# Research on the legal regulation of Generative Artificial intelligence—— Take ChatGPT as an example

Luodongni Yang *

Faculty of law, Qingdao University, Qingdao, Shandong, China

**Abstract.**Generative artificial intelligence is an emerging hot topic in the field of artificial intelligence, with ChatGPT being the most representative worldwide. This article focuses on ChatGPT,.In the first and second parts, we dissect the micro operational structure of ChatGPT and categorize the legal risks/issues it may bring, in order to reduce the risks and solve the current problems brought by generative artificial intelligence, improve user safety during use, encourage the development of generative artificial intelligence industry and protect technological innovation and cultivate entrepreneurial spirit. In the third part, this paper comprehensively examines the legal means and rule systems adopted by the EU, the UK and China to regulate generative AI, and evaluates the respective characteristics of the current Western model and Chinese model. Subsequently, in the fourth part, the macro principles for regulating generative artificial intelligence are proposed..Also, micro rules about regulations are constructed to address specific legal issues. It constructs a reasonable and feasible systematic legal regulatory framework. Finally, we will return to mainland China and provide some forward-looking suggestions for regulating the legal issues related to China that arise during its use.

## 1. Introduction

Artificial intelligence plays an important role in the development of science and technology in the world. In recent years, the generative artificial intelligence represented by ChatGPT has become the latest form of the development of artificial intelligence. When it comes to generative artificial intelligence, we have to discuss the concept of supremacy, namely artificial intelligence. The concept of artificial intelligence was first introduced at the Dartmouth Conference in the United States in 1956.[1] With the development of science and technology, artificial intelligence technology has made new breakthroughs, and the concept system has become increasingly diversified, but it has never formed a formal and unified interpretation. For example, Professor Nils John Nilsson believes that artificial intelligence is the science of how to represent knowledge and how to acquire and use knowledge,[2] while Professor Patrick Winston (Patrick H. Winston) believes that it is artificial intelligence studying how to enable computers to do intelligent work that only people can do in the past.[3] The most popular, Wikipedia defines artificial intelligence as intelligence expressed by artificially made systems. It also refers to the scientific field of studying whether and how such an intelligent system can be implemented.[4] Generally speaking, artificial intelligence can be derived from the appropriate meaning of its text composition, and the artificial intelligence consists of two parts: artificial and intelligence. Among them, artificial refers to the full use of their own subjective initiative, through the way of labor, reflecting human wisdom and strength; intelligence refers to through the perception of the outside world and their internal calculation, so as to make the corresponding behavior, is the expression of human will externalization. Therefore, a safer general concept can be concluded that artificial intelligence refers to the theory, method, technology and application system that uses computer or computer-controlled machines to simulate, extend and expand human intelligence, perceive the environment, acquire knowledge and use knowledge to obtain the best results.[5]

Compared with artificial intelligence, the meaning of generative artificial intelligence (AI Generated Content / AI Generated Media, AIGC) is determined by its directions of usage and ways of utilisation —— generation. That is, compared with other forms of AI, generative AI calculates and processes the learned data, and then produces new output results. Generative artificial intelligence, as a universal concept, does not have a unified definition in various countries. Wikipedia interprets it as short for the production, manipulation and modification of data or media through artificial intelligence algorithms. From the common understanding of current Chinese academic researchers and business practitioners, AIGC refers to a new production mode that automatically generates content with artificial intelligence technology after professional generated content and user-generated content.[6] At the same time, Chinese regulators further define it from the perspective of generating source as generating text technology based on algorithms, models,

* Corresponding author:yangluodongni@163.com

rules, pictures, sound, videos, code and other content.[7] Although the generative artificial intelligence represented by ChatGPT has not been officially recognized at the official level in China, Chinese local products such as Chat intelligent AI, Sheniu oral English, AI creation Max, AIbot and AI Xiaoice have shown a trend of all flowers bloom. To this end, it has sparked a heated discussion in China. Researchers have done a lot of research and proposed a series of relevant theories on generative artificial intelligence and its legal problems. Meanwhile, although there is no effective legal document regulating the generative AI, and there is no section related to generative AI in the current law, the Cyberspace Administration of China has released notices that some provinces and cities issued a series of local pilot legislation with the role of administrative regulators to regulate the legal risks of generative AI technology developments, commercial promotion and usage of users. On the whole, we can see the attitudes, methods, ideas and wisdom of Chinese regulators on regulating generative artificial intelligence.

Therefore, this paper will take the most representative form of generative artificial intelligence in the world, ChatGPT, as the analysis object, dissect its micro-operation structure, and categorize the legal risks / problems it may bring. It is not to find problems than the purpose. It is the proper goal of this paper to reduce and solve the risks brought by generative AI, improve the safety of users and encourage the development of generative AI industry, protect scientific and technological innovation and cultivate the spirit of entrepreneurship. In the end, the third part of this paper will comprehensively examine the legal means adopted by the European Union, the UK and China, and evaluate the characteristics of the current Western model and the China model. Then, in the fourth part, this paper puts forward the macro principle of regulatory generative artificial intelligence, and gives the micro rational rule structure for specific legal issues, so as to construct a reasonable and feasible systematically legal regulation framework. Finally this paper will return to China to provide some forward-looking suggestions on the legal issues related to China in the process of regulating its use.

