



IJAHSS



Copyright@IJAHSS

## An analysis of hot legal issues for Chat GPT products in China

Lei Li<sup>1</sup>, WenLe Kong<sup>1</sup>, RuoHan Li<sup>1</sup>

<sup>1</sup>Guangdong Technology College, Zhaoqing, China

\*Corresponding Author

Lei Li

### ABSTRACT

While various ChatGPT products continue to develop, we need to think about two issues: one is the industry access system. With the continuous advancement of artificial intelligence, Chat GPT products will enter our daily lives slowly and penetrate various fields. From the perspective of laws and regulations, it involves the industry access system in the Internet field implemented in China, that is, before the product is put on the shelf, according to the requirements of China's laws and regulations, how the product should be reviewed, and how the competent authority defines the responsibility for platform governance. The other is the protection of rights and interests. The development of ChatGPT requires a large amount of primary data for artificial intelligence to learn, and the source of this preliminary data is the intellectual property generated by a large number of people investing a large amount of labor. These intellectual property rights are valuable, and then the question is how do we protect them? At present, laws and regulations have yet to have a clear definition. To this end, this article will, based on the existing laws and regulations and related knowledge and combing with the above issues, discuss how Chat GPT products are governed at present, what operational risks exist, and how they should be governed in compliance with laws and regulations to promote the establishment of China's ChatGPT product legal and regulatory system.

**Key Words:** *ChatGPT, legal issues, compliance governance.*

### INTRODUCTION

Artificial Intelligence has developed rapidly in recent years, and Artificial Intelligence-Generated Content (AIGC), a branch of Artificial Intelligence, has gradually become the key to research. Due to the critical breakthrough in natural language processing technology from quantitative change to qualitative change, AI can learn, understand and analyze human language to achieve the output of problem results in natural language. In November 2022, after Open AI published Chat GPT, this product quickly swept the world, setting off a wave of discussion on the Internet and constantly out of the circle. In the context of the era of more and more mature Artificial Intelligence technology and the country's close promotion of the strategy of strengthening the nation with science and technology, many domestic Internet giants have joined it and want to develop Chat GPT products, such as Baidu's Wenxin Yiyuan, JD.com's Chat JD, Tencent's "Mixed Element Assistant" project, 360's Artificial Intelligence development strategy, etc.

#### 1. The compliance governance of products like Chat GPT

##### 1.1 The development history of laws of artificial intelligence products

In 2012, the improvement of audio and video image recognition technology and recognition accuracy made it easier for people to use software to deeply synthesize audio and video images, such as artificial intelligence face-changing technology, which is to cut one person's portrait to another person's head. However, face-changing technology will cause moral and ethical problems; some criminals will even use it to conduct criminal activities. Currently, China's regulations on artificial intelligence have only some sporadic provisions, which have failed to form a systematic legal and regulatory system, making the conviction of criminals more difficult. The following five regulations are the basis for China's current compliance governance of artificial intelligence supervision:

(1) Regulations on the Administration of Internet Audio and Video Information Services were jointly issued by the Cyberspace Administration of China (CAC), the Ministry of Culture and Tourism of the People's Republic of China, and the National Radio and Television Administration (NRTA). They came into effect on 1st January 2020. Article 10, which concerns audio and video information services with media properties or social mobilization functions, and Article 11, involving identifying generated content and prohibiting the publishing of false information, regulate the audio aspect of AI-generated content.

(2) The Cyberspace Administration of China issued the Provisions on the Governance of the Online Information Content Ecosystem, which came into effect on March 1, 2020. Article 23 of the provision states that users of network information content services and producers of network information content and platform platforms shall not use new technologies and applications, such as deep learning and virtual reality, to engage in activities prohibited by laws and administrative regulations. The regulation constrains the ecology of the content generated by network information. The content generated by artificial intelligence using deep learning, machine learning, and virtual network technology must comply with the regulations of the relevant competent authorities.

(3) Article 1019 of the Civil Code stipulates that no organization or individual may infringe on the right to the portrait by scandalizing, defacing, or falsifying utilizing information technology. Article 1023 of the Civil Code stipulates that the protection of the voice of a natural person shall be governed by the provisions relating to the safety of the portrait right. The above requirements protect the right to a portrait and a person's voice.

(4) The Regulations on the Administration of Algorithmic Recommendation of Internet Information Services were jointly issued by the Internet Information Office, the Ministry of Public Security, the Ministry of Industry and Information Technology, and the General Administration of Market Supervision and Administration. They came into effect on 1 March 2022. The second, ninth, and thirteenth articles of the regulations, all of which provide for synthetic content, in particular, regulate the scope of application of algorithms.

