DOI: 10.18254/S278229070031786-0 Original paper

Legal Regulation of Generative Artificial Intelligence in China

©2024 Tianfang Yang[a]

[a] The Shenzhen MSU-BIT University, China
danilyep@mail.ru

Abstract. The rapid development of generative artificial intelligence (AI), exemplified by technologies like ChatGPT, has prompted significant regulatory responses in China. This paper explores the legal framework established by China's Interim Measures for the Management of Generative Artificial Intelligence Services, highlighting its regulatory mechanisms and compliance obligations for AI service providers. The measures aim to address various risks associated with generative AI, such as data security, content management, and user protection, by implementing a dual registration system for algorithms and AI models. The Basic Safety Requirements for Generative Artificial Intelligence Services, published in 2024, provided detailed guidelines for ensuring the safety and legality of AI applications. This includes stringent assessments of data sources, content quality, and algorithm safety. By drawing comparisons with existing regulations like the Algorithmic Recommendations Regulation and the Deep Synthesis Regulation, this paper demonstrates China's consistent approach to AI governance, emphasizing the principles of promoting technological development while safeguarding public and individual interests. The findings suggest that China's regulatory framework for generative AI is designed to balance innovation with risk management, setting a precedent for comprehensive AI regulation.

Keywords: generative artificial intelligence, AI iegulation, China, deep synthesis technology

For citation: Yang T. 2024. Legal Regulation of Generative Artificial Intelligence in China. *Law & Digital Technologies 4(1): 25—30.*

CONTEXT OF LEGAL REGULATION OF GENERATIVE ARTIFICIAL INTELLIGENCE

The emergence of ChatGPT has caused a wide resonance and has become a popular subject of discussion about the place of artificial intelligence (AI) in human life. AI technology marks the fourth technological revolution and can profoundly impact society and the economy. Human life and many production processes are significantly facilitated by the expanded application of AI Generated Content (hereinafter referred to as AIGC). For example, in the medical field, AI can help doctors diagnose and treat diseases; in education, AI can provide students with personalized learning experiences; and in entertainment, AI can create unique artistic works, etc. Thus, the application of AI not only increases efficiency in individual areas of human activity, but also contributes to the development of the economy and society (Yao and Li 2023).

However, with the rapid development of AI technologies, especially the application of AIGC, we inevitably face a series of previously unknown challenges and crises. Issues such as personal data leakage, commercial secret violations, and job cuts attract the attention of the global scientific community. Moreover, to some extent, the development of AI may impact human subjectivity and values, causing justified concerns among people. For example, many prominent figures in technology, led by Elon Musk, have jointly called for a halt to the training of AI systems more powerful than ChatGPT-4. They argued that the influence of technology on human society is enormous and deep, considering the structure of "technology development – social evolution" (Pu and Yearning 2023).

Nevertheless, the development of AI technologies, especially AIGC, is an irreversible and growing trend. The autonomy and neutrality of AI technology do not allow for an artificial blockade or outright rejection, such as in the case of cloning technology. Thus, it is necessary to find ways to identify and minimize the risks associated with developing AI systems and to accelerate the development of the AI industry.

As early as 2017, the State Council of the People's Republic of China published the Next Generation Artificial Intelligence Development Plan[18], emphasizing that AI is a strategic technology that will lead in the future and that the development plan for AI should be systematically outlined at the national strategic

18 《新一代人工智能发展规划》 https://www.gov.cn/zhengce/content/2017-07/20/content_5211996.htm

level. By 2025, China is expected to "initially establish legal norms, ethical standards, and a policy system in the field of AI, forming the capability to assess and control AI safety"[19]. By 2030, it is planned that "China's AI theory, technology, and application will generally reach the world's leading level, making China a major global center for AI innovation and producing notable results in developing a 'smart' economy, which will become an important foundation for transforming China into a leading innovative and economic power"[20].

Against this backdrop, from 2022, China has started to sequentially adopt regulatory documents concerning algorithmic recommendations, deep synthesis technology, and generative AI services. This reflects the creation of a regulatory mechanism for AI technologies and services in specific areas.

Regarding generative AI services, it is noteworthy that on July 10, 2023, the Cyberspace Administration of China (CAC, 国家互联网信息办公室) published the Interim Measures for the Management of Generative Artificial Intelligence Services[21] (hereinafter referred to as the Interim Measures). This act clearly defines the concept of generative AI and compliance obligations for providers of related products and services. According to the specific regulatory framework for generative AI services provided in the act, China has implemented a practical "dual registration" mechanism, consisting of an algorithm registration system and an AI (large language model) registration system.

