



INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Ethical Implications Of Generative Ai In Content Creation: A Multidisciplinary Approach

Parnapalli Pushpanjali

Assistant Professor

S.E.A College of Science, Commerce and Arts-Autonomous

Abstract:

Generative Artificial Intelligence (GenAI) has revolutionized content production by automating text, image, audio, and video generation. As it provides innovative functions, GenAI also presents intricate ethical challenges. This paper takes a multidisciplinary approach to investigate the ethical considerations of GenAI in content production, emphasizing problems of misinformation, bias, intellectual property rights, privacy, and accountability.

Another major concern is the spread of misinformation and deepfakes that can be used to manipulate public attitudes and undermine media credibility. Furthermore, AI models built from biased data can reinforce social prejudices with discriminatory results. The paper explores where these biases originate and discusses ways to create more balanced AI systems.

Intellectual property rights and authorship pose additional ethical dilemmas. Leveraging works currently created to train AI models creates controversy regarding ownership of AI-generated works and the rights of initial creators. This paper analyzes the legal frameworks that govern intellectual property and explores the added complexities of AI-generated works.

Data protection and privacy are also important issues, since GenAI models can inadvertently produce material that is a privacy infringement on people's personal information by copying sensitive information in training data. The paper addresses the ethical implications of data usage and consent for generating AI content.

Through the convergence of media studies perspectives, legal systems, pedagogic practices, and technological advances, this paper hopes to achieve a holistic grasp of the ethical environment of GenAI and lay out recommendations for its responsible application.

Keywords: Generative AI, content creation, ethics, misinformation, bias, intellectual property, accountability, multidisciplinary approach, responsible AI.

Introduction:

Generative Artificial Intelligence (GenAI) has quickly transformed the content creation landscape by supporting the automated creation of text, imagery, sound, and video. GPT-4 and DALL-E models have been taken up in numerous industries, such as journalism, marketing, education, and entertainment that provide unprecedented levels of creative potential. That this integration occurs raises profound ethical issues that require careful analysis from a variety of disciplinary standpoints.

One of the main ethical concerns of GenAI is the spread of misinformation and deepfakes. The potential of GenAI to create realistic but false content presents threats concerning the dissemination of false information, manipulation of public opinion, and loss of trust in media sources. This implicates societal trust and democratic process, highlighting the necessity for mechanisms to identify and counter the spread of false information.

The other main concern is the reinforcement of prejudice and discrimination by GenAI-created content. Prejudices embedded in data used to train AI models can then be exaggerated, and existing social prejudices strengthened, resulting in discriminatory processes in many areas, from recruiting to policing and service

delivery. The origins of these prejudices need to be investigated, and means of creating fairer AI systems need to be debated.

Authorship and intellectual property rights pose further ethical issues in GenAI. The reuse of existing works to train AI poses questions regarding ownership of content generated by AI and rights of initial creators. Intellectual property law needs to be examined to resolve the complexities posed by AI-generated works. Privacy and data protection are likewise essential issues, as GenAI systems can incidentally produce content that violates people's privacy by replicating sensitive data found in training data. Ethical issues about the use of data and consent in developing AI-generated content must be raised, focusing on upholding data protection legislation and the use of practices that protect personal privacy.

Transparency and accountability in GenAI systems are key to guaranteeing ethical use. Attributing responsibility for GenAI system outputs is complex, particularly when content is harmful. The lack of transparency in AI decision-making processes makes it difficult to hold others responsible. Arguing for the creation of transparent and interpretable AI systems, along with well-defined lines of responsibility for AI-created content, is imperative.

By embracing an interdisciplinary approach, merging insights drawn from media studies, legal principles, teaching practices, and technological advances, an all-encompassing grasp of the ethical arena surrounding GenAI can be gained. Suggesting guidelines for responsible utilization, such as transparency, accountability, fairness, privacy, and cooperation among stakeholders, is imperative. Interdisciplinary efforts must be employed to navigate the ethical challenges of GenAI and ensure that its integration into content development is beneficial to the public.

