Proceedings of the 4th International Conference on Global Politics and Socio-Humanities
DOI: 10.54254/2753-7048/2026.35266

Generative AI and Personality Rights Infringement in China: Mechanisms and Policy Responses

Siyu Lu

School of Social Development, East China University of Political Science and Law, Shanghai, China
L2025481769@outlook.com

Abstract. Generative artificial intelligence (AI) has rapidly expanded the use of face swapping, digital avatars, voice cloning and deepfake production in short-form video, virtual livestreaming and online marketing. These applications improve content production efficiency but also create new risks for personality rights, especially where personal images, voices and identity features are reproduced without consent. This paper examines how generative AI infringes personality rights in China through three scenarios: AI face swapping, AI digital humans and platform dissemination. It argues that such infringements differ from traditional cases because they are low-cost, scalable, highly realistic and difficult to trace. Although China's Civil Code and rules on deep synthesis and AI-generated content provide a legal basis for protecting portrait rights, reputation rights, privacy rights and personal information, gaps remain in liability allocation, authorisation boundaries and platform governance. The paper proposes clearer rules for the use of personal characteristics, stronger labelling and algorithmic transparency duties, and more effective platform redress mechanisms.

Keywords: generative AI, personality rights, deepfakes, platform liability, AI-generated content

1. Introduction

Artificial intelligence (AI) is increasingly visible on social media and short-video platforms. From one-click face swapping to virtual presenters that can imitate a person's voice, intonation and appearance, generative AI is reshaping digital life at high speed. These technologies offer entertainment value and industrial convenience, but they also create serious legal risks when a person's likeness, voice or identity characteristics are reproduced and circulated without authorisation.

With innovations in algorithmic technology, generative AI has gradually been adopted across industries to reduce production costs and expand production scale. This is particularly clear in short-form dramas and animation, where production barriers and costs were once high. As generative AI becomes easier to use, personal characteristics can be replicated at low cost and on a large scale. Numerous cases have emerged in which the facial features of ordinary people or celebrities, as well as intellectual property (IP) characters, are used in AI-generated short dramas and animations. Some

DOI: 10.54254/2753-7048/2026.35266

photographs of unsuspecting women have also been used to spread pornographic videos. The risk of infringement has therefore evolved from isolated disputes into a systemic problem.

Based on representative cases and policy materials, this paper examines how generative AI challenges China's existing personality rights framework through face swapping, digital avatars and platform dissemination. It focuses on three legal risks: the reproduction of personal characteristics, the commercialisation of identity and the spread of infringing content through online platforms. The analysis also considers the shortcomings of China's current policies and proposes possible improvements.

The paper adopts a case-study approach. AI face swapping is selected because it directly concerns the misuse of personal characteristics such as likeness and voice. AI digital avatars are examined because they reveal the legal risks arising from the long-term commercial exploitation of identity characteristics. Platform liability is analysed because the dissemination of infringing content is often accelerated by recommendation systems and weak complaint mechanisms. These scenarios help identify practical challenges in the protection of personality rights in the context of generative AI.

2. Current legal and policy framework

2.1. Personality rights and personal information protection

Within the institutional design of China's Civil Code, the right to one's image is established as an absolute right of control, with both active and defensive functions. Article 1018 grants natural persons the right to produce, publish and authorise others to use their image, while also defining image rights by reference to identifiable external characteristics. To balance rights protection and freedom of action, identifiability should be assessed through strict and objective criteria. Whether the relevant features involve facial appearance, body contours, make-up, clothing, silhouettes or other non-physical elements, the key question is whether the public can directly associate the external image with a specific natural person.

In the context of AI face-swapping technology, even where a generated image does not directly use a real photograph, portrait rights may still be implicated if the system recombines facial contours, features and other elements in a way that allows the public to identify the person. Article 1019 prohibits defamation, falsification and unauthorised use of another person's image. The exclusiveness of this protection requires clear boundaries; otherwise, the scope of personality rights may expand without limit and unduly restrict lawful uses by others [1]. If a third party inserts an individual's image into advertisements, short dramas or fabricated videos without authorisation, the conduct may constitute unlawful use of the portrait even where the content is automatically generated by algorithms.

The Civil Code also protects related personality interests. Article 1024 prohibits infringements of reputation through insult or defamation, so the use of generative AI to place a person's portrait into fabricated events may damage reputation. Articles 1032 and 1033 protect privacy and prohibit unauthorised intrusion into or disclosure of private spaces, activities and information. Article 1034 protects personal information, including biometric information and location data, and requires processing to be lawful, legitimate and necessary. Where facial features, voices or behavioural profiles are forged, distorted or used in false propaganda, the conduct may infringe portrait rights, reputation rights, privacy rights and personal information rights at the same time.

