

The Analysis of Technological Ethical Issues in Generative Artificial Intelligence

Liu Weijia, Li Miaomiao

Qufu Normal University, Rizhao, Shandong, China

Keywords: Ethics of science and technology, Generative AI, bias and algorithmic discrimination

Abstract: At present, generative AI represented by ChatGPT has great potential and development prospects in education ecology, academia, media and public domain, etc. However, it also faces many ethical challenges, such as the weakening of human subjectivity, bias and algorithmic discrimination, privacy and data security issues, and impact on the value system. Actively exploring the good recipe for ethical governance of science and technology will help promote agile governance and make science and technology develop for the better.

1. Introduction

Currently, the discussion on generative AI focuses on education ecology, academia, media and public domain, etc. This emerging technology has not only triggered a wide range of discussions in various fields, but also reshaped our social landscape and cognitive approach to a certain extent. With its powerful algorithms and models, it brings us unprecedented opportunities for innovation, however, its potential risks and challenges should not be ignored.

2. Current status of the development of generative artificial intelligence

2.1. Generative artificial intelligence and the ecology of education

Education has always been wary of technology, but ChatGPT/Generative Artificial Intelligence, launched by US-based Open AI, is making huge waves in the education sector. The rapid momentum of generative AI in education has brought mixed emotions among educators, ranging from complete optimism to apprehension.

Generative AI presents many opportunities for teachers' work. Firstly, generative AI can promote the personalised development of students. Each student has different personal experiences, interests and hobbies, and there are large and small differences in learning styles, cognitive abilities and learning levels. Generative AI's massive information base can quickly provide students with a wealth of information, which can stimulate students' interests. Secondly, it can reduce the workload of teachers. Generative AI, which can provide teachers with teaching resources, can enable teachers to look for and screen teaching resources when preparing lessons, and can have more inspiration in teaching design and methods. The third is that generative AI also builds a communication platform

among teachers. Teachers can learn about the latest developments and research results in the field of education and improve their professionalism. Some scholars say that generative AI has great potential in the field of education, and generative AI will change the education ecology. For example, Zhou Hongyu and Li Yuyang pointed out that ChatGPT can fully intervene in the education ecology by redefining autonomous learning, transforming teacher teaching, school curriculum and talent cultivation standards at the level of teaching practice. Gao Qiqi and Yan Wenfeng, from the perspective of knowledge production, regard Chat GPT as a kind of efficiency tool and technology "prosthetic limb", and believe that it will lead to a new round of knowledge revolution, and ultimately change the education ecology.

The involvement of generative AI in the field of education raises many questions. Jeremy Weissman, a professor of philosophy at Nova Southeastern University, has stated that "ChatGPT will probably bring a plague to education." (Weissman, 2023) This statement reminds us that we must carefully consider the risks and challenges that Chat GPT may pose in the process of driving innovation in education. The first is the role of generative AI shockers teachers. Their anxiety in the face of AI is on the high side. Generative AI will change the teacher's monopoly of knowledge, and can be free from the teacher's intervention in the teaching and learning process, independent of the existence of the role of "teacher". This poses a challenge to the existing teaching process and teaching methods. Secondly, the use of generative AI is more advantageous than traditional teaching models in terms of knowledge transfer and skills training, but we have to be wary of technological dependence or even addiction, which would deviate from the original purpose of education. Generative AI can still not replace the teacher's teaching in the cultivation of students' creative and emotional qualities, moral sense, and the cultivation of values. In addition, the development of generative AI will lead to academic abuse, plagiarism and plagiarism by students, cheating on exams, and teachers' insecurity in preparing lessons and lesson plans. Some scholars call the crisis of educational alienation as the crisis of educational vulgarization, which is mainly manifested as the distortion of educational spirit, the superficialization of cognitive structure and the homogenization of cognitive thinking.

