Operator Best Practices

AI Large Model Empowering Verticals Use Cases

Public Administration and Management | Customer Service and Smart Assistants | Social Services and Infrastructure

GSMA

Foreword

Sihan Bo CHEN
Head of Greater China
GSMA

As the mobile communications industry enters the era of 5G and 5G-Advance, AI is revolutionizing sectors at an unprecedented pace. Amid this sweeping wave of change, Chinese operators are leading and exploring new business directions and models, shifting from traditional mobile operators to high-tech information service providers. I am honored to witness and participate in this historical evolution, particularly in the exciting journey of empowering vertical industries with 5G and AI technologies.

Since the 5G boom, we have released the "5G in Verticals in China" globally for three consecutive years. These cases highlight innovative 5G applications across various sectors, showcasing China's strategic insights in harnessing the full potential of 5G. This offers valuable benchmarks for the global advancement and integration of 5G technology in industry applications.

The application of AI large models in vertical industries saw explosive growth in 2024. This growth is driven by two key factors. Firstly, the maturation and optimization of large models have led to unprecedented accuracy and efficiency in practical applications, laying a robust foundation for deep industry integration. Secondly, the digital age has spurred a sharp increase in demand for intelligent, automated, and personalized services across various sectors, creating a vast market and limitless possibilities for the widespread adoption of large models.

In this context, the role of operators is undergoing a profound transformation. Moving beyond their traditional roles as network connectivity and infrastructure providers, they are evolving into information service innovators, data providers, and intelligent service providers. With advantages in data resources, network coverage, and extensive user bases, operators are actively exploring and implementing large model technology in vertical industries, infusing new vitality into their digital transformation.

Recognizing this shift, GSMA Greater China has released the "Operator Best Practices: AI Large Model Empowering Verticals Use Cases", analyzing nearly 20 typical cases in fields such as government services, customer services, industrial applications, healthcare, education, cultural tourism, and urban governance from the operators' perspective. This document not only highlights the remarkable application outcomes of large models across multiple industries but also delves into the industry challenges, technical solutions, key innovations, and core values underpinning these successful cases. It offers valuable experience and insights based on Chinese wisdom for global operators and vertical industries.

AI large models drive modernization of governance through smart customer service and smart search, enhancing the intelligence of government services. Technologies such as precise user intent recognition and multimodal interaction significantly improve customer service efficiency and quality. In the industrial sector, AI large models enable real-time monitoring and smart warning, reducing safety risks and lowering labor costs. In healthcare, education, and cultural tourism sectors, AI facilitates intelligent transformation and enhances user experience through smart diagnostic assistance, personalized teaching, and the construction of smart tourist attractions. Moreover, AI large models show immense potential in areas like urban governance, flood management, dialect preservation, and smart fraud prevention, supporting social governance, cultural heritage, and security assurance.

The futuristic cover of this report, depicting digital messengers racing on a digital track, was created by our GSMA Greater China members using AI tools. It exemplifies the dual nature of AI large models in empowering vertical industries. AI is advancing rapidly, with core elements like algorithms, computing power, and models continuously breaking technological boundaries and demonstrating powerful potential. But as we look at the practical applications of AI, it resembles a long marathon. Despite rapid technological advancements, operators face countless challenges in empowering all industries with large models, such as effective implementation and deep integration with industries, necessitating long-term commitment and strategic vision.

This report is critical in promoting the application of large models in vertical industries. It offers practical guidance and case references for the application of large models, while inspiring new ideas and directions. We believe that as this report reaches its audience, more operators and vertical industries will recognize the value and potential of large models and commit to their application and innovation.

We extend our gratitude to all the teams and individuals involved in compiling and publishing this report. Your efforts and dedication have enabled the sharing of these valuable cases and experience with a wider audience.

Despite the long road ahead, we remain excited about the deep integration and innovative development of large models within vertical industries. As technology progresses and application scenarios expand, large models will increasingly play a vital role in driving digital transformation and upgrades across more industries. We're looking forward to seeing more innovative applications and business models emerge, bringing greater value and contributions to industries and society.

Contents

Heilongjiang Digital Government Project

China Mobile focuses on addressing urgent and difficult issues faced by both enterprises and the public, constructing an efficient digital governance ecosystem. China Mobile's independently-developed Jiutian Haisuan Government Service Large Model incorporates AI technologies to realize "centralized online services, administration, and operation", making government services in Heilongjiang people-centric, convenient, and efficient.

Wang Yun
Deputy Secretary of the Party Committee and General Manager
China Mobile System Integration Co., Ltd.

SOLUTION PARTNERS

Digital government is a basic and guiding project in China's endeavor to build a digital powerhouse with Internet power. It is a crucial step in promoting the modernization of China's governance system and capacity. In January 2024, the State Council issued the Guiding Opinions on Further Optimizing Administration Services to Enhance Administrative Efficiency and Promote "Efficient Services". It explicitly requires the exploration and application of technologies like large language models to improve online customer services in intent recognition and precise answering capabilities. With optimized intelligent question-answering, search, and guidance, enterprises and the public can get their things done more efficiently.

Heilongjiang started building a digital government in 2023 and has since pursued data-driven innovation to realize faster AI application. With a wealth of experience in digital government construction, China Mobile has developed the Jiutian Haisuan Government Service Large Model to empower the digital government endeavor in Heilongjiang throughout the process and at deeper levels. As a result, the local government can provide smarter and more convenient public services and realize intelligent social governance. In 2023, an AI platform built by China Mobile was launched to integrate with ten business systems, including the Heilongjiang unified government service network, intelligent customer service, and service centers. It can support at least 100 users (citizens, government departments) at the same time. An intelligent addition to Heilongjiang's digital government initiative, the platform has made simpler and faster government services possible.

CHALLENGES

Yet AI-empowered digital government construction is faced with some pain points:

Uneven distribution of existing capabilities and lack of a unified platform

The disparity in digital government construction across different cities necessitates a centralized AI platform for the sharing and joint development of AI capabilities among existing and new systems.

No suitable algorithms to support extensive intelligence demand

Current AI demands in various digital government applications, including machine vision, speech semantics, and natural language processing (NLP), are substantial. Existing capabilities are inadequate to meet these demands, and an integrated AI platform is necessary to provide foundational capabilities.

Extensive digital government applications that are not innovative and intelligent

Digital government applications require improvement in innovation and lack fundamental AI capabilities such as facial recognition, text recognition, and natural language understanding. There is a deficiency in innovative applications and decision-making assistance based on new AI technologies like large models, hindering the provision of digitalized and intelligent universal services.

SOLUTION AND VALUE

Based on years of experience in digital government construction, China Mobile has developed the industry-specific Jiutian Haisuan Government Service Large Model by leveraging a general large model, incorporating government data for fine-tuning, and introducing constraint models in the government domain to limit outputs. With in-depth industry intelligence, government information field, and diverse interaction, the model can empower the three typical application scenarios of "centralized online services, administration, and operation" as shown in Figure 1.

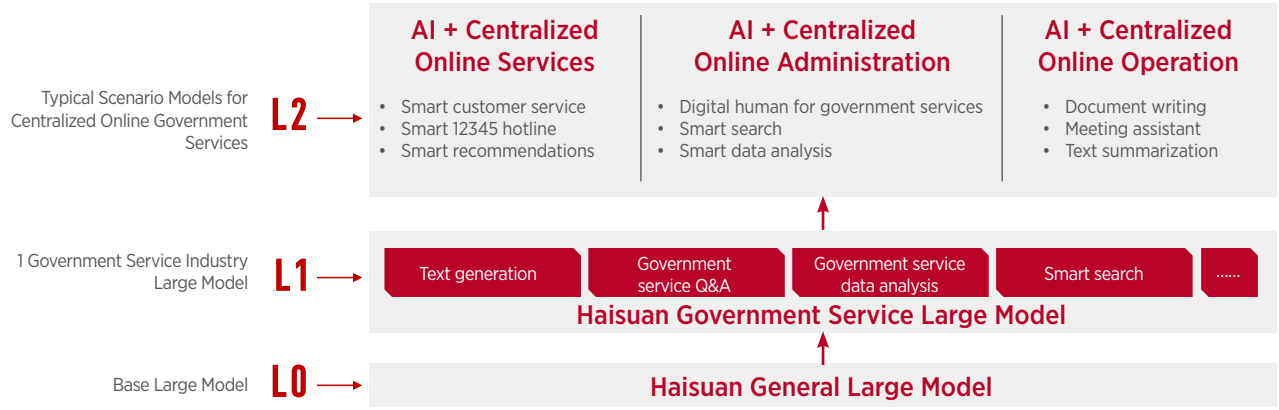


Figure 1: Architecture of China Mobile Jiutian·Haisuan Government Service Large Model

01 Overall structure/solution

Built on the training and inference platforms, the government intelligence system combines large and small industry models to support the three typical application scenarios. As seen in Figure 2, the integrated "AI + government" architecture provides smart capabilities for information input, understanding, analysis, and output. The large industry model is directed toward the government service large model, and the small industry model encompasses more than 50 visual capabilities, 10 speech capabilities, and 20 natural language processing capabilities.

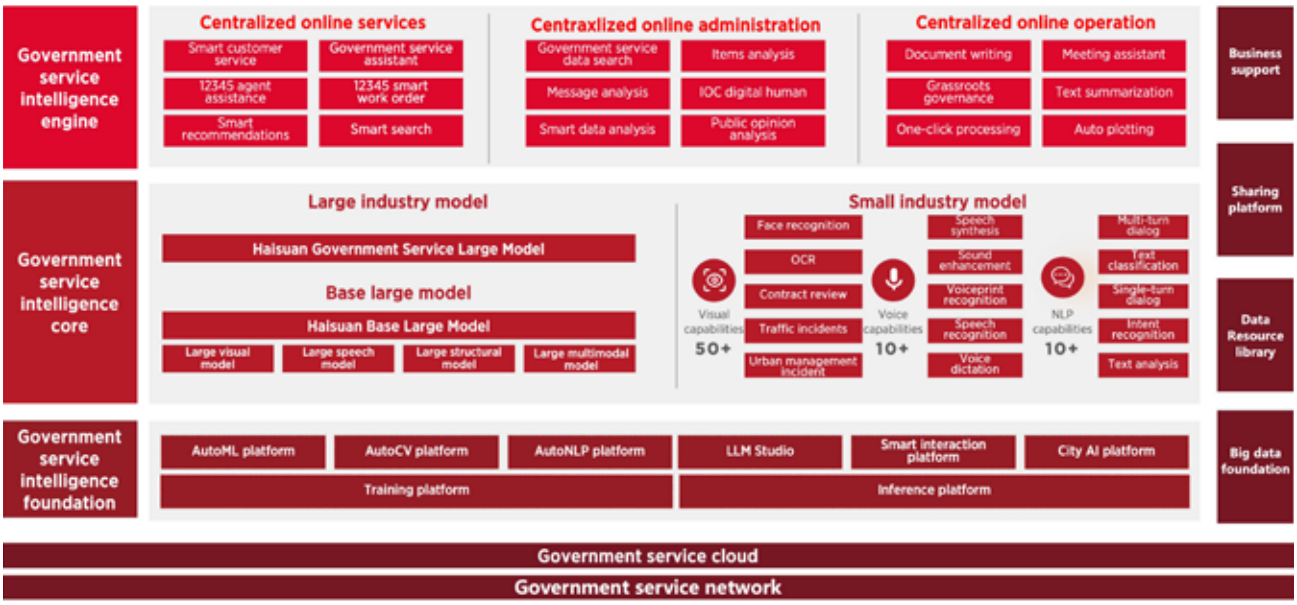


Figure 2: China Mobile "AI + Government" Intelligence System

02 Application scenarios

Centralized online services

Smart customer service for government affairs is developed to address public complaints about inquiry difficulties and cumbersome procedures, as shown in Figure 3. A service funnel system is created based on interviews with government staff to provide more personalized conversations with less jargon for smart customer service, making the interaction between users and customer service more human-like; and to enhance the understanding of government service terminology by customer service, providing citizens with easier-to-understand services. In this way, the government service assistant is truly smart.

Figure 3: Government Service Customer Service Interface

Centralized online administration

An intelligent search engine, as shown in Figure 4, is created to address the challenges such as difficulty in data retrieval due to government data aggregation in large amounts, difficulty in topic-based analysis due to complex and dispersed data types, and difficulty in fast data fetching and use due to lack of data tools. Data from departments across the province is aggregated through channels such as the digital government index database, data resource library, and the policy and regulation database. A tag system is constructed for intelligent search of the aggregated data. Natural language queries are enabled with large language model technology, allowing users to easily search for desired data anytime, anywhere.

