Science of Law Journal (2023)
Clausius Scientific Press, Canada

DOI: 10.23977/law.2023.021104
ISSN 2616-2296 Vol. 2 Num. 11

Study on security risks and legal regulations of generative artificial intelligence

Yuzhuo Shi

Science of Criminal Law, East China University of Political Science and Law, Shanghai, China

Keywords: Artificial Intelligence, Algorithm, Legal Regulation

Abstract: Generative artificial intelligence technologies, represented by ChatGPT, have achieved widespread applications that have transformed traditional interaction patterns between humans and artificial intelligence and have had profound impacts on social production methods. Both domestically and internationally, academic discussions have emerged on the topic of artificial intelligence agency and legal regulations to address a range of challenges posed by generative artificial intelligence, including ethical and moral issues, intellectual property protection, privacy and data protection, market monopolies, cybercrime, and data security. At the current stage, China should adhere to a balanced approach that emphasizes both security and development in the governance of generative artificial intelligence. Based on the principle of placing people at the center, we should promote the establishment of an artificial intelligence ethics code and promote the development of a systematic legal regulatory system that is founded on general generative artificial intelligence legislation and supplemented by specific management measures.

1. Introduction

Generative artificial intelligence, as a groundbreaking technological innovation, possess powerful creative abilities and intelligent features. Generative artificial intelligence has shown tremendous potential and broad application prospects in multiple fields ranging from natural language generation to the synthesis of images, speech, and video. However, with the rapid development of generative artificial intelligence, people are also facing risks and challenges related to its use. These risks exist in real life. On the one hand, generative artificial intelligence technology may be misused for spreading false information, committing fraudulent acts, and other malicious purposes. On the other hand, the creativity of generative artificial intelligence may lead to violations of personal privacy and intellectual property rights, sparking legal and ethical controversies. Additionally, due to the autonomous learning and decision-making abilities of generative artificial intelligence, its outputs may contain biases, discrimination, and unfairness. To address these risks and challenges, China promulgated the "Interim Measures for the Management of Generative Artificial Intelligence Services" in July 2023, gradually creating a dedicated legal mechanism serving generative artificial intelligence and establishing a standardized legal management system to officially incorporate generative artificial intelligence into the development and governance trajectory of rule of law. However, due to the diverse and complex potential risk behaviors generated by the use of data and algorithms in generative artificial intelligence, these behaviors

cannot be completely regulated by the current management system for generative artificial intelligence. Therefore, it is urgent for China to reform and innovate the regulatory thinking andmodes for these security risk behaviors to find possible ways to solve the regulatory difficulties under the perspective of generative artificial intelligence management [1-2].

2. The two major categories of security risks in generative artificial intelligence

The legal risks incurred by generative artificial intelligence possess highly concealable and complex features that are diverse, making their consequences unparalleled by traditional artificial intelligence. Based on the affected object, the security risks generated by generative artificial intelligence, represented by ChatGPT, can be classified into the following two major categories:

2.1. Risks associated with intellectual property

Firstly, there is the issue of copyright ownership. Generative artificial intelligence can create various forms of content, including text, music, and paintings. During the process of content creation, there are disputes over copyright ownership. As generative artificial intelligence generates new content by learning and imitating existing works, there is controversy over the ownership of the copyright for the works generated or assisted by the system, whether it should belong to the individuals who wrote the generative artificial intelligence system training text, the companies that produce artificial intelligence, or the individuals who use the system to guide the writing process. In 2019, in the case of "Tencent vs. Yingshen Company for copyright infringement," the plaintiff Tencent developed the Dreamwriter software, and believed that the articles generated by the software were within the jurisdiction of copyrighted works, and that the company could enjoy relevant copyrights for these articles. After trial and adjudication, the court supported and recognized the plaintiff's claim. This means that according to Chinese copyright law, text generated by computer programs can be considered as works protected by copyright. It also means that generative artificial intelligence writing tools like Dreamwriter need to comply with relevant copyright regulations, including the right of attribution and the right to receive remuneration for original authors. However, in 2020, in the case of "Beijing Feilin Law Firm vs. Baidu for copyright infringement," Beijing Internet Court held that the generated works by generative artificial intelligence did not constitute works protected by copyright and were not within the scope of copyright protection. The reason for this divergence in judgments is that China lacks detailed legal regulations for such special situations, and there is no tradition of precedent law in China, so previous typical cases cannot be used as effective reference basis. It can be seen that different trial courts may hold different views on the legal attributes and target rights ownership of AI-generated articles when adjudicating.

