The Impact of Large Language Models on Public Security Intelligence Work and Countermeasures Research

Ruiyan Zhao*, Tuo Shi
Department of Public Security Management, Beijing Police College, Beijing, China
**Corresponding Author.*

Abstract: During the process of promoting the integrated construction of "information, indicators, and operations," Chinese police department faced the real challenge of accelerating the development of new productive forces in public security. Coupled with the rapid development of Large Language Models and AIGC, driven by data and algorithms, public security intelligence work was constantly pressured to incorporate evolutionary technologies, internalizing the characteristics of Large Language Models tools into the core driving force for the improvement of institutional systems and business capabilities. Influenced by Large Language Models technology, public security intelligence work faced issues and risks such as data security, algorithm "black boxes," and lack of legal regulation. To avoid these risks, this paper, based on an in-depth analysis of the technical logic of Large Language Models and exploring their impact on public security intelligence work and application risks, focused on exploring governance strategies. It proposed constructing a governance system with public security organs as the main body leading and other diverse entities cooperating, creating a "human-machine collaboration" operation mode, to promote Large Language Models to better adapt to the application scenarios of public security intelligence work.

Keywords: Large Language Models; Public Security Intelligence Work; Artificial Intelligence; Human-machine Collaboration

1. Introduction

Since the emergence of ChatGPT, various types of Large Language Models have emerged one after another, continuously attracting global attention and usage. Various industries are seeking the landing applications of Large Language Models in various vertical industries through continuous exploration. As a typical representative in the field of Large Language Models, Large Language Models, with their powerful text processing, analysis, and understanding capabilities, can significantly improve the efficiency of public security intelligence collection, analysis, and writing. However, they may also bring risks and consequences such as privacy ethics and intelligence misinformation due to model illusions. Therefore, clarifying the technical characteristics and application scenarios of Large Language Models in the field of public security intelligence, carefully examining the potential risks they bring, and providing reliable realistic guidelines for the research of risk governance strategies.

2. Related Research

Currently, scholars' research on Large Language Models in the field of intelligence is mainly focused on three aspects:

Firstly, how Large Language Models promote the development and improvement of intelligence work. Cao Shujin, Cao Ruye discussed the impact of AIGC on intelligence studies from the perspectives of research questions, data sources, and research paradigms [1]. They deeply analyzed the changes brought about by AIGC to intelligence practice from four levels: comprehensive knowledge services, academic information services, decision intelligence services, and social information services. Yang Qian, et al. explored the application effects of Large Language Models in intelligence scenarios such as scientific reports, professional reports, and web searches through experiments, demonstrating that Large Language Models effectively assist in key aspects of intelligence analysis, improving quality and efficiency [2]. Zhao Bang, Cao Shujin conducted tests and control experiments on domestic ChatGLM-6B and foreign GPT-3.5-Turbo Large Language Models, evaluating the nine abilities of Large Language Models [3]. They concluded that fully tapping the potential of Large Language Models will improve the efficiency of intelligence work. Stavros Demetriadis and others explored the potential of GPT-3, proving that by implementing appropriate "extraction tasks," knowledge modeling can be conducted to enhance the efficiency of Large Language Models [4].

Secondly, the risks of Large Language Models to intelligence work and their governance paths. Yuan Zeng analyzed risks under the new risk of traditional social architecture and governance risks under the digital society, proposing the construction of a chain-based allocation mechanism for AIGC responsibility to improve the specific responsibility mechanism [5]. The responsibility is allocated based on the state and degree of participation of the subject when the large model causes damage, establishing a risk responsibility system. Zhang Yue, et al. proposed to build a human-machine collaboration relationship and establish a diversified negotiation mechanism to achieve knowledge co-construction by analyzing the characteristics of Large Language Models, their knowledge application capabilities, and their value in knowledge co-construction[6]. Zhang Linghan proposed that the governance of AIGC based on Large Language Models should be based on the current hierarchical classification governance structure of existing regulations [7]. It should combine technology, industry, and applications to establish an organic system and set specific rules to form a more influential deep synthetic governance legal system globally. Su Yu conducted a study on the legal risk governance of Large Language Models based on the paths of inclusive prudent regulation, hierarchical governance, deep governance, and agile governance [8]. Cheng Xuejun proposed to improve the governance path of the algorithm black box of large model platforms under the wave of AIGC from the aspects of algorithm ethics, transparency, and accountability by comparing and studying domestic and foreign algorithm black box governance experiences [9].

Thirdly, how to promote technological improvements of Large Language Models to better serve intelligence work. Hu Changping, et al. looked forward to the development of intelligence studies by combining multimodal information processing, knowledge ideology transformation, and interactive artificial intelligence [10]. Shujaat Mirza, et al. introduced the Global-Liar dataset for evaluating algorithm biases aimed at Large Language Models' information reliability [11]. The analysis found that the latest large model technology does not necessarily have higher accuracy, and the benefits of model upgrades are not equally beneficial to different regions. This emphasizes the necessity of cultural diversity and regional inclusiveness in model training and evaluation. Abhika Mishra, et al. proposed an automatic fine-grained illusion detection task to address the problem of model illusions [12]. They constructed a new evaluation benchmark, FAVA, and experiments showed that FAVA has strong advantages in fine-grained illusion detection, improving the authenticity of generated text, which can be used in future intelligence work.