Figure 4: Longzheng Zhisou Interface

Centralized online operation

Government officials are faced with lots of writing tasks but they are done at a low efficiency. An AI-powered document writing assistant "Longzheng Zhiwen" as seen in Figure 5 is developed by fine-tuning the large model with government text data accumulated by China Mobile for years and leveraging the large model's text generation capabilities. The tool provides functions such as policy analysis reports, drafting of meeting speeches, and generation of meeting minutes to meet the daily writing needs of government officials.

Figure 5: Document Writing Assistant Interface

03 Key innovations

Deep industry intelligence

The model deeply integrates "government policies, government affairs, and government data storage," making the overall business process flexible and easy to use. When natural language commands are sent to the large model, the model can reach deep in the database, link multiple sources of complex and heterogeneous data tables, and quickly obtain intuitive data analysis results.

Government information field

Professional knowledge in the government domain within the information field is utilized for learning-based enhancement and generalization alignment, and private domain data is combined as the final feedback to users. Thanks to the scheduling capability of the government information field, scattered data is aggregated to solve all user inquiries within the field. The scope of government services is expanded for proactive service. It also realizes trustworthy responses to government issues without leaving the "field," and guarantees secure and controllable government services.

Diverse interaction

Diverse interaction modes like task-based dialog, large model generative dialog, and graphical interaction are integrated to make the dialog system more flexible and convenient.

04 Business models

Jiutian Haisuan Government Service Large Model provides standardized products and services and customized implementation services. Standardized products and services include the implementation of existing standardized products such as the large model management platform, Haisuan Government Service Large Model, and large model scenario applications. Customers can choose desired modules based on their needs and China Mobile charges fees based on the selected functional modules.

The customized implementation service provides customized development services for large model training, tuning, fine-tuning, and scenario applications based on actual needs, with charges based on specific workload.

05 Core values

Built on the Jiutian Haisuan Government Service Large Model, the digital government construction project in Heilongjiang includes a system of "platform + algorithm + application". Overall and incremental improvement has been seen in the province's government services. Applications like smart customer service, Longzheng Zhisou, digital personnel, and document writing with local characteristics have been developed to empower intelligent processing optimization and content generation in the government domain. As a result, the Heilongjiang government can effectively fulfill its duties, increase the quality and efficiency of public services, and enhance public satisfaction and trust. By making government services more standard, convenient, and digital, the Heilongjiang government can better serve the public, thus injecting new energy and vitality into their high-quality development.

REFLECTION

The Jiutian Haisuan Government Service Large Model has effectively promoted digital government construction in Heilongjiang, providing more intelligent and convenient government services to the public. The model was included in the blue book "Research Report on Governance Modernization in the Digital Age: Practice and Prospects for the Application of Large Models in the Government Domain (2023)" as a typical case in the application of government service large model. It has set a benchmark in this respect in China. It was also selected in IDC's "Analysis on Application of Large Models in Digital Government in China, 2024" as an exemplary case in the government domain.

Taking the opportunity of the project in Heilongjiang, we will expand into the large model market. The Jiutian Haisuan Government Service Large Model has been selected in the "2024 Priorities for Digital Government Construction in Heilongjiang". The focus next will be on customer contract signing and project implementation. We are also actively tapping business opportunities in other provinces. The projects in Hubei Province, Anlu of Hubei Province, Huangdao District of Qingdao, and the Xi'an 12345 Hotline have been confirmed to include large model capabilities, with the preparation of implementation plans underway.

Intelligent Application of Government Services Based on the Xingchen Large Model

The government is promoting the intelligent transformation of government services to enhance efficiency and quality. The public and enterprises seek efficient, convenient, personalized, and transparent services. China Telecom leverages large model technology to deeply learn government data, accurately identify public inquiry intents, and respond in a friendly, personalized manner, significantly improving the effectiveness of government services and meeting the needs of various sectors.

Qing Fei
Chief Expert of Large Models, China Telecom Artificial Intelligence Technology Co., Ltd.

SOLUTION PARTNERS 中国电信 CHINA TELECOM

With the Xingchen Semantic Large Model, China Telecom was entrusted by a city's Government Services and Data Management Bureau to build a smart government service project that serves a citizen service hotline. It encompasses three main areas: the construction of a citizen request application system, the development of smart citizen request capabilities, and the management of the citizen request and government service operations team.

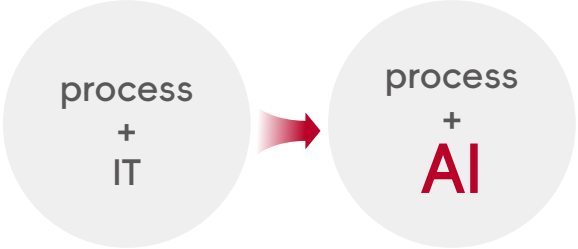
Building on the earlier 12345 hotline project and phase one of the livelihood demand system platform undertaken by China Telecom, the Government Services and Data Management Bureau planned 25 large model application scenarios based on 8 role types, 17 service processes, and 44 business nodes. Through the open competition mechanism to select the best candidates, the plan is to advance eight pilot application scenarios in two phases. Additionally, the project aims to guide and drive the complete process reengineering of demand services using large model technology, upgrading the demand platform from "process + IT" to "process + AI" to meet the optimization requirements of 12 business capabilities for livelihood demands.

8
role types

17
service processes

44
business nodes

25
large model application scenarios



CHALLENGES

The 12345 government service hotline has completed the first phase of transitioning from traditional systems to AI upgrades. Some scenarios have been implemented using small model-based AI technologies, such as smart Q&A customer service, smart IVR (interactive voice response), and smart work order extraction. AI assistants have improved the efficiency of the process. However, the limited capabilities of traditional small models mean that these implemented scenarios face challenges in effectiveness enhancement. New scenarios, such as smart knowledge compilation and smart report generation, are not possible with small models.

Challenges in terms of the application scenarios:

- **Smart customer service:** Users often express complex intents, engage in multi-turn dialogs, and convey vague intents during policy Q&A conversations (such as with smart acceptance assistants and smart policy Q&A). Small models struggle to achieve satisfactory results in handling complex tasks.
- **Smart acceptance assistant:** For long digital strings (such as phone numbers, ID numbers, etc.) and information like addresses or entity names in work orders, small models find it difficult to extract complete and accurate digital information from users with varying expression habits.
- **Smart compilation:** The workload for interpreting, organizing, and compiling documents generated from government policies, business upgrades, or adjustments is vast. Traditional small-model AI has a low processing efficiency, making it time-consuming.
- **Smart summarization:** After transitioning to manual processes, summarizing the content of human-machine dialogs and quickly forming work order summaries from agent-customer conversations is challenging.
- **Smart report generation:** Automatically generating suitable analysis reports from the operational results of the government service 12345 hotline is nearly impossible with small models, making it difficult to save manpower.

Challenges in terms of technical implementation:

- **Difficulty adapting to large datasets: Small models struggle to process large datasets and cannot fully utilize all the information they contain.**
- **Demanding customization requirements:** Small models have weak generalization capabilities, requiring a separate model for each scenario, which increases management and application complexity.

Generative large language models have demonstrated strong intent understanding and natural dialog capabilities in multi-turn conversational scenarios. Establishing a large industry model for government services, encompassing Q&A, work order extraction, and classification, can significantly enhance citizen satisfaction and reduce labor costs for agents.

SOLUTION AND VALUE

01 Overall structure/solution

The smart customer service robot is central to the city's 12345 quick response service for citizen requests. It aims to utilize large models and other AI technologies to train and learn government service knowledge, enabling the robot to deeply understand the intentions and needs behind citizens' requests. The goal is to provide correct, complete, secure, and controllable responses based on defined government service knowledge. The smart customer service robot is used in scenarios like "housing fund", "social security", "medical insurance", enhancing the accuracy, completeness, and security of intelligent customer service responses, and improving service capabilities. The China Telecom Institute of Artificial Intelligence (TeleAI) has built a large model smart Q&A system based on its self-developed Telechat 52B large model, which has been deployed in the production environment of a Government Services and Data Management Bureau to assist agents in resolving citizens' requests.

The traditional smart customer service process includes organizing the Q&A knowledge base, configuring it into the dialog system, and matching citizens' inquiries online to recommend suitable answers. During knowledge base configuration, operators face a significant workload in structuring, interpreting, organizing, and compiling documents generated from government policies, business upgrades, or adjustments. During online matching, various technologies, including dialog intent recognition, semantic similarity matching, and answer recommendation, are used for smart responses.

This project employs Multi-Agent smart Q&A technology based on large models to implement the intelligent government service Q&A system. The overall technical process includes modules for knowledge base construction agent, document understanding and retrieval agent, and Q&A agent, as illustrated in Figure 1.

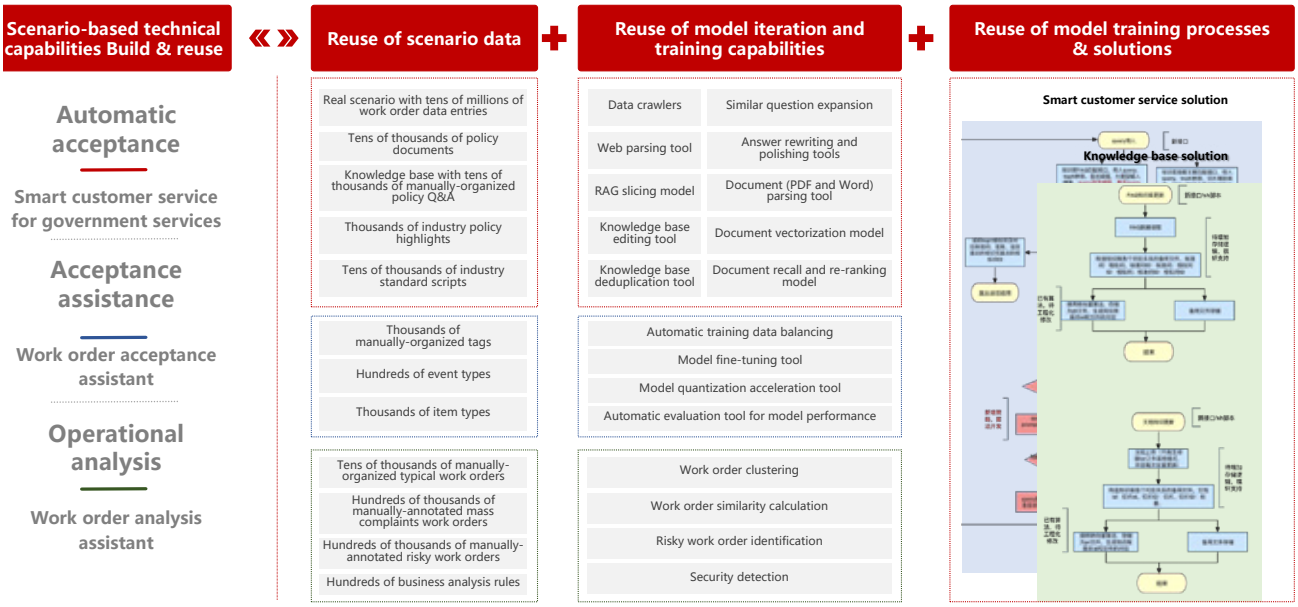


Figure 1: Multi-Agent Q&A Solution

Knowledge base construction agent

The agent collects policy-related knowledge from various sources, including encyclopedias, books, and government websites such as the Housing and Urban-Rural Development Bureau, service halls, and Healthcare Security Administration. At the same time, it supports experts from the Government Services and Data Management Bureau in organizing a Q&A knowledge base based on response data. To reduce the cost of organizing Q&A pairs in government service centers, a large model knowledge compilation approach is adopted. The large model first suggests Q&A pairs from the original policy documents, which are then reviewed by quality inspectors. This method helps efficiently accumulate thousands of high-quality Q&A pairs.

Document understanding and retrieval agent

This project collects policy documents in various formats, including PDF, Word, and PPT. The diverse layouts and elements of these documents, such as graphics, text, and tables, significantly affect the quality of Q&A outcomes. A large model-based document layout analysis strategy is adopted to identify key policy points and segments within the documents while filtering out irrelevant URLs and advertising text. Techniques like MinHash and text similarity calculations are used to deduplicate documents at the chapter and paragraph levels, resulting in high-quality policy document knowledge.

Q&A agent

The government service Q&A system faces challenges due to frequent policy updates and diverse information sources. A key challenge is enabling the large model to learn and understand the latest policy knowledge for accurate responses. This project utilizes a two-phase post-training approach combined with knowledge base enhancement to inject policy document information into the large model, reducing its hallucinations. Phase One: The large model is continuously pre-trained using the original policy texts, allowing it to fully understand the policy information. When given directives to explain policy points, it can accurately reproduce corresponding segments from the policy documents. Phase Two: A Q&A corpus is constructed by using the original policy texts and fine-tuning the large model with specific instructions, enabling it to comprehend citizen inquiries and provide relevant policy answers. To enhance the quality of responses, this project involves multiple rounds of detailed annotation to ensure answers are correct and relevant, further improving their structure and logic.

