MINDWELL SCIENCE PRESS SDN. BHD.

Journal of Emerging Technology Governance
Volume 1, Issue 2

RESEARCH ARTICLE

Ethical Problems Arising from the Academic Application of Generative Artificial Intelligence and Their Responses

Hongpeng He[1,*]

[1] *Institute for Advanced Study in Humanities and Social Sciences, Beihang University, Beijing 100191, China*

*** Corresponding author:** Hongpeng He, E-mail: hehongpeng@buaa.edu.cn

Published: 28 September 2026

Abstract

Although generative artificial intelligence (GenAI) has been available only a short time, it is already widely used in the academic field, and the academic ethical problems it gives rise to have attracted close attention from society and especially from the scholarly community. Drawing on the academic ethical norms already issued by academic institutions, four main categories of academic ethical problem induced by generative artificial intelligence may be distilled: content plagiarism that is difficult to establish, concealed data leakage, systematic knowledge bias, and self-reinforcing disinformation. Compared with traditional academic ethical problems, the root causes of these lie in a reduced sense of moral guilt, the consolidation of systematic tendencies, illusions of understanding, and the non-disclosure of algorithmic mechanisms. In response, and building on the relevant academic ethical norms, citation conventions should be innovated, services should be used with caution, mutual human–machine understanding should be fostered, and data should be supervised throughout, so as to meet the academic ethical shock that generative artificial intelligence brings.

Keywords: generative artificial intelligence, large language models, academic ethics, research misconduct, ethical governance

1 The Problem Posed and the Current State of Research

1.1 The Problem Posed

Generative artificial intelligence is currently being applied ever more widely in academic activity. Surveys show that nearly half of Cambridge University students admit to having used ChatGPT in their studies[1], and nearly three quarters of Zhejiang University students use generative artificial intelligence tools in research activities or coursework[2]. Rising research talent can use the technology to find answers to questions quickly, and can even use it to supply ideas, generate text, and revise and edit, with the possible consequence that the academic thinking capacity the process was meant to train is replaced by technology. At the same time, generative artificial intelligence can become an efficient research assistant. Surveys show that one third of postdoctoral researchers worldwide use ChatGPT every day[3], which has also raised concern about such academic ethical problems as plagiarism and disinformation. In September 2023, for example, an academic paper was retracted because the phrase "Regenerate response" appeared in it — the first SCI paper retracted for the use of ChatGPT.

The academic ethical shock that generative artificial intelligence brings places academic institutions in a predicament. In February 2023 the University of Hong Kong announced a ban on the use of ChatGPT or other generative artificial intelligence tools on campus, while Harvard University encouraged the use of generative artificial intelligence technology for responsible research[4]. The contradictory attitudes among academic institutions show that they are responding to the academic ethical challenge of generative artificial intelligence passively, and that there is an urgent need for the scholarly community to conduct deeper research and respond better. The application of generative artificial intelligence in academic activity has become an irreversible trend of the age; the key is to clarify what academic ethical problems generative artificial intelligence will bring, and to regulate the use of the technology.

1.2 The Current State of Research

As an auxiliary tool for scientific research, generative artificial intelligence holds considerable potential for application in the academic field, chiefly in the following respects. First, it inspires research ideas and designs research schemes: Rahman, for example, verified by putting questions to ChatGPT that it can supply some unexpected and promising research ideas[5].

Copyright: © 2026 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), permitting distribution and reproduction in any medium, provided the original work is cited.

Second, it analyses literature and data and raises research efficiency: Zhu Yongxin and Yang Fan show by case example that ChatGPT can assist researchers in programming, greatly raising the efficiency of conventional programming and thereby accelerating the pace of research innovation[6]. Third, it assists academic writing — improving grammar, translating automatically, generating sentences and so on: Berg points out that researchers may either passively obtain content quickly from generative artificial intelligence and perfect or develop an argument with the least expenditure of effort, or engage it in dynamic dialogue by the Socratic method, iteratively developing and refining their arguments[7].