## 2. Risks and hidden dangers of generative artificial intelligence: taken ChatGPT As a typical example

### 2.1 The microscopic mechanism and super function of ChatGPT

The full name of ChatGPT is Chat Generative Pre-trained Transformer. As an AI chatbot program developed by Open AI, an artificial intelligence research institute in America, was launched in November 2022. This program uses a large language model based on the GPT-4 architecture and trained with reinforcement learning. Currently ChatGPT interacts with users in a written way. As a typical example of generative artificial intelligence, ChatGPT has super capabilities in many aspects.

First, the super-strong ability to process language. With robust text generation capabilities and strong scalability, it can automatically generate text that adheres to grammatical specifications and logical coherence. It can also finish different tasks such as text generation, summary and translation. At the same time, ChatGPT is a multilingual expert, supporting multiple natural languages, and can process translation and communication between different languages.

Second, the strong plasticity of user experience. ChatGPT can be adjusted according to the different dialogue scenes, so that users can form deja vu. ChatGPT has a high immediacy, responding rapidly. It can immediately answer the user's inquiry. ChatGPT is also highly integrated so that developers can arbitrarily integrate it into their own developed software.

Third, super-adaptability in learning and training. ChatGPT has powerful language processing and generating functions, which are mainly derived from the reinforcement learning and training of large-scale text generated and fed back by human’s data. Specifically, learning and training of ChatGPT can be divided into two stages. In the pre-training stage, ChatGPT grasps the basic structure and semantic rules of the language by learning a large amounts of unlabeled text data. These data mainly come from the Internet, including news articles, blogs, forums, books and so on. During training, the model randomly masks some words using a method called Mask Language Model , and the model needs to predict these masked words based on the context information. In this way, ChatGPT learns to capture the semantic and syntactic relations in the text. Then during the fine-tuning phase, the model is trained using task-specific label data to adapt to different application scenarios better. These labeled data often include high-quality conversations generated by humans and question-and-answer conversations relevant to specific tasks. During the process of the fine-tuning, the model learns how to generate more accurate and relevant responses based on the input.

### 2.2 Legal risks and hidden dangers of ChatGPT

As a new generative artificial intelligence product, ChatGPT not only brings convenience to people, but also inevitably issues some risks and hidden dangers, while the legal risks and hidden dangers in the process of usage and development mainly focus on rights protection and supervision mode.

#### *2.2.1 Rights protection*

First, the production process of generative AI software and the product may infringe on the legal rights of others. As mentioned above, ChatGPT relies on massive information of human language databases, including a large amount of information entered by Internet users themselves. Therefore, when users input information such as personal data or trade secrets, ChatGPT may incorporate it into its own corpus then cause the risk of leakage. Although the developers of ChatGPT promise to delete all personally identifiable information, it does not explain the standard and methods of deletion. In the case that it cannot evaluate

effectively and safely delete the information and data sources, these personal information still has the risk of disclosure, which may pose a threat to the individual privacy right. In addition, when using data in existing databases for ChatGPT training, it is possible to use some data involving improper behaviors and materials, including but not limited to data about personal attacks, gender discrimination and racial discrimination and so on. When ChatGPT generation directly or indirectly reflects these data, it is likely to violate the human rights and human dignity of specific individuals or groups. More seriously, the authenticity of the information presented by the ChatGPT generation is questionable. Take China as an example, in February 2023, a news that the Hangzhou municipal government of China will cancel the restriction policy on the final number of motor vehicles on March 1 went viral on the Internet. But Hangzhou police later confirmed that it was a fake news written by ChatGPT. [8]It can be seen that as ChatGPT will make serious nonsense, users need to have certain professional knowledge to distinguish the authenticity, otherwise they will believe the false information easily.

Second, the infringement process of generative AI software is relatively hidden, so it is difficult to allocate the tort liability and the principle of attribution is difficult to be clear, leading to the difficulty of accountability and the high protection cost. Due to the ability to learn of generative artificial intelligence, it is difficult to judge whether the infringement results are implemented by AI after autonomous learning. For example, in the AIGC application scenario, if the code developer is responsible for creating algorithm, which changes through self-learning, and then in a specific application scenario, the generated content ultimately leads to damage. Should the developer bear tort liability? Obviously, the developer did not directly commit the infringement, and had no subjective intention of negligence. However, due to the limitation of technical standard, most of the existing machine learning of artificial intelligence generation tools is not high, and its generated content is mostly subject to the input of operators. On this basis, the corresponding text, image, audio and other content are generated (for example, the ChatGPT application is an AI generation tool based on input of users). In this kind of AI generation scene, if the generated content objectively causes damage results, how to confirm the responsibility subject of it? After all, it is also said that the infringement is completely and independently implemented by AI through autonomous learning. Therefore, this means that the constituents of the tort liability of the generative artificial intelligence are different from those of the general tort liability, but the specific differences need to be further discussed.