Due to the particularity of public security work, in order to ensure the reliability and authenticity of public security intelligence acquisition in actual combat, it relies more on manual identification. At the same time, public security work has strict confidentiality requirements and higher requirements for intelligence products. Therefore, the application of Large Language Models in public security intelligence work requires "personalized" services, which are less discussed in the former research. As the core of AIGC, Large Language Models will always be at the forefront of digital technology development for a certain period in the future. They are the key to innovative breakthroughs in the field of artificial intelligence in the future. This article attempts to explore the impact, risks, and governance strategies of the application of Large Language Models on public security intelligence work, providing ideas and experiences for further research in the future.

3. The Technical Logic and Features of Large Language Models

3.1 The Technical Logic of Large Language Models

Large Language Models (LLMs) is a type of language model composed of neural networks with many parameters. It is trained on a large amount of unlabeled text using self-supervised learning or semi-supervised learning to understand and generate human language. The model uses a Transformer architecture and pre-training objectives similar to smaller models, such as Language Modeling, with the main difference being the increase in model size, training data, and computational resources. The Transformer model is based on attention mechanisms to accelerate deep learning algorithms. It consists of a set of encoders responsible for processing input of arbitrary length and generating its representation, and a set of decoders responsible for converting the new representation into the target words. Based on pre-training on a large amount of text data, Large Language Models can perform a wide range of tasks, including text understanding, translation, and sentiment analysis.

In terms of technical logic, Large Language Models can achieve "super simulation" of human beings from generation to creation through empirical learning and technical imitation. The generation models used mainly include generative adversarial models, autoregressive models, variational autoencoder models, flow models, and diffusion models. They use deep learning neural networks to spontaneously find connections in massive text data libraries and understand users' questioning intentions through human-computer dialogue, providing reasonable answers. They can automatically write software, generate papers, synthesize images, provide consultations, and handle various abstract, specific, or even strange user requests, showing "human-like" intelligent characteristics. Unlike traditional intelligent language models, this technology adopts a new training method of "reinforcement learning from human feedback." Developers use human-computer dialogue patterns to update the database on which the system relies through a cycle of "training-feedback-correction," and then use the updated database to optimize the AI model, thereby strengthening the system's autonomous learning ability and better aligning with users' questioning intentions.

3.2 The Features of Large Language Models

3.2.1 Large Language Models have high costs

Training data is the cornerstone and fuel for training Large Language Models. Current research in the technical field shows that Large Language Models from various companies have similar algorithms at the algorithmic level, showing a trend towards homogenization. Therefore, training data has become one of the important factors that truly differentiate and affect the performance of Large Language Models. Good general-purpose models need to consume a large amount of GPU resources on high-quality platforms, train a large number of parameters, and invest a sufficient amount of training time. Taking the training of GPT-3 as an example, assuming a GPU price of $10 per hour, a FLOPS utilization rate of 46.2%, GPT-3 has 175 billion parameters, requires 3.14E23 FLOPS, and is trained using 45TB of data, the training cost is approximately $9 million. In addition, OpenAI has publicly stated that training the GPT-4 model with over a trillion parameters has an initial cost of up to $2-3 million. The feasibility of fine-tuning consumer-grade models is much higher than retraining Large Language Models. It can be seen that as the number of training parameters and complexity increase, the demand for computational training power grows rapidly, which not only further improves performance but also exacerbates the training costs of Large Language Models.

3.2.2 Large Language Models are highly efficient

Efficiency upgrades in Large Language Models are mainly achieved through two ways: (1) pre-training models. With the emergence of a large number of large model products, the maturity of general artificial intelligence application technologies provides pre-trained models for specific domain business tasks; (2) completing unsupervised learning. Early language models required a large amount of manually annotated data to learn the rules and make inferences about unknown things, which means that labeled data is needed for learning. However, Large Language Models have general knowledge and do not require manual labeling. They only need to be fine-tuned based on the scenario using pre-trained models to achieve performance improvements.

3.2.3 Large Language Models have the ability to fuse multimodal information

In addition to the processing ability of the original single-modal models, Large Language Models have evolved the ability to fuse multiple heterogeneous sources in the process of evolution, achieving cross-modal information interaction. This means that information can be transmitted through different modes or media, such as integrating knowledge from single text recognition into different modalities such as images, audio, and video. For example, OpenAI recently released the Sora model, which can generate a 60-second video based on a text description, simulating the movement and interaction of objects and characters in three-dimensional space. The training of Sora requires a large amount of video data, which means that general artificial intelligence technology has further achieved the fusion of heterogeneous information, better understanding, and simulation of the real world.

3.2.4 Large Language Models have the ability to emerge

As the scale of Large Language Models continues to grow, the effectiveness of task completion is also increasing. For complex tasks consisting of multiple steps, Large Language Models have shown the ability to emerge in downstream tasks, which means that the effect will only increase significantly when the model size reaches a certain level. Before the model size is smaller than a certain critical value, the model basically does not have the ability to solve the task. Currently, there are two types of tasks that can stimulate the emergence ability of Large Language Models: (1) In Context Learning (few-shot prompt), when the model size is not large enough, the completion effect of various tasks is poor, but when it crosses a certain critical value of model size, the large model suddenly performs well in handling tasks; (2) Chains of Thought (CoT), essentially a special few-shot prompt, as one of the core technologies of large model logical reasoning ability, CoT decomposes tasks into clear logical chains to solve them step by step, effectively enhancing the common sense reasoning, mathematical operations, and symbolic reasoning abilities of Large Language Models. At the same time, the deep autonomous feedback learning ability of Large Language Models, under the triggering of the emergence ability, learns a large amount of human-like thinking and expression abilities, gradually generating "human-like expressions," enhancing credibility.

