Scientific Research Publishing

Beijing Law Review, 2025, 16(3), 1503-1519
https://www.scirp.org/journal/blr
ISSN Online: 2159-4635
ISSN Print: 2159-4627

Generative Artificial Intelligence Data Risks and Governance Pathways

Bo Jiang[1*], Xiongbiao Ye[2]

[1]School of International Law, China University of Political Science and Law, Beijing, China
[2]Law School, Central China Normal University, Wuhan, China
Email: *ruclawxiaobo@163.com

How to cite this paper: Jiang, B., & Ye, X. B. (2025). Generative Artificial Intelligence Data Risks and Governance Pathways. *Beijing Law Review, 16,* 1503-1519.
https://doi.org/10.4236/blr.2025.163075

Received: January 4, 2025
Accepted: August 8, 2025
Published: August 11, 2025

Abstract

Generative AI, exemplified by ChatGPT, offers societal benefits while posing challenges to data governance. Addressing data risks is vital for its healthy development. This paper examines the technical framework and pre-training data system of generative AI, identifying risks such as personal information, data property, and national security. These risks amplify traditional network data concerns. To mitigate them, the study suggests categorizing risks and applying relevant legal frameworks, including China's Personal Information Protection Law and Cybersecurity Law. Acknowledging the limitations of existing laws, it advocates for supplementary ethics rules, stricter penalties for violations, and an enhanced sanctions system to strengthen data governance in generative AI.

Keywords

Generative Artificial Intelligence, Technological Ethics, Data Risks

1. Introduction

Generative AI, epitomized by ChatGPT, represents a milestone in artificial intelligence. Its powerful content generation and interactive capabilities have rapidly attracted millions of users worldwide and prompted global tech companies to enter the generative AI race, developing their own products. Currently, OpenAI, the U.S.-based AI research lab behind ChatGPT, has enhanced training parameters and optimized algorithms to introduce GPT-4, a multimodal pre-trained large model, which combines content generation with image recognition, frontend design, and other functionalities. GPT-4 is regarded as a transformative technology, capable of revolutionizing human thought and creativity, and is expected to replace humans in certain tasks (Zhu & Wang, 2023). Google has also launched its

conversational AI system, Bard, while Baidu and Huawei have developed Chinese versions of generative AI products, "Wenxin Yiyan" and "Pangu α", respectively, based on Chinese data. These large-scale language models, developed using "big data, big models, and big computing power," have had an impressive performance in areas such as literature summarization, academic paper drafting, and computer code generation. Scholars predict that the application of large language models will extend to experimental design, manuscript writing, and peer review in the future. However, while generative AI is sparking a technological revolution, it also brings social risks and governance challenges, such as mass unemployment, information pollution, and ideological risks (Shang, 2023). Although these issues have been widely discussed in academia, the data risks inherent in generative AI have not yet received sufficient attention.

In fact, with the rise of computer networks and the iteration of digital technologies, data has gradually become the "new oil" driving technological progress and economic development, while privacy protection and data security have become global challenges for governance. Compared with traditional network products, next-generation generative AIs like ChatGPT rely on even more extensive data aggregation, complex algorithms, and pose greater potential data risks, necessitating further research and legislative response. In this context, the National Cyberspace Administration of China and other government bodies jointly formulated the Interim Measures for the Administration of Generative AI Services (hereafter referred to as the "Interim Measures") to regulate the research, development, and application of generative AI and mitigate data risks. However, the Interim Measures primarily serve as an emergency response to generative AI risks, with a regulatory framework that is more advisory and referral in nature. The provisions regarding data processing behaviors, legal liabilities, and protection of user information interests remain unclear, offering little guidance for the industry's development. Therefore, this paper focuses on the data security issues of generative AI, using ChatGPT as a model to analyze the technical framework of large language models and explore potential data governance frameworks based on existing regulations and related theories. Specifically, the paper addresses the following issues: first, the technical framework and corpus system of generative AI, using ChatGPT as a case study to analyze the technical principles and data system of large language models; second, the analysis of data risks in generative AI, examining the role of data in the development and application of large language models and the associated potential risks; and third, governance pathways for generative AI data risks, exploring regulatory strategies for managing these risks.

2. Technical Framework and Corpus System of Generative AI

Since the concept of AI was introduced in the last century, it has received significant attention, with scientists and engineers aiming to automate human intelligence, assisting or replacing human labor. With the development of electronics and the accumulation of (Vaswani, 2017) massive amounts of data, general AI has

advanced rapidly. The 2016 victory of AlphaGo over world champion Lee Sedol further solidified the confidence of the industry in AI, leading to the release of more mature AI products, with ChatGPT being a prime example. In the quest for general AI, pre-trained foundation models (PFMs) have been widely recognized as a feasible tool. Currently, various models in the field of natural language processing (NLP) have evolved from large Transformer models. OpenAI, founded in 2015, has been dedicated to the research and development of large language models (LLMs). Thanks to advances in neural networks and deep learning technologies, as well as the exponential growth of available network data for training, OpenAI has released several generative pre-trained models (GPT). The GPT-3 model, released in June 2020, used 45TB of text data for training, with 175 billion parameters, significantly enhancing its transfer learning capabilities (Brown et al., 2020). ChatGPT is a fine-tuned version of GPT-3.5, trained on large-scale pre-training data, allowing it to learn human language patterns and rules from the data without supervision, showcasing superior context understanding and content generation capabilities. It is regarded as a marker of the maturation of large language models and the entry of AI into the AI 2.0 era. Six months after the release of ChatGPT, OpenAI fixed several errors and improved the model's theoretical foundation, launching the multimodal model GPT-4, which is now in commercial use.