Third, the intellectual property rights of generative artificial intelligence products have become the main legal risk and problem about ChatGPT, including (1) the copyright ownership of the products (2) whether the generation infringes the intellectual property rights of others. The two issues related to intellectual property rights will be systematically introduced respectively.

The first question is the copyright of ai-generated content belonging. To answer this question, first we should discuss its upper concept —— work. According to Article 3 of the Copyright Law of the People's Republic of China, the works mentioned in this legislation refer to intellectual achievements that are original and can be expressed in a certain form in the field of literature, art and science.[9] Article 2 of the Regulations for the Implementation of the Copyright Law of the People's Republic of China stipulates: The works mentioned in the Copyright Law refer to the intellectual achievements with originality in the field of literature, art and science and that can be reproduced in some tangible forms.[10] It can be concluded from the definition of work in two official legal documents that the work is the achievement of the author's independent intellectual and creative activities. But when it is applied to unscramble AI products, it seems impossible to give a proper interpretation. After all, the process of AI generating content cannot be interpreted as intellectual activities of human. Artificial intelligence generating content is essentially a process of computing for computers based on big data and algorithms. In contrast, human intellectual activities are more complex. It includes not only rational thinking activities (such as the mastery and application of writing rules), but also perceptual thinking activities (such as thoughts, emotion and inspirations). Even if the content generated by artificial intelligence has the superficial features of the achievements of human intellectual creations, it is difficult to be included in the category of works in the meaning of human intellectual creation activities. Therefore, whether there is copyright in the content generated by AI (that is, whether it can be called works) and if there is copyright, where is the copyright belonging? Current law in China cannot answer these questions.

The second question is whether the ChatGPT product may involve infringement of intellectual property of others. The learning and training process of ChatGPT is described in detail above. Obviously, ChatGPT will definitely use the existing human language data in the network database during the training process, and some language data may be written works where others enjoy intellectual property rights. Although ChatGPT officials have constantly claimed that the materials used for its training come from Open Data Bases（ODB）to avoid risks of violating others' intellectual property rights. After all, open databases are often seen as an updatable platform for technicians (developers, users) to share and reuse code among themselves. Anyone can modify it to meet the needs of their own specific application scenario. In addition to providing software code, open databases can also provide documents, such as requirements analysis documents, design documents, test documents and so on. In addition, open databases include data structures, device drivers, sample projects, and language supplement application components and toolkits. Besides, many open databases also include examples to help users with specific tasks to adopt a certain technology in real applications. However, it is hard to say that the data of ChatGPT users comes entirely from open source databases. At the same time, even if the open source database comes from the open source database, ways of utilizing data of the open source database is generally stipulated in advance,

so the production process of using the data beyond the permission undoubtedly infringes the intellectual property rights of others, because it goes beyond the scope of reasonable use. Therefore, the process of AI learning and its products do indeed have the possibility of infringing intellectual property rights.

#### *2.2.2 Risks and hidden dangers of absent supervision*

First of all, the generative artificial intelligence represented by ChatGPT has been applied in some important industries. If its application is not reasonably and effectively regulated, it may hinder the development of the industry under the new trend of widespread application of artificial intelligence. Taking the legal industry as an example, generative artificial intelligence is mainly used to help legal workers to deal with a large amount of information and effectively complete various tasks. It has replaced a part of the judicial auxiliary personnel who do repetitive work. However, the existing generative artificial intelligence software still lacks the ability to create, understand emotion and think comprehensively. Whether it can give professional legal advice like lawyers or make fair judicial decisions like judges is still uncertain, but this possibility cannot be ruled out in the future. It can be predicted that with the deep development and popularization of artificial intelligence, the legal profession will be subversively shaped by the power of science and technology. The career function and development direction of the legal industry under the new trend still need to be further discussed.

Secondly, generative artificial intelligence presents great challenges to transaction security. Generative artificial intelligence technology products may be used to mining transaction data on the network because it has ability to capture data and generate text according to the scene of usage. If the network was attacked by hackers, or received Trojan virus, the information stored in the database of the program would be stolen without awareness of users and developers. For example, in application scenario of marketing, interest, which is closely related to the finance, if users use such an artificial intelligence technology, the content of the process of chatting with ChatGPT may be disclosed in other unsafe scenarios, in which the content is collected and become learning materials of algorithm training. It may cause financial data leakage accidents and threaten the financial security, data security and regulatory security.