DOI: 10.54254/2753-7048/2026.35266

2.2. Regulatory requirements for deepfakes and generative AI

According to the Measures for the Labelling of AI-Generated and Synthetic Content, content generated by generative AI, including text, images, audio, video and virtual scenes, must carry mandatory dual labelling. The first form is explicit labelling, such as a prominent notice stating that the content is AI-generated. The second form is implicit labelling, such as metadata or digital watermarks. Deleting or tampering with such labels is prohibited, and distribution platforms must verify labels and provide additional warnings for suspected AI-generated content [2].

For AI-generated face-swapping content, clear labelling is particularly important because ordinary viewers may not have strong ability to distinguish synthetic content from real recordings. Visible labelling can help the public identify whether a video or image is AI-synthesised and can reduce the impact of false information on portrait rights and reputation rights. Labelling also provides an initial basis for later platform review, evidence preservation and liability allocation.

In determining the liability of generative AI service providers, the Measures for the Labelling of AI-Generated and Synthetic Content and related policies create a regulatory framework covering the process from data input to content output. Service providers should review the legality of training data and ensure that data capture and model training do not infringe portrait rights, reputation rights, privacy rights or other personality rights. They should also fulfil the primary responsibilities of online information content producers by establishing filtering mechanisms, blocking unlawful information and providing users with convenient complaint channels.

2.3. Shortcomings of existing regulations

2.3.1. Complexity of liability

There are numerous short-video and short-drama platforms, and the same content may be uploaded by multiple users or reposted across different platforms. In such circumstances, it is difficult to trace the origin of AI-generated content that infringes portrait rights or copyright, or to identify the actual producer. Given the broad dissemination of such content, determining the liability of those who distribute it is also challenging. Current Chinese policy has not yet clearly stipulated how to allocate liability among the producers, users and distributors of infringing AI-generated content.

The problem is intensified by the many-hands structure of generative AI. The AI supply chain may involve core developers, fine-tuners, prompt writers, platform operators and downstream distributors. This complex socio-technical system causes a severe diffusion of responsibility, making it difficult for traditional infringement rules to identify a direct liable party [3]. Model developers are responsible for data sourcing and model training; content generators are responsible for the immediate production of outputs; platforms bear responsibilities in dissemination and review; commercial users may be liable where they use unauthorised likenesses, voices or styles for marketing; and ordinary users may also bear responsibility if they knowingly repost or repackage false AI-generated content.

2.3.2. Unclear boundaries of authorisation

With the widespread adoption of generative AI, the traditional notice-and-consent framework has become less effective. Many AI-generated works incorporate the likenesses or works of original creators without their knowledge, and the boundary between display use and training use is often

DOI: 10.54254/2753-7048/2026.35266

unclear. The public may find that AI-generated works contain real people or existing IP, yet the content is published without the knowledge of the affected rights holders [4].

At the same time, it is difficult for affected rights holders to trace and pursue legal action against AI-generated content. This difficulty arises partly from the broad dissemination of such content and the high cost of litigation, and partly because the infringing material may already have been incorporated into model databases. In practice, when a user licenses facial or voice data to a specific platform, such data may be used covertly to fine-tune generative AI models or to create synthetic content unrelated to the individual's original intention. This secondary use of data far exceeds the boundaries of initial consent and makes the licensing of personality rights ambiguous in the face of AI's large-scale data processing.

2.3.3. Inadequate remedial mechanisms

Although existing Chinese regulations require generated content to be labelled clearly, this technical requirement has important limitations. Research has shown that relying solely on watermarks or similar marks is insufficient as a remedy or protective measure for AI-generated content. Watermarking technologies may suffer from weaknesses in implementation, accuracy and robustness. Visible or invisible audiovisual watermarks can be manipulated, removed or attacked, which casts doubt on the authenticity and traceability of AI-generated content.

The ex post notice-and-takedown mechanism is also weak when applied to AI face-swapping and deepfake content. Legal research indicates that it is often difficult to hold users responsible for creating deepfake content because they may be difficult to identify and may lack the financial capacity to pay compensation. At the same time, platforms through which deepfake content spreads may rely on existing liability exemptions in some legal systems. Existing ex post remedies and platform moderation mechanisms therefore struggle to protect personality rights at the same speed as viral AI-generated content [5].

3. Case studies of typical scenarios

3.1. AI face swapping and misuse of personal images

In early 2026, with the popularity of AI-generated micro-dramas, several incidents involving ordinary people having their faces "stolen" by AI attracted public attention. In the AI short drama Peach Blossom Hairpin, several ordinary creators, including commercial model Christ Qihai, discovered that their facial features and original photographs had been captured without permission and replaced onto villainous characters through AI face-swapping technology. According to media reports, the victims' facial features were not only misappropriated but also linked to negative characters. This caused serious damage to their social reputation and image and generated fear and anger among the affected individuals [6].