(5) The "Regulations on the Administration of Deep Synthesis of Internet Information Services" was jointly issued by the Office of the Netizen, the Ministry of Industry and Information Technology, and the Ministry of Public Security and came into effect on 10 January 2023. The document outlines complete compliance requirements for deep synthesis services and products. It is one of the central bases for compliance governance of China's current AI content and AI publishing platforms, i.e., a management regulation for in-depth synthesis content.

## 1.2 Licensing and filing of products

China has a system of industry access, and within the Internet industry, Internet products need to meet the basic licensing requirements when listed. What are the basic requirements to be completed by platforms or software operators of Chat GPT products, as the authorities in China impose high requirements on the governance of Internet platforms? In addition to the main business of the Internet platform, other side businesses will continue to expand and develop, deriving more business lines. Still, the premise of the basic requirements is that the content generated is relatively simple text-based explanations [1]. The nature of its questions and answers can not have cultural properties, so the business upgrade and what requirements need to be met.

### 1.2.1 Basic requirements for product launch

(1) Basic license. The content required for artificial intelligence learning and the content output belongs to Internet information. Following Article 7 of the Measures for the Administration of Internet Information Services to engage in operational Internet information services, one should apply for a value-added telecommunication business operation license for Internet information services from the telecommunications administration or the competent department of the information industry under the State Council. ChatGPT products, on the other hand, are subject to B21 (online data processing and transaction processing business) and B25 (information service business, including information publishing platform and delivery services, information retrieval and query services, information community platform services, and information services) under category B of the Telecommunications Business Classification Catalogue [2]. B25 (information service business, including information publishing platform and delivery service, information retrieval and query service, information community platform service, information instant interaction service, information protection, processing service, etc.) licenses. One purpose is to be on file with the two regulators so that they know who is doing this website, what they are doing, and how to contact them; the other purpose is to reserve space for future business development.

(2) Algorithm filing. There are clear regulations on the content of profound synthesis of artificial intelligence. According to Article 19 of the Regulations on the Administration of Deep Synthesis of Internet Information Services, providers of deep synthesis services with the attributes of public opinion and or social mobilization capacity should fulfill the procedures of filing, changing, and cancellation of algorithms. Therefore, as long as the Internet platform has the nature of social opinion or has the ability of social mobilization, the products online and operations need to go through this provision for filing.

(3) Safety assessment. Before the Internet platform goes online with its products, in addition to the need to comply with the Data Security Law, the Network Security Law, and the Personal Information Protection Law, it is also necessary to follow Article 3 of the Provisions on Security Assessment of Internet Information Services with Public Opinion Attributes or Social Mobilization Capability, in which the provider of networked information services shall conduct its security assessment following these provisions and be responsible for the results of the evaluation.

In summary, when China's Internet platforms go live with ChatGPT products, they need to first apply for the related Internet information services value-added telecommunications business license, followed by filing the algorithm, and finally, a security assessment according to the three effective regulations and assessment provisions, all of which must be passed before the product can go live.

### 1.2.2 Progression requirements for business upgrades

(1) Network Culture Operation Licence. Driven by the development of Internet technology and media, Internet literature and arts have flourished, with the creation of online literature, online film and television, online variety shows, online competitions, online animation, online music, short videos, and other cultural and artistic categories bursting forth. Suppose platforms of ChatGPT products carry out business Internet literary and creative activities for profit, charging Internet customers or obtaining rights and interests utilizing network technology, advertising, title sponsorship, etc., and exhibiting Internet cultural products and service items on the subject. In that case, they must apply for a Network Culture Business Licence.

(2) Licence for online publishing services. Traditional books and audio-visual products existed in physical form and were distributed through publishers, newspapers, periodicals, and record companies and were copyrighted and protected by law. With the development of the Internet, most books and audio products no longer need to be distributed through other channels and can be uploaded directly to the platform. The carrier is transformed from physical to data and requires access to the copyright for online publishing. This is why developing a platform for Chat GPT products can only be done with an Internet Publishing Service Licence.

(3) Licence for information network transmission of audio-visual programs. Online audiovisual programs cover a more comprehensive range than online culture, including online dramas, movies, documentaries, literature, entertainment, science and technology, finance, sports, education, original audition programs, etc. Nowadays, the reception and output of Chat GPT are also developing from graphic to video, so it also needs a License for Audiovisual Programmes for Information Network Distribution to support its development. This license is challenging to apply for in China but is extremely valuable.