Finally, some of the actions implemented by using generative AI, which potentially do harm to the public interest and subvert social ethics. In 2022, China issued the Internet Information Service Algorithm Recommendation and Management Regulations. It has declared: Algorithms should adhere to the mainstream value orientation, actively spread positive energy, and shall not use algorithmic recommendation services to endanger national security and social and public interests, disrupt economic order and social order, and infringe on the legitimate rights and interests of others.[11] As an artificial intelligence technology, algorithmic architecture of ChatGPT is planned and programmed by human engineers. It is still a black box model, and it is unknown whether its algorithm can sufficiently disclose information. It cannot be excluded that some illegal elements take advantage of the technical loopholes of their algorithm architecture or database review loopholes to carry out illegal activities, disturbing public order and endangering public interests. For example, some criminals convey false information to the public; deliberately fabricate inflammatory speech; intentionally influence political ecology; maliciously incite violence and crime. Some people will make use of this kind of generative artificial intelligence program technology vulnerabilities, collect different types of user's privacy data, and then suit the remedy to the case by providing phishing site links to directional fraud. These illegal and criminal behaviors by using generative artificial intelligence have greatly challenged the social public interest and social ethical standards.

## 3.Comparative observations of generative AI regulation: the European Union, the UK and China

### 3.1 The UK has a regulatory system centering on the Artificial Intelligence Regulatory Methods to Promote Innovation

The UK has established a regulatory framework for the regulatory needs of specific high-risk scenarios in the use of AI to address the problems of how to regulate the usage of AI applications in such scenarios. Earlier, there was no special law in the UK to regulate AI products and services, and the relevant content of the regulation was scattered in various departmental laws. Its current laws related to AI include the Equality Act 2010, the Product Safety Act and the UK Consumer Protection Act. However, some official documents have set forth requirements for the application and development of AI in related fields, such as the Ethics Principles and Policies of AI released by the Ministry of Defence. In April 2023, for specifically high-risk sences of applications, the British department for science, innovation and technology has issued an official document —— A pro-innovation approach to AI regulation white paper (hereinafter referred to as The Approach). [12]This document has put forward its artificial intelligence management methods based on the five principles. The purpose is to provide enterprises and the masses with confidence of using artificial intelligence, provide industries with certain and consistent regulatory methods. At the same time, the white paper noted that the AI industry will not be strictly regulated, so as to encourage innovation and ensure that the response to challenges is in a timely manner. In addition, the white paper particularly emphasizes the importance of collaborative governance by the government, industry, enterprises and other entities, for the purpose of strengthening global cooperation and interoperability in AI governance, so as to achieve the UK's global position as a leader in AI as soon as possible. It should be noted that the Approach emphasizes the pilot

implement, observing the results of the proposed regulatory methods, and then further legalized the regulatory principles of regulatory agencies, such as the Office of the Information Commissioner (ICO), the regulatory Agency (the Equality and Human Rights Commission (EHRC), and the Financial Conduct Regulatory Authority (FCA)).

In terms of the level of regulatory power, The Approach lays more emphasis on the decentralized regulatory structure and methods, emphasizing the participation of existing regulatory agencies and the functions of regulatory agencies. Firstly, the document stresses that there is no need for a dedicated AI regulator, and the responsibility for regulating AI should be borne by existing regulators. Secondly, The Approach emphasizes the need for coordination of regulation. While abiding by their profession and fulfilling their duties, they should be consistent with law enforcement. At the same time, they should not abuse power of regulation. Thirdly, if the regulators need supports, the central government will provide these regulators with multifaceted support. But in cases where regulators are unable to handle more than two conflicts of principles, the government wants regulators to work with the government in a timely manner to identify potential actions. Finally, the government also hopes that regulators can effectively coordinate with multiple departments to jointly formulate regulatory guidelines to clarify their own understandings of regulatory principles. We also want to develop a cross-economic and whole-social assessment framework of risk, so as to ensure that risks fall in the responsibility range of diverse regulators.

In terms of regulatory subjects, The Approach proposes to establish a special department to develop a centrally reasonable monitoring and evaluation framework (M & E); develop a reasonable and multi-channel feedback framework, and extensively solicit opinions and suggestions from government, regulators, AI industry, civil sector, and society; provide central support to regulatory agencies; supervise and evaluate the regulatory framework mentioned above.

In terms of regulatory professionalism, The Approach proposes that all regulatory agencies should try their best to learn professional knowledge of artificial intelligence to narrow the gap among diverse agencies in professional knowledge and ensure that regulatory behaviors are scientific and correct. For example, the government invites regulators to pay attention to the risks arising from the application of new basic models (such as LLM large-scale language models); the government wants to work with existing regulators to develop sandbox models to assist innovators in the AI industry better. The government also wants regulators to develop generic technical standards. The Approach also mentions that the above measures need to be piloted.

In terms of regulatory coverage, The Approach specifically mentions the need for effective linkage and response to the international regulatory framework. While formulating the domestic regulatory standards and technical standards, we should also pay attention to the international standards, ensuring that these principles of risk management are consistent with international standards. By confirming opportunities for regulatory interoperability, supporting cross-border coordination and collaboration, so as to achieve comprehensive coverage of regulation inside and outside the region and prevent problems caused by the absence of cross-border regulation.