Ethical challenges in GenAI-Driven Content Creation:



figure 1: ethical challenges in gen ai

Bias and Discrimination:

Generative AI (GenAI) technologies, though providing revolutionary features in content generation, are vulnerable to algorithmic prejudices that can reinforce and intensify social inequalities. Algorithmic biases usually have their roots in the training data for AI algorithms, as these result from past prejudices and systemic imbalances.

Causes of Algorithmic Bias

Algorithmic bias occurs when AI systems generate outputs that consistently disadvantage individuals or groups. This frequently results from:

Historical Data Bias

Historical biases in training data can lead AI models to imitate and perpetuate these biases. As an example, if historical data of hiring practices preferred male candidates, an AI system trained using this data may also favour male candidates.

Underrepresentation

Some groups will be underrepresented in training data, and as a result, AI systems will not perform well for these groups. For instance, facial recognition systems have exhibited greater error rates for darker-skinned people because they are underrepresented in training data.

Labelling Bias:

Human biases can affect how data gets labelled, with subjective judgments that distort AI outputs. In 2025, India has seen numerous real-world cases where algorithmic prejudice in generative AI systems resulted in

discriminatory consequences, emphasizing the pressing need for ethical regulation and inclusive AI development.

Caste Bias in Language Models

A recent work, DECASTE, showed that there are huge caste-based biases in large language models (LLMs). It was proven through the study that these models consistently replicate stereotypes against disadvantaged castes like Dalits and Shudras, especially in socio-cultural and political realms. Such AI-generated content biases can reinforce societal prejudices and make social equity efforts difficult.

Accent Neutralization in Call Centers

Tele performance, an international call center company, used AI technology created by Sanas to "neutralize" Indian accents in real-time. Though aimed at improving customer communication, this practice has been controversial due to cultural erasure and linguistic discrimination concerns, as the practice in effect assumes Indian accents are less desirable in global communication.

Intellectual Property and Authorship:

The speedy development of Generative AI (GenAI) has brought forth intricate issues regarding intellectual property (IP) rights and authorship, especially in the Indian scenario. When the AI models are trained over huge datasets, usually existing works of creativity, there is a question of legality and morality behind using such material without a direct go-ahead. The subsequent current examples from India in 2025 reflect the varied nature of these issues:

ANI vs. OpenAI: Unauthorized Use of News Content

In a first-of-its-kind case, Asian News International (ANI), a leading Indian news agency, sued OpenAI, claiming that ChatGPT was developed using ANI's copyrighted news content without permission. ANI argues that this unauthorized use not only violates its IP rights but also resulted in the creation of false news reports with fake quotes attributed to the agency, hence causing damage to its reputation. OpenAI, on its part, argues that it uses publicly available information and follows fair use practices. This matter has compelled the Indian government to form a committee to examine and possibly amend the country's copyright laws in the wake of AI developments.

Indian Media Houses Challenge AI Training Practices

After ANI's legal challenge, major Indian media outlets like NDTV, Indian Express, and Hindustan Times have complained about their use in training AI systems such as ChatGPT. They contend that their copyrighted content has been used without permission, affecting their business model and content sovereignty. OpenAI insists it does not employ the content of these media houses for training and its operations are in conformity with current legal frameworks.

Bollywood Music Houses Join the Legal Battle

In a major turn of events, major Bollywood music companies like T-Series, Saregama, and Sony have moved to become part of the ongoing copyright case against OpenAI. These companies claim their sound recordings are being utilized without permission to train AI systems, which might violate their IP rights as well as earnings. OpenAI argues that its utilization of publicly accessible material is safeguarded by fair use provisions.

Absence of Lucid Legal Frameworks

India's existing copyright regime, in particular the Copyright Act of 1957, makes no clear provision for the reuse of copyrighted content for training purposes in AI. Compared with jurisdictions like the European Union, where there are provisions under which creators can exclude their works from being used in AI training datasets, India is without such provisions. Such legal uncertainty puts Indian creators at a disadvantage, as they have little control to prevent the unauthorized reuse of their works in AI.