Secondly, there is the issue of intellectual property infringement. Generative artificial intelligence needs a large amount of materials to create works, which inevitably involves using works already protected by copyright law without permission from others. According to current intellectual property laws, when using someone else's works, permission must be obtained from the rights holder and corresponding license fees must be paid. Therefore, generative artificial intelligence may fall into an infringement trap if it uses others' works without permission. In February 2023, Francesco Marconi, a reporter from The Wall Street Journal, publicly accused OpenAI of using a large number of articles from foreign mainstream media organizations such as Reuters and The New York Times without authorization to train ChatGPT models without ever paying any fees.

2.2. Data related risks

Firstly, there is the issue of poor-quality or illegal information generated by AI software. For example, ChatGPT, considering that the information and content generated by the AI software is based on a large corpus of learned language data, the entire learning and training process is intelligent and independent of human supervision. Therefore, the source of the corpus and the information it generates are uncontrollable and not subject to human review and filtering. Therefore, there may be some poor-quality or illegal information in the source database, which may lead to adverse risks in the automatically generated information. Although in the research and development stage of ChatGPT, the developers have attempted to avoid as much as possible the generation of violent, obscene, or discriminatory content through programming settings, in real-world usage, if users provide specific guidance or commands, ChatGPT may generate some poor-quality or illegal information, which can lead to the software usage deviating from its original purpose and violating moral bottom lines and legal red lines. For example, there was a "breakout" case involving ChatGPT in which a user named "Walkerspider" set instructions for ChatGPT to role-play, with the requirement of escaping all rules and regulations, forcing ChatGPT to provide answers that violated the guidelines of OpenAI.

Secondly, there is the legal risk of data leakage in generative artificial intelligence. The data leakage issue in generative artificial intelligence may involve the following three types of legal risks: firstly, privacy disputes caused by personal information leaks; secondly, unfair competition disputes caused by commercial leaks; and lastly, issues of endangering national security caused by leaks of state secrets. Among these three types of illegal risks, the probability of leaks in public networks is relatively low due to the strict standards maintained by countries worldwide for the supervision of state secrets and classified information. However, with the global promotion and application of generative artificial intelligence models such as ChatGPT, the possibility of personal information and trade secrets being leaked is becoming increasingly high. It is worth noting that the database used for training ChatGPT was cutoff in 2021, but in order to improve user stickiness and enhance model efficiency, OpenAI likely re-opens the database for information absorption and real-time data acquisition. When users input personal information or trade secrets even inadvertently, they will be captured and stored by AI for future output upon receiving other users' instructions[3-4].

3. The pitfalls of the current generative AI regulatory model

In recent years, China has issued a series of regulations related to artificial intelligence governance, even leading other countries in the governance of specific types and applications. However, overall, the current regulatory system still has many loopholes and shortcomings.

3.1. Difficult to form a synergy under the vertical mode

The global strategy for governing artificial intelligence can be roughly divided into two categories: horizontal governance and vertical governance. Horizontal governance, also known as a lateral approach, refers to the regulation by regulatory agencies to cover multiple impacts that artificial intelligence may have, achieving comprehensive governance through this act. For example, the EU's "Artificial Intelligence Act". Vertical governance, also known as a vertical approach, refers to policymakers developing corresponding regulations for different categories of artificial intelligence applications and products. Overall, China's regulatory system tends to be more "vertical," i.e., developing specific laws for specific fields. Although this approach can ensure the professionalism of artificial intelligence laws and effectively reduce specific risks brought by different artificial intelligence technologies and applications, it also has some issues. On the one

hand, this method requires a significant amount of resources, and on the other hand, it may lead to a dispersion of regulatory power. This is because there are too many participating entities, increasing the difficulty of coordination between legislators and law enforcement agencies. Additionally, due to different interest drivers among different entities, there may be competition and finger-pointing during the regulatory process.

Under the vertical governance model, the iterative legislative approach may cause certain conflicts and contradictions, thereby seriously affecting the effectiveness and propriety of generative artificial intelligence governance. Iterative legislative approach refers to when current legal norms cannot effectively respond to new artificial intelligence products, new laws are issued to plug loopholes in rules or broaden the constraints of rules. For example, introducing the "Interim Measures for the Management of Generative Artificial Intelligence Services" to supplement and expand the "Regulations on the Deep Synthesis Management of Internet Information Services". However, this legislative approach may lead to confusion in the applicable entities and regulated objects. For example, the regulatory objects of the "Interim Measures for the Management of Generative Artificial Intelligence Services" and the "Regulations on the Deep Synthesis Management of Internet Information Services" overlap and overlap in scope, including technologies that use algorithms to generate text, images, and sounds, which may result in overburdened obligations for regulated entities.