### 3.2 The regulatory system formed by the EU cetreing the Artificial Intelligence Act

The EU's regulation of generative AI is mainly reflected in the issuance of the Artificial Intelligence Act (hereinafter referred to as the Act), so as to ensure that development of AI in Europe is people-oriented and moral. Specifically, the Act, which was amended and passed in May 2023, responds to AI infrastructure models, responsibilities of value chain and the recent popularity of generative AI. At the same time, it studies exemptions for businesses, especially small or medium-sized enterprises and start-ups. The Act mainly puts forward regulatory requirements for generative AI from the following aspects, the specific contents of which can be seen in the following table:

**Table 1.** The Construction and essence of the Artificial Intelligence Act[13]

| Regulatory aspects | Law serial number | Core essence |
|---|---|---|
| Definition of the regulated objects | Section 3 of Part I of the Act | The AI system and its related usage, developments, regulatory subjects and their behaviors are clearly defined. |
| Introduction of the regulatory principles | Section 4a of Part I of the Act | It proposes six general principles which are applicable to all AI systems: human subject and supervision, and technology robustness and security, Privacy and data governance, and transparency , diversity, non-discrimination and fairness, and social and environmental well-being |
| Setting-up of the regulatory authorities | Section 3 (42) of Part I of the Act | It proposes to set up a centrally regulatory agency (the European Committee on Artificial Intelligence) and a nationally regulatory authority in each member state |
| AI accomplishment | Section 4d of Part I of the Act | It are requirements for regulatory authorities, providers and deployers of AI systems to strengthen the training and education of AI expertise, requiring the above subjects to ensure that their staff and other personnel who handle the operation and use of AI systems on their behalf have sufficient AI knowledge level |

| | | |
|---|---|---|
| risk management | Section 6,7,8, and 9 of Part III of the Act | The AI system is divided into three risk levels: unacceptable risk; high risk; non-high risk. According to different risk levels, the Act regulates AI systems at that level differently, especially high-risk AI systems, and the Act hopes that evaluate high-risk AI systems in terms of both rights impact and technical standards |
| The behaviors of the supervised subjects | Chapter III, full chapter of Part III of the Act | Detailed guidance and requirements are put forward for the behaviors that are prohibited and encouraged to implement. It also introduced a new penalty of 6% or 30 million euros (whichever ever) |

### 3.3 Summary of the regulatory model in the UK and Europe

The UK and the European Union belong to Europe, and Europe is leading the world in the development of artificial intelligence. Comparing the UK and the legal regulation of the AI system with the EU, we can find something meaningful that the regulatory direction chosen by the UK is clearly different from the EU, and shows the differences in the following five aspects:

First, a broader definition. The UK White Paper defines AI as any system with a function of combining adaptability and autonomy. The broad definition ensures that regulations can be flexibly applied and modified to accommodate evolving technologies. In contrast, the EU's AI regulations use a fixed and narrow definition of AI.

Second, it concentrates on specific situations. The UK White Paper avoids a classification of specific products, instead favors outcome-based regulation. The application of AI will be analyzed on a case-by-case basis, and considered risk Regulators which will balance the benefits and risks. In contrast, the Act proposes a formal risk classification (from low to high) and bans the use of AI technologies that are deemed to pose an unacceptable risk.

Third, there is no new regulator. A decentralized UK system will not create new regulators; instead, it will rely on existing regulators. By contrast, the EU proposes a central regulator (the European Council on Artificial Intelligence) and national regulators in each member state.

Fourth, there is no new punishment. The white paper did not introduce new penalties for violations. Instead, UK regulators will use their existing enforcement powers to punish relevant offences. By contrast, the EU system introduced new fines of up to 6% of the global annual turnover or 30 million euros (whichever highest).

Fifth, the role of the government in the regulation is different. The government will provide central support functions to UK regulators, including monitoring effectiveness and consistency, assessing economic impact, and promoting interoperability with foreign institutions. The white paper recommends that the government will work with regulators to implement its recommendations. Its roles will include identifying risks to AI applications, working with regulators to identify central risks across sectors, and updating laws and regulations to address risks marked by regulators.

It can be concluded that while both are generally designed to promote the adoption of people-oriented and trustworthy AI and to ensure a high degree of protection of health, safety, fundamental rights, democracy, rule of law and the environment from the harmful effects of AI systems within the effectiveness of the law domain, supporting innovation in the meanwhile. However, the two main regulatory concepts are not the same. In terms of the overall concept of legal regulation, the UK prefers the open-regulatory concepts, without imposing too many restrictive provisions on AI systems and their related subjects. While the EU prefers the regulatory concepts which are relatively rigorous, putting forward a large number of prohibitive provisions on AI systems and their related subjects.