To better integrate the government service knowledge base, a hybrid technology framework combining large and small models is introduced. For questions that match the knowledge base, the corresponding answers are directly provided. For citizen inquiries with low similarity to existing Q&A pairs, the large model generates answers.

02 Application scenarios

Smart classification and distribution

After citizens submit work orders online, distributors need to classify and allocate these work orders to the responsible departments. The large model assists distributors in making the classification and distribution process more efficient and accurate.

Smart annotation

After citizens submit work orders, distributors must annotate them for subsequent statistical analysis. The large model aids in making the annotation process more efficient and accurate.

Smart report writing	Business personnel need to regularly write "Citizen Demand Operation Analysis Reports". The large model assists in report writing, improving work efficiency.
Smart customer service robot	Leveraging the large model's capabilities enhances existing intelligent customer service robots, improving government service capacity and citizen satisfaction.
Smart customer service robot	For complaint hotlines, operators need to submit work orders based on citizen feedback. The large model automatically identifies and extracts key information from citizen voice feedback to improve operator efficiency.
Smart acceptance assistant robot	Safety content review is required for smart responses and feedback on request handling, to detect and filter illegal, harmful, or inappropriate content, such as pornography, violence, and hate speech.
Smart knowledge base	1. It needs to provide the intelligent generation of similar questions and corresponding answers based on prompts. 2. It supports plugins that integrate specialized knowledge from a knowledge platform, allowing the large model to understand and respond to citizen questions based on professional articles, Word documents, PDFs, etc.

03 Key innovations

Key technical innovations of the Multi-Agent smart Q&A system:

The project introduces a large model-based document-targeted structured knowledge generation method, capable of extracting multi-level (such as chapters, paragraphs, and events) and multi-granularity structured knowledge from policy documents. This enhances the accuracy and comprehensiveness of knowledge representation.

To address the input length limitations of large models, it innovatively applies rotary positional encoding technology to extend the input of re-ranking networks, allowing for longer document segment inputs. This ensures that segments encompass sufficient semantic information, enabling efficient and precise reordering of policy document segments to improve retrieval effectiveness.

Through continuous pre-training and instruction fine-tuning, we inject the latest policy knowledge into the large model, reducing its hallucinations and improving the accuracy of policy understanding and responses. To tackle the semantic differences between citizens' colloquial expressions and the formal language of policy documents, it introduces a question rewriting mechanism and employs a parallel document slicing recall strategy, significantly enhancing the recall rate of documents. For the government service knowledge base, it designs a hybrid technology framework that efficiently integrates the strengths of both knowledge bases and generative models, providing flexible and accurate Q&A services.

An innovative semantic slicing algorithm for policy documents utilizes the large model for fine-grained key point extraction at the paragraph level. Based on these key points and summaries, it performs semantic segmentation, ensuring coherence within segments and adaption to the frequent updates of policy documents.

These innovations collectively form an efficient, precise, and highly adaptive Multi-Agent Q&A system, bringing new technological breakthroughs and practical application value to the field of government service Q&A.

04 Business models

Providing a unified government service large model capability base

By leveraging the data and capabilities accumulated in projects, we will build core government service capabilities across five scenarios: smart dialog, knowledge management, smart work orders, smart analysis, and smart quality inspection. This unified government service large model capability base will comprehensively innovate existing government service processes, empowering all government services.

Providing standalone interface call for government services

We offer a method for standalone capability calls, targeting newly established government service systems with services for individual government service scenarios. Government agencies can add calls to the large model service interface at business process nodes without replacing their existing systems, directly empowering their application systems. This approach also supports the joint application of multiple capabilities to empower client business scenarios and improve effectiveness.

05 Core values

Government services face numerous challenges today. The "collaboration of large and small models to enhance government service upgrades" solution integrates the strengths of large and small models to innovatively address these key issues.

Core values:

1. Significantly improving the handling of complex inquiry scenarios:

The collaboration of large and small models can quickly and accurately understand the questions posed by citizens and provide friendly, personalized responses. Large models excel at understanding complex intentions, such as multiple intents from a single question or long contextual intents, and generating friendly and personalized responses, while small models focus on accurately identifying simple questions in specific scenarios and quickly locating user intent at a low cost. The collaboration of the two types of models ensures accuracy while offering friendly, personalized services.

2. Enhancing user experience in intelligent government services:

The collaboration between large and small models provides a more intelligent and personalized service experience. Small models accurately identify simple questions, and the large model processes the feedback in a friendly manner, making government services more approachable and improving the inquiry experience for citizens.

3. Reducing government service costs

Most inquiries in current government services involve simple, repetitive questions. With the large and small model solution, simple questions can be addressed by small models, completing the business loop without relying on the more expensive large models. This ultimately enhances service effectiveness and experience without significantly increasing service costs.

4. Driving innovation and growth:

The solution that integrates large and small models enhances the efficiency and quality of government services. It not only addresses existing issues but also leaves room for future technological innovation, promoting further development and upgrading of business models and administrative management practices.

In summary, the core values of the "collaboration of large and small models to enhance government service upgrades" solution lie in its efficient, personalized, and low-cost service processing capabilities. It effectively addresses pain points in government systems while improving user satisfaction, leading the government service sector toward a smarter and more efficient future.

REFLECTION

Reflections

- The relationship between parameters and capabilities: When designing models, appropriately increasing the number of parameters generally leads to significant improvements in model performance and generalization ability. However, the key is balance. Too many parameters require a larger dataset and can lead to computational resource demands.
- Advantages of the Multi-Agent strategy: Combining multiple agents (including but not limited to large models) often yields more stable output than relying solely on a single large model. This approach effectively reduces risk and enhances overall performance, especially in complex tasks and highly uncertain scenarios.
- The core role of data quality: Building high-quality datasets is fundamental to ensuring that models perform well. Accurate, comprehensive, and high-precision data not only significantly enhances the efficiency of model training but also improves the reliability and stability of prediction outcomes.
- The importance of team collaboration: Establishing a unified team across departments (including but not limited to delivery, product, data science, algorithm development, and engineering practices) is crucial for optimizing models and achieving rapid deployment. Close collaboration ensures consistency in strategy, effective implementation of technical solutions, and timely problem resolution, thereby increasing the overall success rate of projects.

Follow-up plans

To quickly promote the existing collaborative solution of large and small models across more regions and business scenarios, continuous advancement in standardization and productization is needed. Additionally, for various government service scenarios, ongoing iterative innovation is required for existing solutions, focusing on capability building and effect optimization in the following areas:

- Technical optimization and standard establishment: First, standardize government service scenarios to enable quick cold starts and direct reuse of existing service scenarios. Second, during the implementation of services based on the collaboration of large and small models, standardize the deployment processes so that more teams can rapidly apply the collaborative methodologies to other government service projects. Last, develop unified service interfaces and technical specifications when connecting to multiple systems, ensuring compatibility and consistency among platforms from different locations.
- Productization accumulation: Collect specific needs and user feedback from various levels of local government through surveys and interviews. Based on these needs, develop specialized modules or plugins to create generalized products that meet the requirements of different regions and fields (e.g., social security, taxation).
- Platform promotion and training system establishment: Collaborate with provincial and municipal government departments to share technological achievements and implementation experiences. Establish a comprehensive training plan with online courses and seminars, provide training resources for AI applications and platform operations, organize offline training meetings or workshops, and set up training centers in various locations.
- Continuous iteration and technological innovation: Based on exploratory research into the business needs of various regions and fields, continuously refine the technical architecture solutions for both large and small models. Utilize data generated by the platform for in-depth analysis to identify areas for optimization and technical bottlenecks.

Government Service AI Agent Platform Built on the Yuanjing Government Service Hotline Large Model

Traditional government service hotlines often suffer from low human efficiency, high line congestion, and cumbersome documentation processes. To enhance the efficiency of government services and social management, China Unicom has introduced the Government Service Large Model that leverages the advantages of digital intelligence technologies tailored to government service scenarios, offering capabilities like agent knowledge assistance, smart task dispatching, and smart data inquiries. These features enable comprehensive and intelligent upgrades across all aspects of interaction with citizens, customer service, and government service centers to boost the efficiency of government management and services while increasing citizen satisfaction with business processing.

Shi Shuming
Chief Scientist of Large Language Modelsof China Unicom AI Innovation Center

SOLUTION PARTNERS

Backed by China Unicom's advanced AI capabilities and Yuanjing Government Service Large Model, the Government Service AI Agent Platform built on the Yuanjing Government Service Hotline Large Model is mainly designed for government scenarios and offers smart work order filling, smart task dispatching, knowledge assistance for agents, digital intelligence analysis for hotlines, and smart data inquiries. It empowers applications through smart robots and allows for flexible integration and customization of robotic solutions to meet specific needs. This initiative provides basic AI capabilities and product solutions tailored for the government service sector.

The Yuanjing Hotline Large Mode and Government Service AI Agent Platform have been implemented in many provincial and municipal government hotline scenarios. Represented by the Liaoning 12345 Hotline, as shown in Figure 1, the scenarios of knowledge assistance for agents and smart work order filling have been fueled by features like automatic work order filling, smart task dispatching, and assisted Q&A. The accuracy of work order filling reached 90%, with completeness improved by over 30%. Knowledge assistance accuracy has risen by 60% compared to traditional methods, while the average customer service call duration has decreased from 186 seconds to 133 seconds. For smart analysis and inquiries for hotlines, as shown in Figure 2, the analysis of over 800,000 previous work orders and 120,000 community complaint records has produced six major report categories, offering the hotline center with conversational data querying, analysis, and visualization.

Figure 1: Liaoning 12345 Hotline Platform

Figure 2: Liaoning 12345 Hotline Digital Intelligence Analysis Platform

CHALLENGES

National policies

The central government has proposed initiatives to integrate government service hotlines to enhance service convenience and efficiency. This includes improving the effective interaction between the 12345 Hotline and the 110 Emergency Call Service Counter, advancing digital government construction, and promoting the digitalization and intelligence of public services. The application of 12345 hotline data is expanding into areas like market regulation and administrative law enforcement. Additionally, various provinces and cities nationwide are introducing policies to significantly boost the intelligence of government service systems.

Industry trends and needs

Government service hotlines play a crucial role in understanding urban, social, and public sentiments, essential for more effective governance. Various regions are accelerating the consolidation of multiple hotline numbers into a unified system, enabling multi-channel integration and single-number responses. These hotlines maintain constant "connections" with citizens, capturing public sentiment in real-time and addressing governance blind spots promptly. At the same time, government service hotlines are crucial for connecting different levels, enabling cross-departmental and cross-regional data flow through platform interconnection, data sharing, and information exchange. This facilitates coordinated interaction among government departments, allowing them to respond to public demands promptly and efficiently, thereby enhancing collaborative governance.

Application scenarios

The Government Service AI Agent Platform built on the Yuanjing Government Service Hotline Large Model mainly addresses the following user pain points:

Self-service for citizens: Current dialog systems rely on predefined processes, with room for improvement in intent recognition accuracy and smart human-like interactions. They also lack emotion recognition, making communication feel less personalized.

Frontline human services: There's a lack of smart assistance tools to support agents in their tasks. The large volume of information makes it difficult for agents to quickly access needed data, and lengthy text entry for work orders reduces service efficiency.

Operational management: The data reporting system takes a long time to process and the data retrieval system is not user-friendly.

SOLUTION AND VALUE

01 Overall structure/solution

The Government Service AI Agent Platform offers various types of smart robots, integrating a wide array of plugins, tools, and knowledge bases. Built on the Yuanjing Government Service Large Model, it rapidly develops robots that empower the smart upgrade of government services and enhance efficiency and citizen satisfaction. Its architecture is illustrated in Figure 3.