2.1. Training Paradigm of ChatGPT

The development of ChatGPT can be simplified into two stages: the first stage involves pre-training the algorithm and building the model, during which the goal is to enable the model to learn highly generalized parameters. The second stage involves fine-tuning the algorithm and refining the model, where the knowledge gained from the first stage is applied to downstream tasks. At this stage, the model only needs to fine-tune the already learned parameters to achieve high-efficiency transfer and adapt to specific tasks (Zhao et al., 2023). It is important to note that, although large language models can extract information from vast network data, convert human information into machine code, and mimic human thinking patterns and language habits to some extent, the content they generate still lacks preference control. Additionally, due to the varied quality of input data, the accuracy, reliability, and reasonableness of the content generated by ChatGPT are not guaranteed. For example, OpenAI's reports show that GPT-4 still risks generating harmful content, defective code, or inaccurate information[1]. Therefore, in the training process of ChatGPT, both labeled and unlabeled training data are mixed, with automatic machine learning, human supervision, and algorithmic fine-tuning all being simultaneously employed.

2.2. Corpus System of ChatGPT

The analysis of ChatGPT's technical framework shows that the intelligence of gen-

[1]See GPT-4 Technical Report.

erative AI primarily depends on two aspects: the algorithm and the training data. Much of the current research on ChatGPT focuses on regulating the automation algorithms, such as the requirement in Article 4, clause 2 of the Interim Measures that AI algorithms should avoid discrimination, while paying less attention to data risks. Compared with traditional intelligent products, generative AI relies more heavily on data, which manifests in two key ways: first, the volume of training data is much larger. Pre-trained language models are becoming the foundation for general AI, and their vast application prospects are prompting research institutions to invest heavily in product design and innovation, with the amount of training data increasing exponentially. For example, GPT-3 used 750 GB of training data. Second, the role of training data is more significant. The development of generative AI requires close coordination between data and algorithms, with algorithms running, testing, learning, and refining the pre-training data to ensure the model outputs the desired results. Data thus becomes the "nourishment" and experience source for AI's growth.

Human knowledge and language expression are derived from both direct and indirect experiences. However, generative AI is entirely reliant on data from the human knowledge base, i.e., indirect experience. For ChatGPT, the data used for training comes from a wide range of sources, covering various fields and industries of human life. This broad corpus system ensures that the model can learn a wide range of knowledge and meet the differentiated needs of users (Shen & Zhu, 2023). According to publicly available information, ChatGPT's training data includes publicly available information from the internet, data acquired from third parties, and information provided by users and human trainers[2]. Data from the internet can be categorized into several types: 1) data from Wikipedia, an online multilingual encyclopedia maintained by volunteers and covering more than 4 billion words; 2) books and journals, including both fiction and non-fiction, which help the model improve its narrative and response capabilities; 3) the Common Crawl dataset, a large collection of raw webpages, metadata, and text excerpts collected by web crawlers; 4) Reddit links, which help the model understand human social behavior and psychology; and 5) other data, including forum conversations and video subtitles. The following table (**Table 1**) shows the training data volume of several major models[3].

Table 1. The pre-training data for some large models (Unit: GB).

Types	Wikipedia	Books	Journals	Reddit Links	Common Crawl
GPT-3	11	21	101	50	570
Megatron-11B	11	5	/	38	107
MT-NLG	6	118	77	63	983
Gopher	12	2100	164	3450	4823

[2] See How ChatGPT and our foundation models are developed. https://help.openai.com/en/articles/7842364-how-chatgpt-and-our-foundation-models-are-developed.

[3] See Alan D. Thompson (2023). What's in My AI? https://s10251.pcdn.co/pdf/2022-Alan-D-Thompson-Whats-in-my-AI-Rev-0.pdf.

3. Analysis of Data Risks in Generative AI

In the digital economy era, network products are closely tied to data, and generative AI is no exception, directly relying on network data. At present, generative AI is rapidly developing, yet it has not yet drawn significant regulatory attention. The network data collected through techniques like web scraping does not inherently possess legitimacy or legality, particularly in an era where there is widespread dispute over data property and increasingly stringent personal information protection policies. As such, generative AI, which relies on massive datasets, faces legal risks concerning data usage in its development process. Furthermore, under the backdrop of globalized networks and enterprises, the applications of generative AI products launched by tech giants have transcended geographical limitations, and with the expansion of business operations and industry layouts, issues surrounding cross-border data flow and national security risks have become increasingly pronounced.