### 3.4 The legal means adopted by China and the rule system formed

In recent years, China has mainly issued administrative regulations and local regulations to supplement the current laws for the lack of AI systems and product sectors. Meanwhile, China puts forward new requirements for the development of the AI industry through releasing policy documents. There are two policy documents concerning the regulatory and normative Generative AI systems or products issued by Chinese regulators: one is the Measures for the Administration of generative ARTIFICIAL Intelligence Services (Draft) (hereinafter referred to as the Measures) and the Regulations for the Administration of Internet Information Services jointly issued by the Cyberspace Administration of China (hereinafter referred to as the Regulation) , the Ministry of Industry and Information Technology and the Ministry of Public Security. Specifically, the Measures set certain obligations for users and developers of generative AI technology, including but not limited to not to release illegal content, no discrimination; respect the rights and privacy of others. In view of the practical harm and potential risks brought by deep synthesis technology, the Regulation puts forward scientific, refined and systematic governance requirements, delimit the bottom line and red line of deep synthesis services, and promote the effective use of deep synthesis technology in accordance with the law and regulations. Generating a pre-trained transformation model (GPT) is a kind of deep synthesis technology, so the ChatGPT developed based on the GPT model is also within the regulatory scope of this regulation. In addition, the local legislation issued by Chinese regulators includes the Regulations of the Shenzhen Special Economic Zone on Promoting the AI Industry and the Regulations of Shanghai Municipality on Promoting the Development of the AI Industry. The two local regulations are actually a manifestation of pilot studies on

AI regulation by Chinese regulators in some developed regions of China.

Overall, Chinese regulators have a relatively positive regulatory awareness. Through the above official documents, we can see that on the one hand, the regulators actively promote the continuous development of the AI industry in China, and on the other hand, they also conduct preventive regulation on the scientific ethics, public interests and other issues that may be involved in the development process. Compared with the UK and the European Union, China covers more comprehensive industries in the legal means and the rule system of regulating generative AI. However, there are also some deficiencies. For example, for regulatory generative AI technology, the regulatory suggestions put forward by Chinese regulators are normally relatively vague, which cannot be implemented directly.

## 4. The future improvement of the legal regulation of generative artificial intelligence

As a new thing, the supervision of generative artificial intelligence must appropriately learn from the advanced legislative experience of various countries. On the premise of promoting the development of generative AI that is more conducive to people's work and lives and developments, we should establish and gradually improve the legal regulatory framework of generative artificial intelligence. In general, a scientific, complete regulatory system should be able to clarify respective rights, obligations and relationships of the developers, users, the third party and related subjects, effectively protecting the legal rights of the relevant subject. Meanwhile, some special behaviors implemented by related subjects of generative AI should be effectively regulated.

### 4.1 Regulatory measures aimed at protecting the rights of relevant subjects

Regulators must clarify the legal status of generative AI. For example, does generative AI have an independent personality at this stage? Should there be a certain capacity of rights and behaviors? In other words, in some situations, it has partially independent rights and (or) capacity. Only when their legal status is clarified can regulators discuss whether the current laws can be applied to restrain, regulate and even punish them under this premise. In addition, regulators should type the infringement issues in specific scenarios of generative AI, and then issue guiding official documents such as legal applications, legal interpretations and guiding opinions for these issues. In particular, the concept of the intellectual property infringement should be clarified, otherwise it will lead to the overlap and confusion of the regulatory system. For instance, regulators should set rules for using code, documents, images and other resources in the open source database about the entry and range. To further illustrate, if resources from a specific open source database are commercially developed, it may be involved in infringing copyrights of others, so these resources can only be used for personal studies and researches, instead of being used for developing commercial systems, products, etc.

### 4.2 Regulatory measures aimed at the reform of regulatory methods and regulatory standards

First of all, regulators should clarify the risk level of the target. Regulators should assess the risk of the generative artificial intelligence existing at the present stage, and then classify these AI. They should adopt different measures with different priorities to regulate the generative artificial intelligence with different risk levels. But they should always pay attention to the aforementioned premise, otherwise it will hinder the development of generative AI. At the same time, regulators should promote the joint participation of multiple departments in special legislative activities and the formulation of cross-departmental technical standards. The different functions of the regulatory authorities lead to their different regulatory priorities, and it is also different that targeted professional knowledge related to generative AI mastered by these regulatory authorities. If multiple regulatory authorities make different nature judgments on the same behavior and adopt different regulatory behaviors, it will cause the regulated objects themselves and the innovators within the artificial intelligence industry to be confused, which may weaken the innovation power of the industry. Therefore, at present, the more urgent goal is to establish a cross-departmental regulatory system and specific technical standards as well as to improve the professional knowledge level of various regulatory departments. Finally, the regulators should work hard to promote the implementation of the regulatory measures. The current official regulatory documents and standards are more designed and regulated in the general concept, lacking of implementation rules that can be implemented. For some phenomena of generation in generative artificial intelligence in actual production and life, grassroots regulatory authorities may know they need to regulate, but do not know what regulations should be followed. Therefore, the grass-roots regulatory departments should conduct more coordination with the superior regulatory departments to further standardize the regulatory behaviors in actual operations.

### 4.3 To refine the constructions of specific rules as the core of the regulatory measures

At present, the regulatory suggestions put forward by regulators are often relatively vague, so these suggestions are difficult to be implemented by the regulatory enforcement agencies. For this reason, it is necessary to further refine regulatory objects more clear, the regulatory scope is more comprehensive, and the regulatory basis is clearer. An ples will be given below.

