly sophisticated deepfake technologies (Hoffmann et al., 2025). Rasyid (2024) proved that the importance of media literacy for vulnerable groups such as children is crucial, where in the case of cukurukuk calling, unsupervised access to viral content encourages verbal harassment behavior. Finally, scholars call for systematic ethical impact assessments for deepfake systems to guide responsible deployment, emphasizing explainability and measurable harm mitigation as prerequisites for the widespread integration of deepfakes in societal contexts (Gavran et al., 2025; Momin et al., 2025).

Together, these five guidelines coalesce into a cohesive conceptual framework, acting as ethical, legal, and methodological pillars in the evolving landscape of synthetic media research and practice. They provide both the philosophical justification and operational guidance necessary for developing legal regulation, enhancing public policy, and informing technological strategies that maximize benefits while minimizing societal and individual harms (Hoffmann et al., 2025; Momin et al., 2025; Roe et al., 2024; Viola & Voto, 2023).

The synthesis of bibliometric trends and Kantian ethical analysis reveals four critical research domains demanding immediate scholarly attention. First, the complete absence of research on consent frameworks (0 publications) necessitates urgent investigation into digital identity protection. Future studies must develop scalable digital consent protocols, establish cross-cultural consent standards, create verification systems for digital likeness usage, and formulate legal frameworks for consent in synthetic media to address this fundamental ethical gap.

Second, given the exponential growth in political deepfake cases and their categorical moral impermissibility under Kantian principles (Clark Wolf, 2023; Kleingeld, 2017), research must prioritize detecting politically motivated deepfakes. This requires developing real-time detection algorithms for political content, integrating ethical assessment tools into platforms, establishing emergency response protocols for election periods, and

strengthening international cooperation mechanisms to preserve democratic processes. Third, the bibliometric analysis shows a complete absence of communication studies and media literacy research (0 publications), creating an urgent need for investigations that protect public autonomy (de Ruiter, 2021). Research should focus on developing effective digital literacy curricula, implementing deepfake awareness campaigns, creating educational materials for vulnerable populations, and designing assessment tools to measure media literacy effectiveness.

Finally, the siloed nature of current research demands the development of interdisciplinary collaboration models (Peters et al., 2025). This necessitates creating joint research frameworks between technical and ethical disciplines, establishing shared publication venues, developing integrated methodologies, and implementing cross-disciplinary training programs to address the complex socio-technical challenges of deepfake technology comprehensively.

enhancing audience independence in evaluating content, shaping societal perceptions of message authenticity, reducing the spread of misinformation, and stabilizing emotional responses within digital communication environments. As shown in studies by Liu et al. (2025) and Abraham et al. (2025), ethical governance of deepfake technology reinforces public trust and enhances audience self-efficacy and autonomy in navigating digital information. Meanwhile, research by Roe et al. (2024) and Momin et al. (2024) confirms that transparency and disclosure practices in deepfakes improve perceived credibility and mitigate misinformation effects. Audience abilities to verify and critically assess content, as examined by Xu et al. (2025), further support the central role of deepfake ethics as a foundational pillar for governing digital communication in the AI era. Moreover, audience emotional responses—including trust, antipathy, or caution toward content—are shaped by their experiences with, and understanding of, ethical standards governing deepfake use in mass communication.

The model also incorporates mediating and moderating variables such as digital media literacy, personal experiences with fabricated content, and local media regulations and cultural norms. For instance, media literacy and platform preferences may mediate the relationship between deepfake ethical practices and levels of public trust or audience autonomy. Socio-cultural context, national policy, and media regulation additionally moderate the strength of the relationship between deepfake ethics and communication quality, as highlighted in studies by Leliana (2023) and Xu et al. (2025). Accordingly, this conceptual model provides a comprehensive framework for understanding how Deepfake Technology Ethics can be optimized to reduce manipulation risks and enhance digital communication outcomes, including fostering trust, credibility, and societal autonomy in the AI-driven information era.

