# A Study of Copyright Governance of AIGC Academic Norms in a Higher Education Environment

Yu,Shiyu

East China Jiaotong University, Nanchang, Jiangxi, 330000, China

**Abstract:** Generative AI represented by ChatGPT brings new challenges to the development of higher education field in China. There is a lack of empowerment between the application and regulation of generative AI. This paper analyzes the current situation of the application of generative AI, and clarifies the two hidden risks embedded, namely, the risk of use and the risk of identification. It vividly explains the impact and influence of generative AI on academic norms and copyright governance in the input, operation, and output ends, and proposes that a clear attitude, policy regulation, and technological autonomy are needed to monitor the use of generative AI and standardize its use, so as to better realize the principles of human-computer synergy and intelligence for the good.

Since the launch of the Big Language Model (ChatGenerativePre-trainedTransformer, Chat Generative Pre-trained Transformer) by the US Open Artificial Intelligence Research Center (OpenAI) in 2022, the global generative AI has entered the culmination of the development of the Digital Intelligence Era. Especially in the past two years, with the emergence of various AI tools such as GPT4.0, StableDifussion, Sora, etc., the level of people's previous knowledge of AI big language models has been subversively changed.

In this context, the AlforScience model of Al-driven scientific research has made unprecedented breakthroughs in broadening the boundaries of scientific discovery, improving the efficiency of scientific research, and enhancing the speed and accuracy of numerical simulations. As a result, generative Al has an application prospect and potential in the higher education environment that cannot be ignored, which can promote the innovation of higher education to a certain extent, whether it is scientific research or enabled education.

Generative Artificial Intelligence (GenerativeArtificialIntelligent) is widely used in higher education because of its multi-source and open multi-modal information aggregation and conversion capability, efficient and convenient cross-domain information integration and generation capability, as well as its wide-area innovation and strong discursive information screening and simulation capability. However, it has also brought about controversial issues such as academic misconduct and copyright infringement, which have profound impacts on the development of China's higher education environment at multiple levels. Therefore, it is of great significance to study the copyright governance of generative AI in the application of academic norms in the field of higher education in China.

This paper discusses from multiple perspectives, such as education, law, and technology, and provides an in-depth summary as well as an analysis of the research literature related to the copyright governance of AIGC academic norms in the higher education environment published in the China Knowledge Network database, with a view to contributing to the development of the relevant research in the field and the construction of the rule of law of generative AI.

Keywords: Artificial intelligence; Copyright; Higher education; Academic norms

DOI: 10.62639/sspjiss05.20250203

**About the Author** 

<sup>(</sup>Manuscript NO.: JISS-25-3-62009)

Yu, Shiyu (2004-), female, Han, Yantai, Shandong, student, undergraduate, East China Jiaotong University, intellectual property.

# 1. The Current Application Situation of Generative Artificial Intelligence in the Field of Higher Education in China

Generative artificial intelligence can show strong human nature and high efficiency. Through massive data capture and model training, it can imitate human language habits and expression styles to a high extent. At the same time, combined with the keyword requirements given by the user, it generates text, images or videos that meet their set goals. It plays an important role in the fields of automatic writing, language translation and academic research assistance, so it is mostly used in the daily teaching activities of college teachers and students to improve efficiency and save costs. According to a study on the application of generative AI by teachers and students in Chinese colleges and universities in 2024 initiated by Maxis, nearly 20% of teachers and students in college students and 56% of teachers will frequently use generative AI software in their study and work life, and the proportion of college teachers and students who have never used generative AI is only 1%. From this, it can be seen that generative AI is bound to become an indispensable part of digital education in the future, and may even become a new education model, empowering education innovation with AI.

But with it come challenges and risks. More than 30% of undergraduate students mainly use AI to write papers or assignments, and 58% of teaching and research personnel most often use generative AI for academic research. However, as the development speed of generative AI technology has far exceeded human expectations, many hidden dangers that users cannot detect in a short time are also more hidden and diversified. Especially in the field of copyright, generative AI has brought a great impact on the existing copyright system, and brought a lot of uncertainty and potential risks to the application of copyright law.