4. The Logic Framework and Impact Analysis of Large Language Models Embedded in Public Security Intelligence Work

4.1 The Logic Framework of Large Language Models Embedded in Public Security Intelligence Work

The development and application of Large Language Models can achieve the full-process integration of artificial intelligence technology into the entire workflow of intelligence rapid collection, analysis and judgment, report generation, and practical application, profoundly affecting the working mode and operational efficiency of public security intelligence. The logical framework of Large Language Models embedded in the workflow of public security intelligence work is shown in Figure 1.

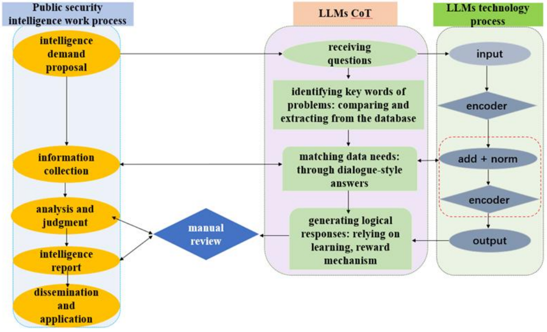
Public security intelligence work process
LLMs CoT
LLMs technology process
intelligence demand proposal
information collection
analysis and judgment
intelligence report
dissemination and application
manual review
receiving questions
identifying key words of problems: comparing and extracting from the database
matching data needs: through dialogue-style answers
generating logical responses: relying on learning, reward mechanism
input
encoder
add + norm
encoder
output

Figure 1. The Logical Architecture of Large Language Models Embedded in the Workflow of Public Security Intelligence Work

4.2 Impact Analysis of Large Language Models Embedded in Public Security Intelligence Work

4.2.1 Impact of Large Language Models on public security intelligence collection

The integration of Large Language Models enables effective improvements in understanding collection intent, expanding collection scope, broadening retrieval methods, and enhancing processing efficiency in public security intelligence collection. Large Language Models possess powerful natural language processing capabilities, able to understand and analyze large amounts of text data. Through deep learning of text, the model can automatically extract key information, thereby improving the efficiency and accuracy of public security intelligence collection. Furthermore, Large Language Models can acquire data from multiple sources and integrate and analyze it. With a vast knowledge base, large models can effectively expand the scope of public security intelligence collection. Additionally, Large Language Models can gather information from a wider range of public security intelligence topics by invoking external knowledge base interfaces. They can also identify and filter valuable information, providing more comprehensive and accurate results for data collection. Moreover, the retrieval capability of Large Language Models surpasses the limitations of traditional stock-based retrieval, reducing platform turnover time. Furthermore, the fine-tuning of large models can rely on human feedback to strengthen the reinforcement learning mechanism, guiding efficient human-machine dialogue to recognize the task intent issued by humans and quickly provide more accurate intelligence information in the shortest time.

4.2.2 Impact of Large Language Models on public security intelligence analysis

Public security intelligence analysis is a core part of intelligence work, mainly human intellectual output. It is also an important basis for facilitating final decision-making, thus requiring high reliability and effectiveness of intelligence products. Large models transform the traditional knowledge production mode, mainly relying on human experts or teams for research, analysis, and organization, into a new knowledge production mode integrating machine autonomous learning [6] 6.Large Language Models can process and understand large-scale text datasets, which is significant for analyzing public security intelligence from multiple sources. They can extract information from various formats of data such as news articles, social media posts, and government reports, helping intelligence analysts quickly obtain the required data. Large Language Models have advanced natural language processing capabilities, understanding complex language structures and implicit meanings. They can also generate precise intelligence summaries, helping analysts quickly grasp key information and reduce the time spent reading large documents. Additionally, Large Language Models typically support multiple languages, assisting intelligence analysts in understanding the context and details of different situations. By analyzing large amounts of historical and real-time data, Large Language Models can help intelligence agencies identify potential risks and hazards, greatly improving the efficiency and effectiveness of intelligence root work.

4.2.3 Impact of Large Language Models on public security intelligence writing

The carrier of intelligence products is public security intelligence reports, which reflect the value density of the overall quality of public security intelligence work. The impact of Large Language Models on the writing of public security intelligence reports is profound and multidimensional. Through its highly developed natural language processing capabilities, Large Language Models change the traditional mode of intelligence writing. Firstly, Large Language Models have powerful language understanding and generation capabilities. They can analyze massive text data, extract key information, and express it in a natural and fluent manner, providing valuable references and insights for public security intelligence writing. Secondly, Large Language Models can quickly generate various writing schemes. Writers can choose the most suitable scheme according to actual needs, improving the relevance and effectiveness of intelligence. Additionally, Large Language Models can help intelligence writers better grasp trends and patterns. By analyzing a large amount of data, they can predict future development trends, providing proactive viewpoints and suggestions for intelligence writing. Large Language Models can generate different text types according to the different needs of intelligence personnel in the shortest time, meeting various practical requirements such as research, analysis, and prediction. Furthermore, in the future work of public security, they have strong expansion learning capabilities. Intelligence departments can provide specialized training for large model technology, provide them with public security intelligence writing format templates, and create large models with specific features of public security business scenarios.