Generative artificial intelligence, however, also induces a series of academic ethical problems, including at least the following. First, authorship and intellectual property: researchers have reached a preliminary consensus that generative artificial intelligence should not become an author of a paper, and Van Woudenberg and others, proceeding from such perspectives as the nature of authorship and the function of attributing authorship, deny generative artificial intelligence the status of author[8]. Second, the authenticity of information: Holly points out that ChatGPT will create realistic and intelligent-sounding text in response to a user's prompt, and can generate false paper abstracts, and that people find it hard to tell the difference[9]. Third, the problem of fair access: as a useful research tool, unequal access to generative artificial intelligence will aggravate academic injustice; Kasneci and other scholars point out that non-English-speaking users have less access to generative artificial intelligence technology than English-speaking users[10]. Fourth, the problem of plagiarism, concerned mainly with whether and how content produced by artificial intelligence may be used: Jarrah and others, by a literature review, find that originality, citation and other matters involved in the use of generative artificial intelligence content remain disputed[11]. The problems above mean that generative artificial intelligence strikes at the existing system of academic ethical governance; Wang Shao, for example, holds that ChatGPT strikes at the connotation of the concept of academic misconduct at the levels of originality, error and fairness[12].

Existing research proceeds mainly from the standpoint of the user, revealing problems through practical operation and the summation of experience, and emphasising that the technology aggravates traditional academic ethical problems. Its understanding of the distinctive character of the academic ethical problems of generative artificial intelligence is, however, insufficiently deep, and it pays relatively little attention to the practical concern of academic institutions about this problem. The present study, by investigating the ethical norms formulated by academic institutions and researchers' cognition of the ethical risk of generative artificial intelligence, summarises the four categories of academic ethical problem that concern academic institutions, analyses in depth the characteristics and roots of those problems in the light of the relevant literature, and proposes means of response.

2 The Academic Ethical Problems of Generative Artificial Intelligence

Before discussing the concrete academic ethical problems, a preliminary question must be settled: ought generative artificial intelligence to be introduced into academic activity at all? Ethics offers many explanatory frameworks for assessing the legitimacy of an act, and these fall mainly into two domains: deontology (concerned with the act and the process) and teleology (concerned with the result and the objective). From the standpoint of teleology, the criterion for whether generative artificial intelligence should be used in academic activity is whether it better realises the objective of academic activity — that is, what Robert K. Merton called "the extension of certified knowledge". Supporters evidently believe that the technology can raise the efficiency of knowledge production and ultimately benefit human knowledge as a whole; but since generative artificial intelligence also supplies disinformation and produces biased results, whether it is conducive to the realisation of the academic objective may depend on the circumstances. From the standpoint of deontology, researchers should observe the moral obligations of academic activity, which appear as the normative requirements of that activity: the requirement of fairness in academic activity, for example, demands that researchers acknowledge the contributions of other scholars by citation and other means, and this is a foundational requirement for the operation of the academic system. Generative artificial intelligence integrates literature quickly and outputs it as "new knowledge", which may give users an unfair advantage and erode academic fairness; but with the construction of new norms there is a chance that the fairness problems arising from the use of generative artificial intelligence may be resolved. In short, the discussion of whether generative artificial intelligence should be introduced into academic activity must still be reduced to concrete ethical problems.

2.1 Content Plagiarism That Is Difficult to Establish

Under the definition of United States federal policy, plagiarism means "the appropriation of another person's ideas, processes, results, or words without giving appropriate credit"[13]. Generative artificial intelligence induces two new problems of plagiarism. The first is the user's plagiarism of content output by generative artificial intelligence: the act of using content output by generative artificial intelligence without labelling it satisfies the definition of plagiarism above. The second is plagiarism by generative artificial intelligence of its training data: since generative artificial intelligence can

Copyright: © 2026 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), permitting distribution and reproduction in any medium, provided the original work is cited.

produce new content only on the basis of training data, and that content may be highly similar to the training data in idea and in content, publishing such content publicly may constitute plagiarism.

The traditional governance of plagiarism relies on technical detection and on researchers' own professional ethics, but the plagiarism that generative artificial intelligence induces poses a challenge to these traditional modes of governance.

