

AI Ethics in Chinese higher education: From “Anxiety” to “Innovation”

Yuke Zhang^{a*}

^aUCL Institute of Education, University College London, London, UK

* yuke.zhang.23@ucl.ac.uk

AI Ethics in Chinese higher education: From “Anxiety” to “Innovation”

Abstract

This critical review explores the ethical dimensions of artificial intelligence (AI) in Chinese higher education, tracing a trajectory from "anxiety" to "innovation." It examines the roots of ethical concerns such as data bias, privacy, and transparency, grounded in the rapid integration of AI technologies into teaching and learning. The paper discusses these results in the context of existing ethical frameworks, highlighting limitations such as unequal access and cultural challenges. It concludes by proposing actionable insights for policymakers and researchers, emphasizing the need for inclusive AI governance and empirical ethics-driven strategies to balance technological potential with humanistic values.

Keywords: AI ethics; Anxiety; innovation; Chinese higher education

Introduction

The integration of artificial intelligence (AI) in Chinese higher education has sparked both excitement and apprehension. While AI tools promise to lead to transformation in teaching and learning through personalized education, efficient assessment, and resource optimization (e.g. Chea & Xiao, 2024; Jia, 2024; Wu et al., 2024), they also raise critical ethical concerns. Existing research often frames these concerns in a relatively limited manner: either emphasizing AI's transformative potential (Wu et al., 2020) or critiquing its ethical pitfalls (e.g. Li & Gu, 2023; Hung & Chen, 2023). However, such discussions fail to capture the subtle interaction between these dimensions.

Initial concerns about the AI implementation in Chinese universities stemmed from various factors, including academic integrity, privacy, and the potential erosion of traditional teaching values. However, as higher education institutions gain experience and confidence in using AI applications, a noteworthy transition from anxiety to innovation is emerging. Moving beyond surface-level debates, this paper critically reviews the literature on AI ethics in Chinese higher education, discussing gaps and contradictions that hinder meaningful progress. By positioning ethical anxiety as a driver of innovation, this analysis reveals the transformative potential of AI when guided by inclusive, empirically grounded ethical frameworks.

Ethical Concerns in AI Use: The "Anxiety" Stage

The integration of AI in Chinese higher education has led to ethical dilemmas, predominantly centred around data bias, privacy, and transparency. Data bias emerges as a critical issue, as illustrated by algorithms favouring certain demographics over others in automated admissions or grading systems. For instance, Wang, Zechner, & Sun (2018) noted discrepancies in AI-driven assessments across nationalities, raising concerns about systemic inequities. Similarly, privacy invasions through facial recognition tools in Chinese universities highlight risks of excessive surveillance and data misuse (e.g. Wu, 2024; Huang, 2023; Zhang, 2025). These issues amplify student and staff ethical anxieties, reflecting broader societal concerns about technological overreach.

However, existing literature often focuses narrowly on technical drawbacks or individual experiences without addressing social, structural, and systemic contributors to these anxieties. For example, Zhu et al. (2025) link AI ethical anxiety to a lack of regulatory clarity, yet fail to propose actionable governance models. Additionally, Tang et al. (2024) explore teachers' cognitive dilemmas when adapting to AI technologies but overlook how institutional support systems might relieve such challenges.

Even though biases in AI systems are evident, they also provide opportunities to survey potential inequities within higher educational frameworks. For instance, addressing algorithmic transparency (Alfaleh, 2026) by making the data sources, decision criteria, and limitations of AI systems more visible, could help redefine accountability in Chinese higher education. Here, accountability refers to the responsibility of universities and educators to explain, justify, and, where necessary, correct AI-mediated decisions affecting students and teachers. In this sense, transparency could improve accountability by making such decisions more traceable, contestable, and open to institutional review (Liu et al., 2026). From this perspective, the "anxiety" stage is not solely a barrier but a transformative opportunity, challenging stakeholders to combine technological advancements with humanistic goals.

Transformative Practices: The "Innovation" Stage

While ethical anxieties surrounding AI in Chinese higher education are well-documented, they have also motivated innovative practices aimed at addressing these concerns. Adaptive learning platforms and AI-powered assessment tools exemplify how technology can enhance higher education when ethically guided (e.g. Fu & Weng, 2024; Yang & Weng, 2023; Hao, Fang & Peng, 2024; Ouyang & Zhang, 2024; Dai, Liu, & Lim, 2023). These innovations illustrate how AI can drive equity and efficiency when aligned with

ethical objectives. Nevertheless, transformative practices often reveal potential contradictions. Adaptive platforms, while addressing regional disparities, have been criticized for perpetuating biases embedded in their training data. Similarly, automated grading tools, though efficient, lack the ability to account for nuance and creative responses, sparking debates about their validity in humanities higher education. Such critiques highlight the need for transparent AI design processes that incorporate diverse stakeholder viewpoints. Although AI innovations demonstrate potential to reshape education, innovative practices need to move beyond technical aims to engage in broader dialogues about Chinese higher education's core objectives.

Conclusion

The transformation of AI ethics in Chinese higher education reveals both promising developments and critical challenges that demand attention. Notably, the rapid pace of AI adoption often exceeds the development of ethical frameworks, creating a concerning gap between technological implementation and ethical governance.

