Research and Exploration of an Intelligent Interrogation Transcription System for Public Security Based on Large Language Models

Junquan Zhou, Shanting Song, Shuifeng Zhang*, Qingyang Gao, Yuekai Ma, Zhengnan Tian, and Buyun Chen

Nanjing Police University, Nanjing, 210023, China
Email: 270847034@qq.com (S.F.Z.)
*Corresponding author
Manuscript received September 20, 2025; accepted November 7, 2025; published December 22, 2025.

***Abstract*—Aiming to address issues in the transcription of police interrogations such as inefficiency, lack of standardization, and poor information coordination, this paper designs an intelligent interrogation transcription system based on a large language model. By constructing a hierarchically decoupled system architecture that integrates the data layer, model layer, service layer, and application layer, the system achieves intelligent processing of the entire interrogation workflow. It innovatively implements core functions such as adaptive transcription generation and dynamic completion for cases, multimodal interaction and active interrogation assistance, compliance review, and analysis of evidential contradictions. The system ensures professional reliability through fine-tuning of judicial large language models and secure deployment techniques. Experimental results demonstrate that the system can reduce transcription time by over 50% while significantly improving transcription completeness and procedural compliance rates. This provides a feasible technical solution for the "technology-driven policing" initiative and the standardization of law enforcement in public security organs. Furthermore, the "human-machine collaborative" smart case handling model offers a replicable practical pathway for the digital transformation of policing operations.**

***Keywords*—intelligent interrogation transcription system for public security, large language model, judicial domain fine-tuning, human-machine collaboration**

I. Introduction

As the core component of criminal investigation and evidence collection, police interrogation transcripts directly impact judicial fairness and evidentiary validity. Traditional transcription methods face three major challenges: inefficiency, lack of standardization, and poor information coordination. Manual recording is time-consuming and prone to omitting crucial information and procedural flaws, while information barriers across departments lead to frequent repetitive inquiries. With the advancement of the Smart Policing strategy and breakthroughs in large language model (LLM) technology, this study develops an intelligent interrogation transcript system based on LLMs. Through core technologies such as case-adaptive generation, multimodal interaction, and compliance review, the system achieves automatic conversion from speech dialogue to standardized documents. It not only reduces transcription time by over 50% but also significantly enhances transcript quality and evidentiary value through element integrity verification and procedural compliance assurance. Theoretically, this research explores the application paradigm of artificial intelligence in the judicial vertical domain; practically, it provides a feasible technical solution for the standardization of law enforcement and the digital transformation of policing, holding significant potential for widespread application.

European and American countries have established relatively mature technical systems in legal natural language processing, yet their application scenarios differ significantly from China's criminal interrogation context. Commercial platforms like IBM Watson Legal and ROSS Intelligence primarily focus on civil and commercial law domains, specializing in contract analysis and case retrieval. In terms of research innovation, the Defense Advanced Research Projects Agency's "Mind's Eye" project explored temporal logic-based confession consistency detection algorithms, achieving cross-modal verification between video and textual statements. However, these studies face two main limitations: first, their research focus remains predominantly on civil and commercial scenarios, lacking specialized investigation into intelligent transcript generation for criminal interrogations; second, fundamental differences in legal systems and law enforcement procedures—such as the adversarial litigation system in common law countries versus China's inquisitorial system, coupled with disparities in privacy protection standards—render foreign technical solutions difficult to directly apply in China's public security interrogation contexts.

Domestic research exhibits application-driven characteristics but has yet to form a complete systematic solution. At the practical level, systems like the "AI-assisted Interrogation" in Liuzhou, Guangxi, and commercial solutions such as "Lujing Zhixun" already utilize large models to analyze interrogation content, recommend follow-up questioning strategies, and automatically generate draft transcripts, reducing transcription time by approximately 30% [1]. Technically, domestic scholars have explored key information extraction based on pre-trained models like BERT and utilized legal knowledge graphs for compliance review. However, existing technical solutions mostly represent "isolated breakthroughs": early systems relied on fixed templates and lacked flexibility, while newer models face challenges such as poor adaptation to legal terminology and difficulties in modeling multi-turn dialogues, failing to construct an end-to-end system covering speech recognition, content generation, compliance review, and data collaboration. Although platforms like "Cloud Sword" have achieved partial data interoperability,

DOI: 10.7763/IJET.2025.V17.1331

intelligent collaboration remains hindered by data format and security barriers [2].