It is worth noting that the algorithm registration system was first established in the Regulation on the Management of Algorithmic Recommendations in Internet Information Services[22], effective from March 1, 2022 (hereinafter referred to as the Algorithmic Recommendations Regulation), according to which businesses can apply for algorithm registration through the CAC's algorithm registration system. The procedures and rules for registration are standardized based on practice. The registration of large language models requires even more effective interaction between regulatory authorities and service providers to accumulate regulatory experience and develop clear and specific rules that will encourage businesses to fulfill their obligations to register algorithms, especially concerning their safety assessment in the field of large language models.

THE INTERIM MEASURE FOR THE MANAGEMENT OF GENERATIVE ARTIFICIAL INELLEGENCE SERVICES IN CHINA

The Interim Measures can be regarded as China's first "response sheet" in the field of generative AI regulation. This act is the first and most prompt attempt to respond to emerging scientific, technological, and ethical challenges of using modern generative AI. This article will examine the main provisions of the Interim Measures and, through a comparative analysis with the existing Chinese algorithm management system, will analyze the logic of algorithm regulation.

Regarding the structure of the Interim Measures, it should be noted that they primarily explain the basics of superior legislation, the scope of its application, and the state's main position concerning generative AI services. In addition, the Interim Measures establish the legal status of the subjects of legal relations, the grounds for accountability, and the corresponding administrative penalties.

Article 2 of the Interim Measures defines generative AI as "technology that generates text, images, sounds, videos, codes, and other content based on algorithms, models, and rules". Based on this definition, the Interim Measures apply to the development and use of generative AI products and the provision of services to users within the PRC territory, and they have general extraterritorial force on the global Internet.

According to the provisions of Article 22 of the Interim Measures, the responsible entity is "the providers of generative AI services such as chat or generation of text, images, and sounds" According to the literal interpretation, this does not include organizations and individuals engaged only in the research and development of generative AI technology.

The Interim Measures establish obligations for the relevant parties concerning algorithms, content, users, regulatory mechanisms, etc., following superior laws such as the PRC Cybersecurity Law[23], the PRC

19 Ibid.
20 Ibid.
21 《生成式人工智能服务管理暂行办法》 https://www.gov.cn/zhengce/zhengceku/202307/content_6891752.htm
22 《互联网信息服务算法推荐管理规定》 http://www.cac.gov.cn/202 2-01/04/c_1642894606364259.htm\
23 《网络安全法》 https://www.cac.gov.cn/2016-11/07/c_1119867116.htm

Data Security Law[24], the PRC Personal Information Protection Law[25], the PRC Science and Technology Progress Law[26], and the Regulation on the Security Assessment of Information Internet Services with Public Opinion Attributes or Social Mobilization Capabilities[27] (hereinafter referred to as the Security Assessment Regulation), the Algorithmic Recommendations Regulation, and the Regulation on the Management of Deep Synthesis in Internet Information Services[28] (hereinafter referred to as the Deep Synthesis Regulation). These regulatory acts together form a closely connected structure for managing algorithms in various sectors, especially in providing generative AI services.

The key content of this act – the obligations – can be conditionally divided into four types: (a) obligations related to oversight mechanisms; (b) obligations related to algorithm training; (c) obligations related to content management; and (d) obligations of service providers concerning users.

The Interim Measures develop the oversight mechanisms outlined in the Algorithmic Recommendations Regulation, such as the obligation to assess the safety and register algorithms (Article 19) and the user complaint mechanism to regulatory authorities (Article 12).

In the field of algorithm training, the Interim Measures establish a comprehensive control mechanism at all stages – before, during, and after, including ensuring the legality of data sources for algorithm training (Article 7); conducting quality assessments of data annotations and spot-checking the accuracy of annotated content (Article 8); and taking corrective measures in case of inappropriate content (Article 15).

Content management obligations include the duty of content creators to produce positive, healthy, and uplifting[29] content (Article 5) and the obligation to annotate generative content (Article 12).

Considering the potential impact of generative AI on users, the Interim Measures require service providers to take several measures to protect user rights, promote a correct attitude towards generative AI, and guide users on its reasonable use. Service provider obligations concerning users include protecting user personal data (Article 9), preventing user addiction (Article 10), safeguarding user-entered data (Article 11), ensuring service stability (Article 13), a mechanism for responding to content infringing on another's rights (Article 14), and a user complaint mechanism (Article 15).

Overall, as mentioned above, the Interim Measures continue the approach similar to previously adopted legal acts in China regarding generative AI service providers, demonstrating consistency and continuity in China's approach to legal regulation in algorithm management.

Firstly, in terms of governance, the Interim Measures inherit the spirit of "developing science and technology for good"[30], one of the main principles of algorithm management in China. The Algorithmic Recommendations Regulation first included the goal of "for better and good" at the ministerial and departmental rule level, requiring algorithmic recommendation service providers to adhere to basic social values and promote the use of algorithms "for better and good" while taking measures to prevent and combat the spread of harmful information (Article 6). Subsequently, numerous incidents involving AI, such as AI face replacements, attracted attention to deep synthesis technologies like Deepfake. Article 4 of the Deep Synthesis Regulation also emphasizes that deep synthesis services should aim "for better and good". Currently, the Interim Measures aimed at regulating generative AI encourage the creation of high-quality content that benefits "for better and good" (Article 5).