3.1. Data Property Risks

The proliferation of the internet has digitalized the world, with both the physical world and human activities being represented as data on the network. Massive volumes of network data have become the essential "nourishment" for the development of generative AI. The previous discussion of generative AI's technical framework has clearly demonstrated the deep reliance of this technology on network data. According to OpenAI's reports, the pre-training data used by ChatGPT primarily comes from publicly accessible online information. However, due to concerns such as trade secrets and corporate competition, the report does not provide detailed information on the construction of the training datasets or specific methods used. In the context of strengthening network data governance and safeguarding data rights, the legitimacy of the pre-training data used by generative AI is not without controversy. Directly obtaining publicly available data still faces ownership disputes. Article 7 of the Interim Measures holds the generative AI service providers responsible for the legality of data sources, ensuring compliance with laws like the Cybersecurity Law of the People's Republic of China (CLPRC) and prohibiting the infringement of others' intellectual property. This provision serves as a foundational compliance framework for the generation of AI training data and has become an important regulatory entry point for addressing data risks. Before the global attention that ChatGPT attracted, the training and research of generative AI were largely in a "gray area", with few existing legal norms addressing this field. Moreover, theories and discussions on such highly concentrated data-driven technologies, which involve diverse risks and significant conflicts of interest, have been relatively underdeveloped. A closer examination of the development process reveals data property risks and doubts regarding the legitimacy of pre-training data.

Determining the ownership of network data and safeguarding the interests of enterprises in the digital economy is a significant challenge. In China, efforts are

being made to explore feasible governance solutions for network data. The government policies have explicitly stated the need to "establish a data property system that protects rights and ensures compliance,"[4] aiming to balance the efficient flow of the data market with the protection of the legal rights of all participants. Theoretical perspectives on data property include the theory of data control, data property rights, data intellectual property, and data as a new type of property (Mei, 2019a). These theories attempt to define the legal status and attributes of data, incorporating increasing instances of infringement, such as unauthorized web scraping, into the legal framework to protect the interests of data processors. In the case of generative AI, one of the key challenges lies in determining whether the data acquisition behaviors of developers—such as web scraping—are legal and compliant. Under current law, the CLPRC (Article 27) and the Data Security Law of People's Republic of China (DSL) (Article 32) explicitly prohibit the illegal acquisition of network data, while Articles 7 of the Data Security Law and 127 of the Civil Code safeguard data rights. These provisions establish the basic legal framework that generative AI developers must follow when obtaining pre-training data. From existing judicial decisions, Chinese courts have already started addressing the unauthorized or overstepping data acquisition behaviors, categorizing them under violations of trade secrets or unfair competition (Wang & Ye, 2023). Even in the United States, network data is not considered universally open or public, and unauthorized or excessive data retrieval falls under the violations of the "Computer Fraud and Abuse Act" (Xu, 2021). Therefore, the claim by OpenAI that its training data sources are publicly available and legally compliant needs further scrutiny. Other generative AI development projects may also be at risk of infringing upon data property rights.

Moreover, the training data used by generative AI includes books, journals, and images, all of which are protected by copyright. Even if these materials exist in binary code form on the internet, their unauthorized use still constitutes a violation of copyright. For instance, books and journal articles account for about 15% of ChatGPT's training data, which helps the model develop human language expressions and scientific theories. This part of the data forms the basis for ChatGPT's academic functions, such as literature review and paper writing, but whether this data acquisition was legitimate remains unclear. For example, OpenAI has been sued by multiple newspapers, including The New York Times, The Intercept, and Raw Story, for using their copyrighted works without authorization to train its large-scale models.

3.2. Personal Information Risks

Personal information protection is a core issue in network governance. The enactment of Personal Information Protection Law of the People's Republic of China

[4]See Opinions of the Central Committee of the Communist Party of China and the State Council on Building a Data Foundation System to Better Leverage the Role of Data as a Factor of Production. https://www.gov.cn/zhengce/2022-12/19/content_5732695.htm.

(PIPL) and other relevant regulations has, to some extent, provided a legal basis for addressing personal information misuse. However, whenever new network products are launched, society is concerned with whether they pose personal information infringement risks, and generative AI is no exception. Although the pre-training corpus system of generative AI models does not directly reveal personal information, this does not mean that generative AI development and applications are entirely free of involvement with personal data. For example, OpenAI's official website explicitly states that it collects and uses user personal information for the development of ChatGPT and other language models. However, in 2023, sixteen individuals filed anonymous lawsuits against OpenAI, alleging that the company used and leaked their personal privacy data without permission. The plaintiffs are seeking damages of up to $3 billion. While data property risks mainly concern the development phase of large language models, personal information risks span the entire lifecycle of generative AI, including both the development and application phases.

First, Personal Information Issues in the Development and Optimization Phase Low-intelligence chat AI typically relies on simple dialogue scenarios and common conversational templates, lacking the ability to adapt to diverse conversations, thus struggling to provide a genuinely engaging and realistic communication experience. On the other hand, large language models acquire near-human conversational capabilities by learning from the vast amount of information available on the internet. The development and optimization of generative AI requires collecting a wide range of data from network users to train the model's understanding of human communication, consumption, entertainment, and social interaction norms. These data not only include common identity information such as names and phone numbers but also involve personal behavioral data, chat records, and social interaction data.[5] With the ongoing tightening of personal data protection and privacy policies, it becomes necessary to evaluate whether the collection of network user data during the development and optimization of generative AI complies with relevant legal requirements. Under the current legal framework, data processors must obtain user consent for the collection, processing, and analysis of personal information. Moreover, the rules surrounding the "identifiability" of personal data have broadened the scope of what constitutes personal information, leading to stricter data processing regulations. Accordingly, there are legitimate concerns about whether generative AI developers have obtained user consent for the collection of personal information and whether their data handling practices comply with relevant laws.