4.2.4 Impact of Large Language Models on public security intelligence information systems and applications

With the establishment of public security intelligence information platforms, the improvement of public security databases, and the implementation of projects such as the

"Golden Shield Project" and "Bright Project," early issues such as intelligence barriers and information silos have been partially alleviated. However, there are still security risks in the construction, operation, and maintenance of current intelligence information systems, as well as difficulties in system management and storage. Intelligence personnel in operational units in many places have reflected that the intelligence platform has not achieved complete coverage, and there are situations where the intelligence information system cannot be used or cannot be "used by multiple parties." The use of Large Language Models can further solve the above-mentioned problems.

Firstly, the application of large models in public security intelligence can break information barriers and achieve cross-domain integration. Facing the widespread situation of multi-headed management and repeated construction, Large Language Models have more diverse data organization forms, which are conducive to eliminating barriers between systems, better coordinating intelligence work through more diversified knowledge organization forms, realizing the interconnection and integration of different intelligence platforms, and improving application efficiency. For example, in the construction of the "Information-Pointer-Action" integrated platform, the large language model can be used as a technical intermediary to comprehensively integrate information from various platforms, achieve technological interoperability, and better promote the construction of the integrated "Information-Pointer-Action" information platform.

Secondly, it promotes the reuse of experience and facilitates communication. The large language model technology realizes the sharing of intelligence information, making up for the shortcomings of intelligence information systems not being able to be shared by multiple regions or used by one region in multiple ways. For example, the large language model can collect and learn intelligence information A, and mark similar data samples. When it recognizes that intelligence B belongs to the same or similar context as intelligence A, the analysis content of intelligence A can be used as a reference for experience reuse. Based on the large language model, information sharing is realized, so that open-source intelligence from different periods, regions, and types can serve as the basis for decision-making among different police forces and departments, promoting experience exchange.

Thirdly, it enhances self-learning ability and draws on the past to create the future. The application of Large Language Models does not mean that public security intelligence information systems are no longer used but rather provides a new technical support and reference, using machine autonomous learning knowledge to replace manual output experience, continuously learning from experience, autonomously discovering, reasoning, and generating new knowledge based on this, utilizing generative responses to disseminate knowledge to various departments, continuously using new knowledge to produce new knowledge, promoting intelligence systems to be more intelligent and specialized, and further enhancing data storage, responsiveness, and learning capabilities.

5. Research on the Risks and Issues of Applying Large Language Models in Public Security Intelligence Work

Large Language Models represented by ChatGPT are essentially commercial artificial intelligence products, distinct from public security intelligence products. On the one hand, the technology of large models still needs improvement, with issues such as biased outputs, lack of real-time autonomous learning ability, and over-reliance on datasets, making them unable to truly replace human independent analysis of intelligence. On the other hand, the audience of large models is society at large, with lower requirements for data security and confidentiality, while public security intelligence products place a high emphasis on data security, especially for some investigative intelligence and police-related information, which must be kept absolutely confidential. Based on this, starting from the technical logic of Large Language Models, this paper constructs a risk system of Large Language Models for public security intelligence work (as shown in Figure 2), to analyze the potential risks of Large Language Models for public security intelligence work.

5.1 Risks of Public Security Intelligence Data Security

Large Language Models are not "large in security." The security risks of public security intelligence information are reflected at both the individual and national levels. On one hand, personal information is an important part of public security intelligence, and control over personal information sometimes determines the success or failure of decisions. On the other hand, public security intelligence contains a large amount of sensitive information about the country, society, and internal public security, and once leaked, it can easily endanger national security and social stability.

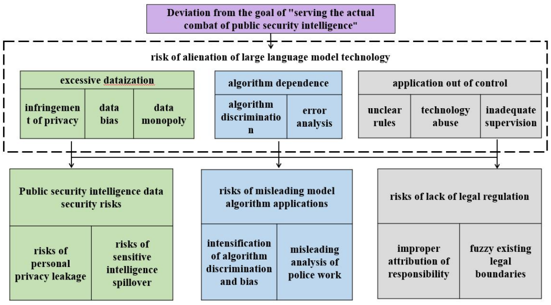
Deviation from the goal of "serving the actual combat of public security intelligence"
risk of alienation of large language model technology
excessive dataization
infringement of privacy
data bias
data monopoly
algorithm dependence
algorithm discrimination
error analysis
application out of control
unclear rules
technology abuse
inadequate supervision
Public security intelligence data security risks
risks of personal privacy leakage
risks of sensitive intelligence spillover
risks of misleading model algorithm applications
intensification of algorithm discrimination and bias
misleading analysis of police work
risks of lack of legal regulation
improper attribution of responsibility
fuzzy existing legal boundaries

Figure 2. The Risk System of Large Language Models for Public Security Intelligence Work

5.1.1 Risk of personal privacy leakage

During the pre-training stage, Large Language Models rely on massive data through two channels: first, they use openly available intelligence resources on the internet as training samples; second, they collect and learn dialog information in real time, continuously enriching the generated content using deep autonomous learning feedback capabilities. Although China's "Personal Information Protection Law" clearly stipulates that personal information processors should moderately handle legally obtained personal information while fully respecting user privacy, at the current stage, large model technology still cannot accurately distinguish personal privacy data to avoid its inclusion in the training scope. The process and mechanism of data collection have not been universally recognized, and public awareness of protecting personal information is still lacking, leading to a significant increase in the risk of intelligence information leakage.