2.2. Generative artificial intelligence and academia

Generative Artificial Intelligence has had a significant impact on human knowledge production activities and the ecology of knowledge dissemination in academic research. The use of generative AI in the process of knowledge production has produced a tremendous thrust on academic research. Some scholars summarise it as a multi-directional help such as help in reading, help in knowing, help in writing, help in thinking, etc., i.e. the ability of generative AI to acquire, browse, summarise, condense and summarise massive academic resources as well as the ability to instantly capture the whole range of network materials can provide a powerful help in the preparation of academic research in the link of reading materials. Generative AI can generate the answers people need with a single click through a dialogue with a human being regardless of time and faculty constraints, which is both convenient and timely. Generative AI can provide more inspiration for human beings, so that the work can be more valuable, expanding the creative thinking and innovation ability of human beings, and improving the efficiency of people's writing. [1]With this help, researchers can be more efficient in conducting academic research. The eyes are wider and the results are better. In this human-computer collaborative knowledge production method, users interact with AI to make knowledge production more democratic and decentralised.

The application of generative AI in academia can also have some negative effects. The first is that the results generated by generative AI can have wrong answers or misleading answers due to the limited nature of the technology. On the other hand, there are legal and ethical issues regarding

the ownership and attribution of the content generated by generative AI, especially the question of whether Chat GPT can have authorship. At the same time, there are also issues of plagiarism, copyright protection and privacy protection. Scholar Zheng Quan pointed out that generative AI has triggered the change of knowledge production and dissemination paradigm in the fields of scientific research, education, finance, and medical care, and then analysed the risk challenges of the change of knowledge production and dissemination paradigm of generative AI in three aspects: epistemology, pragmatism, and value theory. [2] On this basis, it is more important to correctly understand and assess the roles and responsibilities of generative AI and human beings in knowledge production in order to promote knowledge justice in the human-machine collaboration model. In summary, the uncertainty of generative AI-generated content is step by step touching legal regulations such as intellectual property law and criminal law. People are also continuously exploring the moral subject status of generative AI and its corresponding responsibility.

2.3. Generative artificial intelligence and others

In addition to this, generative AI is also used in a variety of fields such as governmental affairs and military operations. Generative AI's powerful text analysis and generation functions can assist the government in dealing with some routine matters and promote the flattening of the government structure. Relying on massive databases, generative artificial intelligence can also provide a basis for government scientific decision-making. At the same time, generative artificial intelligence can also be loaded onto drones and military equipment to sense battlefield information in real time. For example, ChatGPT can efficiently connect different battlefield information sensing systems, create intelligence sharing terminals by means of sensing summary processing technology, and the AI based on the GPT model architecture can continuously filter a large amount of battlefield information data and mark its important information, improve the effectiveness of artificial information processing and analysis, and provide accurate information data and decision-making basis for the commanding department. However, this also brings a series of security risks, political governance and military operations are very complex and sensitive areas. Some scholars have pointed out that generative artificial intelligence technology embedded in digital government governance embodies a high value, can effectively control the scale of digital government, and can help digital government governance decision-making, improve the efficiency of digital government governance. The convenience of generative AI brings many risks to digital government governance, such as technical risks, security risks, administrative risks and other new problems .[3]

3. Ethical Value Challenges of Generative Artificial Intelligence

While generative AI is developing rapidly in various fields, it is also facing a number of ethical risk challenges in science and technology. Thinking further about these risks and challenges on a philosophical level, generative AI will result in the weakening of human subjectivity and impact on the human moral and cultural system, as well as discrimination due to algorithmic bias, privacy and data security issues, and so on.