## 2. Risk Analysis in the Application of Generative Artificial Intelligence

#### (1) Risks of technique : algorithmic black box leads to recognition problems

Algorithms in generative AI technology extremely lack transparency and interpretability at present. The essence of the traditional black box problem has changed from "information asymmetry between human beings" to "common ignorance of human beings in the face of strong artificial intelligence".

In the new era where algorithms and data can control information, the original existence and continuous evolution of the algorithmic black box has increased or even cut off the possibility of knowing the internal operation and decision-making process of generative AI. Even developers can't fully analyze the content generated by deep neural networks that utilize multiple levels of neurons to build recognition patterns. This complex non-linear operation mechanism makes it impossible for universities to regulate and investigate the infringement behavior of generative AI on an accurate and effective basis. In the field of copyright, there is a technical "Safe Harbor Rule" and "Technological Impossibility", network service providers cannot be recognized as the bearer of copyright infringement liability.Hence, in the event of infringement disputes and academic misconduct, it is difficult for universities to identify the work and responsibility, such as whether the content generated by AIGC can be a work protected by the copyright law, and who should bear the responsibility after infringement disputes occur. At the same time, because China's copyright law and other related copyright protection laws have not yet given specific identification of AIGC-generated content, so universities in practice, there are difficulties in evidence, insufficient legal basis, lack of expertise in the reality of the dilemma. This not only makes it difficult for university supervisory departments to obtain comprehensive and accurate information to avoid supervision.

Artificial intelligence has not only its technical and technological attributes, but also its political attributes. Due to the uncontrollability and potentiality of the technology, it is easy to generate misleading information, algorithmic

discrimination and algorithmic bias in the process of output, and invisible consciousness control on college students, which may even lead to ethical risks and ideological permeation and other major issues affecting national security. Especially for students who have not yet formed stable and independent values, this phenomenon is likely to negatively influence the Western mindset, stereotypes and sensitive topics in a subtle way. However, due to the limited development of current technology, it is difficult for universities to identify and capture negative information in a timely manner, especially the potentially large amount of data, and it is difficult to avoid the impact of AI on the ideological security of universities.

At present, there is no official, authoritative and clear regulation on the use of AIGC in academic norms at either the legislative level or the university level, and therefore, it is not possible to effectively solve the problem of identification at the practical level of universities. However, Article 3 of the Measures for the Prevention and Handling of Academic Misconduct in Colleges and Universities issued by the Ministry of Education (hereinafter referred to as "the Measures") emphasizes that "colleges and universities shall adhere to the principles of prevention and combination of education and discipline in the prevention and handling of academic misconduct". Therefore, for AIGC, which is still in the initial stage of development in China, we should keep an open, tolerant and prudent attitude, and strictly abide by the legal bottom line on the basis of adhering to academic integrity.

#### (2) Risks of use: training data empowerment leads to infringement disputes

At this stage, the development of AI still relies heavily on human intellectual work, and needs to use the rich data base as a source to replicate and "learn" the rules, patterns, structures, and trends of human language and artistic creation through data encoding of content, in order to generate content that conforms to language habits and has aesthetic value. In the academic specification of universities, the high quality and large volume of academic data has high requirements, often the user requirements of generative content will be generated around the core of copyrighted intellectual property results, at this time for the use of data whether the authorization is often the key to the problem of copyright infringement. At present, generative artificial intelligence has formed a fixed research and development mode that includes data input, model training, model adjustment, content output and other steps. On the input side, data infringement often involves the reproduction and adaptation rights of the right holders. In addition to manual "Feeding ", AIGC technology will also use crawlers to crawl the data. In this process, it is very likely that the results of other people's intellectual property rights will be included in the information database without their permission, which is highly uncontrollable in terms of scope, quantity and quality. However, due to the massive amount of data crawling, it is extremely difficult to monitor such infringement behaviors in a timely manner, and the results of infringement are often lagging behind, which may also lead to "copyright market failure".