Secondly, in terms of regulatory methods, the Interim Measures continue to use mechanisms adopted in the Algorithmic Recommendations Regulation and the Deep Synthesis Regulation – besides general obligations for personal information protection, obligations such as algorithm safety assessment, algorithm registration, and annotation of AI-generated content are also provided for generative AI service providers.

---

24 《中华人民共和国数据安全法》 http://www.npc.gov.cn/npc/c2/c30834/202106/t20210610_311888.html

25 《中华人民共和国个人信息保护法》https://www.cac.gov.cn/2021-08/20/c_1631050028355286.htm?eqid=b7a9c7a1000acbc7000000026465ed77

26 《中华人民共和国科学技术进步法》 https://www.gov.cn/xinwen/2021-12/25/content_5664471.htm

27 《具有舆论属性或社会动员能力的互联网信息服务安全评估规定》https://www.cac.gov.cn/2018-11/15/c_1123716072.htm

28 《互联网信息服务深度合成管理规定》 https://www.gov.cn/zhengce/zhengceku/2022-12/12/content_5731431.htm

29 In the act the rule maker uses the term "向上向善," which literally translates as "for the better and good."

30 Chinese term – "科技向善".

Using the above-mentioned measures, on the one hand, regulatory authorities can, through algorithm registration, learn the properties of algorithms used in generative AI services, such as algorithm data, algorithmic strategy, and algorithm risks, and through safety assessment, understand other relevant information about the object of regulation, including the state of safety system implementation and technical measures. On the other hand, users of generative artificial intelligence services can also mark content created using a specific algorithmic technology through labels, be fully informed about the use of the respective technologies, and thus reasonably select and evaluate relevant informational content on the Internet.

In addition, to facilitate the enforcement activities of state agencies, such as Chinese cyberspace authorities, and to enhance user rights protection, the Interim Measures continue to provide the powers of "algorithmic service inspections" granted to the relevant authorities by the Regulation on Algorithmic Recommendations Management and the Regulation on Deep Synthesis Management. On this basis, the Interim Measures further specify that service providers need to provide necessary information that may influence user trust and choices. For example, the sources and volumes of data used for algorithm training, manual labeling rules, and so on. This way, based on the most critical information and data, regulatory authorities can manage the complexity, uncertainty, and variability of artificial intelligence effectively.

In practice, cyberspace regulatory authorities, in collaboration with relevant departments, carry out systematic work on the registration of generative artificial intelligence services in accordance with the requirements of the Interim Measures. The registered information is published on the official website of the Cyberspace Administration of China[31]. Entities providing artificial intelligence services with public opinion attributes or social mobilization capabilities can apply for registration with the local cyberspace regulatory authority. The local cyberspace regulatory authority should promptly publish the registered information, and the Cyberspace Administration of China will regularly update and summarize it on its official website without the need for additional notifications.

BASIC SAFETY REQUIREMENTS FOR GENERATIVE ARTIFICIAL INTELLIGENCE SERVICES

On March 1, 2024, the Basic Safety Requirements for Generative Artificial Intelligence Services[32] (hereinafter referred to as the Basic Safety Requirements) were officially published by the National Committee of China for Cybersecurity Technology Standardization. The Basic Safety Requirements detail the requirements for the implementation of relevant provisions of the Interim Measures, such as the legality of data sources, content security, etc., and provide effective methods for service providers of generative artificial intelligence to conduct safety assessments in practical activities. This not only contributes to improving the internal capabilities of enterprises in the field of generative artificial intelligence service safety but also provides standards for regulatory authorities to evaluate the safety level of specific generative artificial intelligence services.

The fundamental provisions proposed in the Basic Safety Requirements can be summarized as follows:

(1) Clear definition of key terms.

In the Basic Safety Requirements, the concept of a generative artificial intelligence service provider is clearly defined – this is an organization or individual that provides generative artificial intelligence services in the form of an interactive interface, a programmable interface, etc. Additionally, other key terms are defined. For example, "training data"[33] includes all data directly used for algorithm training, including those used during pre-training or optimization training, regardless of the training stage. The quality metric of the sample should be calculated considering 31 types of safety samples listed in Appendix A of the Basic Safety Requirements. Clear definitions of these terms help service providers of generative artificial intelligence better understand compliance requirements and properly fulfill their responsibilities.

(2) Clear definition of five types of safety risks.