From the technical point of view, plagiarism of content output by generative artificial intelligence is difficult for detection tools to identify accurately. Compared with traditional or basic plagiarism, which simply transcribes another's content, generative artificial intelligence tools possess a powerful paraphrasing function: they can alter the syntax and grammatical structure of an entire paragraph while still adopting the ideas contained in the source paraphrased, and they omit the citation of the reference source when presenting it to the user[14]. This amounts to artificial intelligence processing the text in place of the user, and that processing is merely a meaningless splitting and recombining of words and phrases without adding new ideas, yet it evades duplicate-checking detection very well, so that this new form of plagiarism is hard to discover. Although detection tools such as Turnitin and GPTZero, used to judge whether text has been generated by artificial intelligence, have already appeared, these tools are at present unreliable; after complex rewriting in particular there is no clear basis of comparison by which plagiarism may be judged, so that detection results cannot serve as evidence of a violation of academic ethics. The use of such detection tools also carries the potential problems of privacy and security, and they have therefore not been widely adopted. While detection tools remain undeveloped and the relevant norms have yet to be established, the immediate "cut and paste" function of modern digital technology makes the risk people can perceive far smaller than the benefit to be gained, which increasingly induces researchers to use generative artificial intelligence to plagiarise in order to complete academic tasks.

From the subjective point of view, plagiarising content output by generative artificial intelligence produces a lower sense of moral guilt. In this study's survey of whether plagiarism gives rise to guilt and unease, the number of respondents who would not feel uneasy about plagiarising artificial intelligence exceeded by 60 per cent the number who would not feel uneasy about plagiarising a published paper. A sense of moral guilt is an important emotional foundation for sustaining moral behaviour, and research on moral disengagement theory has shown that, absent suffering, individuals decouple their internal moral standards from immoral and even criminal conduct[15]. Through the paraphrasing done by artificial intelligence, the plagiarist does not come into direct contact with or trace the original author and the original work, which weakens and even circumvents the thought that an injustice has been done to the person plagiarised, lightens the insecurity and guilt felt about the plagiarising act itself, and weakens moral sensitivity and self-supervision.

2.2 Concealed Data Leakage

In academic activity, data entered by the user may be accessed and even used by the service provider or its partners, leading to the disclosure of confidential data and private information. The data leakage that generative artificial intelligence induces is rooted in its operating principle. That principle may be divided into two processes. The first is the formation, through machine learning and deep learning after being fed a large corpus of human language, of the capacity to "predict the next token using the correlation between tokens", with the model being continuously trained and optimised through a two-way interaction from input to output and from output back to input. The second is that, on this basis, generative artificial intelligence products will moreover continuously adjust themselves according to the linguistic habits users display in actual participation[16].

Data leakage may occur in the course of both the training and the use of generative artificial intelligence. Take the process of use: in order to ensure that the output answer is better adapted to the user, the content entered into a generative artificial intelligence product will ordinarily become training data; and because the algorithmic mechanism is not disclosed, it is hard to say whether the entered content will be presented in other forms to persons other than the user. A survey of the user agreements of generative artificial intelligence products shows that at present such products carry the risks of entered content being automatically used to train the model, of a lack of risk warnings, and of the excessive use of personal information.

2.3 Systematic Knowledge Bias

"Bias" manifests itself as partiality, preference or deviation in respect of a thing, a person, a group or an idea; it is not merely an odious belief held by isolated individuals of defective character, but a widespread phenomenon in society, and a necessary product of the human cognitive structure and of the forms of social organisation[17]. "Bias" in research activity denotes the possibility that a researcher's subjective beliefs, values, attitudes or preferences may affect the design or conduct of the research, or the interpretation of its results. When a researcher's personal bias unintentionally or intentionally affects the research process or the reporting of its results, the results may be distorted or inaccurate.

The problem of bias has always existed in research activity, including bias against the identity of the researcher — discrimination on grounds of sex or race, for example, which violates the norm of "universalism" — and bias concerning

Copyright: © 2026 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), permitting distribution and reproduction in any medium, provided the original work is cited.

knowledge outcomes, such as conflicts of interest, which violates the norm of "disinterestedness". These traditional forms of bias have all been brought within the consideration of academic ethical norms and have become part of research on the professional ethics of scientists.