In the face of the problems arising from the use of AIGC that affects academic standards, colleges and universities are trying to implement applicable solutions, and the mainstream practices of overseas colleges and universities are mainly in the two models of "total ban" and "conditional permission". For example, in January 2023, the New York City Department of Education (DOE) announced a ban on the use of ChatGPT on school equipment and networks. Jenna Lyle, a spokesperson for the New York City Department of Education, suggested, "While this tool may be able to provide quick and easy answers to questions, it does not foster the critical thinking and problem-solving skills that are essential for academic and lifelong success." However, due to the powerful tool attributes and convenient performance of ChatGPT, in the era of continuous technological development and progress, the use of AIGC is prohibited and cannot be eliminated, and will gradually change into a more secretive and difficult to regulate the way of use. On the one hand, it is more likely to aggravate the occurrence of academic misconduct and other matters, and it is not in line with the essence of copyright law, which is to stimulate innovation through dissemination. As a result the New York City Department of Education removed its blanket ban on ChatGPT in May 2023 after revisiting its policies and announced the formation of the Artificial Intelligence Policy Laboratory that same year to ensure colleges and universities were properly utilizing the technology.

The "conditional license" path is the approach adopted by more universities at home and abroad, such as Harvard University, the University of California, Deakin University and other well-known overseas universities, as well as the first domestic university to issue an "AI ban" - Fudan University. -Fudan University. In the "Regulations of Fudan University on the Use of AI Tools in Undergraduate Thesis (Design) (for Trial Implementation)" (hereinafter referred to as the "Regulations") issued by Fudan University, six standardized management rules on the use of AI tools in undergraduate thesis (design) are clearly stated: The School of Journalism and Communication of Beijing Normal University, in conjunction with East China Normal University, jointly released the Guidelines for Students' Use of Generative Artificial Intelligence (AIGC); The School of Continuing Education of Communication University of China as well as the Academic Affairs Office of Tianjin University of Science and Technology have also successively stipulated the scope of application of AIGC to undergraduate theses. In a study conducted at the University of Chinese Academy of Sciences, it was shown that there were significant differences in critical thinking and selfdirected learning skills among college students who used GenAI (a series of tools represented by ChatGPT) tools and who persisted in the use of such tools compared to those who did not use them, or did not persist in their use. That is, the use of GenAl tools by higher education students exercises critical thinking and enhances selfdirected learning. Therefore, we should not completely confront and prohibit science and technology on the basis of keeping the right thing. At present, although universities have explored the regulation of the use of AIGC with the advancement of technology, there is still a lack of a legally compliant basis for its wide adoption. Existing legal norms may be difficult to effectively protect intellectual property rights, and there is a high risk of copying and misappropriation of academic ideas and research results, violating the legitimate rights and interests of copyright holders, and causing an inherent academic "crisis of confidence" in colleges and universities.

## 3. Suggested Countermeasures

#### (1) Clarify the concept of academic norms and uphold a prudent approach to governance

With the development of science and technology driving the gradual development of artificial intelligence to the "strong artificial intelligence" era, the discussion of its legal status has also triggered a series of controversies: Part of the affirmative scholars believe that artificial intelligence can be given "electronic person" "anthropomorphic person" civil subject qualification; the negative claim that artificial intelligence is only a civil object, can not be given the status of the subject; compromise is to recognize the limited subject qualification of artificial intelligence, to enjoy the limited power, bear limited obligations. In the higher education environment, "denying the civil subject status of artificial intelligence is the fundamental compliance with the spirit of humanism, safeguarding the anthropomorphic status and respecting the dignity and value of human beings." It is more in keeping with the fundamental aims of our academic norms to view generative AI exclusively as a tool, and its generators as the yield of things.