After the incident was exposed, several victims sought redress through social media and complained to the platform. The platform then reviewed the work. Public information indicated that the production company failed to provide sufficient evidence of the lawful source of the material, and the platform ultimately determined that the work violated content compliance rules. As a result, Peach Blossom Hairpin was removed, and the relevant production company was suspended from uploading further episodes. However, the victims remained in a passive position during rights enforcement. Comments left by victims below the videos were removed by the production company, and complaints to the platform were met with evasion.

DOI: 10.54254/2753-7048/2026.35266

The harm caused by AI short-form video infringement differs from simple misappropriation. AI face swapping and voice cloning can be produced quickly, while victims face high evidentiary, time and financial costs when asserting rights. AI-merged faces may also fail to meet traditional visual recognition standards. More importantly, generated content can forcibly integrate innocent individuals into pre-determined plots, creating a destructive and viral false image in the public consciousness. Compared with text-based rumours, such generated content is more likely to evoke public empathy or revulsion and therefore can seriously damage the victim's reputation [7].

3.2. Commercialisation of AI digital humans and real identities

In addition to AI face swapping, AI digital humans, AI actors and virtual brand ambassadors have raised public debate. Digital humans can imitate a natural person's voice, facial micro-expressions, body language and performance style. Once a person's core biometric characteristics are bound to an AI model, users can transcend the temporal and spatial constraints of the real person and generate a wide range of fictional narratives, actions and commercial scenes at very low cost [8]. This technical capacity may allow a person's identity features to be exploited continuously and even after death.

The replication of digital personas has already moved from theory into practice. In 2026, following the death of Zhang Xuefeng, a well-known tutor for China's postgraduate entrance examinations, an AI project named Zhang Xuefeng.skill sparked broad online discussion. By using his published works, interview transcripts and linguistic style, developers constructed an AI system capable of simulating his thought processes and mode of expression. Supporters regarded the project as a continuation of his knowledge and experience, while opponents argued that it might infringe posthumous personality interests and exceed the boundary of authorisation [9].

From the perspective of personality rights protection, although the civil rights capacity of a natural person ends upon death, the use of identity characteristics without authorisation from the person, a relevant institution or close relatives may still raise legal concerns. If such use exceeds commemoration or public-interest purposes and enters commercial exploitation, it may infringe protected interests associated with the deceased's image, reputation and personal dignity. This scenario shows that generative AI challenges not only the rights of living persons but also the legal treatment of digital personality and posthumous personality interests.

4. Directions for improving China's policy response

4.1. Clarifying authorisation rules for the use of personal characteristics

Faced with new issues created by generative AI, the traditional model of authorisation and consent is no longer sufficient. The use of information by generative AI has become increasingly complex and diverse. Future rules should not merely increase the burden of proof on victims. Instead, they should strengthen the recording and filtering obligations of platforms and service providers. For generative AI services involving personal characteristics, platforms and technology providers should establish comprehensive mechanisms for recording data usage and authorisation, including data sources, scope of authorisation, model training processes and content generation records.

When relevant data is used for model training, digital avatar development or other commercial purposes, the specific purpose, period and application scenario should be clearly documented. In the event of infringement, victims should only be required to provide preliminary evidence, such as highly similar images, sounds or other identifiable features. Platforms and service providers should then provide generation records, model invocation records and authorisation proofs. If the relevant

DOI: 10.54254/2753-7048/2026.35266

parties cannot provide a lawful authorisation basis, or if records are missing or have been tampered with, they should be presumed to have failed to fulfil corresponding management duties [10]. This approach can reduce the technical burden of proof on victims and enhance the traceability of infringing content.

4.2. Strengthening AI content labelling and algorithmic transparency

In an era of rapid technological development, relying solely on hidden digital watermarks is insufficient. Whenever core material has undergone substantial synthesis by a generative large model, platforms and publishers should affix a mandatory label stating "AI-generated" or "non-realistic imagery" in a prominent position. This requirement should apply to prank-style face-swap videos, composite photographs that imitate news footage and other content that may mislead viewers.

In consumer scenarios, clear disclosure should also be required where AI customer service, virtual presenters, personalised recommendations or dynamic pricing are used. Such disclosures can prevent practices including big data price discrimination and false advertising. Deployers of AI systems should bear strict obligations regarding transparency and information disclosure, particularly in high-risk commercial scenarios where user behaviour may be manipulated, consumer decisions influenced or discriminatory pricing may occur. Concealing algorithmic intervention mechanisms may violate basic commercial integrity and anti-fraud principles [11].