**Example 1: For the using permission and scope of the code, documents, pictures and other resources in the open source database.**

Article 7 of China's Measures for the Management of Generated Artificial Intelligence Services (Consultation

Paper) stipulates the following provisions: Providers shall be responsible for the legitimacy of the pre-training data of generative AI products and the optimization of the source of training data. The pre-training and optimization training data used for generative AI products shall meet the following requirements: (1) meeting the requirements of the Network Security Law of the People's Republic of China and other laws and regulations; (2) not containing the infringement of intellectual property rights; (3) obtaining the consent of the personal information subject for the personal information or other circumstances complying with the laws and administrative regulations; (4) ensuring the authenticity, accuracy, objectivity and diversity of the data; (5) meeting the other regulatory requirements of the national cyberspace department on generative AI services. [14]But there are some problems in this article alone, such as the fourth article, which is stricter. Large models are inseparable from the training of using existing data. If too strict regulations are made for the training data, it means that the development speed of large models will be seriously hindered. The scale of data training will determine the quality of the content generated by AI, and the quantity and quality of data are more important than algorithms and models. Therefore, if the relevant training data cannot be used due to the authenticity, objectivity and other problems, it will have a great negative impact on the development of artificial intelligence technology. At the same time, the first model also suggested adding but not including synthetic data sources for AI-generated content. That is, the provider should be responsible for the legitimacy of the pre-training data of the generative AI products and the optimization of the training data sources, but not including the synthetic data sources of the AI-generated content. It is required to explain that pre-training and optimized training data for generative AI products are divided into natural data and synthetic data. The natural data includes personal information in the physical world; public data acquired directly and data purchased through data transactions. According to the five requirements stipulated in this article, it can be understood that the pre-training data and optimized training data stipulated in this article exclude the source of synthetic data, and only refer to the natural data of the three sources. Otherwise, requiring the provider to take responsibility for the authenticity, accuracy and objectivity of the synthetic data, which obviously aggravates the obligation of the provider. It is difficult to ask the provider still take responsibilities to the legitimacy of data resources of the generated content reproduced by generative AI from the previous content that as the origin of synthetic data. Technically, it is also impossible to achieve.

**Example 2: the risk level classification for the regulatory objects.**

Article 15 of the Measures for Generated Artificial Intelligence Services (Consultation Paper) stipulates: For the generated content found in operations and reported by users that does not meet the requirements of the Measures, besides taking measures like content filtering, it shall be prevented from recurring within 3 months through model optimization training and other methods.[14] ChatGPT has been on line for several months, which obscure problems still cannot be correctly understood and not correctly answered. At present, China's online AI is still in the private beta stage, the time of implementation is not long enough, the obscure problems expose insufficiently. Therefore, it is required to prevent recurrence through model optimization and training within 3 months, the mentioned period is too short. In the large-scale model of artificial intelligence, it is difficult to accurately predict and control most of the generated content. Some experts believe that it is difficult to prevent the regeneration through model optimization training in 3 months technically. It is advisable to firstly stipulate the remedial measures that should be taken for violations, and a reasonable time limit should be determined after a lot of practices after the stable development of artificial intelligence. At the same time, the article does not specify the legal liability of the same problem still occurring three months later. The legislation is not clear whether three months is a period of exemption or increased liability.

**Example 3: The generative artificial intelligence is classified according to its functions and uses.**

For example, according to the European Union's AI Liability Directive (AILD), only high-risk AI applications can bear AI strict liability and assume different responsibilities according to the level of risk. In view of the fact that generative AI products and services are still in the development stage, the relevant regulations should not be too strict to hinder the development of artificial intelligence. The production should be encouraged and too much restrictions also should not be imposed on it, so as to improve the international competitiveness of AI industries. In addition, the scope of the regenerated content can be clarified, such as preventing the recurrence of the same kind content or the content with the same characteristics. If the model optimization fails, or it is difficult to avoid regenerating, it can be further clarified how to regulate.

**Example 4: For the joint participation of multiple departments in special legislative activities and the formulation of cross-departmental technical standards.**

Article 20 of the Measures for Generated Artificial Intelligence Services (Consultation Paper) has regulated: the provider in violation of the provisions of these measures shall be punished by relevant departments in accordance with this law, or Cybersecurity Law of the People’s Republic of China, or the Data Security Law of the People's Republic of China, or the Personal Information Protection Law of the People's Republic of China or other revelant laws and administrative regulations. In the absence of laws and administrative regulations, the cyberspace department and relevant competent authorities shall, according to their duties, give a warning and circulate criticism, and order them to correct within a time limit; if they refuse to correct or if the circumstances are serious, they shall be ordered to suspend or terminate their services provided using generated artificial intelligence and impose a fine of not less than 10,000 Yuan but not more than 100,000 Yuan. If the case constitutes a violation of public security administration, it shall be given administrative punishment for public security according to law; if the case constitutes a crime, criminal responsibility shall be