When regulating generative AI, most scholars believe that an "Inclusive and Prudent" attitude should be adopted: "Inclusive and Prudential Regulation requires regulators to provide sufficient development space and fault-tolerant environment for generative AI while maintaining a high degree of vigilance on potential risks and taking effective preventive measures in a timely manner when implementing regulation." In addition to accommodating the foreseeable shortcomings and limitations of the development of generative AI, the regulators should take a prudent attitude towards the regulation and handling of its risks, continuously optimize and adjust their strategies according to the actual situation, and dynamically evolve with the development of the times and technological progress.

The potential consequences of the loss of independent thinking and critical ability caused by relying on generative AI and the various risks caused by the unrestricted use of AI technology are the solution objectives that colleges and universities should prioritize when responding to the challenges posed by generative AI. While regulating, they should first clarify their own cognitive concepts of generative AI and regulatory attitudes towards

academic norms, in order to better formulate norms, solve practical problems, and promote the progress of legislation.

#### (2) Strengthening policy and regulatory instruments and establishing protective mechanisms for constraints

At present, most of the academics in China tend to amend the laws and regulations to include the use of training data into the category of "fair use" in the copyright protection of training data. However, in a horizontal comparison, the European Union "Artificial Intelligence Act" and "Copyright Directive" provisions, although set up the reasonable use of machine training provisions, but retains the copyright holder's "opt-out", that is, without the authorization of the copyright holder's permission of the data, the machine training is not allowed to arbitrarily use and crawl. "In fact, the copyright choke point for training in generative AI works is not in paying for them, but in obtaining authorization on a case-by-case basis." Despite the operational cost of obtaining authorization one by one, based on the broad market environment for the future development of generative artificial intelligence and the considerable economic benefits, coupled with the extremely rapid progress of technological development, this cost can be gradually simplified in the future. And in the higher education environment, the data required for machine learning is more concentrated in the academic achievements of universities and research institutes, and the awareness of intellectual property protection is stronger, and the possibility of copyright infringement disputes is also higher. For enterprises and universities, the practicability and defensibility of applying licensing is higher.

In addition to determining the legal regulation of data authorization, it is also necessary to combine a variety of measures to jointly participate in the specific governance of generative artificial intelligence in the higher education environment. For example, colleges and universities should introduce comprehensive and specific regulatory policies, clarify the regulatory red line, adhere to the concept of human-computer cooperation with human subjects in teaching, and distinguish the auxiliary functions of generative AI for academic paper touch-ups, review feedback, editing optimization, etc. from academic misconduct such as academic copying, plagiarism, etc., and put an end to the latter. At the same time, attention should also be paid to ideological supervision, to implement the guiding ideology of Marxism into the construction of ideology, to correctly play the leading role of Marxist values, and to prevent phenomena such as the alienation of Western thought. Secondly, colleges and universities should establish professional technicians and technical equipment, and set up a collaborative supervision model for faculty and students and an integrated top-to-bottom supervision system, so as to ensure that the information is open and transparent, and the treatment is fair and reasonable.

In addition, scientific research norms for generative AI should be reformulated to guide researchers to conduct responsible research. There are already a number of documents, such as the Guidelines for the Use of Generative Artificial Intelligence in Education and Research issued by UNESCO 2023, the Guidelines for the Code of Conduct for Responsible Research (2023) issued by the Chinese Ministry of Science and Technology, and the Guidelines for the Boundary of AIGC Use in Scholarly Publishing published by the China Institute of Scientific and Technological Information, which indicate scientific research compliance with the use of generative AI in scholarly norms and regulations. In order to ensure the accuracy of academic research and the innovativeness of research results, we should combine "self-discipline" and "other-discipline", and pursue excellence, integrity and innovation while rationally using generative AI.

#### (3) Utilize AI technology autonomy to achieve co-win technology empowerment

"The opportunities that generative AI brings to education are mainly manifested in five aspects: assisting teachers in education, facilitating students' personalized learning, reforming the education evaluation model, optimizing the learning environment, and promoting the development of students' critical thinking." Under the specific application environment of generative AI, technological means should be actively adopted to build a technology-enabled copyright governance pathway with government and university supervision as the main focus and platform autonomy as a supplement.