The bias of a human researcher is sporadic and individual, whereas the bias of an algorithm is systematic. Algorithmic bias denotes the bias produced by an algorithm in the course of its operation, whose essence is the mapping of existing social and academic bias onto the artificial intelligence system; it appears in several respects, including the "built-in bias" of the algorithm's operating rules, bias in the input data, and bias in computation[18]. Human researchers hold different "biases" — that is, bias itself is diverse; generative artificial intelligence, however, carries a "selection of bias", and because that bias is formed in the course of data labelling and training, it affects all users, thereby amplifying machine bias, until in the end the bias is consolidated and made "correct" while the truth is obscured. This is the great divide between algorithmic bias and the bias of a human researcher.

The algorithmic bias of generative artificial intelligence will give rise to academic regression, producing academic mediocrity and homogeneity. First, it suppresses innovative views. Artificial intelligence algorithms have their own preferences and will weaken, distort and even eliminate the voices of certain groups whose representation is already insufficient; they will in particular neglect innovative views that have not yet become mainstream, while conservative views or already consolidated biases are continuously reinforced, producing the phenomena of "garbage in, garbage out"[19] and "bias in, bias out"[20]. It should be noted, however, that generative artificial intelligence will not only consolidate existing views by way of algorithmic bias; in some cases it may also, through the "emergence" characteristic of complex systems, autonomously create new combinations, inferences or ideas to a certain extent. Second, non-Western scholarship is further marginalised. At present the data used in model training are mostly English texts and Western images, and data from non-Western or non-English cultures are seriously insufficient in training and unrepresentative in content[21]. When biases already existing in the real world are output as truth and, in the course of academic research, passed on to new audiences for exploitation and dissemination, such bias is not only hard to eliminate but will be further consolidated and aggravated by the power of technology. Third, it induces illusions of understanding and lowers researchers' cognitive capacity. The algorithmic bias of generative artificial intelligence also involves partiality in the comprehension of information — the presentation of incomplete information. In using generative artificial intelligence in academic activity, for example, scientists may mistakenly believe that they have a deep understanding of the subject of their research, that artificial intelligence tools are objective, or that they are exploring all testable hypotheses, when in fact they are exploring only within the narrower range the artificial intelligence tool is able to test[22].

2.4 Self-Reinforcing Disinformation

Generative artificial intelligence readily "fabricates information", or creates realistic and intelligent-sounding text in response to a user's prompt; in academic research this appears as the generation of false experimental data, the generation of false references, and the direct generation of articles. "Disinformation" in academic activity has the following two principal sources.

The first is "fake in, fake out": training on disinformation fed in. Academic research is an organised activity that stands on the shoulders of giants, and new progress can be made only on the premise of the reliable results of predecessors. The rigour and reliability of knowledge generated by artificial intelligence still fall considerably short of traditional academic research. Because the data currently fed in for model training come mainly from the internet, the information is unavoidably false, and the failure to label the source of information makes this problem worse. Take ChatGPT: when answering a user's question it does not label the source of the information or of the answer, and the user finds it hard to verify the authenticity of the content generated. If a scholar applies content generated by generative artificial intelligence (such as ChatGPT) to academic research and writing, the failure to label the source of information will make it all the more difficult to verify the authenticity and reliability of that content, and will subject academic integrity to a further challenge[23]. Even requiring users to state the source of the material at the same time as supplying it is unlikely to solve the problem, since it is not rare for generative artificial intelligence to make erroneous or false citations.