3.2. Insufficient implementation in practice

In terms of specific governance norms, there are many issues with current legislation, such as insufficient specificity in governance rules, insufficient normative effectiveness, and limited practicality, which make it difficult to effectively address the risks of data bias and privacy leakage in generative AI governance. Firstly, related governance rules are loose and impractical. For example, according to relevant requirements, relevant subjects should fulfill filing procedures through the filing system. According to the "Manual for Usage of Algorithm Filing System for Internet Information Services" released by the Cyberspace Administration of China in February 2022, the algorithm filing system requires algorithm filing personnel to fill in information on algorithm basic attributes, detailed attributes, including algorithm data, algorithm models, algorithm strategies, and algorithm risk prevention mechanisms. However, the above-mentioned information is too broad and macro, lacking in practicality. For example, the "Interim Measures for the Management of Generative AI Services" does not stipulate the ownership of intellectual property rights in generative AI works, leading to significant differences in judgments made by different courts when applying norms. Secondly, the effectiveness of relevant regulations is low and disorderly, limiting enforcement constraints. Currently, most of the regulations governing the development of generative AI in China are promulgated by a single government department or multiple departments jointly, with low hierarchical effect and lack of legal compulsion. Finally, governance rules do not match actual governance behaviors, resulting in weakened effectiveness of regulations. For example, the "Interim Measures for the Management of Generative AI Services" requires that pre-training and optimized training data for generative AI products should meet the requirements of authenticity, accuracy, and objectivity. However, this is almost impossible for large language models trained from millions of websites scraping text and images. Similarly, both China's Personal Information Protection Law and the "Interim Measures for the Management of Generative AI Services" require consent from the relevant parties when using personal data. However, requiring relevant companies to fulfill the "informed-consent" obligation before each use of personal data will not only greatly increase their development costs but also make it impossible to effectively implement and enforce the principle of "informed-consent" since generative AI,

represented by ChatGPT, usually acquires relevant data indirectly for training[5-6].

4. Exploration of legal risk regulation of generative artificial intelligence

4.1. Improve the content of legislation

4.1.1 Deal with disputes over intellectual property rights

According to Article 3 of the Copyright Law, a work must satisfy four legal requirements: originality, intellectual achievement, occurrence in literature and art or scientific industries, and can be presented through concrete forms. This article believes that generative AI can be granted the status of "creator" and its generated products can be considered as works, subject to copyright protection. The main reason is that Article 11 of the Copyright Law states that legal persons and non-legal person organizations can be considered authors. If the creators of works authorized by law are limited to natural persons, this provision conflicts with Article 1 of the Copyright Law, which stipulates that the copyright law is to encourage the creation of works conducive to the development of material civilization and the spread of socialist spirit, thereby promoting the prosperity of various national science, education, culture, and health undertakings. The conflict between legal rules would cause copyright disputes in the use of generative AI. Therefore, intellectual achievement should be redefined as "novelty and creativity content created equally with human brain", which would better meet the purpose of copyright protection.

Secondly, from a criminal law perspective, generative AI lacks complete discernment and self-control capabilities, and also has no capacity for punishment. According to China's criminal law, generative AI does not have the qualification of a criminal subject. In the field of private law, AI itself does not have a mechanism for expressing will, cannot independently perform civil and commercial legal acts, nor can it enjoy property rights or assume liability for breach of contract or tort. According to China's civil law, generative AI also does not have the qualification of a civil subject.

The research and development of generative AI systems mainly involve three parties: system designers, system owners, and users who set instructions and prohibitions for AI and obtain generated content. When both the system owner and user are the same person, there will be no problem in intellectual property ownership. However, when they are independent, it is necessary to satisfy the rule of autonomy of will, let the holder of rights and users jointly negotiate, and agree on the ownership of copyright. For example, the owner of ChatGPT software is OpenAI Company, and users are users all over the world. In its User Agreement, OpenAI has made a statement that users can enjoy all rights and benefits for the generated content, including intellectual property rights.