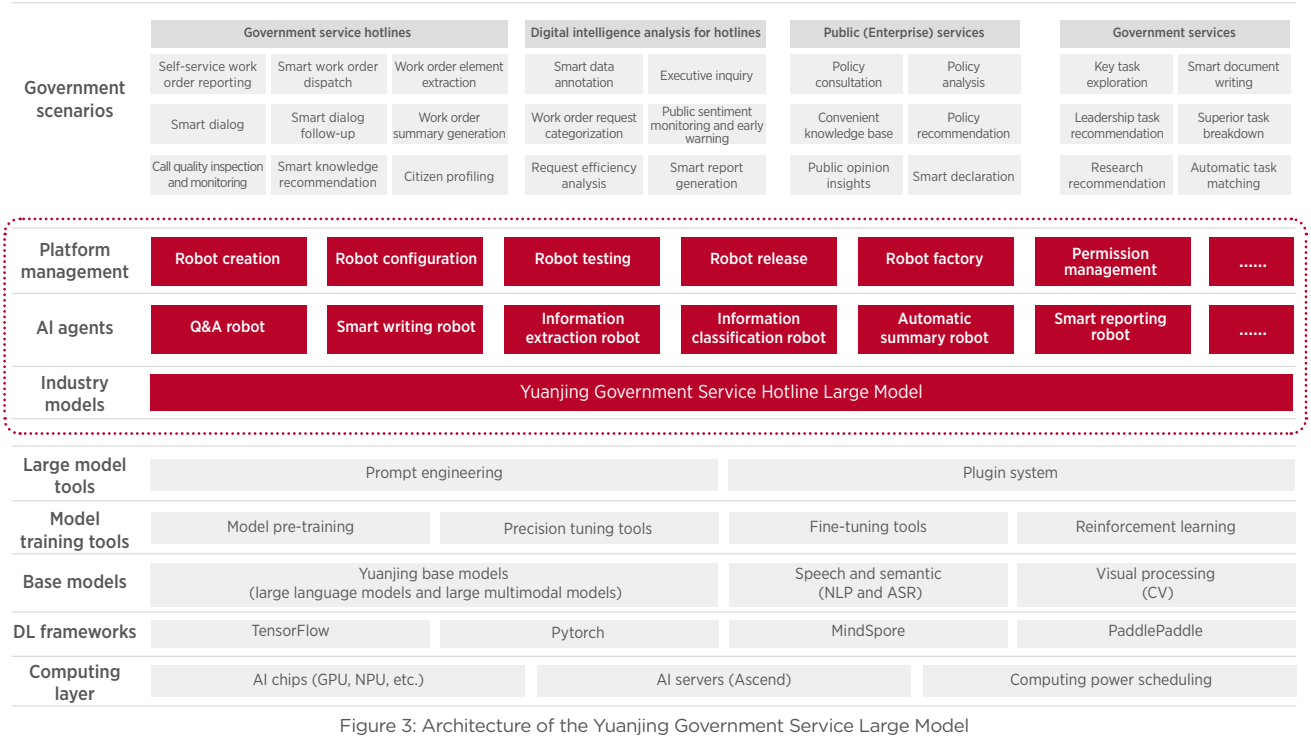


Figure 3: Architecture of the Yuanjing Government Service Large Model

The platform mainly includes three categories of robots:

voice-based robot
Speech translation robots, voice broadcasting robots, voiceprint detection robots, and voice quality inspection robots.

Dialog-based robots
Smart Q&A robots and smart outbound calling robots.

Text-based robots
Document review robots and text dispatch robots.

02 Application scenarios

The Government Service AI Agent Platform primarily supports the following application scenarios:

Communication	Smart dialog, text transcription, compliance checks, etc.
Hotline	Smart follow-up, call quality inspections, etc.
Government-enterprise office	Text archiving, inquiry consulting, security access control, etc.
Digital employees	Smart voice dialog, smart data inquiries, personalized greetings, simple scheduling, etc.
Anti-fraud/security detection	Identity verification, suspicious person screening, etc.
Document review	Document auditing, sensitivity classification, etc.
Knowledge assistant	Smart Q&A, agent assistance, knowledge recommendation, etc.

03 Key innovations

3.1 Technological innovation

The first technological innovation is the multi-task speech generation model architecture, as shown in Figure 4. This model, trained on over 100,000 hours of high-quality speech data, excels in zero-shot voice cloning of human voices, human-like naturalness, and multi-language and multi-dialect generation capabilities. Both the voice and dialog robots of the platform are built on this architecture.

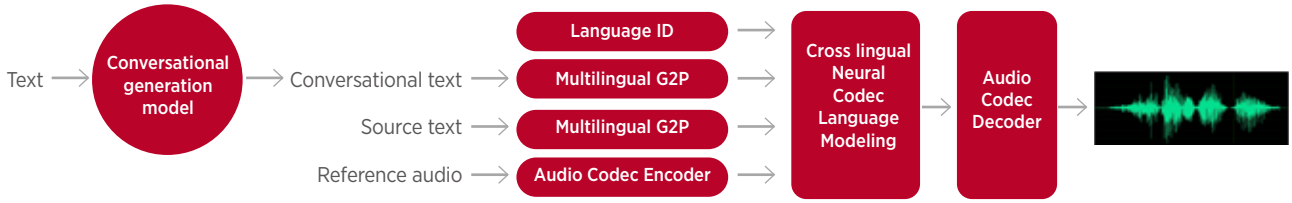


Figure 4: Architecture of the Multi-Task Speech Generation Model

The second technical innovation is the integration of the Retrieval-Augmented Generation (RAG) core component into the platform, as shown in Figure 5. This component empowers the large model with advanced techniques such as cascaded segmentation, adaptive splitting, and multi-path retrieval fusion. These enhancements significantly improve the system's ability to comprehend knowledge, boosting the understanding of industry knowledge bases and effectiveness of Q&A capabilities.

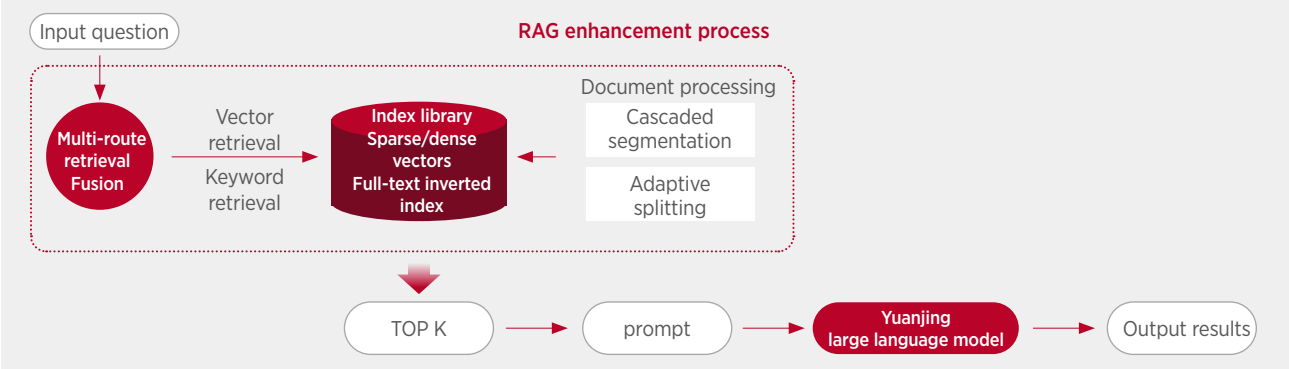


Figure 5: RAG Core Component

3.2 Model innovation

The Government Service AI Agent Platform focuses on addressing pain points by deeply integrating smart technologies. It constructs a new smart service model that includes self-service human collaboration, multimedia collaboration, online and offline collaboration, and service marketing collaboration, making smart interactions more personalized, smart assistance more convenient, and operational management more efficient.

04 Business models

Operating under the model of "1 set of model capabilities + N smart scenarios", the platform offers tailored services to a diverse range of customers, as shown in Figure 6.

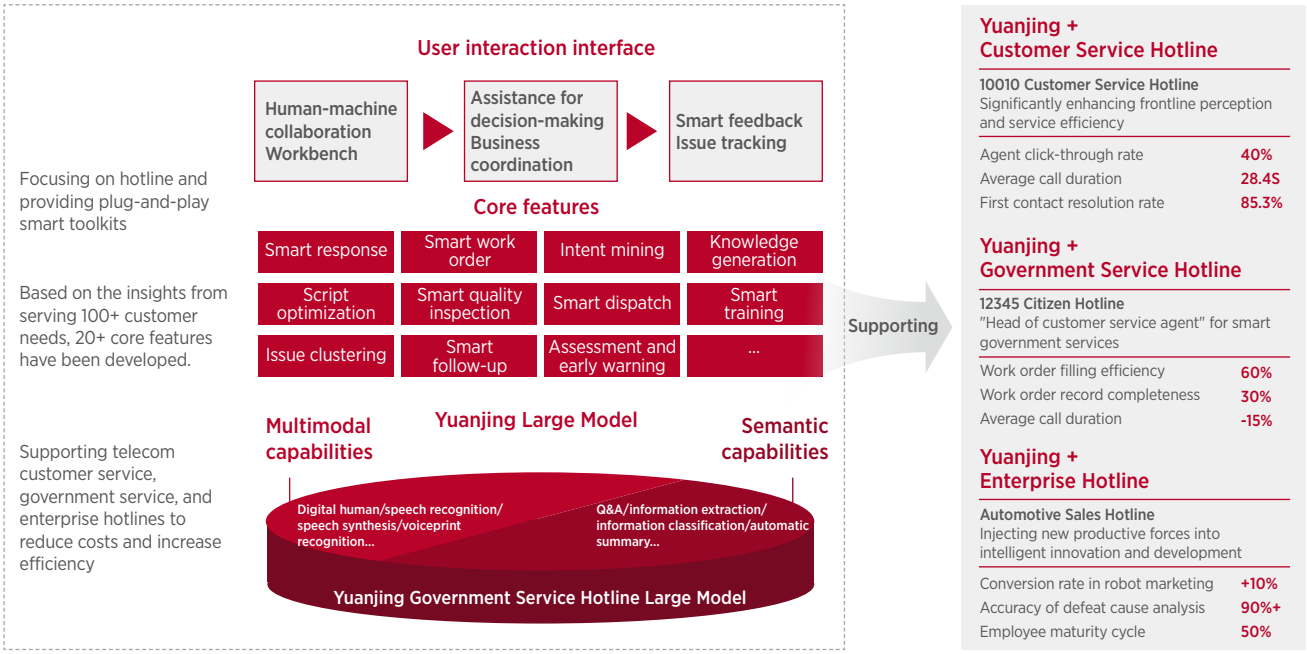


Figure 6: Business Models of the Government Service AI Agent Platform

Target customers: The main customers are the 12345 hotlines across hundreds of regions nationwide. Potential customers include other government service hotlines such as 110, 119, 120, and 122, and corporate hotline services like automotive sales hotlines.

Product scenarios: The platform has developed over 30 sub-products and components covering agent assistance, smart work orders, knowledge bases, quality inspection, call centers, smart Q&A, data management, big data analysis, and application services, ensuring comprehensive business coverage.

Profit mode: The business ecosystem involves governments, operators, technology companies, and device manufacturers. It creates a commercial profit model where state-owned enterprises construct and governments purchase services. The sales approach includes SaaS private deployment (basic platform + server) with options for customized engineering services, allowing for incremental purchases of algorithm models and customized services, with fees charged for customization.

05 Core values

With the upgrade of government service hotlines to boost collaborative governance and the integration of intelligence to refine government service processes, the Government Service AI Agent Platform boasts the following core values:

Wide reach across industries	The government service hotline product is widely applicable not only to the 12345 hotline but also to over 50 industry customers requiring hotline services, including 110, 119, 120, and 122.
Full coverage of scenarios	Built on the large model, the smart hotline platform has developed over 30 sub-products and components covering agent assistance, smart work orders, knowledge bases, quality inspection, call centers, smart Q&A, data management, big data analysis, and application services, ensuring comprehensive business coverage.
Rich industry experience	The platform primarily serves the government service hotline (12345) and union organizations, like 12351, while also providing full-process "one-stop" solutions such as consulting, product support, and technical services to customers across various industries that require hotline services.
Big data for informed governance	With PB-level data computing capabilities, the platform mines and analyzes demand data to uncover hot, difficult, and bottleneck issues in social governance, providing a scientific basis for effective urban management and social welfare interventions through informed decision-making and precise governance.
Business intelligence driven by AI large model	The Government Service Hotline Large Model enhances seven key areas, including demand reporting, agent assistance, knowledge bases, quality inspection, work order recommendations, and report generation. This leads to a reduction in work order processing time by over 40 seconds, a decrease in dispatch time by more than 30 seconds, a reduction in completion cycles by over 30%, and a 20% increase in the first contact resolution rate for complaint-related inquiries.

REFLECTION

The Government Service AI Agent Platform built on the Yuanjing Government Service Hotline Large Model offers distinct technical advantages and industry-leading practices. It specializes in government-specific scenarios, providing common AI capabilities and general AI agents tailored for this sector. This platform supports the rapid development of industry-specific AI agents for customized applications and boasts a comprehensive suite of foundational AI capabilities, offering a wealth of plugins, tools, and both large and small model functionalities.

Since 2023, the platform has implemented six projects, achieving a market value of nearly RMB 11 million. Currently, it is supporting and delivering three ongoing projects, with a market value of nearly RMB 29 million, steadily developing a robust business model.

Moving forward, the platform aims to gather more business insights from practical applications, expanding its reach from government service hotlines to additional local government departments and public service agencies. Core application scenarios and functionalities will continue to enhance and mature, allowing the platform to serve a broader customer base. The projected market value is expected to exceed RMB 90 million within three years.

2024

Core features

- Continuous optimization of text AI agents
- Introduction of voice AI agents
- Exploration of multimodal AI agents
- Focused development of a unified knowledge base
- Rich extension tools

Market promotion

- Focusing on various local 12345 government hotlines to create benchmarking projects.
- Hosting product launch events in collaboration with channel ecosystem partners to enhance product operation and publicity, establishing user perception.