In summary, foreign research has established relatively mature technical systems in legal natural language processing and privacy computing, yet their direct applicability to China's criminal interrogation context is limited due to legal system differences and their civil-commercial law orientation [3]. Domestic research has achieved remarkable progress in practical applications but still faces challenges such as technological fragmentation and insufficient system integration. Current research is at a critical transition phase from isolated technical verification to holistic system architecture design, urgently requiring the construction of end-to-end solutions integrating speech recognition, content generation, compliance review, and data collaboration. Future development trends will exhibit three main characteristics: first, deep integration of multimodal information to achieve precise transcript generation; second, construction of dynamic compliance review and contradiction detection mechanisms based on deep semantic understanding; third, promotion of cross-departmental operational collaboration while ensuring data security. Against this backdrop, this study is dedicated to building a fully functional intelligent interrogation transcript system, driving the in-depth development of smart policing through technological innovation.

II. System Overall Architecture

A. Design Philosophy

The system adheres to the core design principles of "data-driven, intelligent collaboration, and business closed-loop," employing a hierarchically decoupled architecture pattern that divides the system into the data layer, model layer, service layer, and application layer. This layered design ensures clear responsibilities for each module, ease of maintenance, and independent scalability. It not only guarantees the stable and efficient operation of current functions but also reserves elastic space for integrating new AI capabilities and business requirements in the future. The overall architecture of the system is designed to build an end-to-end solution from multi-source data integration to intelligent application output.

B. Layered System Architecture

1) Data layer

As the foundation of the system, the Data Layer is responsible for the unified storage and management of multi-source data, providing comprehensive data support for upper-layer intelligent applications. This layer adopts a structured storage solution to systematically integrate four categories of data resources: raw data (audio and video files from the interrogation process), process data (speech recognition transcripts, basic case information), structured data (standardized transcript documents, evidence graphs constructed based on case elements), and knowledge data (laws and regulations database, historical case database). Through data integration interfaces with case management and intelligence systems, the Data Layer enables real-time acquisition and intelligent correlation of case-related information, establishing a comprehensive case data ecosystem [2]. In terms of data management, a tiered storage strategy is implemented, with strict masking procedures applied to sensitive data. Data security is ensured through encrypted transmission and storage mechanisms. Meanwhile, a unified data retrieval service is established, supporting rapid queries based on case elements and correlation analysis. This provides an accurate and complete data foundation for intelligent transcript generation, compliance review, and evidence analysis, effectively supporting the full-process business closed-loop.

2) Model layer

As the intelligent core of the system, the Model Layer adopts a collaborative architecture of "backbone model + specialized models" to provide comprehensive AI capability support for intelligent transcript processing. The backbone model employs a large language model (such as DeepSeek) deeply fine-tuned with massive judicial documents and legal regulations, [4] undertaking core tasks including semantic understanding, text generation, and logical reasoning to ensure generated content complies with legal norms and case logic. Around the backbone model, a cluster of specialized analysis models is constructed, including: text comparison models (enabling cross-document statement consistency verification), emotion detection models (analyzing emotional characteristics in speech texts), and evidence correlation analysis models (mining intrinsic connections between clues), forming a multi-level, multi-dimensional intelligent analysis system [5]. All models achieve efficient collaboration through a unified scheduling mechanism and are flexibly deployed within a hybrid cloud architecture. This approach not only ensures the high-performance computing requirements of core models but also meets the practical needs of grassroots units regarding response speed and resource consumption, providing a stable and reliable intelligent foundation for the system.