Second, Personal Information Risks in the Application Phase. GPT-4 has already started its commercial application, providing a useful lens through which to examine personal information risks in the application phase of generative AI. The risks of personal information in the development and optimization phase stem from the legality of data collection from the internet, while the risks in the appli-

[5]See OpenAI Privacy policy. https://openai.com/policies/privacy-policy.

cation phase relate to how personal information is processed and used. For example, ChatGPT's user privacy agreement outlines the types of data collected, how personal information is used, information disclosure, user rights, and relevant legal policies. These provisions reveal certain problems in the data processing practices of generative AI. Firstly, the existing personal data protection model gives individuals absolute control over their personal data, requiring explicit authorization for any data processing actions. Any processing beyond the scope of consent would constitute an infringement. If data processors obtain broad consent from users through vague user agreements, this could result in illegal data collection agreements, and non-compliant data processing or sharing practices. Secondly, as data protection policies continue to tighten, recent data governance standards are centered on the principle of necessity (Jin & Zai, 2023), which requires network service providers to reduce data collection from users. However, ChatGPT collects a broad array of user data, including account information, user content, communication data, social data, log data, usage data, and device information. This extensive data collection conflicts with the current logic of data governance frameworks, creating tension between the broad scope of data collection and existing regulatory principles. Finally, as a human-AI communication system, generative AI like ChatGPT collects a wide range of user information during interactions. The integration of this data with algorithms could lead to the analysis of user preferences, and in extreme cases, result in discriminatory outcomes. Although OpenAI's user agreement states that it does not retain user conversation data for model training or preference analysis, the implementation of this commitment still requires legal oversight. This issue will also be a challenge for domestically developed generative AI models in China.

3.3. National Security Risks

The global interconnectedness of the internet facilitates the cross-border flow of data, and with this flow comes the globalization of information. This inevitably creates national security concerns regarding the control and management of data. In the current international landscape, where competition among nations is increasingly driven by technology and information, data outflows due to corporate and technological expansion raise critical issues about how to regulate the cross-border movement of data and ensure national security. The development of generative AI further amplifies the risks to national security by challenging the goal of "ensuring the security and controllability of key domain information systems and data." These challenges manifest primarily in the following areas:

First, the threat to Data Sovereignty. The 2003 Geneva Principles Declaration proposed that "the decision-making power regarding public policy issues related to the internet rests with national sovereignty." In 2013, the United Nations General Assembly further confirmed that "international norms and principles related to national sovereignty and territorial sovereignty apply to the use of information and communication technologies." Since then, the principle of network sover-

eignty has been widely recognized (Cai, 2018). Data sovereignty, as a subset of network sovereignty, refers to a nation's ownership, control, jurisdiction, and usage rights over its data and the data of its citizens (Meili & Yu, 2022). This can be categorized into two main aspects: "data localization" and "data retention" (De Jong-Chen, 2015). In practice, there are three models for "data localization": "local backup mode," "accessible local storage mode," and "absolute local storage mode" (Bu, 2021). China, while insisting on the principle of local data storage, allows exceptions for cross-border data transfer under specific security assessments, as outlined in the Measures for the Security Assessment of Outbound Data Transfer. It allows for cross-border data transfers after security evaluations but still uphold the positions of protecting data security, public interests, and the legitimate rights and interests of individuals and organizations. However, the development and application of generative AI could disrupt the already fragile system of data flow regulation. This is because the development of generative AI requires access to data from different languages, cultures, and countries. This necessitates the collection of data from various countries, which may conflict with the principles of data sovereignty. Furthermore, generative AI products are designed for global application, meaning they will collect data from multiple countries, with data storage, processing, and circulation potentially occurring in the developer's home country or in foreign nations. This inherently requires cross-border data flow, thus conflicting with Data Sovereignty[6] protection principles. The rise of cloud storage further complicates this issue, as it decouples the collection and storage locations of data, and the uncertain physical locations of cloud storage servers give rise to data property challenges, intensifying the debate on national data sovereignty (Guo, 2022).