5.1.2 Risk of sensitive information leakage

During the training data collection process, large models may collect important sensitive information such as public safety intelligence, national core data, and military intelligence, which, due to the inherent risk of data leakage in the technology itself, can threaten national security once mined. Large models also carry a high risk of being transformed into tools for illegal activities. If exploited by criminals, they can generate false, violent, and other illegal and criminal information, which, once disseminated, can disrupt information dissemination, incite ideological conflicts, and even endanger national political security and stability. Due to such overflow risks, various countries and regions have enacted laws to regulate the use of large models to ensure data security. For example, the European Union has issued the "General Data Protection Regulation" to strictly regulate the training and use of data for large models. Italy has announced a ban on the use of ChatGPT to protect social order, national ideology, and political security. The United States has introduced the Federal Data Privacy and Algorithm Accountability Act to protect data security while preventing algorithmic discrimination and improper decision-making.

5.2 Risk of Misleading Application of Model Algorithms

Algorithms have inherent technical flaws, leading to an increasingly serious "algorithm black box" problem. The analysis process of Large Language Models in human-machine dialogues is not visualized. The lack of clarity in data selection standards will affect the effectiveness of the generated products, resulting in risks such as algorithmic illusions, analysis errors, and biases, further increasing the difficulty of management and services for public security agencies.

5.2.1 Algorithmic discrimination and bias fueling social conflicts

Looking at the decision-making process of large models, from the pre-training stage, the data collected by large models is mixed with discriminatory and biased information. In the data analysis stage, the algorithm's technical barriers are high, relying mainly on developers to make decisions, which involves a lot of human judgment and subjective emotions. This leads to the attachment of subjective initiative to algorithmic technology, resulting in algorithmic discrimination, bias, and bullying in sensitive issues such as race, religion, and gender, which can easily escalate conflicts between special groups and society. Additionally, due to the emergent capabilities of large models, algorithmic technology generates a large number of unknown and irreversible new codes during continuous self-learning. After unlimited program runs deepen algorithmic discrimination, facing retraining will incur higher costs. Developers often choose minor adjustments instead of achieving zero discrimination in algorithms.

5.2.2 Misleading conclusions in error analysis output for police work

In the "ChatGPT Chinese Performance Evaluation and Risk Response,"[13] the authors found risks such as error confusion and inconsistency in facts in ChatGPT's performance evaluation. Specifically, on the one hand, ChatGPT-like Large Language Models may not always correctly answer academic and some common sense questions. On the other hand, they may fabricate false information and output it in a mixed manner with true and false information, misleading intelligence agencies to make correct decisions. Combining this with police intelligence work, there are two reasons for generating erroneous analysis: first, the time mismatch between the problem and the text database. Large models cannot use existing data to respond timely to new issues and can only provide experiential references instead of making decisions for the present or future, which may lead to misleading or erroneous responses. Second, the asymmetry between the classified and secret nature of police intelligence and the intelligence data controlled by large models. The limited police intelligence data collected by large models significantly reduces the rigor and professionalism of the generated responses. Currently, there is no performance evaluation standard for the use of large models in police intelligence work. Intelligence personnel cannot quantitatively evaluate whether the generative responses are objective and accurate, which may lead to investigative misjudgments and increase the cost of law enforcement and case handling by public security agencies.

5. 3 Risk of Legal Regulation Vacancy

The law is the cornerstone for ensuring the safe and standardized operation of large language model technology and is the action guide for public security organs to enforce the law. The widespread application of large models has brought tremendous challenges to society, with issues arising regarding whether AI-generated products constitute the subject of rights and the subject of technological crimes. However, there are still loopholes in legal regulation.

5.3.1 Improper attribution of responsibility

Under the current legal system, there is controversy over whether the information products generated by Large Language Models constitute criminal objects and the subject of attribution. On the one hand, some scholars believe that Large Language Models themselves lack free will and subjective purposes, and their infringement behavior stems from the subjective manipulation of users or developers, making it meaningless to attribute responsibility to them. At the same time, imposing heavy responsibility on technology suppliers may greatly increase regulatory costs, ultimately weakening the market's willingness to update and invest in technology [14]. Therefore, large models themselves cannot be the subject of attribution, and the subject of attribution can only be represented by relevant natural persons. On the other hand, the generated products of large models meet the definition of knowledge products for copyright, possessing a certain "originality." Therefore, some scholars agree that they can constitute criminal objects and bear subject responsibilities. Globally, the EU's "Digital Services Act" assigns responsibility for AIGC to technology suppliers; the "Interim Measures for the Administration of AIGC Services" jointly issued by the Cyberspace Administration of China and others stipulates that technology providers should fulfill the obligation to ensure that training data does not contain content that infringes on intellectual property rights, similarly assigning responsibility for generated products to technology providers. Considering the practical situation, AIGC has given rise to areas such as large models and autonomous driving. The emergence of new technologies has posed the issue of the attribution of responsibilities for non-human criminal subjects to existing laws. Although relevant regulations have been issued, the law still needs to be further clarified in balancing the relevant subjects of large model application and national supervision.

5.3.2 Ambiguity in existing legal boundaries

The "Data Security Law" enacted in China requires the establishment of a data classification and grading protection system. The "Personal Information Protection Law," "Interim Measures for the Administration of AIGC Services," and "Ethical Review Methods for Science and Technology" provide practical guidance and basis for ethical review of technologies such as large models in the fields of science and technology.