3) Service layer

The Service Layer adopts a microservices architecture, encapsulating the core AI capabilities of the Model Layer into a set of standardized, independently deployable service interfaces, ensuring the system exhibits high cohesion and low coupling characteristics. This layer provides four core services: the Speech Transcription Service enables high-precision real-time recognition of front-end audio streams; the Intelligent Generation Service automatically generates standardized transcripts by invoking large language models based on transcribed text; the Content Review and Compliance Service performs format validation and logical checks on generated content; the Association Retrieval Service supports intelligent case database queries and clue correlation based on semantics. Through the microservices architecture, each service module can be independently upgraded and elastically scaled, effectively handling business peak loads with load balancing mechanisms, significantly enhancing system reliability and maintainability. The standardized interface design facilitates subsequent functional expansion, providing stable and efficient technical support for upper-layer applications, and ensuring the smoothness and business continuity of intelligent transcript processing.

4) Application layer

As the direct interaction interface between the system and frontline police officers, the Application Layer is closely designed to align with actual case handling procedures. By integrating core functions such as speech transcription, intelligent Q&A, transcript generation and review, and clue association, it establishes an efficient, intuitive, and operable human-machine collaborative working environment. This layer provides an interrogation recording and real-time transcription interface that dynamically displays speech recognition results. The real-time follow-up questioning suggestion interface, based on semantic analysis of the dialogue content, actively recommends next-step interrogation questions to assist officers in constructing a solid factual chain. The transcript editing and review interface supports convenient revision of the system's auto-generated draft transcripts and highlights risk points and procedural flaws identified during compliance reviews. The case clue visualization interface intuitively presents the correlations among multidimensional elements such as involved persons, evidence, time, and location through knowledge graphs, aiding officers in conducting in-depth case analysis. Through the organic integration of the above functions, the Application Layer seamlessly embeds the intelligent capabilities of large language models into the entire daily case handling workflow of police officers, ultimately achieving the comprehensive goals of enhancing production efficiency, ensuring procedural standardization, and strengthening evidentiary support.

III. Core System Technologies

A. Intelligent Transcript Generation and Dynamic Completion Technology

Leveraging the deep semantic understanding capabilities of large language models, this technology achieves efficient automatic conversion from unstructured interrogation dialogues to standardized transcript documents. Based on a standardized template library covering various case types, the system can dynamically match the optimal template framework according to the case category and preliminary case description. Simultaneously, it accurately extracts key information such as the "Seven Key Elements" (Who, What, When, Where, Why, How, With Whom) from the real-time transcribed dialogue text and automatically populates the corresponding positions in the template, generating a standardized draft transcript containing core facts and significantly reducing repetitive manual entry [6]. Furthermore, the system possesses context-aware intelligent completion capabilities. When officers use abbreviations or markers, it can intelligently predict and complete unfinished sentences or standardize legal terminology based on semantic analysis of the current transcript. For example, it automatically expands the abbreviated "informed of their rights and obligations" into the complete rights and obligations notification statement. This technology not only effectively addresses the issues of missing elements and template rigidity in traditional transcript creation but also greatly enhances production efficiency while ensuring the standardization and normative quality of the transcript text. It provides crucial technical support for improving policing efficiency and ensuring law enforcement standardization.

B. Multimodal Interaction and Active Interrogation Assistance Technology

This technology elevates the system from a traditional passive recording tool to an intelligent interrogation assistant with active analytical capabilities, realizing a new intelligent interrogation model of human-machine collaboration. The system supports real-time semantic interaction and intelligent retrieval based on natural language. Officers can directly query using natural questions such as "What was the suspect's first mentioned criminal tool?" The system rapidly locates and highlights relevant passages in the transcript through deep semantic comprehension, significantly improving the efficiency of case detail review. More importantly, based on real-time semantic analysis of the interrogation dialogue, the system can dynamically assess the completeness and logical coherence of the current testimony [7]. When it identifies temporal gaps, plot jumps, or missing key elements, it proactively generates and recommends core follow-up questions, such as prompting "It is recommended to inquire about their specific whereabouts at XX time," thereby assisting officers in constructing a more rigorous and complete factual chain and evidence system. This capability transition from "passive response" to "active assistance" not only significantly enhances the intelligent level of the interrogation process but also contributes to deepening case investigation, ensuring that critical case information is comprehensively and accurately documented.