Second, Impact on National Security Systems. While the internet has removed physical borders, it has not eliminated the political and economic rivalries between nations. The competition in areas like technology, finance, and culture continues to intensify, and data and information have become critical components of national security. Generative AI, beyond merely sparking debates over data sovereignty, poses more concrete threats in terms of cross-border data flow and the resulting national security risks. This was evident when a foreign organization tampered with DNS of routers at a northwest Chinese energy base to block data transmission in January 2025, aiming to cause a regional blackout. These incidents show AI-powered cyberattacks have become strategic real-world threats. Cur-

[6]According to Professor Guo Shuo from China University of Political Science and Law in his article *Maintaining Data Sovereignty in Cloud Storage: Taking the Regulation of "Long Arm Jurisdiction" by Blocking Statutes as an Example, in an era when information technology was underdeveloped, data did not yet exist and did not play a significant role. With the development of information and network technology, economic and trade exchanges and judicial assistance among countries have also promoted data flow. The importance of data for industrial development and even national security has been increasingly recognized. Just like the Internet, data is inherently attached to computer hardware devices. These devices must have a placement location, and the data stored in devices at a certain location is often generated by local people, events, and things. Therefore, data naturally belongs to the country where the relevant devices are located - this is how the concept of "data sovereignty" emerged.

rently, countries' assessments of the risks posed by cross-border data transfers are no longer limited to violations of personal information rights or privacy. Instead, such concerns have escalated to the level of national security. As data intermingles with sectors such as technology, finance, politics, and military, the extension of cyberspace leads to an increasing overlap between the virtual and the physical realms (Yan, 2022). The cross-border movement of data is emerging as a new battleground that tests the ability of nations to safeguard their national security. For example, generative AI systems like ChatGPT, which are trained on English-language data and shaped by Western discourse, may have a profound impact on China's ideological landscape, potentially jeopardizing national cultural security (Zhi, 2023). Moreover, the global proliferation of generative AI products controlled by foreign tech giants introduces the risk of data leakage, as the data may be stored, processed, and circulated outside of China, posing significant threats to the country's overall security system.

4. Governance Pathways for Data Risks in Generative AI

As a pioneer of the digital revolution, ChatGPT demonstrates the immense creative power that results from the aggregation of algorithms and data. However, it also challenges the existing fragmented network governance system. In addressing the data risks inherent in generative AI, there is an urgent need to integrate and apply existing data governance regulations, while also refining technological ethics rules and strengthening enforcement measures.

4.1. Integration and Correspondence of Data Governance Norms

The data risks triggered by generative AI include personal information risks, data property risks, and national security risks, among others. However, these risks are not unique to generative AI: the spread of the internet has long been accompanied by the risk of personal information leakage, the development of big data technology and the commercialization of data have led to issues of data infringement, and national security risks in data are an inevitable result of the global network and cross-border data flow. The data risks within generative AI are essentially the manifestation of existing network risks and societal issues in new technologies. What distinguishes them is that this technology aggregates various data risks and significantly amplifies their severity. Therefore, the call for entirely new legislation specifically targeting generative AI may be an overreaction. A more effective approach is to identify and decompose the data risk types in generative AI, matching each type of risk with existing legal rules. In this regard, the Interim Measures make an effort to align these risks with existing laws.

Although the Interim Measures represent a specialized form of legislation for generative AI, the regulations are more of a collection of referral provisions. The Interim Measures decompose the risks within generative AI and refer each risk type to existing legal norms, relying on current legislation in the network field to regulate data risks in generative AI. For example, Article 7 of the Interim Measures

specifies that data processing activities related to training data must comply with the CLPRC and other relevant laws and regulations. Article 12 requires providers to label generated content according to the Provisions on the Administration of Deep Synthesis of Internet-Based Information Services, and Article 17 mandates compliance with the Provisions on the Administration of Algorithm-generated Recommendations for Internet Information Services. Concerning data risks such as data property risks, personal information risks, and national security risks in generative AI research and application, the network legislation introduced in recent years can to some extent address these concerns: 1) Data Property Risks: The materials used by generative AI, such as books, journals, and images, are not devoid of property. The use of such materials without authorization constitutes an infringement of intellectual property rights (Chen, 2023). These issues are entirely within the scope of existing regulations such as the Copyright Law and do not require new legislation. The debate about whether generative AI should be held liable for infringing copyright or who owns the intellectual property rights to AI-generated works involves philosophical questions of legal subjectivity. These questions are not the distribution of social risks, and other jurisprudential factors (Mei, 2019b). Thus, it is premature to delve into these questions at this stage. Regarding data scraping by generative AI developers, which often involves techniques like web scraping, this undoubtedly infringes upon the enterprise data rights. This has been a major point of discussion in academia and remains an area where legislation is lacking. Although the principle of data rights protection in Article 127 of the Civil Code provides some guidance, it does not offer sufficient practical direction. The Interim Measures' Article 7 requires that generative AI use legally sourced data, but it does not define what constitutes "legally sourced" data. While the generally accepted models of trade secrets and anti-unfair competition law offer some protection for business data rights (Xu, 2018), more explicit legislation is still urgently needed. 2) Personal Information Risks. The use and protection of personal information are central to the digital economy. Generative AI amplifies the risks to personal information, but these risks can still be addressed under existing regulations such as the Civil Code and the Personal Information Protection Law. Specifically, the Civil Code establishes personal information as a fundamental personality right and provides the legal basis for the private protection of personal information. The Personal Information Protection Law outlines the conduct guidelines for data processors, the right of individuals to be informed and to consent, as well as their rights to deletion, correction, and portability, which are sufficient to address the risks related to personal information in generative AI applications. 3) National Security Risks. The National Security Law provides a macro framework for overall national security, encompassing network information and data security, thus offering a legal foundation for addressing national security issues arising from the development and application of generative AI. The CLPRC, as the core law for network security, emphasizes the controllability of network infrastructure and information technology security, regulating internet infrastructure

and data flow order, and setting basic requirements for generative AI providers to conduct activities in compliance with laws, improve infrastructure, and enhance data security.