(4) Internet News and Information Service License. Users using Chat GPT will be given some news events to report or have exported, equivalent to providing news and information services on the Internet to the public. An Internet News and Information Service License should be obtained to carry out Internet news and information service activities.

(5) Data security assessment: ChatGPT is a significant language model, which can be output as natural human language because it relies on the rules of human language derived from analyzing large volumes of content. Because there is data interaction, there is an issue of data security here. When users use the product, they leave behind a lot of personal information, such as the actual name of the phone number they fill in, the IP address of the network when they access ChatGPT and their language habits. The data security risks associated with the activities of network operators need to be analyzed and assessed following the requirements of the Cyber security Act standards.

In summary, when upgrading the business of ChatGPT products on China's internet platforms, the corresponding qualification licenses and filings should be issued according to the expanded business type. The data security assessment should be conducted before the development can continue.

## 2. ChatGPT products operational risk prevention and control

ChatGPT products contain two main parts, one for the input side and the other for the output side. The input side consists of two types: the first is the interactive input from the user; the second is a crawler that actively crawls the web for various information, the primary input source. ChatGPT analyses the information on the input side to retrieve information. It uses crawlers to crawl the web for multiple information data and process the information in natural language and output it on the output side. For example, suppose a user asks a question about the history of Qin. In that case, ChatGPT uploads the user input to the back-end to retrieve information about the development history of Qin, related events, key figures, and allusions to the times for comprehensive analysis, forming a clear and complete vein and outputting the answer you need after processing in natural language. Without these essential prerequisites, it is impossible to meet users' needs. ChatGPT constantly cycles through the above process, storing the content of each retrieval and user feedback in a database for learning. We focus on the following two points: what are the problems and operational risks associated with obtaining information on the input side through the content crawled by the crawler? Is there any infringement or illegal information in the content output?

### 2.1 Inputs

The user's interactive input analyzes the input side from a risk perspective. It is a better way for users to supply relevant content to ChatGPT products to learn knowledge, especially in professional papers, so this input method is less risky. Based on the basic concept of Internet interconnection, interoperability, and openness, the Robert Agreement does not explicitly prohibit crawlers from crawling, so in principle, search engines can crawl data (e.g., text data, social media

data, question-and-answer website data, news website data, literature data, etc.) through crawlers to establish an ecology of coexistence and integration.

Although the above concept is perfect, it appears Internet platform traffic cannot realize the problem, which destroys the risk of Internet ecology. The conventional search engine does the information positioning for the Internet platform, and users visit the platform through the positioning of the search engine, bringing traffic to the latter. Any virtual industry needs to be supported by the entire industry. The actual sector increases its sales volume through many advertisements placed by the Internet platform, which then returns profits to the Internet platform and the search engine, thus forming a co-existing and co-integrating industrial chain. However, the emergence of ChatGPT products has broken this industry chain, as it directly accesses the data of the Internet platform through crawlers. Although it enters the website through the search engine, it does not lead the user to it. So the result is that the search engine is unable to direct real traffic, and the physical industry is unable to increase sales through advertising on the internet platform, which ultimately leads to a reduction in the total funding of the chain, creating a situation where all three are lost.

In addition to the problems mentioned above, the risk of the Internet platform breaking the Internet's concept of interconnection, interoperability, and openness, blocking its information and building its information silos, ultimately destroying the Internet ecology, is that it cannot access the dividends of public traffic. For example, QQ and WeChat are two communication products from the same company and are closely linked. As long as WeChat can use the program, there is no need for additional channels to interface with it, and it can be used directly in QQ and vice versa. However, the same cannot be said for WeChat and the short video platform. Suppose a short video user wants to share a video with a WeChat user. In that case, the short video user needs to download the video, or the WeChat user needs to open the corresponding short video app to watch it. An additional channel is required to connect the two platforms. This is the risk in the operation of ChatGPT products, which accelerates the formation of information silos, dramatically reduces the efficiency of Internet transmission, and defeats the purpose of Internet interoperability.

## 2.2 Outputs

The issues on the output side revolve around the content of its output.

The first is copyright infringement, and the patterns used by ChatGPT products must constitute infringement. This involves an analysis of the back end, which means that ChatGPT crawling for learning will, in principle, be combined with the back end and is not an isolated concept. If you analyze it in isolation, it is whether or not it affects your convertive use. For example, suppose a publisher includes a paper book in a book management system for online sale. In that case, it is not an infringement because the publisher owns the book's copyright. However, ChatGPT products are not copyrighted when they reference these books, so the content (text, audio, images, video) and the way they are output will largely determine whether they are copyright infringements, and this is where the commercial risk lies.