3.1. Bias and algorithmic discrimination

Regarding "bias", the philosopher of technology Andrew Feenberg has a clear explanation. Feenberg's critical theory of technology proposes the concept of "techno-code", which refers to an unconscious incarnation of human society and culture, a coding system based on shared ways of thinking and behavioural patterns in different social and cultural contexts. Techno-code refers to an unconscious incarnation of human society and culture, a coding system based on shared ways of

thinking and behavioural patterns in different social and cultural contexts. Critical theory of technology implies that technological neutrality and bias are not opposites in a given situation, but merely different aspects of a single concrete object. Feinberg borrows from Max Weber's framework of rationality theory to delineate two kinds of prejudice, one substantive and the other formal. Fenberg's substantive bias focuses on the potential bias brought about by the social, political, and economic benefits and influences gained through technology. Formal bias, on the other hand, refers to the biases and limitations brought about by the form, function and characteristics of the technology itself. As a world-leading advanced technology, AI plays an important role in influencing the development of digital economy, social and cultural progress and even the international landscape of various countries. To a certain extent, generative AI has become a tool for hegemonic countries to monopolise technological advantages and impose technological oppression on other countries and gain unequal political and economic socio-economic, cultural, military and other fields. Scholar Jochen Hartmann, after conducting a study on Chat GPT ideology in political elections, also argues that it is important to pay attention to Chat GPT's left-liberal political tendencies.

Formal bias in generative AI mainly manifests itself in the form of algorithmic discrimination and technical characteristics, which is partly due to the limitations of the algorithmic "black box" under certain historical conditions, and partly due to the influence of engineers' values. Engineers can implant their own preferences into the algorithm design through the algorithm setting, and if the engineers have values such as racial discrimination, cultural bias, gender discrimination, religious discrimination, etc., then these concepts will subconsciously have a negative impact on the user interaction, forming algorithmic discrimination. To date, several studies have confirmed that natural language processing technologies may capture, propagate, and amplify social discrimination, including racial, gender. Research at the Institute of Cross-Information Studies at Tsinghua University has also shown that Chat GPT's model will determine professions such as teachers and doctors as male with a high probability. When Chat GPT evaluates whether someone can be a scientist, it only suggests white males. As a result, "Chat GPT text output appears neutral and objective, but is in fact the 'spawn' of an algorithm's complicity with human bias.[4]" Meanwhile, in the training and evaluation test of the algorithm, the historical pattern of injustice in the pre-training is retained in the training data, on the other hand, there are errors in the algorithmic simulation and practical operation, and some under-represented groups are marginalised and excluded during the data training process, deepening the bias. For example, Open AI's risk test of its generated content in April 2020 showed that some of the image generation results may reinforce stereotypes and biases, for example, when inputting words such as "wedding" and "restaurant", the programme produces more images of Western-style scenes; when inputting the words "wedding" and "restaurant", the programme produces more images of Western-style scenes; when inputting the words "wedding" and "restaurant", the programme produces more images of Western-style scenes. For example, when entering words such as "wedding" and "restaurant", the programme produces more images of Western-style scenes; and when entering keywords such as "CEO" and "lawyer", DALL-E-2 tends to produce images of men and white people. Generative AI technologies are inherently susceptible to value preferences. In the interaction of generative AI is a collision between two cultures, or worse, ideologies and values. It is worth paying attention to the fact that in this kind of interaction between "positive feedback" and "negative feedback", it is easy for people to fall into an "information cocoon" and be labelled by the machine in a more obvious way.

3.2. Privacy and data security

The leakage of user privacy issues in generative AI mainly exists in the pre-training stage, the