investigated according to law.[14] However, in the second paragraph of this Article, the responsibility of providers is too strict, and even for the current laws and regulations that have not been clearly stipulated, the practice of ordering suspension does not comply with the provisions of the Administrative Punishment Law. Article 14 of China's Administrative Punishment Law stipulates that if no laws and administrative regulations have been formulated, the departmental rules of The State Council may impose administrative penalties of warning, notice of criticism or a certain amount of fine for acts violating the administrative order. Visibly, with regard to unspecified violations of the management order in laws and administrative regulations, the highest penalty limit is just a warning, criticism, and fines.[15] However, the measures also stipulates that certain subjects shall be ordered to suspend or terminate the utility of generative artificial intelligence services. This belongs to some subjects that should be restricted in production and business and be ordered to suspend production and shut down. They may also be restricted from becoming practitioners. The penalty limit is significantly higher than that stated in the law on administrative punishment. This is not friendly to the subsequent development of the emerging generative artificial intelligence service industry. It is suggested that the limit of responsibility should be consistent with the Administrative Punishment Law. The suspension or termination of its services using generative artificial intelligence should be deleted, and the fine should be retained.

## 5. Conclusion

The use and development of generative artificial intelligence has facilitated to the user's work and life, greatly promoting the development of the AI industry. However, the technical complexity of generative artificial intelligence technology leads to legal risk of some process of use, strengthening the complexity of the relevant legal issues, such as generated content infringement of intellectual property rights and the difficulty of proof, damage of the network transaction security owing to data leakage. These things can be concluded as algorithm infringements. Their occurrence also helps the legal regulation of generative artificial intelligence technology which becomes an important issue and a common problem in the field of researching artificial intelligence. Overall, if Chinese leaders want to develop generative AI in an orderly, rapid and vigorous manner in China, it is necessary to refine the specific rules of regulatory measures related to generative AI. In the process of detailed rules, we can find many hidden, controversial legal issues and form a more comprehensive legal regulation advice in the collision of thought. At the same time, we can refer to foreign legislation, regulatory ideas and measures, helping Chinese regulators with the development of generative artificial intelligence constantly adjusting and modifying regulations, so as to better promote the development of Chinese native generative artificial intelligence related industries.

## Reference

1. Yongyu,2021.Overview of the Development History of Artificial Intelligence.https://www.zhihu.com/column/c_1637861790871711744
2. Nilsson, N. (1983) Principles of Artificial Intelligence. Science Press, Beijing.
3. Winston, P. (2005) Artificial Intelligence. Science Press, Beijing.
4. Anonymity,2023.Artifical Intelligence.https://encyclopedia.thefreedictionary.com/Artifical+Intelligence
5. Standing Committee of Shanghai Municipal People's Congress,2022.Regulations of Shanghai Municipality on Promoting the Development of Artificial Intelligence Industry.https://law.sfj.sh.gov.cn/#/detail?id=63ac062e60b224bab986e200
6. Qingfeng Gan,2023.On the Legal Regulation of Artificial Intelligence Jurisdiction.https://mp.weixin.qq.com/s/VriLWFYzzMaNJk-prx0XhQ
7. National Internet Information Office,2023.Management Measures for Generative Artificial Intelligence Services (Draft for Soliciting Opinions).http://www.cac.gov.cn/2023-04/11/c_1682854275475410.htm
8. JIMU NEWS,2023.Hangzhou will lift travel restrictions on March 1st? The media claims it was written by ChatGPT, and the police will release the investigation results.https://baijiahao.baidu.com/s?id=1758049381599427211&wfr=spider&for=pc
9. the National People's Congress (NPC),2020.Copyright Law of the People's Republic of China.https://www.gov.cn/guoqing/2021-10/29/content_5647633.htm
10. the State Council,2013.Regulations for the Implementation of the Copyright Law of the People's Republic of China.https://www.ncac.gov.cn/chinacopyright/contents/12230/346377.shtml
11. Ministry of Industry and Information Technology of the People's Republic of China,National Internet Information Office,Ministry of Public Security of the People's Republic of China,2023.Regulations on the Management of Deep Synthesis of Internet Information Services.https://www.gov.cn/gongbao/content/2023/content_5741257.htm?eqid=fb80e86900094abc00000004645e095c
12. the Secretary of State for Science, Innovation and Technology ,2023.A pro-innovation approach to AI regulation.https://www.gov.uk/government/publications/ai-regulation-a-pro-innovation-approach/white-paper
13. European Union,2023.EU Artificial Intelligence Act.https://techonomy.com/the-eus-artificial-

intelligence-act-a-pragmatic-approach/

14. Cyberspace Administration of China, 2023. Measures for Generative Artificial Intelligence Services (Consultation Paper). http://www.cac.gov.cn/2023-04/11/c_1682854275475410.htm
15. the National People's Congress (NPC),2021.Law of the People's Republic of China on Administrative Penalties.https://www.gov.cn/xinwen/2021-01/23/content_5582030.htm