4.3. Establishing more effective platform liability and redress mechanisms

Future policies should require online platforms to establish dedicated redress channels for infringements involving AI-generated content. Once a victim submits an initial complaint supported by prima facie evidence, the platform should not only remove and block the content promptly, but also preserve relevant evidence. In practice, many infringers are difficult to trace and may delete videos after generating traffic. Platforms should therefore use technical tools to preserve the relevant AI-generated content and backend data, preventing wrongdoers from deleting posts and avoiding liability.

When artificial intelligence systems no longer serve merely as passive information carriers but actively participate in content generation and targeted distribution, traditional immunity defences lose part of their original legitimacy. In this technological environment, judicial decisions should consider the active role played by platforms in content dissemination and their commercial incentives. Platform liability should be linked to the platform's ability to review, recommend, label, restrict and preserve AI-generated content [12]. A more effective redress mechanism should combine rapid removal, evidence preservation, user notification and compensation channels so that personality rights can be protected before the damage becomes irreversible.

5. Conclusion

Generative AI has transformed personality rights infringement from a relatively individualised legal problem into a complex risk involving identity replication, data reuse and platform dissemination. In traditional infringement scenarios, the actor and the victim are usually easier to identify. In the context of generative AI, the infringement chain may involve data collectors, model developers, prompt users, commercial deployers, platforms and secondary distributors. Although China's Civil Code and relevant policies have established important boundaries for AI-generated content, liability

DOI: 10.54254/2753-7048/2026.35266

allocation remains unclear, victims face evidentiary difficulties, and dissemination platforms often become central actors in the spread of harm.

The protection of personality rights in the age of generative AI therefore requires a more integrated response. Rules on the authorisation of personal characteristics should be clarified so that identity features such as face, voice and behavioural style cannot be reused beyond the original purpose of consent. Labelling rules and algorithmic transparency duties should be strengthened to help the public distinguish synthetic content from real content. Platform liability should also be improved through special complaint channels, evidence preservation obligations and more effective remedies. Future research can further examine public perceptions of AI face swapping, digital avatars and posthumous digital personality, as well as the effectiveness of platform governance in practice. As generative AI continues to expand into commercial, cultural and social spaces, the legal protection of personality rights must develop from fragmented ex post remedies toward a more preventive and traceable governance system.

References

[1] Cheng, X. (2026). Tort liability for generative artificial intelligence infringement of portrait rights. Legal Forum, 41(03), 5-16.
[2] Yuan, J. (2025). Constructing judicial review rules for the labelling of AI-generated and synthetic content: The first case concerning platform determination of user content as AI-generated. Journal of Law Application, (09), 121-135.
[3] Khosrowi, D., Finn, F. and Clark, E. (2025). Engaging the many-hands problem of generative-AI outputs: A framework for attributing credit. AI and Ethics, 5(5), 4495-4513.
[4] Menagh, N.M. (2025). When the Screen Lies: Navigating Privacy and Publicity in an AI World. Fordham Law Review, 94, 337.
[5] Grodon-Tapiero, A., Kaplan, Y. and Parchomovsky, G. (2025). Deepfake Liability. North Carolina Law Review, 104, 377.
[6] Wang, Z. (2026, April 1). Investigation into AI short dramas "stealing faces": Ordinary people become AI villains; where is the boundary of portrait rights? National Business Daily. Retrieved from https: //www.nbd.com.cn/articles/2026-04-01/4321530.html
[7] Kusche, I. (2024). Possible harms of artificial intelligence and the EU AI Act: Fundamental rights and risk. Journal of Risk Research, 1-14.
[8] U.S. Copyright Office. (2024). Copyright and Artificial Intelligence, Part 1: Digital Replicas Report.
[9] Xiao, X. (2026, April 14). Dialogue with the developer of Zhang Xuefeng.skills: Refinement, mythologisation and legal boundaries. 21st Century Business Herald. Retrieved from https: //m.21jingji.com/article/20260414/3c81b81692a74989021a59ecb2210474.html
[10] Rosati, E. (2025). Infringing AI: Liability for AI-generated outputs under international, EU, and UK copyright law. European Journal of Risk Regulation, 16(2), 603-627.
[11] Hacker, P., Engel, A. and Mauer, M. (2023). Regulating ChatGPT and other large generative AI models. In Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency (pp. 1112-1123).
[12] Yang, L. and Guan, T. (2025). From safe harbours to AI harbours: Reimagining DMCA immunity for the generative AI era. Journal of Intellectual Property Law & Practice, 20(9), 605-616. https: //doi.org/10.1093/jiplp/jpaf043