The current approach in Chinese higher education involves adapting dominant Euro-American and international AI ethics frameworks, such as those emphasizing human agency, transparency, fairness, accountability, privacy, and technical robustness in the EU Ethics Guidelines for Trustworthy AI, the OECD AI Principles, and UNESCO's Recommendation on the Ethics of Artificial Intelligence (Güzelergene, Cinar & Nayır, 2026; Ramamoorthy, 2026). If this proceeds without sufficient consideration of Chinese pedagogical traditions and cultural values, this risks creating superficial imitation rather than meaningful integration. Furthermore, the emphasis on technological innovation sometimes covers up significant discussions about the equity and accessibility of Chinese Higher Education, particularly in less-developed regions.

Future research should focus on developing better ways to account for the ethical aspects of AI implementation, explore the range of views of different stakeholders, and investigate the multifaceted ethical impacts of AI use in China.

References

- Alfaleh, M. (2026). Sustainable AI-Driven Assessment in Higher Education: A Systematic Review of Fairness, Transparency, Pedagogical Innovation, and Governance. *Sustainability*, 18(2), 785.
- Chea, P., & Xiao, Y. (2024). Artificial intelligence in higher education: The power and damage of AI-assisted tools on academic English reading skills. *Journal of General Education and Humanities*, 3(3), 287-306.

- Dai, Y., Liu, A., & Lim, C. P. (2023). Reconceptualizing ChatGPT and generative AI as a student-driven innovation in higher education. *Procedia Cirp*, 119, 84-90.
- Fu, Y., & Weng, Z. (2024). Navigating the ethical terrain of AI in education: A systematic review on framing responsible human-centered AI practices. *Computers and Education: Artificial Intelligence*, 7, 100306.
- Güzelergene, E. S., Cinar, D. B., & Nayır, F. (2026). Artificial intelligence and ethics: a global perspective. *Çukurova Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 35.
- Hao, Z., Fang, F., & Peng, J. E. (2024). The integration of AI technology and critical thinking in English major education in China: Opportunities, challenges, and future prospects. *Digital Applied Linguistics*, 1, 2256-2256.
- Huang, L. (2023). Ethics of artificial intelligence in education: Student privacy and data protection. *Science insights education frontiers*, 16(2), 2577-2587.
- Hung, J., & Chen, J. (2023). The benefits, risks and regulation of using ChatGPT in Chinese academia: a content analysis. *Social Sciences*, 12(7), 380.
- Jia, J. (2024). Artificial Intelligence Implementation in Higher Education in China: Case Study of Beijing Technology and Business University. *Journal of Current Social Issues Studies*, 1(1), 17-31.
- Li, S., & Gu, X. (2023). A risk framework for human-centered artificial intelligence in education. *Educational Technology & Society*, 26(1), 187-202.
- Liu, X., Luo, J. H., Yun, T., & Lan, X. (2026). Enhancing Academic Integrity: Drivers, Barriers, and Governance Pathways in China. *Higher Education Quarterly*, 80(2), e70107.
- Ouyang, F., & Zhang, L. (2024). AI-driven learning analytics applications and tools in computer-supported collaborative learning: A systematic review. *Educational Research Review*, 44, 100616.
- Ramamoorthy, L. (2026). Ethical Principles and Global Guidelines for Trustworthy AI Systems. In *Trustworthy AI Systems: Engineering Secure, Scalable, and Responsible Intelligence for Real Applications* (pp. 15-29). Cham: Springer Nature Switzerland.
- Tang, K., Ma, H., & Wu, Y. N. (2024). Sustainability in Professional Development: Manifestations and Reliefs of Technology Anxiety among Mandarin Chinese Teachers in the Era of Digital Intelligence. *Clinical Case Reports and Studies*, BioRes Scientia Publishers. 7(6):1-17.
- Wang, Z., Zechner, K., & Sun, Y. (2018). Monitoring the performance of human and automated scores for spoken responses. *Language Testing*, 35(1), 101-120.
- Wu, F., Lu, C., Zhu, M., Chen, H., Zhu, J., Yu, K., ... & Pan, Y. (2020). Towards a new generation of artificial intelligence in China. *Nature machine intelligence*, 2(6), 312-316.
- Wu, S., Cao, Y., Cui, J., Li, R., Qian, H., Jiang, B., & Zhang, W. (2024). A comprehensive exploration of personalized learning in smart education: From student modeling to personalized recommendations. *arXiv preprint arXiv:2402.01666*.
- Wu, Y. (2024). Facial recognition technology: College students' perspectives in China. *Journal of Research in Social Science and Humanities*, 3(1), 53-79.
- Yang, M., & Weng, F. (2023). AI-powered personalized learning journeys: revolutionizing information management for college students in online

- platforms. *Journal of Information Systems Engineering and Management*, 8(1), 23196.
- Zhang, A. H. (2025). The promise and perils of China's regulation of artificial intelligence. *Columbia Journal of Transnational Law*, 63(1).
- Zhu, W., Huang, L., Zhou, X., Li, X., Shi, G., Ying, J., & Wang, C. (2025). Could AI ethical anxiety, perceived ethical risks and ethical awareness about AI influence university students' use of generative AI products? An ethical perspective. *International Journal of Human–Computer Interaction*, 41(1), 742-764.