The second is "creating something from nothing": the autonomous output of false results. Generative artificial intelligence may fabricate disinformation in the absence of input or reference information. An artificial intelligence language model can create a highly persuasive fraudulent article that resembles a genuine scientific paper in word usage, sentence structure and overall composition. Generative artificial intelligence can also generate false experimental datasets: traditional academic fraud usually requires the manual tampering of data or the forging of experimental results, which is not only time-consuming and laborious but also apt to leave traces and so raise the risk of discovery. Generative artificial intelligence, by contrast, can, through deep learning algorithms, rapidly analyse and imitate the characteristics of real research data, and this highly automated means of fraud makes academic misconduct easier to carry out and extremely low in cost. Because such false data are generated on the basis of training on real research datasets, they simulate the distribution

Copyright: © 2026 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), permitting distribution and reproduction in any medium, provided the original work is cited.

and characteristics of real data very well and are therefore hard to identify. This concealment not only increases the possibility of falsified data being misused but also makes academic misconduct harder to detect and correct. The root of these problems lies in the algorithmic norms of generative artificial intelligence. First, the rationality of algorithmic rules takes precedence over truth: generative artificial intelligence frequently makes factual errors, because these tools tend to give what appears to be the most reasonable answer on the basis of existing data, while lacking the capacity to judge whether information is right or wrong. Some obvious errors are readily found by an expert, but sometimes users put questions to artificial intelligence precisely because they do not understand the content concerned. Second, the algorithmic model has a built-in randomness: even under the same prompt, the content output by artificial intelligence will differ each time, and the accuracy of the content cannot be guaranteed.

Generative artificial intelligence separates the subjective and objective conditions of academic fraud, increasing the difficulty of detecting and establishing it. In the past, establishing academic fraud required that the actor have subjective intent or be seriously negligent; the United States' determination of academic misconduct, for example, refers to the legal concept of fraud, and subjective intent is a constitutive element of the offence of fraud[24]. Generative artificial intelligence, however, has no thought or consciousness and at present does not meet the conditions for being a moral subject, so that no subjective intent exists, and this will challenge the existing basis for establishing academic misconduct.

The false academic information produced by generative artificial intelligence has a "self-reinforcing" character: once disinformation has been generated and is improperly used as training data, it forms a self-reinforcing closed loop of misleading, so that the capacity to produce and disseminate disinformation is continuously strengthened and its misleading character increased, appearing as an "amplification effect" and "irreversibility". On one hand, once new academic literature is "contaminated" by disinformation, that literature will in future become part of the training data, so that the weight of disinformation in the model is continuously increased and it has a higher probability of being selected as the answer to a similar question, with the result that disinformation becomes a constructed "fact". On the other hand, the disinformation generated by artificial intelligence is irreversible. So long as the disinformation remains within the range of the training data, generative artificial intelligence may supply it to users; and even if the disinformation were entirely erased from the internet — something practically impossible — generative artificial intelligence could still output it by training on its own earlier answers containing the disinformation.

3 Responses to the Academic Ethical Problems of Generative Artificial Intelligence

Ethical conduct is part of academic professional activity, but the scope of application of existing academic norms and of the system of academic ethical governance is insufficient to meet the challenge generative artificial intelligence brings, and forward-looking and innovative measures need to be explored.

3.1 Innovating Citation Conventions

The substance of the plagiarism problem is "use without acknowledgement", and the key to solving it is reasonable citation and acknowledgement of sources; but as a new source of knowledge, how the knowledge produced by artificial intelligence is to be cited has become a new problem. An important cause of induced plagiarism is the absence of clearly stipulated rules[25], and ethical norms of good academic practice should be set by way of deontology.

Some academic journals and institutions already require researchers to acknowledge and indicate the use of generative artificial intelligence, and provide specifications for citation format. Imperial College London, for example, provides a clear example: acknowledging the use of generative artificial intelligence in academic work should include at least the name and version of the generative artificial intelligence system used, the publisher, the URL of the tool, and a brief description of the concrete manner of use[26]. The American Psychological Association, the Modern Language Association and the Institute of Electrical and Electronics Engineers have all given suggestions on citation format. New York University further suggests that students keep a copy of their interactions with generative artificial intelligence for later confirmation[27].

The new citation conventions, however, conflict to some extent with traditional norms. Northwestern University (United States), for example, has proposed an innovation in citation format, labelling content cited from ChatGPT in the text as "(OpenAI + year)"[28]. The innovative idea is in line with the "(author + year)" labelling of the APA format, but it amounts to indirect recognition of OpenAI's authorial status and has provoked controversy. Yale University and other institutions hold that current generative artificial intelligence tools rely more on prediction than on originality in producing content, and the European Union's *Guidelines on the Responsible Use of Generative Artificial Intelligence in Research* likewise state that artificial intelligence does not possess authorial status. If labelling of this kind were taken to discharge the responsibility of academic norms, the genuine original author might be quietly struck from the record in the process, and intellectual property problems would also arise.