Appendix A lists five main safety risks, including the violation of core socialist values, discriminatory content, violation of commercial rights, violation of the legal rights and interests of third parties, and the inability to meet specific service needs. It appears that the violation of core socialist values and discriminatory content belong to the high-risk category and require special attention and management. These clear definitions of safety risks help service providers better identify and avoid potential administrative violations.

(3) Compliance obligations list.

The Basic Safety Requirements offer a detailed list of compliance obligations, including language, model, and other safety measures. For example, regarding language data safety, in addition to the requirement

31 www.cac.gov.cn/2024-04/02/c_1713729983803145.htm

32 《生成式人工智能服务安全基本要求》 https://www.tc260.org.cn/front/postDetail.html?id=20240301164054

33 Chinese term – "训练语料".

of preliminary assessment and re-verification of language data sources, service providers are required to conduct a separate assessment for each item of the Basic Safety Requirements according to the safety assessment provided by the Interim Measures to ensure full compliance with safety standards. Furthermore, for model safety, specific requirements are presented, such as using a registered base model when providing services based on a third-party base model to ensure the safety and accuracy of the output data.

It is worth noting that to achieve the dual goals of promoting the healthy development of the industry and preventing risks in content, one can also explore regulatory approaches based on risk classification levels. This approach implies first classifying the risks arising from AI-generated content and then requiring service providers to take necessary measures depending on the risk level. For example, the European Union Artificial Intelligence Act[34] uses a regulatory approach based on the classification of different AI systems by risk levels to manage risks. This act first divides various AI systems into risk levels and then develops corresponding risk regulatory measures for each level, paying particular attention to strict regulation throughout the life cycle for high-risk AI systems. Some scholars note that the EU regulatory approach faces significant difficulties. If generative AI is considered high-risk overall, it may lead to over-regulation and negatively affect industry development. Therefore, generative AI should be classified as having a general risk level while applying risk regulation measures similar to the EU's Digital Service Act[35], requiring service providers to take higher measures to prevent systemic risks (Helberger and Diakopoulos 2023)

The above-mentioned risk regulation scheme is essentially a transition from risk classification for various products to classification of various risks arising from the same product, which is more suitable for technologies like generative AI capable of creating diverse risks. This risk-based regulatory method already exists in Chinese legislation. An important means of preventing illegal or undesirable content on the Internet is content verification by platforms. Chinese legislation establishes public law obligations for verification for internet platforms concerning national and public security risks, which are high-risk. Conversely, for low-risk risks related only to the violation of private rights, China follows international practice in its legislation, exempting internet platforms from general private law verification obligations but requiring due diligence. This risk-based regulatory approach can be applied to prevent risks related to AI-generated content. In cases of illegal and undesirable content related to national and public security, characterized by high risk, the obligation for service providers to take preventive measures should be stricter; in cases of illegal and undesirable content related only to the violation of private rights, characterized by low risk, the obligation for service providers to take preventive measures should be relatively less stringent (Yao and Li 2023).

CONCLUSION

In modern society, algorithms and artificial intelligence technologies are developing rapidly. Detailed regulation of relevant algorithms or artificial intelligence services based on their application in various fields undoubtedly has a positive impact on the development of the modern economy. While generative artificial intelligence services like ChatGPT are very convenient in the field of language translation and content creation, at the same time, there are risks of copyright infringement and privacy breaches since the promotion of algorithms and artificial intelligence technologies requires extensive data training, which involves relevant risks of data compromise during acquisition. Thus, in creating content with the help of artificial intelligence, issues of illegal or harmful content may also arise due to algorithmic discrimination and other factors. Therefore, the development of the Interim Measures and Basic Safety Requirements, introduced after the adoption of special regulations to manage algorithmic recommendations and deep synthesis to address societal concerns in China, is an important step, and deep regulation of artificial intelligence services can be a very useful experiment.

34 https://artificialintelligenceact.com

35 https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/europe-fit-digital-age/digital-services-act_en

REFERENCES

1. Helberger, N., and N. Diakopoulos. 2023. ChatGPT and the AI Act. Internet Policy Review 12(1). https://doi.org/10.14763/2023.1.1682
2. Pu, Q., and W. Xiang. 2023. Opportunities and challenges aroused by ChatGPT as generative AI and strategy for response. Journal of Chongqing University(Social Sciences Edition) 29(3): 102-114 (in Chin.) / 蒲清平，向往. 生成式人工智能一—ChatGPT的变革影响、风险挑战及应对策略[J]. 重庆大学学报(社会科学版), 2023,29(03):103.
3. Yao, Z. W., and Z.L. Li 2023. Legal Regulation of Generative Artificial Intelligence Content Risk. Journal of Xi'an Jiaotong University (Social Science Edition) 43(05): 147 (in Chin.) / 姚志伟, 李卓霖. 生成式人工智能内容风险的法律规制[J]. 西安交通大学学报(社会科学版), 2023, 43(05): 147.