2025

Feature expansion

- In-depth construction of multimodal AI agents
- Co-creating new scenarios and capabilities with new customers

Market promotion

- Reaching out to other potential users such as public service agencies and enterprise hotlines
- Strengthening in-depth collaboration with provincial/branch/subsidiary companies to expand business opportunities (e.g., deep dive into user needs and scenarios)

2026

Continuous product upgrades

- Keeping up with AI technology advancements and implementing more advanced AI generation technologies
- Providing personalized and customized products and services based on previous project experience to enhance user value.

Market promotion

- Strategic alliances: Establishing long-term cooperative relationships with IT enterprises and research institutions to maintain the advanced products and technical capabilities required by the platform.
- Brand upgrade: Upgrading brand image based on market feedback and user needs to enhance brand value and influence.
- Deepening cooperation: Deepening relationships with partners to jointly explore new business models and market opportunities.

Figure 7: Future Plans

"Four Predictions" Flood Control System at East Helan Mountain Area in Ningxia

It is a national strategy to pursue smart and digital development in water resources management. The Ministry of Water Resources has proposed introducing large models to realize smart water conservancy, with four core requirements of making forecasts, issuing early warnings, conducting simulations and drafting contingency plans. With resources in perception connectivity, computational network, and basic large models, we work closely with industry partners to enhance defense against floods and droughts. By tackling key technologies, we have developed capabilities such as smart Q&A, contingency planning, and smart water conservancy systems, making the four requirements much more digital, intelligent, and precise.

Xiong Xiaopeng
Vice President and Chairman of Labor Union, China Mobile IoT Co., Ltd.

SOLUTION PARTNERS

The East Helan Mountain Area, the focus of flood control in the Ningxia Hui Autonomous Region, has built a highly-available smart computing service platform by integrating the four requirements with new-generation information technologies such as digital twins, big data, and AI, and utilizing the western big data center. The platform follows the principle of "data-driven application, digital empowerment, and capability improvement".

As a joint contractor of the flood control project at East Helan Mountain Area in Ningxia (Phase I), China Mobile IoT Co., Ltd. introduces the visualization model at the China Mobile Kunling Digital Twin Platform to establish the digital twin support system for the area. The system can virtually reproduce water conservancy projects in cyberspace, thus supporting applications like flood analysis and warning, and intelligent scheduling. It is the first knowledge platform in the industry that innovatively integrates the water conservancy large model. It extracts real-world source data relevant to water resource management from perceptual data and industry data to construct a database that encompasses contingency plans, practical experience, documents, and historical scenarios for flood prevention. With these efforts, a water conservancy knowledge engine is available to enable rapid retrieval, query, and generation of contingency plans. The platform provides such capabilities as knowledge retrieval, knowledge management, contingency plan generation, and simulation reasoning in practice. It has been a robust support for the Water Resources Department of Ningxia to develop and implement contingency plans and reservoir operation proposals against floods at the East Helan Mountain Area.

CHALLENGES

Major difficulties and pain points

Inaccurate and difficult forecast

Traditional empirical models have short forecast cycles and low accuracy. It is challenging to comprehensively analyze and efficiently process a large amount of heterogeneous data from various sources, such as meteorological, geographical, aquatic ecological, remote sensing, hydrological, ecological, socio-economic data, etc. These data fluctuate significantly over time, making predictions difficult and manual costs high. Real-time perception and monitoring data cannot be effectively combined for analysis and judgment, so timely and accurate forecasts cannot be guaranteed.

Fixed thresholds in traditional warning systems

Traditional warning systems cannot adjust flood and engineering risk thresholds scientifically and dynamically; cannot provide warnings based on specific real-time flood scales, rainfall amounts, inflow, water velocity, and other dynamic trends; and cannot issue targeted warning notifications for different risk levels in specific regions.

Non-real-time and non-intuitive reasoning and simulation: It is not possible to dynamically drive the digital twin watersheds for simulation

Emergency command centers cannot intuitively and effectively understand disaster trends and impact ranges because the traditional system cannot clearly demonstrate the extent and severity of disaster impacts.

Inefficient contingency plan formulation, and long and difficult cross-regional coordination

Traditional knowledge base retrieval methods fail to respond promptly. The generated contingency plans which take a long time to prepare and have a short validity period cannot provide timely decision support. Unified and scientific coordination is not possible in emergency response and resource scheduling.

SOLUTION AND VALUE

01 Overall structure/solution

In this project, a practical water conservancy knowledge base is constructed based on rules and standards in the previous contingency and scheduling plans in the region to promote the synchronous interaction between the physical and digital twin models at the East Helan Mountain Area.

The foundation consists of a data platform that aggregates various heterogeneous spatiotemporal data transmitted by the water information network, including geospatial data, sensor data, and business management data. The knowledge engine extracts and merges knowledge from the source data, and structurally processes it to build a comprehensive water conservancy knowledge base and a knowledge graph. It provides computational data support for model invocation and assisted decision-making.

An industry database is built and fine-tuned to create the water conservancy large model based on China Mobile Jiutian Large Model. This model can drive digital twin watershed to realize reasoning and simulation, and synchronous sharing of historical and real-time data. As a result, the four requirements of the upper flood control applications are satisfied: flood forecast and warning, flood emergency modeling and simulation in all scenarios, and effective generation of visualized contingency plans.

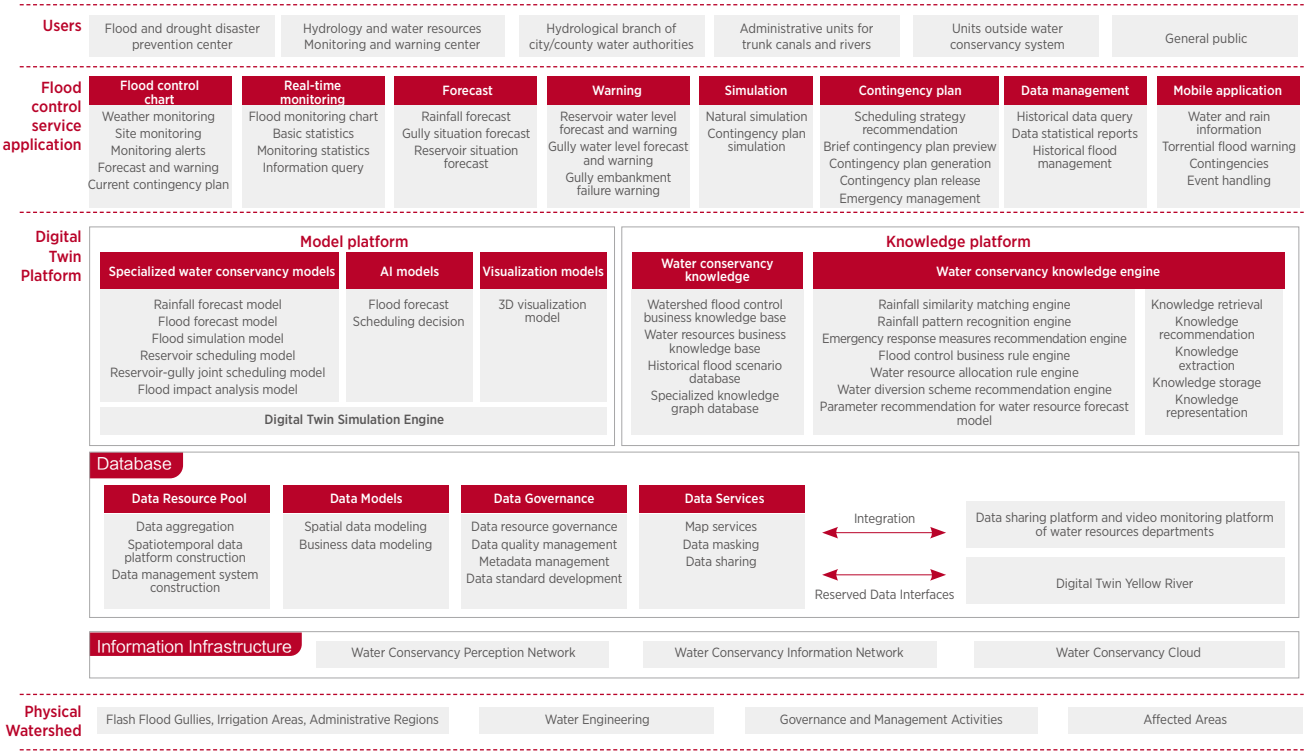


Figure 1: Overall System Architecture

02 Application scenarios

Guided by the technical framework of the four requirements in the water conservancy business, the project realizes full process development for flood control at the East Helan Mountain Area in Ningxia.

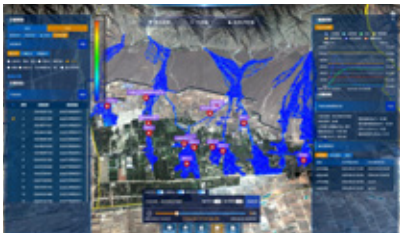


Figure 2: Flood Simulation Effect

Figure 3: Contingency Plan Generated by Water Conservancy Large Model

Forecast

Comprehensive assessment of water conservancy safety elements

The water conservancy large model integrates and processes historical data and real-time perception data extracted from the knowledge base. It drives professional models to quantitatively or qualitatively forecast different lead times of future hydrological situations (such as water levels, flow, rainfall amounts, etc.) and provides qualitative descriptions and explanations to help users and regulators better understand and respond effectively.

Warning

Early alert of water-related disasters

In the new business system, the water conservancy large model incorporates a warning rule library to automatically set warning thresholds and trigger conditions based on predictions and forecasts. When preset thresholds are exceeded or specific conditions are met, the warning mechanism is immediately triggered. It sends warning messages through SMS, calls, mobile apps, etc., to local governments, residents, and relevant personnel, prompting them to take necessary preventive measures to ensure timely evacuation and rescue.

Simulation

Simulating emergency scheduling plans

The water conservancy large model, combined with historical cases, expert experience, and professional knowledge base, automatically generates or optimizes response plans. It synchronizes these plans and data with the digital twin watershed model to reason and simulate conditions such as rainfall, water level increases, and the extent of flooding in the area for the corresponding period, based on time, space, and regional data such as rainfall intensity, river inflow, water levels, etc.

Contingency Plan

Automated generation of emergency response plan

When simulations indicate a potential disaster or when customers seek contingency plans based on warning situations, the large model automatically retrieves historical similar scenarios and experience data based on the existing real-time data as well as historical flood scenario database, expert experience, business rule library, and scheduling plan database to generate and optimize emergency response plans. It also outputs emergency response processes, resource scheduling plans, evacuation and rescue plans on the interactive interface of the large model based on local conditions, government policies, laws and regulations.

03 Key innovations

Multi-dimensional data fusion and processing

The water conservancy large model supports the application of graph databases and vector databases, and the storage of data from different sources and in various forms, thus achieving full-process multi-source heterogeneous data processing. It has built a comprehensive water conservancy knowledge base.

Semantic recognition via knowledge engine

With data integration and association capabilities of the water conservancy knowledge graph and knowledge engine, a professional water conservancy large model is constructed to effectively identify and represent concepts, properties, and relationships in the water conservancy domain. This helps in understanding the issues in water conservancy business scenarios. The precise semantic recognition capability is highlighted.

Second-level analysis and reasoning

With the enhanced RAG architecture for data retrieval, the water conservancy large model can effectively expand output with contextual information and answer complex water conservancy questions. The data retrieval engine can efficiently match and retrieve similar associated data based on user intent, improving service response and generating response plans in seconds.

Continuous learning and performance optimization

The water conservancy large model supports dynamic updates to the knowledge base, enabling real-time addition and update of monitoring data. With a more comprehensive range of historical and real-time monitoring data across time dimensions, it can effectively analyze future hydrological trends in specific application scenarios. Performance evaluation and optimization are carried out through user feedback mechanisms, enhancing user experience and decision support quality.

Virtual reality Integration of digital twin

By constructing a "digital twin" of the physical water conservancy infrastructure, the water conservancy large model can simulate, monitor, predict, and optimize the operational status of water conservancy systems in real time. Mapping of the digital twin watershed enables early prediction and threshold warnings. When highly coupled with the specialized water conservancy large model, the digital twin can realize more accurate flood reasoning and simulation.

04 Business models

The project primarily targets organizations responsible for water engineering planning, design, and operational management, including water conservancy departments, water authorities, and watershed management departments. It promotes the application of AI industry large models and digital twin technology to assist in design optimization, resource management, disaster prevention and control in water conservancy. With optimized operation, inspection, maintenance, and cost control, it allows the implementation of AI large models in various aspects from construction and production to resource management.