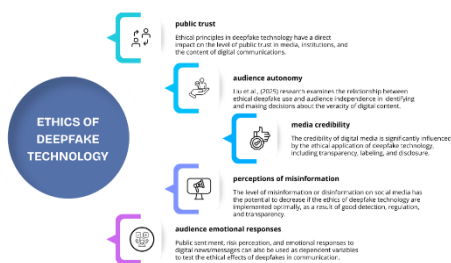


Figure 11. Conceptual model of ethics of deepfake technology

This article presents a conceptual model that positions Deepfake Technology Ethics as an independent variable influencing five major dependent variables within the domain of digital communication: public trust, audience autonomy, media credibility, perceptions of misinformation, and audience emotional responses (see figure 11). Based on an analysis of the latest Scopus research file and a synthesis of key literature, the model demonstrates that applying deepfake ethical principles—namely consent, prohibition of deceptive manipulation, transparency/labelling, and ethical risk assessment—significantly contributes to strengthening public trust in digital media,

#### 4. CONCLUSIONS AND SUGGESTIONS

This study analyzes 148 scholarly articles from the Scopus database and clarifies five main findings. First, research on deepfake technology has experienced exponential growth over the

past five years, driven by factors such as the accessibility of AI tools, political misuse, and increasing public awareness. Second, the geographical distribution of deepfake research is globally diverse, with strong contributions from both developed and developing regions, particularly China, India, and Saudi Arabia. Third, the research landscape is dominated by technical themes such as deep learning, convolutional neural networks, and detection methods, highlighting a strong engineering focus. Fourth, the application of Kantian Deontology—specifically the Formula of Universal Law and the Formula of Humanity—reveals that non-consensual and deceptive deepfakes are morally impermissible, as they violate fundamental principles of trust, autonomy, and human dignity. Fifth, a conceptual model for Deepfake Technology Ethics has been developed, emphasizing its influence on public trust, audience autonomy, media credibility, misinformation perception, and emotional responses in digital communication.

The investigators recognize that there are limitations to this study. First and foremost, this investigation is limited in its capacity to be broadly applied because it only uses papers taken from the Scopus database. Future studies are urged to incorporate results from the Web of Science and other academic databases to improve the thoroughness of research findings. Future research could explore emerging areas within this field, such as the role of digital consent frameworks, the impact of deepfakes on democratic processes, and the development of interdisciplinary ethical guidelines. Furthermore, although the present research has identified important ethical principles and relationships, the proposed conceptual model would be strengthened by additional empirical validation using mixed-methods approaches. Finally, even though the researchers have used a systematic and philosophical approach to minimize bias, future studies could employ advanced methodologies such as big data analytics, longitudinal designs, and experimental studies to enhance and broaden the investigation's conclusions.

## REFERENCES

- Babaei, R., Cheng, S., Duan, R., & Zhao, S. (2025). Generative Artificial Intelligence and the Evolving Challenge of Deepfake Detection: A Systematic Analysis. *Journal of Sensor and Actuator Networks*, 14(1). <https://doi.org/10.3390/jsan14010017>
- Bhandari, R., & Bhandari, S. (2025). Artificial intelligence: understanding deepfakes. *EDPACS*, 70(9), 21–31. <https://doi.org/10.1080/07366981.2025.2484863>
- Birrer, A., & Just, N. (2024). What we know and don't know about deepfakes: An investigation into the state of the research and regulatory landscape. *New Media and Society*. <https://doi.org/10.1177/14614448241253138>
- Chapagain, D., Kshetri, N., & Aryal, B. (2024). Deepfake Disasters: A Comprehensive Review of Technology, Ethical Concerns, Countermeasures, and Societal Implications. In D. S. Jat, N. Z. Jhanjhi, F. B. Shava, J. Gajrani, Z. G.-A., S. Hamunyela, & M. Tripathi (Eds.), *2024 International Conference on Emerging Trends in Networks and Computer Communications, ETNCC 2024 - Proceedings* (pp. 139–147). Institute of Electrical and Electronics Engineers Inc. <https://doi.org/10.1109/ETNCC63262.2024.10767452>
- Chotisarn, N., & Phuthong, T. (2025). Place Branding as a Soft Power Tool: A Systematic Review, Bibliometric Analysis, and Future Research Directions. *International Review of Management and Marketing*, 15(4), 123–142. <https://doi.org/10.32479/irmm.18519>
- Clark Wolf, W. (2023). Kant's Formula of Universal Law as a Test of Causality. *Philosophical Review*, 132(3), 459–490. <https://doi.org/10.1215/00318108-10469512>
- Dagar, D., & Vishwakarma, D. K. (2022). A literature review and perspectives in deepfakes: generation, detection, and applications. *International Journal of Multimedia Information Retrieval*, 11(3), 219–289. <https://doi.org/10.1007/s13735-022-00241-w>
- de Ruiter, A. (2021). The Distinct Wrong of Deepfakes. *Philosophy and Technology*, 34(4), 1311–1332. <https://doi.org/10.1007/s13347-021-00459-2>
- Dean, R. (2006). The Value of Humanity in Kant's Moral Theory. In *The Value of Humanity in Kant's Moral Theory*. Oxford University Press. <https://doi.org/10.1093/0199285721.001.0001>
- Edwards, P., Nebel, J.-C., Greenhill, D., & Liang, X. (2024). A Review of Deepfake Techniques:

- Architecture, Detection, and Datasets. *IEEE Access*, 12, 154718–154742. <https://doi.org/10.1109/ACCESS.2024.3477257>
- Furizal, Ma'arif, A., Maghfiroh, H., Suwarno, I., Prayogi, D., Kariyamin, Lonang, S., & Sharkawy, A.-N. (2025). Social, legal, and ethical implications of AI-Generated deepfake pornography on digital platforms: A systematic literature review. *Social Sciences & Humanities Open*, 12, 101882. <https://doi.org/10.1016/j.ssaho.2025.101882>
- Gavran, I., Honcharuk, S., Mykhalov, V., Stepanenko, K., & Tsimokh, N. (2025). The Impact of Artificial Intelligence on the Production and Editing of Audiovisual Content. *Preservation, Digital Technology and Culture*, 54(3), 223–235. <https://doi.org/10.1515/pdte-2025-0022>
- Hill, T. E. (2006). Kantian Normative Ethics. In *The Oxford Handbook of Ethical Theory*. Oxford University Press. <https://doi.org/10.1093/0195147790.003.0018>
- Hill, T. E. (2015). Kantian perspectives on the rational basis of human dignity. In *The Cambridge Handbook of Human Dignity: Interdisciplinary Perspectives* (pp. 215–221). Cambridge University Press. <https://doi.org/10.1017/CBO9780511979033.027>
- Hoffmann, C. P., Bendahan Bitton, D., & Godulla, A. (2025). The Role of Self-Efficacy and Intellectual Humility in the Relationship Between Perceived Deepfake Exposure and Media Cynicism. *Social Science Computer Review*. <https://doi.org/10.1177/08944393251354977>
- Kangothra, R., & Sehajpal, S. (2025). Voices without faces: The impact of Deepfake technology on authentic communication. In *Deepfake Technology Dilemmas: Mastering Managerial Navigation* (pp. 187–206). Nova Science Publishers Inc. <https://www.scopus.com/inward/record.uri?eid=2-s2.0-105008365079&partnerID=40&md5=9c100d69117b318e7f59445149fea920>
- Kerner, C., & Risse, M. (2021). Beyond Porn and Discreditation: Epistemic Promises and Perils of Deepfake Technology in Digital Lifeworlds. *Moral Philosophy and Politics*, 8(1), 81–108. <https://doi.org/10.1515/mopp-2020-0024>
- Kleingeld, P. (2017). Contradiction and Kant's Formula of Universal Law. *Kant-Studien*, 108(1), 89–115. <https://doi.org/10.1515/kant-2017-0006>
- Kleingeld, P. (2023). A Defense and Development of the Volitional Self-Contradiction Interpretation. *Philosophia (United States)*, 51(2), 505–524. <https://doi.org/10.1007/s11406-023-00641-0>
- Kubanek, M., & Szymoniak, S. (2026). Deepfake Manipulation and Ethical Dilemmas: A Comprehensive Risk Assessment. In I. Alvarez, N. Silva, M. Arias-Oliva, & D. A.-H. (Eds.), *Lecture Notes in Computer Science: 15939 LNCS* (pp. 542–553). Springer Science and Business Media Deutschland GmbH. [https://doi.org/10.1007/978-3-032-01429-0\\_46](https://doi.org/10.1007/978-3-032-01429-0_46)
- Lemos, J. (2017). God's Existence and the Kantian Formula of Humanity. *Sophia*, 56(2), 265–278. <https://doi.org/10.1007/s11841-016-0511-7>
- Loovens, J., & Tinmaz, H. (2025). A systematic literature review of deepfakes in forensic science. *Forensic Imaging*, 43. <https://doi.org/10.1016/j.fri.2025.200647>
- Malik, S., Surbhi, A., & Roy, D. (2024). Blurring Boundaries between Truth and Illusion: Analysis of Human Rights and Regulatory Concerns arising from abuse of Deepfake Technology. In D. Roy & G. Fragulis (Eds.), *AIP Conference Proceedings* (Vol. 3220, Number 1). American Institute of Physics. <https://doi.org/10.1063/5.0234995>
- Maniyal, V., & Kumar, V. (2024). Unveiling the Deepfake Dilemma: Framework, Classification, and Future Trajectories. *IT Professional*, 26(2), 32–38. <https://doi.org/10.1109/MITP.2024.3369948>
- Marzi, G., Balzano, M., Caputo, A., & Pellegrini, M. M. (2025). Guidelines for Bibliometric-Systematic Literature Reviews: 10 steps to combine analysis, synthesis and theory development. *International Journal of Management Reviews*, 27(1), 81–103. <https://doi.org/10.1111/ijmr.12381>
- Misirlis, N., & Munawar, H. B. (2022). From Deepfake To Deep-Useful: Risks And Opportunities Through A Systematic Literature Review. In P. Kommers, I. A. Sanchez, P. Isaias, & L. Rodrigues (Eds.), *Proceedings of the International Conferences on e-Society 2022 and Mobile Learning 2022* (pp. 26–32). IADIS Press. <https://www.scopus.com/inward/record.uri?eid=2-s2.0-85137610892&partnerID=40&md5=17aa0dd8650923c89405ab7b43d4484e>
- Momin, M. S., Sufian, A., Barman, D., Leo, M., Distant, C., & Damer, N. (2025). Explainable deepfake detection across different modalities: An overview of methods and challenges.