Government Initiatives and the Way Forward

While acknowledging the urgent need to deal with these issues, the Indian government has taken the first steps to rethink and possibly revise current IP laws. A committee of legal professionals, industry players, and policymakers has been constituted to look into how the current legislation works in the context of AI-generated material and to suggest appropriate changes. This is an attempt at striking a balance between encouraging technological advancement and safeguarding the interests of creators in the new digital ecosystem.

These live examples highlight the intricacies of intellectual property and authorship in the age of Generative AI in India. With AI advancing further, it is essential that legal frameworks also keep pace with such developments, protecting the rights of original creators while promoting innovation and technological growth.

Misinformation and Deepfakes:

The advent of Generative AI (GenAI) has greatly improved the sophistication and ease of creating synthetic media, resulting in a rapid increase in misinformation and deepfakes. These falsehoods produced by AI create multi-dimensioned ethical issues that affect public confidence, democratic systems, individual privacy, and social cohesion.

Deterioration of Public Confidence and Democratic Integrity

Deepfakes have been used more and more to sway public opinion and erode confidence in media outlets. For example, in the 2024 general elections, the actors Aamir Khan and Ranveer Singh were the subject of viral deepfake videos. The actors were misrepresented in the videos, criticising Prime Minister Narendra Modi and supporting the opposition Congress party. The actors disputed their involvement and lodged police complaints, citing the misuse of their likenesses. Through 2025, the fight against deepfakes continued in the form of sophisticated versions targeting political personalities, such as an image manipulated to link CPM youth leader Srijan Bhattacharya with a YouTuber accused of espionage, which prompted a police investigation. Again, a Vandsa resident was arrested after disseminating a malicious deepfake video of Prime Minister Modi in a WhatsApp group during an attack scenario with disparaging remarks against the Indian Army.

Psychological Impact and Social Harm

The spread of deepfakes goes beyond the political sphere, reaching into the personal lives of individuals as well as societal norms. The unauthorized production and sharing of explicit deepfake material has been aimed at public figures and regular people alike, resulting in emotional trauma and damage to their reputations. Deepfake cyberbullying has also become a disturbing trend, especially among teenagers, where manipulated media is employed to bully and intimidate classmates, causing anxiety, depression, and social isolation.

Privacy and Data Protection

By 2025, the spread of Generative AI (GenAI) in India has raised fear about privacy and data protection. AI models that have been trained on large datasets could unknowingly replicate sensitive personal data, resulting in probable privacy violations.

Tirumala Tirupati Devasthanams (TTD) Facial Recognition Deployment: The TTD board has deployed AI and facial recognition technology to boost security and expedite pilgrim experiences at the Tirumala Temple. Although these technologies are designed to make things more efficient, they have raised eyebrows regarding their collection and storage of biometric information. The board has insisted that strong data privacy controls will be put in place to safeguard pilgrims' personal data.

Pune Mahanagar Parivahan Mahamandal Limited (PMPML) AI-Enabled Cameras: PMPML will install AI-enabled CCTV cameras in all its buses to monitor driver behaviour and passenger demographics by July 2025. Though aimed at improving safety and service quality, the gathering of demographic information, including gender ratios, raises data usage and privacy protection concerns.

Regulatory Developments

IndiaAI Safety Institute:

Founded on 30 January 2025, the IndiaAI Safety Institute seeks to enable the ethical and secure use of AI models. The institute encourages indigenous R&D based on India's rich socio-cultural background and encourages ventures under the Safe and Trusted Pillar of the India AI Mission.

MeitY's AI Governance Framework:

The Ministry of Electronics and Information Technology (MeitY) stresses the importance of responsibility and transparency throughout India's AI ecosystem. The framework mandates regulators to collect sufficient information on the traceability of data, models, systems, and actors through the AI lifecycle.

Accountability and Transparency

Accountability of AI-generated content is challenging, particularly when the content created is harmful or illegal. Clear lines of responsibility between developers, deployers, and users must be established. Transparency in AI systems, explainability, and auditability can facilitate attribution of accountability and trust from the public.

Telangana's Fire Department AI-Based NOC System: Telangana's fire department will be introducing an AI-based system to make No Objection Certificates (NOCs) for fire safety automatic. The system can automatically review architectural fire safety plans, identify violations, and produce detailed reports, increasing transparency and accountability in the approval process.