Based on the actual engineering business needs of customers, the project provides customized services that integrate model deployment training and reasoning. It offers full-process private deployment services from application platform development to model integration, including water data collection and cleaning, knowledge base development, large industry model training and fine-tuning, model deployment, and flood control platform integration. In this way, it can ensure data security and controllability.

05 Core values

Project achievements

The water conservancy large model supports the construction of modules such as the database, water conservancy knowledge platform, knowledge engine, and a unified visualization application, enabling four functions of "forecast, warning, simulation, and contingency plan" for eight flood control systems (54 diversion channels and 15 detention reservoirs). Key technologies such as prompt word templates, enhanced retrieval generation, and efficient fine-tuning using LoRA are incorporated to build capabilities like expert assistant and contingency plan generation. With the capability to train large models with billions of parameters, it provides flood perception, information sharing and communication, reasoning and simulation, and decision-making support against floods around the East Helan Mountain Area.

Value proposition

This is one of the first model projects in China that use the water conservancy large model and one of the first batch of projects of China Mobile in water conservancy large model application. The project helps the Ministry of Water Resources and water authorities to explore replicable and scalable business models and technical systems, and provides a technical reference for other provinces and cities to build digital twin watersheds.

Task breakdown: Provide a unified, standardized, optimized database, knowledge base development and automated data processing workflows for the construction of water conservancy data centers; replace the previous manual data collection and analysis methods to break down flood control and management tasks in emergency command centers; automate contingency plan generation to provide more intelligent decision-making suggestions for regulatory and governance agencies. As a result, flood control efficiency is significantly improved.

Simulation display: Driven by the AI large model, the project combines the digital twin watershed for reasoning and simulation, making textual information more intuitive. It can also identify the scope of affected areas by displaying simulation results, providing more comprehensive information for making resource scheduling decisions.

REFLECTION

Summary

In the early stage, the Open Platform Department of China Mobile IoT Co., Ltd. collaborated with the Northwest Regional Center to conduct detailed on-site inspections, demand analysis, and feasibility studies, thus effectively clarifying project goals. Steady progress was seen in the development of digital twin system for flood control. A thorough analysis of the basics, flood control status, and informatization status in the region was conducted, and key issues and weak areas were outlined according to the technical guidelines of the Ministry of Water Resources on the digital twin watershed and four requirements in small watersheds. During the project, a robust database and data management tools were developed to ensure the quality, timeliness, and security of water conservancy data, providing decision support for the project. A systemic approach was adopted, creating the top-level design for the overall architecture and integrating technologies like AI large models and digital twin with business applications for development. Furthermore, the long-term maintenance mechanism ensured the real-time reliability of data and continuous iterative optimization of models.

Follow-up plans

- To establish an all-in-one AI development platform: We will deepen the practical application of AI large models in water conservancy, create a universal training paradigm, enhance reasoning capabilities, thus creating a flexible standard engine and realizing modular training and deployment.
- To coordinate facilities and equipment: We will explore the application of integrated space-air-ground monitoring and remote sensing technologies, address the challenges of coordination between water conservancy large models and sensing devices, promote the integration of the water conservancy system with new technologies such as remote sensing satellites and precipitation radar. In this way, comprehensive water resources management across the entire basin and at all times can be realized.
- To drive specialized water conservancy models: The water conservancy large model will drive specialized small models for simulation, and achieve adaptive correction with the help of smart perception, making data processing more efficient. In the end, a smart high-precision water conservancy system is available.

Multimodal Sichuanese Dialect Protection

The AI-basedmultimodal dialect protection project is of great significance and has broad prospects. With technological innovation and cross-sector collaboration, this project will promote the protection and development of cultural diversity and push ahead international cultural exchanges and cooperation, contributing to sustainable development goals. To overcome challenges and achieve long-term goals with lasting impact, the project needs the joint efforts of the government, enterprises, and all sectors of society.

Feng Jing
Deputy General Manager of Cloud Network Development Department (Science and Technology Innovation Department)
China Telecom Corporation Limited Sichuan Branch

SOLUTION PARTNERS

Cultural Diversity

The Ministry of Education and UNESCO jointly published the document Protection and Promotion of Linguistic Diversity of the World: Yuelu Proclamation (2019). The document states the goal of language resource protection. It encourages protecting and promoting linguistic diversity in multiple ways, including scientific research, media, curricula, arts, cultural production, and ICT. It also advocates creative transformation, innovative development, and effective dissemination of language culture using AI, ICT, and other technologies.

Language Protection: Build Dialect Voice Datasets

Build data production capabilities: We have carried out research on the commonality and individuality of dialects based on voice, vocabulary, grammar, and pragmatics after analyzing the characteristics of dialect voice and word usage in all regions of Sichuan. First, we collected corpora such as numbers, addresses, and telegraphs related to the business scenarios of enterprises. Then, we manually collected voice data in multiple ways. We adopt deep learning technology to automatically extract audio and video data and generate research results, gradually building the data production capabilities in the voice data field for enterprises.

Build voice datasets: We have conducted research on tag system creation and data retrieval, extraction, storage, and visualization of unstructured voice data. Our goal is to build compliant and high-quality proprietary Sichuanese dialect voice datasets for enterprises. We have collected corpora that comply with the Chinese national voice protection standard for 19 branch companies across the province.

Operational Scenarios Based on Proprietary Technologies

Voice data
Dialect voice about 1,200 hours long and Mandarin voice longer than 1,000 hours

Text data
109.1 million general data entries and 1.44 million communications industry Q&A data entries

The multimodal database

Data collection capability
multiple collection tools, including IVRs, mobile phones, and computers

Data processing capability
preliminary AI annotation system, with an annotation platform, annotation specification, and an annotation team

AI-based Atomic Capabilities That Can Be Commercially Used and Implemented

Develop AI-based atomic capabilities: Sichuanese speech recognition, Sichuanese speech synthesis, speech separation, voiceprint recognition, sensitive sentence recognition, and customer sentiment analysis. These capabilities are implemented to reduce costs and improve efficiency in multiple production scenarios.

CHALLENGES

The multimodal dialect corpora are faced with challenges in the industry in terms of data collection, processing, standardization, privacy protection, and technical implementation.

Data Dispersion and Heterogeneity	Dialect data is often scattered to different regions, communities, and individuals, and therefore cannot be collected at a time. Data sources are stored in diverse forms, including audio, video, and text, with different formats and structures. As a result, data processing is difficult.
Data Quality and Integrity	Errors may occur during the collection and recording of dialect data. Low recording quality and accuracy deteriorate data quality. Data integrity is also unsatisfactory. Some dialects may gradually disappear due to historical reasons or modern language changes. As a result, some data is missing.
Data Standardization and Normalization	Chinese has diverse dialects. Each has its unique phonetics, vocabulary, and grammar rules. It is difficult to formulate unified standards for the dialects. Standardizing multimodal data is even more complex. We need to consider standardizing data in multiple forms, such as audio, video, and text. In addition, data processing must comply with specifications and processes to ensure data accuracy and consistency. However, the unique nature of dialect data makes standardization even more complex and cumbersome.
Privacy Protection and Compliance	Dialect data usually involves sensitive personal data, such as name, address, and contact information. During data collection and processing, strict privacy protection measures must be taken to prevent data leakage. In addition, the collection and use of dialect data must comply with relevant laws and regulations, such as the Data Protection Law and the Privacy Protection Law. How to ensure compliance during actual operations is also an important topic.
Technical Implementation and Application	The processing and analysis of multimodal data need to leverage advanced technical means, such as natural language processing, computer vision, and audio processing. The R&D and application of these technologies require large investments in human and material resources.

SOLUTION AND VALUE

01 Overall structure/solution

Develop a pronunciation dictionary for Southwestern Mandarin according to the unified Chinese national language protection standards. Annotate and expand international phonetic symbols to build a multimodal corpus database for Chinese dialect speech.

Autonomously train AI recognition algorithms to implement automatic speech recognition. For more information, see Figure 1 "Dialect processing flowchart."

Apply text-to-speech technology to protect Chinese dialects.

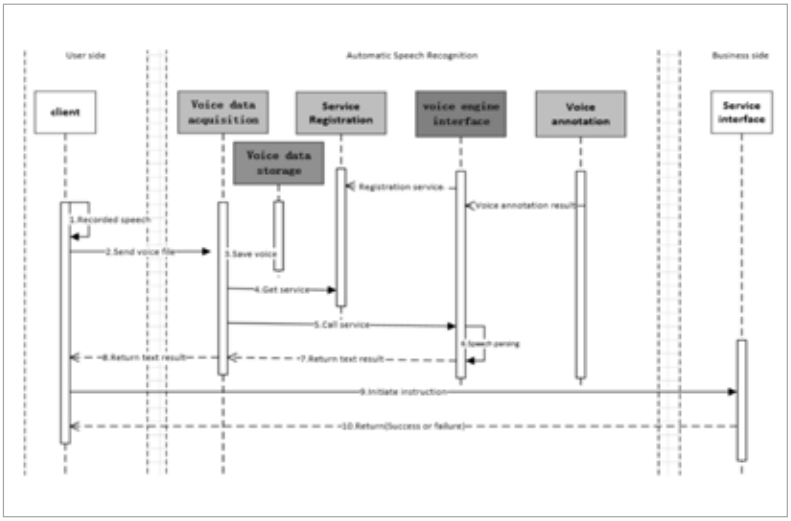


Figure 1: Dialect processing flowchart

02 Application scenarios

Health
Dialects may help spread health information and traditional medicinal knowledge within communities. Protecting dialects ensures that the information reaches more people, boosting public health and well-being.

Education
Dialects are the mother tongue of many people and the initial language in which they are educated. Protecting and promoting dialects make educational opportunities more inclusive and equal. This ensures that all people can receive high-quality education.

Economic growth
Dialects function as a tool to stimulate local economic development and employment. Protecting and promoting dialects enhances the economic vitality and sustainable development of communities.

Value communication
Dialects can help convey information about sustainable consumption and production patterns so communities can better understand and implement sustainable development.

Peace and inclusion
Protecting dialects contributes to cultural diversity and understanding, building a peaceful, just, and inclusive society.

Global partnership
Protecting and promoting dialects are beneficial to global partnership, achieving the common goals of sustainable development.

Application
Share the multimodal dialect database to optimize models.

Launch the atomic digital capabilities on the platform. This helps protect Chinese dialects and build multimodal corpora.

Prepare to build a digital human lab by using dialect voice as a unique cultural communication medium to highlight the characteristics and distinctions of digital humans.

03 Key innovations

This project aims to protect and preserve multimodal dialects with key innovative approaches, which is critical to maintaining and driving cultural diversity. This project is innovative in promoting cross-cultural communication and understanding to break communication barriers, which builds a bridge for cultural exchanges and cooperation. In addition, this project also creates key innovative values in cultivating talents in the AI field, enhancing technological innovation and revitalizing economic growth. This project will significantly benefit economic growth in related industries. Its innovative strategies for cultural protection and cross-cultural exchange will also enhance social stability and environmental protection.

04 Business models

Dialect voice technology assists in the use of IPTV dialect commands to improve IPTV's user experience, and also explores cultural inheritance and protection. For example, dialect-speaking digital humans are developed to livestream and sell local specialty crops in counties (One County One Priority Product initiative, OCOP). While driving the sales of local agricultural products, it is also a new way of preserving and protecting dialect culture. This mode meets market demands and paves a new path to spread and develop dialect culture.

1. Building dialect voice datasets

(1) Build data production capabilities: Carry out research on the commonality and individuality of dialects based on voice, vocabulary, grammar, and pragmatics after analyzing the characteristics of dialect voice and word usage in all regions of Sichuan. First, collect corpora such as numbers, addresses, and telegraphs related to the business scenarios of enterprises. Then, manually collect voice data in multiple ways. We adopt deep learning technology to automatically extract audio and video data and generate research results, gradually building the data production capabilities in the voice data field for enterprises.

Figure 2: Live dictionary for Sichuanese dialects

(2) Build voice datasets: Conduct research on tag system creation and data retrieval, extraction, storage, and visualization of unstructured voice data. Our goal is to build compliant and high-quality proprietary Sichuanese dialect voice datasets for enterprises.

2. Optimizing in-house models

(1) Conduct in-depth research on the regional division, pronunciation characteristics, voice annotation methods, and other related content of Sichuanese dialects. Based on the existing Southwestern Mandarin pronunciation dictionary, build a pronunciation dictionary with the regional voice characteristics of Sichuanese dialects.

(2) Continuously optimize the in-house dialect voice recognition model in internal business scenarios of enterprises according to acoustic principles and the pronunciation dictionary.