While existing laws offer a degree of legal basis for managing data risks in generative AI, there are still gaps. Currently, China is in the process of drafting an AI Law, which could serve as an opportunity to reorganize and integrate data governance issues within the AI domain and establish a comprehensive data governance framework. It is important to recognize that the development of AI technologies also challenges the existing data governance and intellectual property protection rules. For example, the Personal Data Protection Law, which protects personal information, may limit AI companies' use of data, and the strict copyright protection regime may not be fully aligned with AI's need for electronic works during training. As a result, legislators may need to rethink the balance between individual and public interests and create a legal system that is more conducive to AI development.

4.2. Supplementing and Coordinating Technological Ethics Rules

In the face of the challenges and risks posed by new technologies, "our society needs to carefully consider and establish a set of standards and policies to guide the development of these technologies" (Hendler & Muivehill, 2016: p. 11). The creation and updating of such standards and policies is an inevitable requirement as society transitions from the industrial era to the digital era. This includes not only legal rules but also ethical guidelines and industry standards. Legal norms, as the cornerstone of social governance, are essential in regulating data risks in generative AI. Apart from the Interim Measures, China is planning a specialized AI Law, while the U.S. Congress has proposed bills like the Generating Artificial Intelligence Network Security Act and the National Artificial Intelligence Initiative Act, and the European Union has enacted the first AI act, the Artificial Intelligence Act. However, in the rapidly evolving digital age, where risks diversify and technologies change rapidly, a legal-centric rule system is increasingly showing signs of fatigue. The concept and rule system for social governance need to be refined. As data and algorithms continue to integrate, and the influence of code grows stronger, ensuring that "technology serves humanity" has become a key topic in industry discussions. At its core, this involves the ethical values and professional integrity guiding technological development. To a large extent, current laws are struggling to reach deeply into complex algorithmic "black boxes" such as large language models. The shift from punishing harmful outcomes to regulating the compliance of development processes will require assistance from technological ethics rules.

Technological development should serve the common welfare of humanity, not become a tool for capital and power to dominate human progress (Ma, 2016). As the digitalization of the physical world, human life, and economic production activities deepens, the digital world and the physical world are increasingly merging.

The material foundation, social relationships, and modes of communication on which modern law relies have undergone significant changes, making existing regulatory systems inadequate to cope with the rapid pace of technological innovation. As code becomes the core rule of the internet and technology gradually takes control of the networked world, reshaping technological ethics and correcting "code justice" becomes an essential strategy for addressing the data risks associated with generative AI. In 2016, the IEEE (Institute of Electrical and Electronics Engineers) published the Ethical Design: A Vision to Maximize Human Well-being with AI and Autonomous Systems (AI/AS) (First Edition), encouraging researchers to prioritize ethical issues during the development of AI. More recently, policies from various countries have further refined the goals, principles, and rules of technological ethics. For instance, the EU's Trustworthy AI Ethics Guidelines released in 2019 defined the trustworthiness of AI in three dimensions: "lawful, ethical, and robust." It set out fundamental ethical principles that AI product development should adhere to, including "respect for human autonomy," "prevention of harm," "fairness," and "explainability." It also proposed seven aspects that trustworthy AI must meet, such as "human agency and oversight," "robustness and security," "privacy and data governance," "transparency," "diversity, non-discrimination, and fairness," "social and environmental well-being," and "accountability" (European Commission, 2019).

In 2022, China's Opinions on Strengthening the Governance of Technological Ethics emphasized five ethical principles: "enhancing human well-being," "respecting life rights," "adhering to fairness and justice," "reasonable risk control," and "maintaining openness and transparency." Similarly, the Principles of Governance for New Generation Artificial Intelligence, published in 2019, focused on "ensuring the safety, reliability, and controllability of AI," and proposed several technological ethics principles, including "respect for privacy," "fairness and justice," and "safety and controllability," to guide AI product development and industry progress. Technological ethics rules, which impose moral constraints on the developers of generative AI and its algorithms, fill in some of the gaps in legal oversight and help control risks at the source. In addressing the data risks in generative AI discussed in this paper, data property risks fall under the principle of "fairness and justice," personal information risks are covered by the principle of "respect for privacy," and national security risks fall under the principle of "safety and controllability." The Interim Measures have already incorporated technological ethics into their regulatory framework, for instance, Article 4, which states that generative AI development should respect "social morality and ethics."

However, it should be noted that although the Interim Measures and other policy regulations are increasingly strengthening the role of technological ethics in AI governance, the challenge remains to transform these ethical constraints into legal oversight. The critical issue for the future will be how to translate technological ethics into measurable governance content that can be enforced through law. The EU Trustworthy AI Ethics Guidelines, for instance, not only define ethical

principles but also provide evaluation processes for various stages of AI development, such as data collection, initial design, system development, and model training, attempting to transform technological ethics into operational standards for legal regulation and business compliance (Zhao, 2023). This approach is worth emulating. In practice, the concretization and proceduralization of technological ethics is a necessary path to making ethical standards externally visible. The design of trustworthy AI evaluation metrics offers both a reference standard for regulation and guidance for developers. It promotes internal adherence to technological ethics by providing external constraints. When it comes to data security, the evaluation of generative AI should include several key aspects:

1) **Data Source:** Developers must disclose the sources of their training data to ensure its legality.