Copyright: © 2026 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), permitting distribution and reproduction in any medium, provided the original work is cited.

The author further suggests developing academic generative artificial intelligence tools designed specifically for research activity: unlike a general-purpose large model, such a tool should take reliable academic resources as the dataset for training the model and embed citation conventions in its algorithmic rules.

3.2 Using Services with Caution

The problem of data security is to a large extent caused by the model-training strategy of generative artificial intelligence service providers; academic institutions have no authority to prohibit such practices directly, but the following measures are possible.

1. Researchers should raise their vigilance, and in particular avoid entering sensitive personal and confidential information. The University of Oxford advises against sharing sensitive personal data such as financial information and passwords, and one's own or others' intellectual property, with any artificial intelligence tool[29]; Stanford University advises avoiding the input of information that ought not to be made public when using generative artificial intelligence tools[30].

2. Service providers and service options should be chosen with caution. The University of Oxford advises reducing risk by examining a service agreement for the use and ownership of information, selecting an appropriate interface, and making the processing of data transparent[29]. Harvard University requires that generative artificial intelligence tools undergo a risk assessment by the Harvard Information Security and Data Privacy Office before use[31].

3. Data should be processed by grade. Graded processing draws on existing confidentiality management systems, classifying research materials into more detailed levels and making clear which data and information may be disseminated publicly and which may not. Harvard University, for example, provides that confidential data of level 2 and above should not be entered into publicly available generative artificial intelligence tools[31].

Research institutions and universities may moreover respond, through independent technological innovation, to technical problems that are hard to solve at the institutional level. They might develop, for instance, a large model for the use of members of the institution only, which would not only avoid the leakage of confidential data but also allow such data to be used to train and improve the model and help members to solve complex problems.

3.3 Mutual Human–Machine Understanding

If algorithms with "built-in bias" were widely introduced into academic activity, they would give rise to concern about the consolidation of views; and because the model absorbs data from the past, the consequences of bias appear mainly as regression, and analyses of academic activity can hardly reflect the most recent progress. To avoid this, the article proposes two responses.

1. Artificial intelligence should "understand" the value norms of human researchers and preserve the plurality of its output. In operation, artificial intelligence does not genuinely "understand" the value norms of human researchers at the cognitive level; rather, by learning or imitation it internalises the norms of human research as "algorithmic rules", and so observes those norms when analysing data and outputting results. This means that when artificial intelligence selects output according to its algorithm, it should take different academic views into account and have regard for the plurality of human knowledge and values.

2. Human researchers should understand the characteristics and limits of artificial intelligence and use generative artificial intelligence tools appropriately. In view of the fact that the current level of technology cannot yet create a machine comparable to human intelligence, researchers need some cognition of the characteristics and limits of generative artificial intelligence. In the survey on the proposition "I believe the information provided by generative artificial intelligence is value-neutral", 86 per cent of respondents were unaware that an algorithm may be biased. Harvard, Oxford, Cambridge, Imperial College and other institutions therefore advise that teachers acquire artificial intelligence literacy and understand its strengths and limits, so that students strengthen their reflection on artificial intelligence technology, and that the capacity for independent thought and creative thinking be emphasised as valuable qualities of a researcher.

3.4 Supervising Data throughout

The falsity of information is one of the most prominent academic ethical problems that generative artificial intelligence induces. The survey found that 85 per cent of respondents were unaware that the information provided by artificial intelligence might be false. Artificial intelligence gives users an illusion of reliability; to prevent the spread of disinformation in the academic field, supervisory responsibility should be discharged at every stage — data training, data generation and data use.