Utilizing the AIGC tool, measures should be taken for early prevention, verification, as well as timely supervision and checking after the fact, in order to address the scientific research ethical issues arising from the possible fabrication, bias, outdated citation, and opacity that may exist in the technical restrictions of Chat GPT, which may affect the accuracy of the research. At the operational end of the content generation phase, the Regulatory Sandbox mechanism should be applied to address the algorithmic black box problem, "The Regulatory Sandbox provides an effective way of mitigating information asymmetry through the exchange of adequate information among the various participants throughout the access, operation, and exit phases" It establishes a safe, controlled and dynamically regulated environment for generative AI scenarios, applications, and risks at the development and pre-market stages in advance, and tests the performance, safety, and potential impacts of AI systems for a limited period of time under a specific sandbox program agreed upon by the AI system provider or potential provider and the competent authority before these AI systems are put on the market or otherwise put into use. Not only can it provide tech companies and regulators with inclusive trial and error in the early stages of risk, but it can also stimulate the vitality of corporate research and innovation, and accumulate a large amount of data and experience, so as to prepare for the development and risk descent of generative AI in the future. At the same time, there are also flexible and diverse technical regulatory tools such as cue engineering, which "can reduce the risk of uncontrolled output of algorithmic models, such as algorithmic bias and algorithmic discrimination, which are common, and ensure that generative Al-generated content meets the requirements of ethicality."

On the input side of training data, mark recognition can be used to mark the data source after obtaining data authorization in order to leave traces, such as marking watermarks and other conventional means, in addition to Machine Unlearning. Relative to Machine Learning, Machine Unlearning is to remove some specific sample information from the model due to the need for privacy, usability and right to be forgotten, to meet the request of Machine Learning to remove the training samples from the model due to the privacy, intellectual property rights, usability, or other rights requirements. Compared to traditional means of data protection, such as deleting the training set, machine forgetting is able to minutely condition the data to hide the private and secure information about the people involved in the data while maximizing the retention of the underlying data of the original training samples, reducing the huge cost of re-training the model, and also helping to enhance the legal rights of the copyright holders to better meet the regulatory needs.

To summarize, at the level of technical regulation, with the development of generative AI technology, the level and measures of technical countermeasures will also continue to progress. Different technical paths are adopted to regulate the input, operation and output of generative AI, so as to realize the dynamic balance of the final generated content and reduce the possibility of copyright infringement disputes. Especially in the field of higher education, cutting-edge academic resources and scientific research power of universities can be fully utilized to achieve accurate technical regulation of generative AI, to deal with the ethical challenges of scientific research generated by it, and to "build a more inclusive technical value system in order to prevent generative AI from becoming an invisible knowledge cage".

#### 4. Conclusion

The idea is to emphasize the need to regulate the values of AI at the legal, ethical and humanitarian levels through various regulatory channels, so as to ensure that the development of AI is safe and controllable, as well as at the educational level. The development of AI technology is a double-edged sword. While we emphasize the unprecedented innovative heights it provides for scientific research, we should also guard against its potential crises in research ethics, moral ethics and even legal governance. The industrial system formed by the generative AI business operation model of "technology supporter-service provider-AI user" will become increasingly mature and

optimized in the future. "Under the idea of full-cycle governance, the governance of copyright infringement risk of generative AI needs to be formulated according to the different positions of different subjects in the value chain." At the level of academic norms, it is also necessary to regulate and constrain academic misconduct with a full-cycle and comprehensive perspective, paying particular attention to the scientific research values of young student groups. In the process of dynamic development, we constantly adhere to the bottom line, explore innovation, adhere to the principle of People-oriented, Intelligent and Good, and promote the governance framework of human-machine synergistic development and science and technology-enabled innovation in the higher education environment. Improve the regulatory path of tolerance and prudence, technological autonomy, and maximize risk avoidance while realizing the harmonious and progressive development of the higher education environment.