C. Compliance Review and Evidentiary Contradiction Analysis Technology

Driven by the dual engines of legal rules and knowledge graphs, this technology establishes an intelligent quality control system. The built-in legal procedure verification engine, based on knowledge graphs and rule engines, automatically scans draft transcripts for missing procedural requirements-including key steps such as rights and obligations notification, signatures, and fingerprints - triggering real-time alerts to effectively enhance law enforcement standardization. At the evidence analysis level, the system employs temporal analysis and semantic evolution modeling algorithms to achieve cross-document statement comparison [8]. Through deep semantic similarity calculation and logical relationship analysis, it accurately identifies statement discrepancies and contradictory conflicts, automatically generating a "Statement Discrepancy Analysis Report." This assists officers in identifying risks of retracted confessions and evidence collection gaps, establishing an evidence chain integrity assurance mechanism characterized by "risk prediction and dynamic early warning." Consequently, it significantly improves the quality of transcript evidence and its effectiveness in litigation proceedings.

D. Domain Model Optimization and System Secure Deployment Technology

Domain model optimization and system secure deployment technology serve as critical safeguards ensuring the reliable operation of the intelligent transcript system in real-world public security environments. At the model

optimization level, this research conducted deep continuous incremental training of the foundation large language model specifically for the judicial domain [5]. Utilizing high-quality domain-specific data including massive historical transcripts, laws and regulations, and judicial documents, the model has precisely mastered legal terminology, case narrative logic, and document formatting standards. This significantly enhances the professional accuracy of generated content while effectively suppressing "model hallucinations" [9]. In terms of secure deployment, the system implements stringent privacy computing and security reinforcement measures. All interrogation data is processed within the internal public security network, with comprehensive desensitization strategies applied to sensitive case data alongside encrypted transmission and storage mechanisms. Simultaneously, fine-grained permission management ensures data access is granted on a need-to-know basis. Considering the actual hardware conditions of grassroots units, the core model has been optimized through quantization and compression techniques, adapting to computing resources of various configurations while maintaining performance standards. Ultimately, this establishes a professional deployment solution that both meets policing security specifications and demonstrates practical applicability in operational contexts.

IV. Conclusion and Prospects

This research addresses key challenges in police interrogation transcript creation—including inefficiency, lack of standardization, and poor information coordination—by developing an intelligent interrogation transcript system based on Large Language Models (LLMs). Through a hierarchically decoupled system architecture, it establishes a closed-loop process from multi-source data integration to intelligent application output.

At the core technological level, the system achieves four major breakthroughs:

1.Intelligent transcript generation and dynamic completion technology enables efficient conversion from dialogue flow to standardized documents.

2.Multimodal interaction and active interrogation assistance technology upgrades the system into an active interrogation assistant.

3.Compliance review and evidentiary contradiction analysis technology provides intelligent quality assurance.

4.Domain model optimization and secure deployment technology ensures professional reliability.

Experimental results demonstrate that the system achieves remarkable outcomes in two key aspects: enhancing efficiency (reducing transcript creation time by over 50%) and ensuring quality (improving transcript completeness and procedural compliance rates). These findings fully validate its technical feasibility and practical value, providing a concrete technical solution for public security organs to advance "technology-driven policing" and standardize law enforcement practices.

This research has constructed an intelligent closed-loop system covering the entire workflow of "intelligent generation-compliance review-evidence analysis," featuring three core innovations:

1.Technical Pathway: It achieves case-adaptive dynamic generation, overcoming the challenges of missing elements and template rigidity through the LLM's dynamic template matching and element extraction.

2.Model Capability: It builds a vertical domain-specific large model dual-driven by legal rules and knowledge graphs, embedding legal knowledge directly into the model's reasoning process to achieve intelligent reinforcement of law enforcement standardization.