- Image and Vision Computing*, 163.  
<https://doi.org/10.1016/j.imavis.2025.105738>
- Ni, C., & Abdullah, N. L. (2025). Research on absorptive capacity in the green context: a bibliometric and visualization analysis. *Cogent Business and Management*, 12(1).  
<https://doi.org/10.1080/23311975.2024.2435602>
- Pandian, H., & Rawindaran, N. (2025). Redefining Reality in Political Propaganda: Exploring the Impact of Superimposed Deepfakes in Misinformation Campaigns. In *Data Protection: The Wake of AI and Machine Learning* (pp. 47–61). Springer Nature.  
[https://doi.org/10.1007/978-3-031-76473-8\\_3](https://doi.org/10.1007/978-3-031-76473-8_3)
- Pendlebury, T. A. (2024). The Will of All in Kant's Groundwork. *Kantian Review*, 29(3), 423–445.  
<https://doi.org/10.1017/S1369415424000347>
- Peters, C. S., McGinnis, A. J., Im, S., & Hennessey, H. L. (2025). Satire Versus Illicit: Policy Issues Regarding the Creation of Deepfakes Using AI Technology. *Behavioral Sciences and the Law*, 43(4), 462–471.  
<https://doi.org/10.1002/bsl.2726>
- Rasyid, B. A. (2024). Analisis pengaruh perkembangan teknologi komunikasi pada fenomena “cukurukuk calling” oleh anak-anak RW 22 Padukuhan Nogosaren. *Lektur :Jurnal Ilmu Komunikasi*, 7(3), 2988–344.  
<https://doi.org/https://doi.org/10.21831/lektur.v7i3.23071>
- Reepu, R., & Johri, A. (2025). Unveiling deepfake: Ethical concerns in social media discourse. In *Mastering Deepfake Technology: Strategies for Ethical Management and Security* (pp. 211–236). River Publishers.  
<https://www.scopus.com/inward/record.uri?eid=2-s2.0-105007745521&partnerID=40&md5=6c4929547a9fd715baa161a840b4e9f9>
- Roe, J., Perkins, M., & Furze, L. (2024). Deepfakes and Higher Education: A Research Agenda and Scoping Review of Synthetic Media. *Journal of University Teaching and Learning Practice*, 21(10). <https://doi.org/10.53761/2y2np178>
- Rusli, A. N., Aziz, S., & Akbar, M. (2025). MEDIA SOSIAL DAN PERUBAHAN BUDAYA LOKAL: TINJAUAN SISTEMATIS ATAS DAMPAK FACEBOOK. *Lektur :Jurnal Ilmu Komunikasi*, 8(3), 2988–344.  
<https://doi.org/https://doi.org/10.21831/lektur.v8i3.25143>
- Sharma, K., Taneja, S., & Monga, S. (2024). Deepfake technologies: Concepts, methods, and applications. In *Navigating the World of Deepfake Technology* (pp. 59–65). IGI Global.  