AI Pragya Training of UP Government: The AI Pragya program was introduced by the government of Uttar Pradesh to empower citizens and government officials with future skills in AI, machine learning, data

analytics, and cybersecurity. The program is intended to enhance delivery of government services and ensure transparency in administrative functions.

Policy and Governance

India's Foray into AI Regulation: The Indian government has provided goals for AI regulation in terms of reducing harm to users, countering misinformation, enhancing accountability, and balancing commercial interests. The strategy involves enforcing principles of transparency, fairness, privacy, security, and accountability by a system of rights and obligations.

These live examples from India in 2025 emphasize the utmost importance of privacy, data security, accountability, and transparency in the use of Generative AI technologies. As AI keeps changing, it is only appropriate that stakeholders establish strong data governance policies, seek informed consent, and have well-defined lines of responsibility to gain public confidence and protect individual rights.

Multidisciplinary Perspectives on Ethical Implications of GenAI

Multidisciplinary Insights on the Ethical Consequences of Generative AI Content Generation

A multidisciplinary perspective is needed to understand the ethical consequences of Generative AI (GenAI) in content generation, combining insights from media studies, legal systems, education practice, and technological innovation. Each field provides different insights on challenges and issues in GenAI.

Media Studies: Managing Information Integrity and Public Discourse

GenAI's capacity for generating realistic and persuasive content has far-reaching implications for media and public discourse. Research has indicated that AI-generated arguments are more effective than their human equivalents, and this has raised the specter of potential misinformation and manipulation. For example, one Nature Human Behaviour study discovered that GPT-4 and other large language models were more convincing than humans during online arguments 64% of the time, even adapting arguments from very limited demographic information. This highlights the importance of media literacy and critical thinking skills on the part of the public to be able to differentiate AI content.

Legal Studies: The Intellectual Property and Liability Question

The application of GenAI across industries raises sophisticated legal concerns, mostly related to intellectual property rights and responsibility. The application of existing creative content in training AI models challenges issues of ownership and payment for original creators. Laws are already struggling with these matters, attempting to set strong guidelines on the application of copyrighted content in AI training and the identification of AI-created content. In addition, accountability for AI-created content, particularly when it results in harm or violates rights, is an ongoing issue.

Educational Practices: Fostering Digital Literacy and Ethical Use

In education, GenAI presents opportunities as well as challenges. While AI has the potential to enrich learning experiences using tailored content and assistance, it also creates questions regarding academic integrity and misuse. Teachers are urged to include AI literacy in courses of study, educating learners on ethical use of AI tools and the value of critical thinking. UNESCO's recommendations on generative AI in education focus on the necessity for policy that guarantees fair and ethical use of AI technologies in learning contexts.

Technological Development: Developing Ethical and Transparent AI Systems

Technologically, creating GenAI systems with integrated ethical protections is important. This involves adopting mechanisms to avoid generating harmful or biased content and promoting transparency in AI decision-making. Technology developers are encouraged to utilize fairness-aware algorithms and perform periodic audits to detect and eliminate biases in AI output. It requires cooperation between technologists, ethicists, and policymakers to establish AI systems that support societal values and ethical principles.

By synthesizing insights from these disciplines, stakeholders can better navigate the ethical landscape of GenAI, fostering responsible innovation that respects individual rights and societal norms.

Security Issues in GenAI :

Generative AI (GenAI) has transformed content creation but also poses severe security threats. These threats arise from the technology's capability to generate genuine content, which can be used for harmful purposes. An outline of important security threats posed by GenAI and solutions to manage them follows

Deepfakes and Misinformation

GenAI can generate very realistic fake images, videos, and audio clips, referred to as deepfakes. These are dangerous as they can be employed to disseminate misinformation, manipulate public opinion, and destroy reputations. Deepfakes have, for example, been used to impersonate public figures, resulting in the spread of false information.

Mitigation Strategies:

- Develop and release deepfake detection technologies.
- Use digital watermarking to sign content.
- Educate the public about the existence and threat of deepfakes.