3. Use cases

(1) Empowering the Digital Panda IP using dialects

Figure 3: Smart AI-powered digital panda with dialect synthesis for cultural tourism

Digital Panda IP for cultural tourism integrates AI and digital pandas to provide personalized and interactive guide services, enhancing the sense of engagement and experience of tourists. The IP leverages the panda, the world-renowned cultural symbol, to boost the international reputation and image of Sichuan's culture and tourism. This way, cultural dissemination and the development of smart tourism are promoted. "Dialect voice assists in cultural tourism streamers" to help deepen the independent exploration and interactive experience of tourists. The voice of the streamers can be customized to meet local needs, making the accent more familiar and appropriate for cultural tourism settings.

(2) Assisting farmers in livestreaming using dialects for OCOP

Farmers can present and sell agricultural products in livestreams with technologies to read bullet comments and automatically answer the questions on the bullet screen using dialects. This improves interactivity and user experience, saves the costs of hiring professional streamers, and optimizes the livestreaming results with technical means.

05 Core values

The project delivers eight patents that target collaborative and adaptive text error correction, voice extraction, voiceprint reinforcement, digital encryption, and other technologies in the speech recognition system. The content covers speech recognition, natural language processing, red teaming, and so on. The project was nominated for the cultural diversity prize at the World Summit on the Information Society (WSIS) of the United Nations. It was also selected in the success stories in the National Scientific Computer Corpus Conference.

Based on the domain adaptation theory, the system learns the representation codes of other unclassified secondary dialects with no or few labels in the target domain, obtaining voice codes of common representation dialects and accurately classifying primary dialects. Then, we establish a correspondence between basic sound units and basic Chinese character units using a fuzzy search algorithm based on the pre-trained general acoustic model. By integrating the model with the Chinese sentence generation method that depends on syntactic analysis, we establish a universal dialect recognition model where acoustic and semantic spaces are consistent. Ultimately, we enhance the generalization of different dialect recognition models by referring to the few-shot meta-learning theory.

REFLECTION

Reflections

1. Technological innovation promotes dialect protection: The digitalization and intelligence technologies have significantly changed the conventional dialect protection model. This project transforms the obscure academic research path into a feasible, accessible, and interesting digital interactive method. This transformation ensures that dialect data is properly stored and research results are vividly displayed and widely used in real life.

2. Application value drives sustainable conservation: The application value is the key to the sustainability and efficiency of the dialect conservation method. China Telecom Sichuan Branch deeply explores the application scenarios of dialects and provides related services, achieving a win-win between dialect conservation and commercial benefits. This innovation model paves a new path for dialect conservation and significantly improves the quality of service.

Follow-up plans

- Deepen technology research and extend the application: We will further develop an AI-powered Sichuanese dialect speech synthesis engine based on the Sichuanese dialect live dictionary. In detail, we will study the text-speech consistency encoding technology, speech encoder fine-tuning and training technologies based on transfer learning, and zero-shot timbre synthesis technology. By making breakthroughs in these technologies, we will be capable of synthesizing smooth and natural Sichuanese dialect speech from text and imitating the timbre.
- Expose APIs and meet diverse requirements: We will develop and expose programming and visualization APIs based on the speech synthesis engine to meet personalized and customized development requirements and reduce use costs. In addition, we will conduct a pilot test in actual business scenarios and extend the application of dialect speech synthesis. For example, we will extend it to virtual streamers in new media, dialect-speaking customer service agents, and personalized voice assistants. This will further boost the preservation and development of dialect culture.
- Cross-sector collaboration and sustainable development: The AI-basedmultimodal dialect protection project is of great significance and has broad prospects. We will strengthen technological innovation and cross-sector collaboration to promote the protection and development of cultural diversity, pushing ahead international cultural exchanges and cooperation. In addition, we will seek support and cooperation from governments, enterprises, and all sectors of society to jointly contribute to the sustainable development goals.

Yuanjing Urban Governance Large Model Reshaping Urban Governance

The acceleration of urbanization brings up a multitude of complex issues, including the lack of real-time guidance and inefficient incident management. Traditional urban management models, which rely heavily on manual decision-making and rigid data processing, are increasingly inadequate in addressing these issues. More refined and dynamic urban management requires advanced AI technologies to enhance decision-making efficiency and service quality. The China Unicom Yuanjing Urban Governance Large Model has been developed in this context. The Yuanjing Urban Governance Large Model aims to integrate diverse data from urban operations and utilizes AI technologies such as visual recognition, speech recognition and synthesis, intent recognition, text generation, and multimodal interaction to offer such features as event reports via photos or voice, automatic task dispatching, log generation, and smart follow-ups.

Lian Shiguo
Chief AI Scientist and Chief Engineer of China Unicom AI Innovation Center

SOLUTION PARTNERS

The China Unicom Yuanjing Urban Governance Large Model is built on the Yuanjing large model as its technological foundation, as shown in Figure 1. It focuses on the full life cycle of various urban events, supporting comprehensive closed-loop management of citizen requests through the processes of "aggregation - dispatch - handling - closure - analysis". Portable smart work assistants can provide professional knowledge support to enhance urban governance levels and service effectiveness. The model has already been implemented in the Xinjin District of Chengdu.

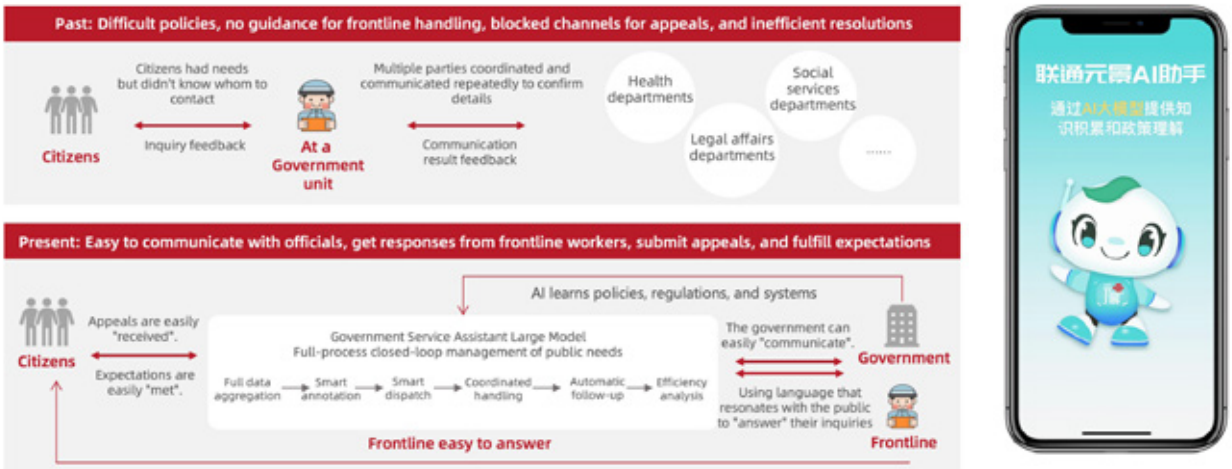


Figure 1: Introduction to Yuanjing Urban Governance Large Model

To improve urban risk prevention and control capabilities, enhance refined management levels, and accelerate the construction of an urban operation management and service platform, a quick response product for citizen requests has been launched based on the urban governance large model. This product aims to address the challenges of low event handling efficiency, insufficient event traceability, inadequate early warning capabilities, and weak event assessment abilities.

CHALLENGES

Traditional urban governance models primarily face the following challenges:

Population density and management

As urbanization picks up pace, issues related to housing, transportation, and public resource allocation due to dense populations are becoming increasingly prominent.

Traffic congestion

Urban traffic congestion not only affects residents' quality of life but also leads to environmental pollution and hinders economic development.

Resource shortages

The high demand for resources such as water and energy in cities often results in shortages, necessitating effective resource management and allocation strategies.

Policy formulation and execution

It is necessary to develop effective policies and ensure their successful implementation to address various issues in urban governance. At the same time, it is a must to increase residents' participation in urban governance and encourage community autonomy and civic involvement in the decision-making process.

Public safety

Ensuring the safety of residents and effectively preventing and responding to natural disasters, public health events, and criminal activities are crucial.

Digital divide

TIn the digital age, it is important to ensure that all residents benefit from the conveniences of information technology, thereby narrowing the digital divide.

SOLUTION AND VALUE

01 Overall structure/solution

The Yuanjing Urban Governance Large Model, empowered by AI large models, reshapes urban governance processes and deepens governance scenarios, achieving "efficient handling, closed-loop circulation, single-screen visibility, and network management" in urban governance. It explores new models to make urban governance more scientific, refined, and intelligent. Figure 2 shows the overall architecture of the Yuanjing Urban Governance Large Model.The Yuanjing Urban Governance Large Model has now implemented AI applications such as "smart data inquiry", "smart government service inquiry", "smart reporting", "smart task dispatching", "smart recommended responses", and "tag warnings" across citizens, grid attendants, and leaders. It offers data-driven suggestions for processing citizen requests, improves efficiency in addressing these requests, reduces dissatisfaction risk, enhances refined governance capabilities, and assists the government in reducing costs and increasing efficiency, ensuring that "the people's calls are properly met with responsive action".

During the pre-training phase, the model utilizes various categories of urban governance data to enrich its industry knowledge. The data used includes open-source urban governance datasets, news and announcements, legal regulations and policy documents, government website Q&A data, and multi-category data related to public services and social affairs. It also incorporates hotline call data and task dispatch data specifically related to citizen requests.

To enable the rapid deployment of applications built on the urban governance large model, a centralized application platform has been established, focusing on urban governance scenarios. This platform offers rapid construction capabilities for smart knowledge inquiry agents and allows for the quick creation and configuration of various agents based on demand, supporting the combination and plug-and-play usage of agents.

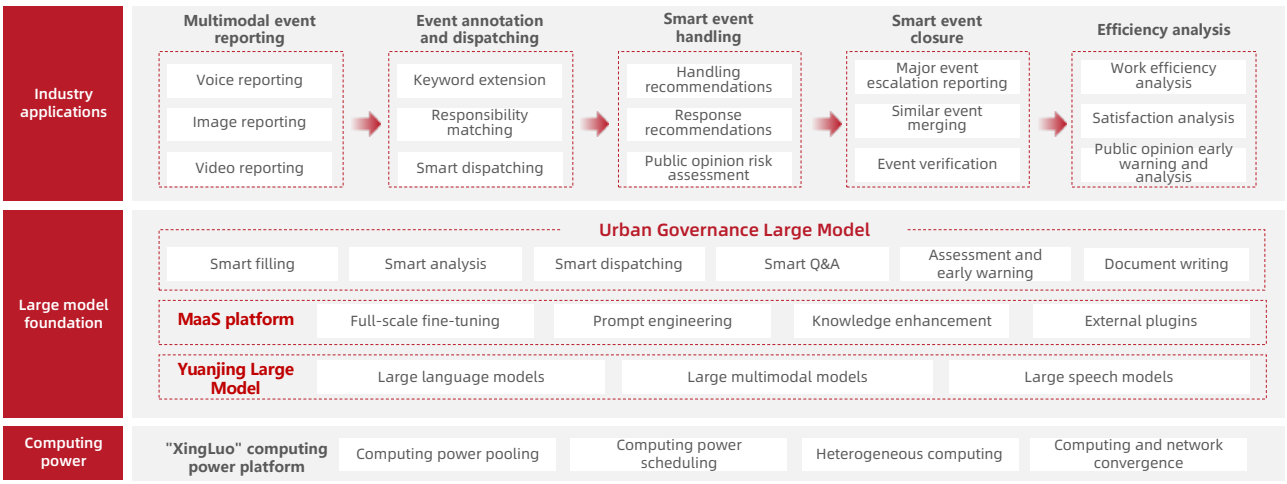


Figure 2: Architecture of the Yuanjing Urban Governance Large Model

02 Application scenarios

The Yuanjing Urban Governance Large Model has been implemented in the Xinjin District of Chengdu, focusing on urban governance and unified management to create a rapid response system for citizen requests. This enables intelligent filling of public demand requests, complete aggregation, smart dispatching, coordinated handling, and automatic follow-ups, forming a comprehensive closed-loop process. The main application scenarios include:

Smart government service inquiry	Users can ask various questions related to government services, and the large model provides targeted, comprehensive answers and recommended documents.
Smart reporting	In the event reporting scenario, multimodal models such as text, speech, and image analysis empower grid attendants or citizens to quickly, efficiently, and intelligently complete event reporting tasks, reducing reporting time and improving work efficiency.
Smart data inquiry	Users can inquire about data related to specific events without manual filtering or matching. By simply asking questions in natural language, they can easily obtain the required data responses.
Smart task dispatching	In the event dispatching scenario, the AI large model intelligently matches responsibilities and accountability based on event type and related information, associates dispatch rules, recommends receiving departments, and provides reminders for pending events.