2) **Data Use:** Developers should clarify how they use data, including training data and user data, to avoid infringing on user rights, such as using data for "user profiling."

3) **Data Storage:** Providers must ensure that data storage is secure and does not pose a risk to users or national security.

4) **Data Outflow:** If generative AI services involve cross-border data flow, developers should evaluate the security of this flow internally and report it according to national regulations to prevent national security risks.

4.3. Strengthening and Coordinating Multi-Layered Sanctions Mechanisms

The data risk mitigation mechanisms discussed above, involving legal rules and technological ethics, focus largely on preemptively determining behavioral standards to prevent risks. However, these rules must be accompanied by punitive mechanisms to ensure compliance. Therefore, further clarification and enhancement of penalties for violations is crucial to preventing data risks in generative AI. Article 21 of the Interim Measures outlines the responsibilities of generative AI service providers but still needs to be applied in conjunction with relevant provisions in the CLPRC, Data Security Law, and other laws to complete the liability framework for generative AI. In addition to legal responsibility, industry rules and punitive mechanisms within the internet industry also play a crucial role in regulating network behavior and development activities, such as platform account bans and the removal of products from e-commerce sites. These internal sanctions can be as effective as legal measures. Internet platforms are gradually evolving into quasi-governmental organizations, forming community standards, dispute resolution mechanisms, and punishment methods (Liu, 2021). Industry norms have thus become a form of "soft law" in network governance. In the context of generative AI data risk governance, flexible civilian sanctions and rigid legal responsibility systems are equally important and practice a new idea of shared societal governance.

Legal Sanctions: The network data governance system is a collection of norms where private law protects private interests in data, while public law (such as criminal law and administrative law) aims to maintain network order for public purposes. Under private law, the Civil Code has clarified the fundamental principles of protecting personal information and business data rights, but the issue of data infringement liability is still unclear. Key aspects like damage recognition, attribution principles, and compensation standards for business data need clarification through legislation. In public law, various illegal penalties are stipulated, including orders for correction, business license revocation, fines, and criminal penalties. However, the punishment intensity remains relatively weak, and the match between illegal behavior and legal responsibility is insufficient, reducing the deterrent effect. To effectively resolve generative AI data risks and ensure stable industry development, future legislation needs to clarify legal responsibility and strengthen penalties.

Civil Soft Law Sanctions: Soft law refers to rules or practices not enforced by state coercive power but still exerting public regulatory influence (Ma, 2016). The rise of the internet has exacerbated social risks, marking the entry of what sociologist Ulrich Beck calls a "risk society". Given the heightened uncertainty and subjectivity of human-made risks (Hu, 2018), social resources and order maintenance can no longer be entirely reliant on national laws. The delayed response and high cost of hard law often lead to governance failures. Thus, soft law, including industry customs, voluntary guidelines, and action protocols, plays a supplementary role in filling gaps where state law is lacking. Various industry standards have emerged in the networked world, which, although lacking the authority of state law, are effective in regulating behavior. For example, user evaluation mechanisms can severely affect a product or business's survival in the market. Apart from market-generated punitive mechanisms, semi-official industry organizations, such as industry committees and research associations, also play an essential role in regulating the conduct of their members and practitioners through industry rules. Soft law's increasing role in regulating AI risks is gaining international attention. For instance, IEEE and ISO are developing AI standards to guide its development. In addressing generative AI data risks, the enhancement of soft law functions and sanctions should include:

1) **Creation of Social Pressure Mechanisms**: For example, AI developers who misuse personal data or improperly acquire network data could face reputational damage through user reviews or industry reports, making it difficult for them to survive in the market.

2) **Formation of Industry Committees**: In addition to national law enforcement agencies, non-official organizations like investigative and executive committees should be established in the industry to discover and punish violators.

3) **Linking Soft Law with Hard Law**: Soft law in the AI sector should act as a proactive guide in cases where national laws are absent and gradually transform into national law to ensure compliance through state enforcement.

5. Conclusion

The development and application of large language models have brought significant benefits to society but also pose risks that challenge existing governance systems. The ability of generative AI to generate information depends on the volume and breadth of training data, and the concentrated use of data, such as personal information, books, journals, and publicly available data, amplifies data risks. Data governance rules that are oriented toward data types and risk areas need to be integrated. The Interim Measures, as an emergency response regulation for generative AI, attempt to parse the types of data risks involved and match them with existing legal norms like the Personal Information Protection Law and Data Security Law, effectively regulating the data risks of generative AI. However, the continued lag of legal frameworks in the face of network governance and technological regulation means that technological ethics rules must play a complementary role in preventing generative AI data risks, while the penalty system must be further refined to promote responsible technological development.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J. D., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., & Askell, A. (2020). Language Models Are Few-Shot Learners. *Advances in Neural Information Processing Systems, 33,* 1877-1901.