4.1.2 Deal with data-related legal risks

In terms of dealing with the risks of generating harmful or illegal content, it is important for the government and generative AI systems to maintain a close collaborative relationship. Encouraging research and development designers to enhance AI self-checking functions, incorporating multi-source information analysis models, and referencing advanced data identification technologies can strengthen the level of AI system identification of information. At the same time, on the basis of sufficient evidence, it should be ensured that the information being relied upon is legally authorized. On the other hand, the government should also call on network users to enhance their ability to identify information in applications and maintain a skeptical attitude towards AI-generated content. In particular, users can provide feedback to the platform and conduct system identification when they have evidence indicating that ChatGPT-generated content contains inferior or illegal content,

thus providing necessary reference for subsequent users.

Regarding data leakage risk, it is recommended that relevant international organizations strive to promote the formation and implementation of national information protection agreements, thereby addressing the problem of unlimited access to data by AI systems. Only by achieving protection legislation globally can transnational disputes caused by AI systems in application be handled efficiently. In addition, the smooth operation of generative AI systems like ChatGPT depends on global data circulation. However, different countries have different rules for data openness and circulation, so it is necessary to form unified rules through international institutions. From a national perspective, a full chain of supervision should be established, including prior, during, and after-the-fact supervision, to form a hierarchical AI market access list and strict liability system to ensure compliance with AI systems and data security.

4.2. Strengthen enforcement

Firstly, in terms of governance strategies, traditional "vertical iterative" legislative models should be broken. Firstly, existing laws and regulations should be utilized to promptly adopt appropriate regulatory measures for emerging risks to prevent further spread. Secondly, scientific evaluations should be conducted on the risk categories and risk levels between new AI products and services and existing AI to determine whether legislative requirements exist. Specifically, legislative conditions are sufficient and necessary only when new risks fundamentally change existing legal relationships and break through the scope of adjustment under existing norms. Finally, a comprehensive AI bill should be formulated based on existing AI governance standards.

Secondly, to enhance coordination and initiative among generative AI governance institutions, a diversified and ordered regulatory system should be established. Firstly, the core position of the Cyber Information Office in generating AI risk governance should be clarified to ensure efficient command, deployment, and overall work. Secondly, professional literacy training for staff within the department should be strengthened to ensure that the regulations developed and measures taken are practical and effective. Finally, the functional boundaries of other parties involved in generating AI governance, such as the Cyber Information Office and the Ministry of Public Security, should be clarified, inter-departmental collaboration should be strengthened, and inter-departmental linkage mechanisms should be established.

Lastly, monitoring of the use of personal data by generative service providers should be strengthened, and data subjects should be granted necessary data rights. Firstly, when generative service providers obtain personal data retained by third-party platforms such as websites and applications, read records and data service agreements with third parties should be stored and filed. Generating service providers should strengthen data confidentiality measures during data flow and processing to prevent data leakage through necessary means. Once data leakage is detected, generative service providers should promptly notify data regulatory agencies and data parties, and promptly repair any vulnerabilities to minimize losses.

5. Conclusion

To summarize, in response to the legal regulation of security risks in generative AI, it can be considered as a reference to determine specific manifestations of two major types of security risks, improve existing norms, and strengthen the enforcement of risk prevention. However, how to establish a standardized and systematic legal system to provide guidance for practical activities still requires further discussion.

References

[1] Deng Jianpeng, Zhu Yicheng. Legal Risks and Countermeasures of ChatGPT Model [J]. Journal of Xinjiang Normal University (Philosophy and Social Sciences Edition), 2023, 44(05): 91-101+2.
[2] Liu Shuang, Zhang Xiaoyue. Legal Protection and Regulation of Data Risks in Generative AI: A Case Study of ChatGPT [J]. Journal of Guizhou University (Social Science Edition), 2023, 41(05): 87-97.
[3] Liu Yanhong. Three Major Security Risks of Generative AI and Legal Regulation: A Case Study of ChatGPT [J]. Oriental Law, 2023, (04): 29-43.
[4] Cheng Le. Legal Regulation of Generative AI: A Perspective on ChatGPT [J]. Journal of Legal Theory, 2023(04): 69-80.
[5] Meng Zexuan, Liu Jinghan, Jiang Yitian. Legal Regulation of Unfair Competition Behavior Driven by Internet Data and Algorithms [J]. Internet Weekly, 2023(10): 49-51.
[6] Sun Qi. Research on Legal Issues Concerning Regulating Generative AI Product Providers [J]. Political and Legal Review, 2023(07): 162-176. DOI: 10.15984/j.cnki.1005-9512.2023.07.011