However, the characteristics of AI crimes in the era of large models, such as strong concealment, low cost, and high destructiveness [15], and issues such as whether the excessive and illegal collection of data by Large Language Models, the use of data by actors for illegal activities, and the continuous feeding of malicious data by criminals to Large Language Models to guide users to commit crimes or provoke ideological disputes and threaten national security, are not clearly defined by national laws. In particular, the "Criminal Law" has not formed a joint force with the preceding laws. Although it has formulated the crime of illegal access to computer information systems, its subject constitution is restricted. One is whether the large model itself constitutes the subject of the crime, and the other is how to prove the legality of the data and channels obtained by the large model, etc., which are not clearly explained, leading to a lack of specific law enforcement basis for public security organs in such crimes. At the same time, intelligent technologies under big data have presented a "black box effect." If public security organs can use them effectively, they can provide positive help for public security operations. However, if the technology is reverse-used by the dark web to quickly generate malicious codes, it will instead pose a threat to public security organs. The legal boundaries urgently need to be clarified.

6. Research on Countermeasures for the Application of Large Language Models in Public Security Intelligence Work

Existing research on how to improve the application performance of Large Language Models in the field of intelligence mainly provides solutions from the perspectives of algorithm transparency, technical preparation, policy regulation, institutional improvement, the establishment of a "model + intelligence data" system, and enhanced user privacy protection. This article adheres to a systematic approach, taking public security intelligence work as a starting point, exploring the construction of a governance system with public security organs as the main body and multiple entities participating in a coordinated manner, creating a "human-machine collaboration" operation mode with humans as the main focus, and promoting the better adaptation of Large Language Models to the application scenarios of public security intelligence work.

6.1 Building a Tight Information Security Protection Network under the Background of Large Language Models

Data security is the foundation of public security intelligence work, and ensuring data security requires technical support. It is necessary to actively utilize a diverse range of social entities to jointly build a tight information security protection network under the background of Large Language Models to prevent the leakage of public security intelligence data.

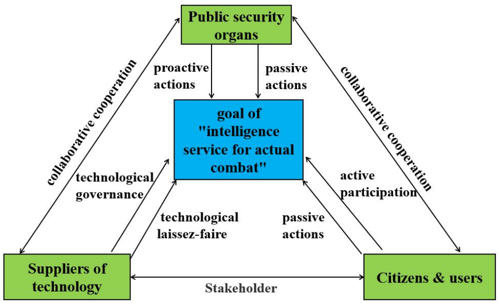

Figure 3. The Tripartite Relationship among Public Security Intelligence Agencies, Technology Companies, and Citizens

Using a three-party game model for strategic exploration, based on the Malthusian equation and Lyapunov's discriminant method, an analysis of the stability of the equilibrium point among the three parties is conducted. It is concluded that under the proactive role of public security agencies, both the large model technology suppliers and citizens, in order to avoid the risk of punishment, will choose strategies of "technological governance" and "active participation". At this point, a unique stable strategy emerges in the evolution game of the three parties (as shown in Figure 3).

6.1.1 At the level of technology developers and platforms: strengthen review and data cleaning

Data review is a prerequisite for improving the effectiveness of large model technology. Developers and platforms of large models should cooperate with the government to promote algorithm transparency, strengthen joint review of human and AI technologies, use intelligent algorithm technology to identify the attributes of generated content, quickly review and prioritize high-quality content for release [16], and facilitate real-time screening and national supervision.

Data cleaning is the follow-up link of data review, mainly dealing with missing values, outliers, and duplicates [17]20. Developers should cooperate with public security agencies to jointly formulate performance evaluation standards for the application of Large Language Models in public security intelligence work, clean sensitive, erroneous, and false data, reduce misleading behaviors, and avoid the overflow of sensitive information.

6.1.2 At the level of intelligence personnel: refine review and manual supervision

In the era of big data, traditional human intelligence gathering and analysis methods should not be abandoned and are still a beneficial supplement to public security agencies' intelligence information gathering [18]. Recently published in Nature, "Five Major Research Questions about ChatGPT" points out that when ChatGPT is used in the scientific community, it must adhere to the principle of human review, further confirming the importance of strengthening the review and supervision responsibilities of intelligence personnel in the context of "human-machine collaboration". Based on practical experience, intelligence personnel need to quickly detect sensitive information involving data security, classified information, bias, discrimination, and other sensitive information generated by Large Language Models, and promptly contact developers for secondary data cleaning, continuously guiding the benign operation of algorithms.

6.1.3 At the level of citizens: emphasize privacy protection and guard against information leakage

Citizens are the group most vulnerable to the infringement of personal data and the most difficult to protect their legitimate rights and interests when it comes to developers and public security intelligence personnel. It is necessary to popularize knowledge of artificial intelligence and relevant laws and regulations to the public, improve users' correct understanding and use of large model technology, learn to protect their rights and interests in accordance with the law; at the same time, when using large model platforms, guide developers to increase the mandatory reading of relevant agreements and procedures for platform information collection and use by users, enhance the importance of both parties in protecting privacy, avoid personal data leakage, and establish a protection network from the source.

6.2 Controlling the Application of Public Security Intelligence Analysis under the Background of Large Language Models

The comprehensive digitization of intelligence work is to digitize the source data, analysis process data, analysis conclusions, etc., which is the basis for the development of intelligence tools, the training of specific domain intelligence large models, and the training of specific task adapters in the future[2]6. Combining with the characteristics of public security work, in the process of achieving comprehensive data digitization in public security intelligence work, attention should be paid to the breadth and depth of the application of Large Language Models, adhere to the leading position of public security agencies, reduce the risk of misleading analysis caused by model illusions, and reduce error costs.