human-computer interaction dialogue stage, and the privacy abuse problem mainly exists in the transactions with third parties. In the pre-training stage, data cultivation requires certain information support, and this information is not explicitly authorised by the user in the interaction with the user. This situation is more obvious in the dialogue stage of human-computer interaction, where the user's private information is continuously acquired through continuous learning to receive positive and negative feedback. The misuse of user's private data is mainly dependent on the development of deep simulation techniques. Deep simulation technology is a technology that uses generative algorithms such as deep learning and simulated reality to produce fake text, images, audio, video and virtual scenes. Common forms are chatbots, synthetic painting, scene sound editing, simulated sound, AI face swap and so on. Such as Alipay face swiping and community entry door swiping, the user's facial recognition information will be commercially collected and misused. There are also some entertainment video software inadvertently collects user's facial and body information, which will endanger user's personal and property safety once obtained by wrongdoers. Unlawful elements can use generative AI to create false information, impersonate organisations or individuals to commit fraud, obtain cybercrime tools, and conduct online "phishing". At the same time, the deep simulation technology simulation can simulate the human language style, voice tone and facial features, the victim and his family members to carry out criminal acts, to the depth of forgery technology as the representative of generative artificial intelligence is the biggest risk is to lower the threshold of criminals to obtain criminal tools.

3.3. Impact on ethics and value systems

Generative AI is impacting people's cultural values. On the one hand, the cross-cultural communication and integration of generative AI may lead to the marginalisation of local cultures in international communication or their domination by Western mainstream culture, which may lead to cultural conflicts and social controversies. On the other hand, the process of generative digital reconstruction and re-creation of traditional art forms and cultural heritage may involve the risk of simplification and distortion and over-modernisation triggering discussions on cultural authenticity, respect for the original context and inheritance of cultural connotations. The information generated by generative AI in the deep learning of massive data is inevitably influenced by cultural concepts in a certain space and time. At the same time it conveys the cultural values he is trained to learn in its interaction with the user. The subtle influence of generative AI-generated content on human thought patterns must and should be given special attention and careful response.

The development of Generative Artificial Intelligence (GAI) is in a booming period, in order to cope with the problems and potential ethical risks at this stage, it is necessary to strengthen the construction of policies and regulations, interdisciplinary cooperation and communication, as well as the enhancement of the public's digital literacy, and, more importantly, the application of the principle of preemptive technological ethics, to make prospective, predictive, and controlling decisions, so as to minimise ethical risks of Generative Artificial Intelligence, and better enable the society at large. In view of the above potential risks and problems of generative AI, the application and development of generative AI will generate certain moral and ethical risks, and how to make generative AI become a qualified "moral intelligence" has become a key issue. Some scholars have proposed that the introduction of value-sensitive design at the early stage of generative AI research and development can solve the endogenous problem of AIGC at the source. [5] Some scholars have also pointed out that under the scrutiny of value-sensitive design research, AI moral bodies are feasible.[6]Value-sensitive design applied to generative AI will reduce the risk of potentially harmful and unjust content, enhance user trust, prevent potential legal liabilities of generative AI due to violation of laws and regulations, etc., ensure that the development of the technology is more

humane and socially friendly, and also help to build a safer, more reliable, fairer and more socially responsible AI ecosystem.

References

- [1] Huang Shijin. "Help" and "Replacement": The Dual Effects of Generative AI on Academic Research[J/OL]. *Journal of Shanghai Normal University (Philosophy and Social Science Edition)*, 2024,(02):65-74[2024-04-16].
- [2] Zheng Quan. Knowledge production and dissemination paradigm change and response of generative artificial intelligence[J]. *Natural dialectics research*, 2024, 40(03):74-82.
- [3] Xiaheng Zhang. ChatGPT-like artificial intelligence technology embedded in digital government governance: value, risk and its prevention and control[J]. *E-Government*, 2023,(04):45-56.
- [4] Zou Kailiang, Liu Zubing. The Ethical Risks of ChatGPT and China's Institutional Arrangements for Response[J]. *Journal of Hainan University (Humanities and Social Sciences Edition)*, 2023, 41(04):74-84.
- [5] HU Bing, ZHONG Siyue. Implications of the Dutch Prejudicial Technological Ethics Thought for AIGC Governance [J]. *Young Journalist*, 2024(02):48-54.
- [6] Zhang Haopeng, Xia Baohua. How Artificial Moral Intelligence is Feasible - Based on an Examination of Value Sensitive Design[J]. *Natural dialectics research*, 2021,37(04):37-42.