AI-Generated Phishing and Social Engineering

Attackers use GenAI to create advanced phishing messages and emails that appear as genuine communications. These computer-generated messages can trick users into divulging sensitive data or downloading malware.

Mitigation Strategies:

- Employ state-of-the-art email filtering systems that identify AI-generated content.
- Provide regular employee training to identify phishing attempts.
- Have multi-factor authentication so as to provide an additional layer of security.

Malicious Code Generation

GenAI tools can create code snippets, which, if tampered with, can inject vulnerabilities or malicious features into software systems. This is a threat to software integrity and security.

Mitigation Strategies:

- Carefully review and test AI-created code before it goes live.
- Utilize code analysis tools to identify possible vulnerabilities.
- Adhere to coding standards and best practices to promote code quality.

Prompt Injection Attacks

Prompt injection is creating inputs that trick GenAI models into generating unwanted or harmful outputs. It can evade safety checks and contribute to the spreading of unsuitable content.

Mitigation Strategies:

- Apply rigorous input validation and sanitization.
- Periodically update and refine AI models to detect and resist evil prompts.
- Watch AI outputs for anomalies and apply feedback mechanisms.

Data Privacy Concerns

GenAI models are built on large datasets, possibly containing sensitive personal data. There's a potential that the models might unknowingly replicate or expose private information.

Mitigation Strategies:

- Use data anonymization methods during training.
- Adopt stringent data governance principles to safeguard personal data.
- Utilize differential privacy techniques to avoid the leakage of sensitive information.

Model Poisoning

Attackers can inject malicious information at the training stage of GenAI systems, resulting in biased or malicious output. This undermines the integrity and trustworthiness of the AI system.

Mitigation Strategies:

- Verify and sanitize training data to eliminate threats.
- Track model performance for poisoning or bias symptoms.
- Enforce rigorous training procedures and regular audit controls.

Intellectual Property Infringement

GenAI is able to create content that is highly similar to original copyrighted work, and this can result in intellectual property infringement. This is causing legal and ethical issues surrounding ownership of content.

Mitigation Strategies:

- Utilize licensed data to train AI models.
- Employ content filters to recognize and block the creation of infringing content.
- Have definitive guidelines and policies around AI-created content and intellectual property.

Shadow AI Usage

Employees can utilize unauthorized GenAI tools, resulting in uncontrollable data flow and possible security breaches. This "shadow AI" use circumvents official IT protocols and monitoring.

Mitigation Strategies:

- Train employees in approved AI tools and their proper use.
- Use monitoring systems to identify unauthorized use of AI tools.

- Set clear policies and repercussions for unauthorized use of AI tools.

Infrastructure Vulnerabilities

Incorporating GenAI into current systems may pose new threats, particularly where the AI models interface with sensitive components of the critical infrastructure. This may result in system failures or intrusions.

Mitigation Strategies:

- Perform complete security evaluations prior to incorporating AI models.
- Segregate AI elements from critical infrastructure where possible.
- Update and patch AI systems routinely to counter known vulnerabilities.

Lack of Transparency and Explainability

GenAI models tend to be "black boxes," and as a result, it is hard to identify the thought process behind their decisions. This transparency issue might impede trust and accountability.

Mitigation Strategies:

- Create and utilize explainable AI models that give insight into how they function.
- Incorporate logging and monitoring to monitor AI decision-making activities.
- Conduct frequent audits to confirm AI models function as expected.
- Solving these security issues involves an integrated strategy, uniting technical solutions with policy and governance actions. Organizations must be proactive and alert to manage the risks of GenAI to responsibly unlock its advantages.

Guidelines for Ethical GenAI Use:

To navigate the ethical landscape of Generative AI (GenAI), several key recommendations have been proposed. These guidelines aim to ensure responsible development and deployment of GenAI technologies. Below is an exploration of these recommendations, accompanied by real-time examples from India in 2025:

Develop Ethical Guidelines

Recommendation: Establish comprehensive ethical frameworks that guide the development and deployment of GenAI technologies.