3.Analytical Method: It introduces an innovative contradiction detection mechanism based on temporal analysis and semantic evolution modeling, establishing an evidence chain assurance system characterized by "risk prediction and dynamic early warning."

However, the system still faces challenges including factual deviations caused by model "hallucinations" [9], legal knowledge lagging behind the latest updates, optimization of human-machine collaboration workflows, and adaptability to complex scenarios [3]. Future work requires continuous domain-specific training, establishing dynamic knowledge updating mechanisms, enhanced personnel training, and scenario-specific optimizations to drive further system refinement.

By leveraging intelligent generation and automated review capabilities, the system liberates police officers from burdensome administrative tasks and significantly reduces transcript preparation time. Simultaneously, through element integrity verification and procedural compliance assurance, it enhances evidence quality at the source, laying a solid foundation for successful case litigation.At the law enforcement standardization level, the system transforms abstract legal requirements into concrete automated checks, providing officers with real-time procedural guidance that effectively reduces procedural violations and flaws, thereby systematically enhancing law enforcement credibility.Within the policing ecosystem, the system promotes the formation of a new "human-machine collaborative" intelligent case handling model [7], facilitating the transition toward "data-driven and intelligently-assisted" policing operations. Its hierarchical microservices-based architecture design possesses high replicability, enabling adaptation to the operational needs of public security organs at different levels while providing technical support for horizontal expansion into judicial scenarios such as prosecutorial inquiries and court trials [2], demonstrating broad application prospects and social benefits.

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

Junquan Zhou conceived the study and drafted the manuscript; Shanting Song collected and organized materials; Shuifeng Zhang (corresponding author) supervised the research and revised the paper; Qingyang Gao investigated technical literature; Yuekai Ma studied core functionalities; Zhengnan Tian researched deployment and security; Buyun Chen conducted the literature review and compiled references; all authors discussed the findings and approved the final manuscript; all authors had approved

the final version.

REFERENCES

[1] "'AI Intelligence+' assists public security organs in law enforcement and case handling," *Guangxi Legal Daily*, 2025-08-27(01). 2025 (In Chinese)

[2] L. N. Tian, "On new challenge to police interrogation based on construction of trans-agency sharing platform of criminal case information," *Journal of Zhejiang Gongshang University*, vol. 01, pp. 26–35, 2019. (In Chinese)

[3] J. H. Zhan and Y. X. Qiu, "Risks of AI-embedded investigative interrogations and regulatory frameworks," *Journal of Zhejiang University (Humanities and Social Sciences)*, vol. 55, no. 5, pp. 112–127. (In Chinese)

[4] Y. Xiang and Y. X. Sun, "Application of the DeepSeek large language model in social risk assessment for arrest," *People's Procuratorial Semimonthly*, vol. 7, pp. 35–40, 2025. (In Chinese)

[5] J. B. Nie, "Architecture, paradigm and security application for the large model of artificial intelligence investigation," *Journal of People's Public Security University of China (Social Science Edition)*, vol. 41, no. 1, pp. 86–94, 2025. (In Chinese)

[6] Z. Z. Xu, "Application of speech recognition technology in audio-visual archives management," *Beijing Archives*, vol. 8, pp. 30–32, 2023. (In Chinese)

[7] S. F. Ou, "Research on trust in investigation and interrogation," *Journal of People's Public Security University of China (Social Science Edition)*, vol. 41, no. 4, pp. 31–39, 2025. (In Chinese)

[8] X. Q. Bi and C. T. Li, "Analysis and research on interrogation discourse from the perspective of conversation cooperation principle," *Journal of People's Public Security University of China (Social Science Edition)*, vol. 41, no. 4, pp. 22–30, 2025. (In Chinese)

[9] Y. Li, "The path and limits of DeepSeek in assisting electronic evidence review," *The Chinese Prosecutors*, vol. 9, pp. 15–19, 2025. (In Chinese)