6.2.1 Adhere to the leading position of public security agencies in the application of public security intelligence under the background of Large Language Models

In the application of large models in public security intelligence work, public security intelligence departments should firmly establish the leading position of public security agencies. Firstly, developers and platforms of large models should establish legal awareness, cooperate with public security agencies in supervision and inspection, not maliciously manipulate data, and have the right and obligation to order them to clean data suspected of harming security and prohibit secondary use; secondly, users of large models have the obligation to protect personal data security, and public security agencies have the right to prompt them for behaviors suspected of leaking privacy and endangering personal information security; for abuse of data, causing illegal crimes, public security agencies have the right to take measures according to law; thirdly, intelligence personnel of public security intelligence should consciously accept internal supervision of public security agencies, strictly review and supervise the generated products according to law.

6.2.2 Prudent application of Large Language Models in public security intelligence work

The application of large models in public security intelligence work should focus on high precision requirements and avoid high-security areas.

On the one hand, large models show advantages in vocabulary extraction and data organization and integration in the intelligence collection stage, and should strengthen their use in intelligence preprocessing work to make up for the shortage of manpower; at the same time, give full play to the auxiliary role of large models, provide new technical support for the intelligence system, and increase the value of open source intelligence. It is worth noting that for personnel warning and monitoring work, the use of large models is prone to information cocooning effects, resulting in incorrect guidance, model biases, and the re-polarization of groups, which can easily lead to secondary risks. Therefore, based on the characteristics of large model technology evolution and the applicable links of open source intelligence activities, it is necessary to formulate basic rules or guidelines for the standardized application of large model technology to prevent the problems of false information or intelligence leakage that may be caused by abuse[19].

On the other hand, to ensure the absolute security of classified data, the application of large models should avoid importing classified data, restrict their use in areas involving national core interests such as criminal investigation and national security, and prevent Large Language Models from learning the business logic of public security, leading to the risk of technology being tampered with, breached, and threatening political stability, etc. For example, when the network security department uses base station analysis to track the movements of criminal suspects, when the large model masters this technology, more privacy of citizens will face the risk of leakage.

6.2.3 Standardize the application depth of Large Language Models in the field of public security intelligence

The "Data Security Law" stipulates that a data security classification and grading protection system should be established. Public security agencies should include large model technology in the existing classification and grading system and further update and improve it. For "general data," such as dog management and other routine data, it is allowed to include them in the data training samples without involving the privacy of the parties; for "important data," such as internal management of public security agencies, the standard for whether they can be publicly available is used as the basis, and its collection of non-public information is closely supervised; for "core data," that is, data that may hinder investigation, threaten citizens, society, and national security, it is prohibited for large models to collect and learn.

6.3 Improve Laws to Provide Legal Basis

To cope with the many challenges brought by the application of Large Language Models, improving laws is the first step. Legislation should clarify the subject of responsibility, establish a accountability mechanism, clarify the legal boundaries, and improve the legal content to provide legal basis for relevant personnel, achieve a closed loop of legislation, law enforcement, judicial, and compliance, and guide the positive application of large models.

6.3.1 Clarify the responsible entities and improve the regulatory framework

Under the promotion of existing laws and regulations, China has initially established a legal framework for regulating large models. To further prevent legal risks associated with large models, the legislative body should work with departments such as the Cyberspace Administration and the Ministry of Industry and Information Technology to determine the responsible entities and improve the regulatory framework: at the national level, establish specialized agencies or organizations for dedicated supervision and enact technical supervision laws; at the technical level, foster cooperation among the government, academia, and technology developers to ensure the security of model training and execution environments; at the legal level, based on the general tort liability system, supplemented by the principle of presumption of fault, dynamically adjust judicial responsibilities for providers of AIGC[20].

6.3.2 Clarify legal boundaries and clarify the basis for public security law enforcement

Taking the "Interim Measures" as a basic reference, it is necessary to redefine the criminal subjects in the Criminal Law, cancel restrictive constitutive elements, and convict them based on comprehensive factors such as data classification and grading standards and the quantity of illegally obtained data[18]20. The Civil Law, Administrative Penalty Law, Public Security Administration Punishment Law, and other laws should be adjusted synchronously to clarify the attribution and penalties of large model crimes in civil, criminal, and administrative cases, providing a basis for public security agencies to enforce the law strictly.

7. Conclusion

Based on the analysis of the logic and operational characteristics of large models, this paper combines large models with public security intelligence work. Taking the collection, analysis, writing, information system, and application of public security intelligence as the starting point, it analyzes the positive and negative impacts of applying large models, focusing on countermeasures. It explores solutions to the data security, model illusions, and legal gaps faced by public security intelligence work under the current large model, and proposes a "human-machine collaboration" and a human-centered operational mode in which public security agencies take the lead, with technology developers and the public mutually participating in governance. This model aims to jointly maintain data security and development in the era of artificial intelligence, and promote the more proactive role of Large Language Models in public security intelligence work.

Currently, the global artificial intelligence industry has made significant developments in the field of large models. Whether it is base models or products developed based on base models, they have to varying degrees promoted the liberation of productivity in various industries. Domestic technology companies have entered a "battle of a hundred models." While the technology is developing, there is still much room for improvement. In the future, in-depth research should be conducted on multimodal translation, interactive technology, low-resource language translation, and precise translation of Chinese to enhance the effectiveness of large models in serving public security operations.