Example: National Association of Software and Service Companies (NASSCOM) published "Responsible AI Guidelines for Generative AI" in 2023. These guidelines seek to establish stakeholder agreement around the essential normative responsibilities of researchers studying, developers creating, or consumers applying GenAI technologies. They will be used as a basis for creating standards, protocols, audit checklists, certifications, and monitoring tools to reduce potential harms from GenAI uptake.

Improve Transparency

Recommendation: Establish systems that enable users to comprehend how content is created by AI.

Example: To address fears of AI-created disinformation, the Indian government made it mandatory for technology companies to obtain permission before releasing AI tools labelled as "unreliable" or in the trial phase. Such tools should be labelled clearly so that they alert users to the possibility of giving wrong answers, maintaining transparency and safeguarding information integrity.

Encourage Fairness

Recommendation: Design AI systems to avoid biases and ensure equitable results.

Example: The January 2025 published AI Governance Guidelines Development report of the Indian government, made available for public consultation, focuses on responsible AI governance principles. These principles, drawing from international guidelines and Indian systems, will address bias and ensure fairness in AI systems.

Safeguard Intellectual Property

Recommendation: Establish legal frameworks that identify and remunerate original creators whose work is utilized in AI training.

Example: Indian book publishers and their overseas counterparts lodged a copyright suit against OpenAI in New Delhi. The suit claims that OpenAI employed copyrighted materials to train AI services without license, preventing such use and claiming compensation for possible effects on creativity and sales.

Protect Privacy

Recommendation: Implement robust data protection mechanisms to avoid the abuse of personal data.

Example: The Indian government published proposed rules under the Digital Personal Data Protection Act (DPDPA) in January 2025. The rules intended to protect digital rights and establish compliance requirements for organizations processing personal data, including requirements for data localization and limitations on cross-border data transfer.

Foster Multidisciplinary Collaboration

Recommendation: Foster inter professional collaboration among technologists, legal professionals, ethicists, and educators to tackle the multi-dimensional ethical issues of GenAI.

Example: NASSCOM's Responsible AI Hub promotes and facilitates strategic programs for stakeholder sensitization and capacity building. These programs entail collaborative efforts among stakeholders such as academia, civil society, and industry professionals to generate responsible AI practices.

These instances reflect the proactive measures being undertaken in India to guarantee the ethical development and implementation of Generative AI technologies. Stakeholders can avoid navigating the intricate ethical landscape of GenAI by following these recommendations.

Conclusion

Generative AI leads the charge in content innovation, but with its introduces complex network of ethical dilemmas that need to be subjected to close examination. Through this paper, a multi-disciplinary approach has been employed to unpick these issues with particular attention given to the key areas of misinformation, bias, intellectual property, privacy, and accountability. The possibility of GenAI generating believable misinformation and deepfakes not only undermines public trust but also jeopardizes the credibility of digital media. Likewise, biases embedded in training data can perpetuate harmful stereotypes and discrimination, attesting to the critical need for fairer data practices and transparent models. Intellectual property issues involving AI-produced content underscore the conflict between creative newness and legal authorship, demanding more transparent models that acknowledge original creators as well as algorithmic contributors. To secure responsible deployment, stakeholders need to get together and configure policies and practices that align technological capacities with ethical demands, ultimately steering GenAI toward a more credible and inclusive future.

References:

- <https://contentbloom.com/blog/ethical-considerations-in-ai-generated-content-creation/>
- Bender, E. M., et al. (2021). "On the Dangers of Stochastic Parrots." FAccT.
- OpenAI. (2023). "GPT-4 Technical Report."
- U.S. Copyright Office. (2023). Policy Statement on AI-Generated Works.
- Crawford, K. (2021). Atlas of AI. Yale University Press.
- Citron, D., & Chesney, R. (2019). "Deepfakes and the New Disinformation War." Foreign Affairs.
- Floridi, L. (2019). Establishing the Rules for Building Responsible AI. Nature Machine Intelligence, 1(6), 261-262.
- Crawford, K., & Paglen, T. (2019). Excavating AI: The Politics of Images in Machine Learning Training Data. International Journal of Communication, 13, 21.
- <https://www.computer.org/publications/tech-news/trends/ethical-concerns-on-ai-content-creation>