The second is unfair competition. The operation of any website or platform requires a large amount of capital investment for development. Many websites need to employ many creators to create content for the website and produce it over a long period to stabilize its growth. However, ChatGPT only needs to use crawlers to quickly capture its creative content at a meager cost, which constitutes unfair competition and alleges copyright infringement.

The third is the infringement of personal information. Based on the Personal Information Protection Act, information subjects can exercise the right to forget and request that the relevant platform delete information published on the internet about themselves that does not match, as well as personal information previously collected. ChatGPT products store data in a database without your knowledge, and it will not output relevant content without a keyword search, so it isn't easy to define whether this right is infringed [3].

## 3. Legal protection of content generated by ChatGPT products

Whether it is the developer of the ChatGPT product or the user who uses it as an aid to creativity, the process involves an input of labor. The premise of protection under our law is that the information has a labor value, whether material or moral loss. It requires the concept of reciprocal equivalence to be protected by the law. We will discuss these key issues: can the content generated by ChatGPT products, such as articles and audio-visual images, be legally protected, is the ChatGPT product itself legally protected, and does the content generated by the ChatGPT product belong to the copyright holder who uses it for creation?

Our law protects human beings' power, not artificial intelligence's. According to Article 2 of the Regulations for the Implementation of the Copyright Law, "works referred to in the Copyright Law refer to intellectual achievements in the fields of literature, art, and science that are original and can be reproduced in some tangible form." This regulation indicates that the object protected by copyright law is a work of originality created by a natural person through intellectual labor. Although artificial intelligence can simulate the intellectual activity of a human being to build a career and generate content similar to that produced by a natural person, the creation is dependent on the mechanical operation of an algorithm and, therefore, cannot be compared to the intellectual work of a human being. As a result, the content generated by ChatGPT products does not fall within the scope of legal protection.

ChatGPT is a product created by human intellectual labor and belongs to the category of computer software works. According to Article 3 of the Copyright Law of the People's Republic of China, "The works referred to in this Law refer to intellectual achievements in the fields of literature, art, and science that are original and can be expressed in a certain form." This includes computer software, so as long as the copyright owner or its assignee enjoys the various rights of software copyright, it also indicates that the ChatGPT product itself is protected by law.

The copyright owner uses ChatGPT products for the creation, which does not conform to the logic of the subject-object attribute, and the generated content is not the subjective will of the creator [4]. ChatGPT products are created by the software creator, then there is a clear subject and object, and the creator creates the object ChatGPT product for the issue. The ChatGPT creation process then considers it the matter of creation, which means that ChatGPT has the attributes of subject and object, forming a dichotomy, thus making it impossible to judge whether it complies with laws and regulations. In addition, according to the basic view of the copyright system, originality should reflect the author's personality and the expression of the author's thoughts and feelings. From the perspective of the creation process of ChatGPT, the user of ChatGPT only gives instructions and does not participate in the formation of the generated content, and the user does not necessarily know what the final generated content is, so the content generated by ChatGPT is not an expression of the thoughts and feelings of the creator or the user. In general, it is difficult to deal with the issues related to the content generated by ChatGPT products in the current legal system.

The birth of ChatGPT and its initial applications show that the large model of natural language already has some of the characteristics of Artificial Intelligence Generated Content (AIGC) oriented General Artificial Intelligence (AGI) and has the potential for a wide range of applications in many industry sectors, and has the potential to become one of the most outstanding achievements in computing and a significant force in technology as the technology continues to grow and is widely used by society. A considerable significance in the field of technology. There is still a long way to go before ChatGPT products can be created in China to go live, but we also need to maintain a cautious and objective attitude toward the future to face it.

#### REFERENCES

1. Cong Lixian & Li Yonglin. (2023). Copyright Risk of Chatbot-Generated Content and Its Governance - A View of ChatGPT's Application Scenario. *China Publishing Journal*, (5):11-16.
2. Zhao Liguang. (2023). Has ChatGPT knocked on the door of General AI yet? *Chinese Journal of Science*.2-22.
3. Zhu Guanghui & Wang Xiwen. (2023). ChatGPT's operation mode, key technologies, and future picture. *Journal of Xinjiang Normal University (Philosophy and Social Sciences)*, (4):113-122.
4. Jiao Heping(2022). Copyright Risks of Data Acquisition and Utilization in Artificial Intelligence Creation and the Path to its Resolution," *Contemporary Law Review*, (4):128-140.