<https://doi.org/10.4018/979-8-3693-5298-4.ch004>
- Shrivastava, P., & Mishra, R. (2025). Systematic Reviews in Epidemiological Studies. In *Epidemiology and Environmental Hygiene in Veterinary Public Health* (pp. 229–244). Wiley.  
<https://doi.org/10.1002/9781394208180.ch18>
- Taneja, S., Bhatnagar, M., Gupta, S., Kumar, P., Kumar, S., & Ozen, E. (2025). Deepfake technology dilemmas: Mastering managerial navigation. In *Deepfake Technology Dilemmas: Mastering Managerial Navigation*. Nova Science Publishers Inc.  
<https://doi.org/10.52305/AQVC8385>
- Verma, A. (2025). Deepfakes and the crisis of digital authenticity: ethical challenges in the age of synthetic media. *Journal of Information, Communication and Ethics in Society*.  
<https://doi.org/10.1108/JICES-04-2025-0083>
- Viola, M., & Voto, C. (2023). Designed to abuse? Deepfakes and the non-consensual diffusion of intimate images. *Synthese*, 201(1).  
<https://doi.org/10.1007/s11229-022-04012-2>
- Visnjic, D. (2025). Generative Models and Deepfake Technology: A Qualitative Research on the Intersection of Social Media and Political Manipulation. In K. S. Soliman (Ed.), *Communications in Computer and Information Science* (Vol. 2300, pp. 75–80). Springer Science and Business Media Deutschland GmbH.  
[https://doi.org/10.1007/978-3-031-79086-7\\_7](https://doi.org/10.1007/978-3-031-79086-7_7)
- Wahab, A. (2025). Futures of Deepfake and society: Myths, metaphors, and future implications for a trustworthy digital future. *Futures*, 173.  
<https://doi.org/10.1016/j.futures.2025.103672>
- Wikum, T., & Wijayanayake, J. (2024). A Systematic Literature Review of Deepfake Literacy, Societal Impacts, and AI-Driven Prevention Strategies for Platform-Specific Vulnerabilities. *9th International Conference on Advances in Technology and Computing, ICATC 2024*.  
<https://doi.org/10.1109/ICATC64549.2024.11025231>
- Zegarow, P., & Bartuzi, E. (2024). Deepfake Influence Tactics through the Lens of Cialdini's Principles: Case Studies and the DEEP FRAME Tool Proposal. *Applied Cybersecurity and Internet Governance*, 3(2), 286–302.  
<https://doi.org/10.60097/ACIG/201147>