1. Generative artificial intelligence used in academic research should be built on a high-standard training dataset. Developing academic-purpose generative artificial intelligence is an inevitable trend: the reference literature of academic research today should be professional academic papers and monographs, not knowledge such as that of *Baidu Baike*, which has not undergone academic peer review; by the same token, general-purpose generative artificial intelligence may inspire

Copyright: © 2026 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), permitting distribution and reproduction in any medium, provided the original work is cited.

ideas, but should not be a reliable or directly citable source of information. Training generative artificial intelligence on professional literature and data will safeguard the authenticity and reliability of the information it generates, and reduce the challenge its disinformation poses to academic ethics.

2. Providers of generative artificial intelligence services should label the source of information for verification. The absence of source labelling means that information cannot be traced and verified, which further increases the difficulty of testing the reliability of information generated by generative artificial intelligence, and makes it harder still for users to verify the authenticity of the content generated. Generative artificial intelligence should therefore label the concrete source of information when giving an answer, so that users may themselves calibrate its reliability and the adverse effects of disinformation are reduced.

3. The review and subsequent supervision of academic papers should be strengthened, so that "contaminated" literature is as far as possible prevented from becoming new training data. In the retracted paper mentioned above, neither the peer reviewers nor the journal editors noticed the obvious trace "Regenerate response". It is imperative to strengthen the awareness of reviewers and intensify the scrutiny of academic papers, and at the same time the subsequent supervision and retraction of academic papers should not be underestimated. As noted above, so long as disinformation remains within the training data of generative artificial intelligence, the system may still supply that disinformation to users; it is therefore particularly important, once contaminated academic literature is discovered, to erase it as far as possible.

4. User responsibility should be implemented. Users are obliged, when citing or referring to information generated by artificial intelligence, to verify its accuracy, authenticity and reliability; even where a researcher marks content generated by artificial intelligence with a strict and standard citation format, the researcher remains responsible for the authenticity and reliability of the cited content. This differs from the responsibility borne in citing an academic paper or monograph, because the authors of academic papers and monographs are subjects who can be held accountable, whereas artificial intelligence cannot be held accountable; the user should therefore become the party primarily responsible for content generated by generative artificial intelligence.

Acknowledgements

Undergraduates of the Institute for Advanced Study in Humanities and Social Sciences, Beihang University — Lin Yixuan, Wen Yuru, Zhang Likai and Yuan Di — also contributed to the collection and collation of materials, which is acknowledged here.

References

[1] Sleator L. Almost half of Cambridge students admit they have used ChatGPT[EB/OL]. (2023-04-21)[2024-07-13]. https://www.thetimes.com/business-money/technology/article/cambridge-university-students-chatgpt-ai-degree-2023-rnsv7mw7z.

[2] Li Y, Xu J, Jia C Y, et al. The current state of and reflection on university students' use of generative artificial intelligence: a survey based on Zhejiang University[J]. *Open Education Research*, 2024(1). [in Chinese]

[3] Nordling L. How ChatGPT is transforming the postdoc experience[J]. *Nature*, 2023, 622.

[4] Harvard University. Generative Artificial Intelligence (AI) Guidelines[EB/OL]. [2024-07-13]. https://huit.harvard.edu/ai/guidelines.

[5] Rahman M M, Watanobe Y. ChatGPT for education and research: opportunities, threats, and strategies[J]. *Applied Sciences*, 2023, 13(9).

[6] Zhu Y X, Yang F. ChatGPT/generative artificial intelligence and educational innovation: opportunities, challenges and the future[J]. *Journal of East China Normal University (Education Sciences)*, 2023(7). [in Chinese]

[7] Berg C. The case for generative AI in scholarly practice[EB/OL]. (2023-04-03)[2024-07-13]. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4407587.

[8] Van Woudenberg R, Ranalli C, Bracker D. Authorship and ChatGPT: a conservative view[J]. *Philosophy & Technology*, 2024, 37(1).

[9] Else H. Abstracts written by ChatGPT fool scientists[EB/OL]. (2023-01-12)[2024-07-15]. https://www.nature.com/articles/d41586-023-00056-7.

[10] Kasneci E, Sessler K, Küchemann S, et al. ChatGPT for good? On opportunities and challenges of large language models for education[J]. *Learning and Individual Differences*, 2023, 103.