# References

- [1] Liu Yaoyao, Liang Yongxia, Li Zhengfeng. Generative Artificial Intelligence and Research Ethics: Changes, Challenges and Prospects[J]. Science Monitor, 2024, 19(04): 1-8.
- [2] Zhang Linghan, Yu Lin. From traditional governance to agile governance: governance paradigm innovation with generative artificial intelligence[J]. E-Government, 2023, (09):2-13.
- [3] Xu Lei. Development and safety at the same time: inclusive prudential regulation of generative artificial intelligence risks[J]. Theory and Reform, 2024, (04): 67-83+176.
- [4] Yang Rencai. Challenges and Responses to Artificial Intelligence-Enabled Ideological and Political Education in Colleges and Universities[J]. Journal of National College of Education Administration, 2020(5):54-59.
- [5] SeeAndrewW.Torrance&BillTomlinson,TrainingisEverything:ArtificialIntelligence,Copyright,and"FairTraining",128DickinsonLaw Review233,242(2023).
- [6] Zhang P. Institutional Response to the Copyright Problem of Generative Artificial Intelligence[J]. Southern Media Research, 2024,(05):21-31.
- [7] WendyJ.Gordon.FairUseasMarketFailure:AStructuralandEconomicAnalysisoftheBetamaxCaseandItsPredecessors[J].ColumbiaL awReview,1982,82(08):1600-1657.
- [8] QI Jia, XU Yanru, LIU Ji'an, XUE Kai. Effects of generative artificial intelligence tool use on college students' critical thinking and independent learning abilities[J]. Research on E-Chemical Education, 2024, 45 (12): 67-74.
- [9] Chen Si-Min, Cai Sheng. On the Negation of Civil Subject Qualification of Artificial Intelligence[J]. Network Security and Data Governance, 2024, 43 (09): 71-76+83.
- [10] Xie Yizhang. Copyright Disputes and Resolutions of Generative Artificial Intelligence Work Training[J]. Chinese Editorial, 2024,(11):38-46.
- [11] UNESCO. Guidelines for the application of generative artificial intelligence in education and research [EB/OL]. (2024-7-17).
- [12] Supervision Department, Ministry of Science and Technology, People's Republic of China. Guidelines on the Code of Conduct for Responsible Research (2023)
- [13] Scholarly Publishing AIGC Boundary Guidelines Released: use of AIGC requires a disclaimer or will constitute academic misconduct[N].(2024-6-19).China Youth Daily.
- [14] Zou Jiahui, Rong Weidong. Research Status and Prospect of Generative Artificial Intelligence Enabling Education in China[J]. Journal of Chengdu Normal College, 2024, 40(06): 99-106.
- [15] Liu Sheng.Jurisprudential Logic and Institutional Expansion of Regulatory Sandbox[J].Modern Law, 2021, 43(01):115-127.
- [16] Liu, X.R.. From misbehavior to regulation: innovation of regulatory framework for generative artificial intelligence[J]. Cybersecurity and Data Governance, 2024, 43 (06): 58-63+71.
- [17] Yu Jianxing, Liu Yuxuan, Wu Chao. Change and Governance of Artificial Intelligence Megamodeling[J]. Chinese Administration, 2023,39(4):6-13.
- [18] Shao H.H.. A Study on Copyright Infringement Governance of Generative Artificial Intelligence[J]. Publishing and Distribution Research, 2023, (06): 29-38.
- [19] Nadine Liv, Neurolaw: Brain- Computer Interfaces, University of St. Thomas Journal of Law and Public Policy, 2021,15(1)p.348
- [20] Tianxiang He, The Sentimental Fools and the Fictitious Authors Rethinking the Copyright Issues of Al- generated Contents in China, Asia Pacific Law Review, 2019, 27(2), p.229
- [21] Wang Qian: "On the Characterization of Content Generated by Artificial Intelligence in Copyright Law" Legal Science (Journal of Northwest University of Politics and Law), No. 5, 2017.