Bu. X. M. (2021). On the Reconsideration and System Construction of Data Localization Mode. *Information Studies: Theory & Application, 44,* 80-87, 79. http://doi.org/10.16353/j.cnki.1000-7490.2021.12.011

Cai. C. H. (2018). Geopolitics in the Cyberspace: A New Perspective on U.S.-China Relations. *The Journal of International Studies, 39,* 9-37.

Chen. Y. W. (2023). Beyond ChatGPT: Opportunities, Risks, and Challenges from Generative AI. *Journal of Shandong University (Philosophy and Social Sciences), No. 3,* 127-143. http://doi.org/10.19836/j.cnki.37-1100/c.2023.03.012

De Jong-Chen, J. (2015). Data Sovereignty, Cybersecurity, and Challenges for Globalization. *Georgetown Journal of International Affairs, 16,* 112.

European Commission (2019). *Ethics Guidelines for Trustworthy AI.* https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai

Guo. S. (2022). Data Sovereignty Maintenance from the Cloud Storage: Taking the Long Arm Jurisdiction Regulated by the Blocking Statute as an Example. *China Law Review No. 6,* 72-85.

Hendler, J., & Mulvehill, A. (2016). *Social Machines: The Coming Collision of Artificial Intelligence, Social Networking, and Humanity.* Apress.

Hu, W. (2018). The Soft Law Governance in Risk Society. *Studies in Dialectics of Nature, 34,* 33-37.

Jin, L. J., & Zai, Y. (2023). Review of the Principle of Minimum and Necessity in Personal Information Processing. *Journal of Beijing Institute of Technology (Social Sciences Edition), 25,* 140-150. http://doi.org/10.15918/j.jbitss1009-3370.2022.0363

Liu, H. (2021). Logic of Platform Power—The Recentralization Mechanism of Online Society. *Beijing Cultural Review, No. 1,* 31-39.

Ma, Z. S. (2016). Issues of 'Rule of Soft Law' and Its Countermeasures in 'Internet +' Age. *Modern Law Science, No. 5,* 49-56.

Mei, X. Y. (2019a). Between Sharing and Control: The Limitation of Private Law and the Building of Public Order in Data Proteciton. *Peking University Law Journal, 31,* 845-870. http://doi.org/10.3969/j.issn.1002-4875.2019.04.001

Mei, X. Y. (2019b). The Abstraction of Civil Subject and Its Impact on the System of Private Law. *Business and Economic Law Review, No. 1,* 86-99.

Meili, W., & Yu, C. (2022). Data Sovereignty in China: Legal Meaning and System Construction. *Journal of Intelligence, 41,* 92-98. http://doi.org/10.3969/j.issn.1002-1965.2022.06.014

Shang, J. G. (2023). On the Meta-Rules for Risk Governance of Generative Artificial Intelligence. *Oriental Law, No. 3,* 4-17. http://doi.org/10.3969/j.issn.1007-1466.2023.03.001

Shen, S. S., & Zhu, Z. T. (2023). ChatGPT-Like Products: Internal Mechanism and Its Impact on Learning Evaluation. *Chinese Journal of Distance Education, 43,* 8-15. http://doi.org/10.13541/j.cnki.chinade.20230223.001

Vaswani, A. (2017). *Attention Is All You Need.* arXiv: 1706.03762.

Wang, D., & Ye, X. B. (2023). Theoretical Expression and Protection Path of Enterprise Data Rights and Interests. *Information Studies: Theory & Application, 46,* 99-106. http://doi.org/10.16353/j.cnki.1000-7490.2023.04.013

Xu, K. (2021). The Legitimacy and the Boundary of Data Scraping. *China Legal Science, No. 2,* 166-188. http://doi.org/10.14111/j.cnki.zgfx.2021.02.008

Xu, S. (2018). An Intellectual Property Protection of Corporation's Data: Current Methods and Beyond. *Oriental Law, No. 5,* 55-62. http://doi.org/10.3969/j.issn.1007-1466.2018.05.006

Yan, H. (2022). Multidimensional Changes in Cross-Border Data Governance System and the Countermeasures. *Pacific Journal, 30,* 57-69. http://doi.org/10.14015/j.cnki.1004-8049.2022.04.005

Zhao, C. Y., Zhu, G. B., & Wang, J. Q. (2023). The Inspiration Brought by ChatGPT to LLM and the New Development Ideas of Multi-Modal Large Model. *Data Analysis and Knowledge Discovery, 7,* 26-35. http://doi.org/10.11925/infotech.2096-3467.2023.0216

Zhao, J. W. (2023). The Theoretical Misunderstanding and Path Transition in the Application Risk Governance of Generative Artificial Intelligence Technology. *Jingchu Law Review, No. 3,* 47-58.

Zhi, Z. F. (2023). Information Content Governance of Large Model of Generative Artificial Intelligence. *Tribune of Political Science and Law, 41,* 34-48. http://doi.org/10.3969/j.issn.1000-0208.2023.04.003

Zhu, G. H., & Wang, X. W. (2023). ChatGPT: Operation Mode, Key Technology and Future Prospects. *Journal of Xinjiang Normal University (Philosophy and Social Sciences), 44,* 113-122. http://doi.org/10.14100/j.cnki.65-1039/g4.20230217.001