[11] Jarrah A M, Wardat Y, Fidalgo P. Using ChatGPT in academic writing is (not) a form of plagiarism: what does the literature say[J]. *Online Journal of Communication and Media Technologies*, 2023, 13(4).

[12] Wang S. ChatGPT and the governance of academic misconduct: challenges and responses[J]. *Science & Technology Progress and Policy*, 2023(23). [in Chinese]

[13] Office of Science and Technology Policy. Federal Research Misconduct Policy[EB/OL]. (2000-12-06)[2024-11-08]. https://ori.hhs.gov/federal-research-misconduct-policy.

[14] Navarro-Dolmestch R. Risks and challenges posed by artificial intelligence generative applications for academic integrity[J]. *Derecho PUCP*, 2023, 91.

[15] Zhang L, Amos C, Pentina I. Interplay of rationality and morality in using ChatGPT for academic misconduct[J]. *Behaviour & Information Technology*, 2024(2).

[16] Chen X P. Large models: artificial intelligence thought and its social experiment[J]. *Cultural Horizon*, 2023(3). [in Chinese]

[17] Allport G W. *The Nature of Prejudice*[M]. Xu J W, trans. Beijing: China Renmin University Press, 2021. [in Chinese]

[18] Yue P, Miao Y. Social governance: the problems of algorithmic bias and their regulation in the age of artificial intelligence[J]. *Journal of Shanghai University (Social Sciences Edition)*, 2021(6). [in Chinese]

Copyright: © 2026 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), permitting distribution and reproduction in any medium, provided the original work is cited.

[19] Citrome L. Artificial intelligence and the potential for garbage in, garbage out[J]. *Current Medical Research and Opinion*, 2024, 40(1).

[20] Wang J. From human bias to algorithmic bias: can bias be eliminated?[J]. *Exploration and Free Views*, 2021(3). [in Chinese]

[21] University of Michigan. Using Generative AI for Scientific Research[EB/OL]. (2024-03-28)[2024-07-13]. https://midas.umich.edu/generative-ai-user-guide/.

[22] Messeri L, Crockett M J. Artificial intelligence and illusions of understanding in scientific research[J]. *Nature*, 2024, 627(8002).

[23] Cotton D R E, Cotton P A, Shipway J R. Chatting and cheating: ensuring academic integrity in the era of ChatGPT[J]. *Innovations in Education and Teaching International*, 2024, 61(2).

[24] He H P, Qi K P, Wang C. The basis for and controversies over the determination of research misconduct[J]. *Journal of Beijing University of Aeronautics and Astronautics (Social Sciences Edition)*, 2022(1). [in Chinese]

[25] Ross G. Plagiarism in philosophy: prevention better than cure[J]. *Discourse: Learning and Teaching in Philosophical and Religious Studies*, 2004, 3(2).

[26] Imperial College London. Generative AI Guidance[EB/OL]. [2024-07-13]. https://www.imperial.ac.uk/admin-services/library/learning-support/generative-ai-guidance/.

[27] New York University. Student Learning with Generative AI[EB/OL]. [2024-07-13]. https://www.nyu.edu/faculty/teaching-and-learning-resources/Student-Learning-with-Generative-AI.html.

[28] Northwestern University. Academic Integrity & Artificial Intelligence[EB/OL]. (2024-04-04)[2024-07-13]. https://libguides.northwestern.edu/ai-tools-research/acad-integ.

[29] University of Oxford. Use Generative AI Services Such as ChatGPT Safely[EB/OL]. [2024-07-13]. https://infosec.web.ox.ac.uk/use-generative-ai-services-such-as-chatgpt-safely.

[30] Stanford University. Course Policies on Generative AI Use[EB/OL]. (2024-05-17)[2024-07-13]. https://tlhub.stanford.edu/docs/course-policies-on-generative-ai-use/.

[31] Harvard University. Initial Guidelines for the Use of Generative AI Tools at Harvard[EB/OL]. [2024-07-13]. https://huit.harvard.edu/ai/guidelines.

Copyright: © 2026 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), permitting distribution and reproduction in any medium, provided the original work is cited.