References

[1] Cao Shujin, Cao Ruyi. The Impact of AIGC on the Research and Practice of Intelligence Studies: A Perspective from ChatGPT. Modern Intelligence, 2023, 43(04): 3-10.

[2] Yang Qian, Lin He. Digital Strategies and Practical Scenarios for Intelligence Research under the Background of Large Language Models. Competitive Intelligence, 2023, 19(03): 2-13. DOI: 10.19442/j.cnki.ci.2023.03.003.

[3] Zhao Bang, Cao Shujin. Test and Analysis of Typical Tasks in the Intelligence Field Executed by Generative Large Language Models at Home and Abroad. Information and Documentation Services, 2023, 44(05): 6-17.

[4] Demetriadis, S., Dimitriadis, Y. (2023). Conversational Agents and Language Models that Learn from Human Dialogues to Support Design Thinking. In: Frasson, C., Mylonas, P., Troussas, C. (eds) Augmented Intelligence and Intelligent Tutoring Systems. ITS 2023.

[5] Yuan Zeng. A Study on the Responsibility of Generative Artificial Intelligence. Eastern Philosophy, 2023(03): 18-33. DOI: 10.19404/j.cnki.dffx.20230505.002.

[6] Zhang Yue, Li Zhengfeng, Qian Wei. From ChatGPT to Knowledge Co-construction in Human-Machine Collaboration. Studies in Science of Science, 1-11[2023-09-23].https://doi.org/10.16192/j.cnki.1003-2053.20230817.002.

[7] Zhang Linghan. The Logic Update and System Iteration of Deep Synthetic Governance: The Chinese Path of Governance of Generative Artificial Intelligence such as ChatGPT. Legal Science (Journal of Northwest University of Political Science and Law), 2023, 41(03): 38-51. DOI:10.16290/j.cnki.1674-5205.2023.03.015.

[8] Su Yu. Legal Risks and Governance Paths of Large-scale Language Models. Legal

Science (Journal of Northwest University of Political Science and Law), 2024, (01): 1-13 [2024-01-24]. https://doi.org/10.16290/j.cnki.1674-5205.2024.01.010.

[9] Cheng Xuejun. Governance Paths for the Algorithm Black Box of Super Artificial Intelligence Platforms under the Wave of AIGC// Shanghai Law Society. Collected Works of "Emerging Rights" Volume 2, 2023. [Publisher unknown], 2024: 9. DOI: 10.26914/c.cnkihy.2024.000892.

[10] Hu Changping, Lu Meijiao. Theoretical Development of Intelligence in the Era of Big Data and Intelligence Environment. Information Studies: Theory & Application, 2020, 43(10): 1-6. DOI: 10.16353/j.cnki.1000-7490.2020.10.001.

[11] Mirza, S., Coelho, B., Cui, Y., Pöpper, C., & McCoy, D. (2024). Global-liar: Factuality of llms over time and geographic regions. arXiv preprint arXiv:2401.17839.

[12] Mishra, A., Asai, A., Balachandran, V., Wang, Y., Neubig, G., Tsvetkov, Y., & Hajishirzi, H. (2024). Fine-grained hallucination detection and editing for language models. arXiv preprint arXiv:2401.06855.

[13] Zhang Huaping, Li Chunhan, Li Chunjin. Evaluation of Chinese Performance and Risk Response in ChatGPT. Data Analysis and Knowledge Discovery, 2023, 7(03): 16-25.

[14] Zeng Xiong, Liang Zheng, Zhang Hui. "Regulatory Path of Artificial Intelligence in the European Union and Its Enlightenment for China: An Analysis of the 'Artificial Intelligence Act'." Published in "E-Government," 2022, Issue 9.

[15] Yang Yanchao.In the era of Large Models, how to govern artificial intelligence crimes?. Qunyan, 2023, (07): 32-34. DOI: 10.16632/j.cnki.cn11-1005/d.2023.07.022.

[16] Legal Issues - Law Reviews; Reports from University of Texas Austin Describe Recent Advances in Law Reviews (Deep Fakes: a Looming Challenge for Privacy, Democracy, and National Security). Journal of Engineering, 2020.

[17] Cai Shilin, Yang Lei. Research on the risks and collaborative governance of ChatGPT intelligent robot applications. Information Theory and Practice, 2023, 46 (05): 14-22. DOI: 10.16353/j.cnki.1000-7490.2023.05.003.

[18] Chen Chengxin, Zeng Qinghua, Li Lihua. Innovative development path of public security intelligence work in the big data environment. Information Theory and Practice, 2019, 42 (01): 10-15. DOI: 10.16353/j.cnki.1000-7490.2019.01.002.

[19] Li Rong, Wu Chensheng, Dong Jie, et al. The impact of ChatGPT on open source intelligence work and countermeasures. Information Theory and Practice, 2023, 46 (05): 1-5. DOI: 10.16353/j.cnki.1000-7490.2023.05.001.

[20] Yan Yaru, Luo Xiaochun. Legal risks and determination of infringement liability of generative artificial intelligence. Journal of Yancheng Teachers University (Humanities and Social Sciences Edition), 2024, 44 (02): 54-65. DOI: 10.16401/j.cnki.ysxb.1003-6873.2024.02.020.