03 Key innovations

The Yuanjing Urban Governance Large Model utilizes China Unicom's Yuanjing base model and the MaaS platform to create industry-specific models for various applications. The key innovations include:

Building an industry-specific urban governance large model based on the base model.	This model leverages the multi-domain knowledge of a general large model to handle various tasks across different fields, such as image and text generation. It is specifically designed for the urban governance sector, with optimization and adjustment using a vast amount of urban governance data to meet specific industry needs.
Enhancing scenario applications using the Yuanjing Urban Governance Large Model and scenario data	The industry-specific large model possesses greater professionalism and specificity, allowing it to better meet the unique demands of specific sectors. By utilizing scenario data, the model parameters can be fully or partially updated to improve the performance of the pre-trained model in new scenarios, thereby enhancing the service for scenario-based applications.
Providing large model application services based on scenarios	The model supports various applications for citizen request scenarios, including AI features such as "smart data inquiry", "smart government service inquiry", "smart reporting", "smart task dispatching", "smart recommended responses", and "tag warnings." These applications integrate with actual business needs to provide AI application services.

04 Business models

The "citizen request quick response" product developed based on the Yuanjing Urban Governance Large Model serves as a unified platform for efficiently addressing various issues. It is an important scenario for expanding smart city applications and an effective tool for delivering government services and social governance. Figure 3 illustrates the product's business model. The sales model combines hardware and software, offering low costs and acceptable inference performance. It enables rapid upgrades and private delivery (deployment) without commercial ties, creating a controllable supply chain for acceptable and reliable large model-enabled products in the government and enterprise sectors.

Figure 3: Business Models of Yuanjing Urban Governance Large Model

05 Core values

The Citizen Request Fast Response product developed based on the Yuanjing Urban Governance Large Model has generated a revenue of RMB 3.25 million. It has quickly established a commercial promotion strategy in such dimensions as product development, pilot applications, and scaled promotion. Over the next three years, starting next year, six new pilot projects will be added, expected to deliver a profit of approximately RMB 6.54 million. The Yuanjing Urban Governance Large Model provides innovative solutions for optimizing resource allocation, improving urban operational efficiency, promoting sustainable development, and enhancing residents' quality of life.

The construction of the urban governance model in the Xinjin District of Chengdu has shown significant results. After three months of trial operation, the response and handling time for citizen requests has reduced from six days to within three days, with a noticeable downward trend in the total number of requests. The system's smart task dispatching rate and accuracy both exceed 95%, significantly reducing the workload of manual task dispatching and improving work efficiency.

The Yuanjing Urban Governance Large Model has delivered the following social benefits:

Enhanced citizen experience	Increased efficiency for grid attendants	Decision support by leaders
Citizens can conveniently report urban issues via the "Xinjin Super Green Leaf Code" mini program, which is empowered by the urban governance model. They can use voice, photos, or videos for reporting and quickly access policy notifications, service guides, and information about the 15-minute community living circle through an AI assistant.	Community grid attendants utilize the "Report App", empowered by the urban governance model, to respond rapidly to public demands through features like smart processing, smart reply, and smart Q&A, allowing for quick resolution of community issues.	Leaders from various departments and townships in Xinjin District use the "Accelerated Smart Governance Cockpit", empowered by the urban governance model. Features like the Smart Governance Index, smart annotation, enhanced governance, and analysis of citizen dissatisfaction risks enable them to grasp public pain points in real time and quickly identify and effectively supervise major and urgent issues.

REFLECTION

Reflections

Technology-driven innovation
The implementation of the Yuanjing Urban Governance Large Model benefits from China Unicom's continuous investment and innovation in AI large model. By integrating diverse data, it provides comprehensive insights and predictions, offering robust technical support for urban governance.

Scenario-based applications
The application of large model technology in practical scenarios, such as smart government service inquiry, smart data inquiry, and smart reporting, not only enhances the efficiency of urban governance but also increases citizen engagement and satisfaction.

Multi-party collaboration
Throughout the project's advancement, we have closely collaborated with local governments and various departments, such as the Xinjin District of Chengdu, to promote the implementation of the urban governance large model. This collaborative approach helps us better understand the actual needs of urban governance, allowing us to deliver more precise services.

Follow-up plans

As generative artificial intelligence continues to mature, AI has become a significant driving force in today's societal development. A phase where intelligent methods are central to urban governance has already started. In the future, China Unicom will continue to actively fulfill its responsibilities as a central enterprise by developing core capabilities of large models and applying them in practice, providing comprehensive and precise intelligent services to enhance the efficiency and quality of urban governance.

Strengthening multimodal data integration
China Unicom will further expand data sources and enhance the deep integration of multimodal data, including text, audio, and images, to provide more comprehensive and accurate smart services. This will help improve the efficiency and quality of urban governance, offering more scientific decision support to governments and relevant departments.

Expanding application scenarios
Building on existing application scenarios, China Unicom will continue to explore and expand new application scenarios to further enhance the intelligence level of urban governance. At the same time, we will pay attention to actual citizen needs and introduce more convenient and beneficial service features.

Promoting standardization and regulation
China Unicom will take an active role in the standardization and regulation efforts in the field of urban governance, promoting the widespread application of large models. By establishing unified standards and norms, we aim to enhance interoperability between different systems, thereby improving the overall efficiency of urban governance.

China Mobile's Customer Service Large Model

Digital technologies represented by AI large models are flourishing, with accelerated technological iterations, increasingly diverse participants, and a wide expansion of application scenarios. The application of large model technology in the service sector is diverse, significantly enhancing the precision, intelligence, and response speed of services. It can even drive process transformations and reshape the service chain. China Mobile is actively embracing this technological wave, leveraging its in-house Jiutian Large Model to create an efficient and smart large model for the customer service industry. This approach breathes new life into the service sector and propels it into a new phase.

Zhao Fang
General Manager of Customer Service Department of
China Mobile Communications Group Co., Ltd.

SOLUTION PARTNERS

中国移动 China Mobile

China Mobile's Customer Service Large Model, built on the "Jiutian Large Model", enhances specialized service capabilities through advanced and efficient deep learning algorithms, strengthening multimodal interaction processing and enabling personalized human-like interactions. In terms of specialized service capabilities, a high-value professional dataset has been constructed, utilizing RAG (Retrieval-Augmented Generation) to enable the large model to effectively use internal knowledge and data. With AI agents, the large model can deeply understand user intent, orchestrate processes, and appropriately utilize internal resources. For multimodal interaction processing, different modalities of data, including text, images, and audio, are mapped to the same representational space. Through cross-modal learning and generation, voice, image, and video interaction modes are made possible. For personalized human-like interactions, the system seamlessly integrates with video customer service, online customer service, and customer service agent systems to enhance capabilities in customer insights, demand understanding, smart recommendations, and emotional support. Additionally, by generating innovative content and stylistic settings, it presents a personalized smart service model to customers. With the capabilities of the Customer Service Large Model, China Mobile has developed innovative applications targeting customers, frontline staff, and management, such as enhanced smart customer service, enhanced agent assistance, and enhanced smart coaching, providing a more thoughtful, user-friendly, and personalized service experience.

Since the commencement of research and application, China Mobile's Customer Service Large Model has completed six major version iterations and has been commercially promoted. It has been showcased at several major public exhibitions, such as the Global Partners Conference in October 2023, the Digital China Summit in Fujian in May 2024, and the World AI Conference in Shanghai in July 2024. Overall results show that the enhanced smart customer service based on the Customer Service Large Model demonstrates exceptional understanding and generative capabilities in handling various customer service inquiries. Comparative tests indicate that the enhanced smart customer service significantly outperforms existing smart customer service in terms of accuracy, completeness, and user-friendliness.

CHALLENGES

In the digital economy era, the widespread adoption of digital technologies in the service sector has resulted in applications like smart customer service and customer experience management systems, which have improved service quality to some extent. However, due to technological limitations, persistent issues such as irrelevant responses, slow service response times, and inefficiencies in problem resolution remain challenging, making it difficult to provide satisfactory customer service. In terms of customer interactions, smart customer service often struggles to accurately grasp contextual semantics and handle complex and nested logical questions. It also faces challenges in recognizing customer emotions and managing complex issues. Its language can be awkward and formulaic, making it hard to adapt to the actual needs or preferences of users. Regarding frontline responses, the skill level of service personnel varies significantly, leading to delayed responses, inaccurate resolutions, and difficulties in maintaining consistent service quality. Moreover, the application of digital technologies relies heavily on pre-set knowledge bases and algorithms, which cannot keep pace with rapidly changing business needs. This reliance impacts the overall service quality.

By harnessing the capabilities of the Customer Service Large Model in intent understanding, text generation, and data analysis, it can accurately understand customer needs, rapidly resolve customer issues, and deliver "smart companionship, emotion-driven, and multimodal interactions", providing a smart and user-friendly service experience.

SOLUTION AND VALUE

01 Overall structure/solution

The Customer Service Large Model is built on a standard network interconnection model, leveraging an AI + X systematic AI development approach. It employs a training paradigm of "pre-training + instruction fine-tuning + reinforcement learning with human feedback". This model also integrates data application resources from customer service information systems. By focusing on data insights, technological capabilities, and industry applications, it promotes the training & inference and application of the Customer Service Large Model through an efficient collaboration between the online marketing service center and the Jiutian Innovation Research Institute, thereby enhancing smart customer service capabilities. The overall project architecture is illustrated in the figure below:

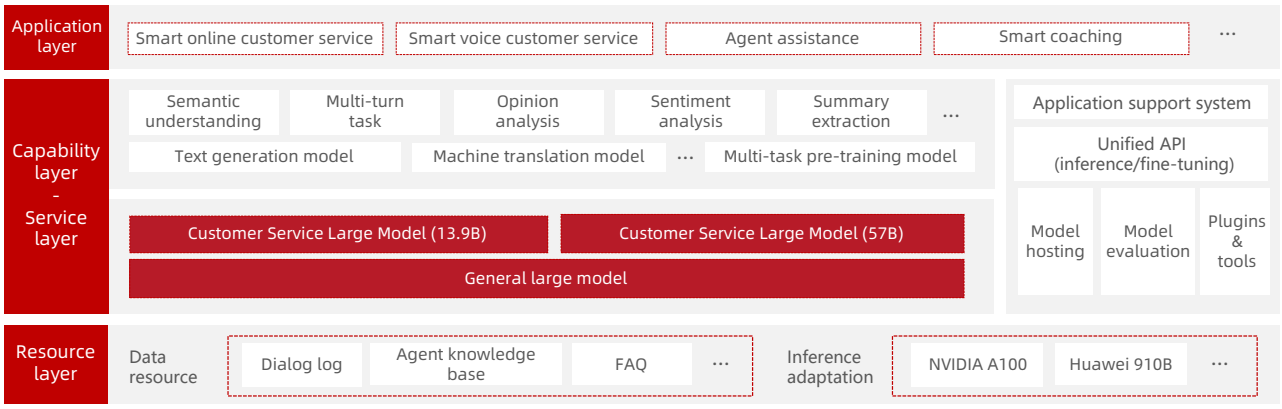


Figure 1: Architecture of the Customer Service Large Model

02 Application scenarios

In addition to existing explorations of customer-oriented smart services, the large model's strengths in multimodal and human-like interactions will be adopted to enhance the intelligence of smart customer service, agent assistance, and smart coaching. The goal is to transform and upgrade service experience and management production models, enriching smart and user-friendly services through smart companionship, emotion-driven, and multimodal interactions.

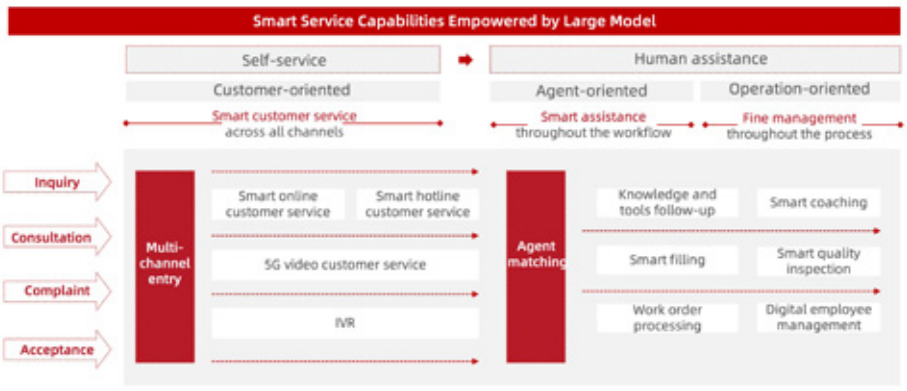


Figure 2: Application Scenarios

Enhanced smart customer service - online customer service

By incorporating China Mobile's customer service knowledge base, the large model constructs retrieval-augmented generation capabilities, improving the professionalism and accuracy of online customer service responses. Dialogs transition from process-based configurations to on-demand execution based on scenarios combined with the large model. The collaboration between the large model and small models enhances capabilities like predicting inquiries and emotion recognition, ultimately improving the smart interaction experience. As of 2024, the application has been made available across the entire network, with a cumulative testing customer base of 17,979,100.

Figure 3: Online Customer Service Scenario

Enhanced smart customer service - voice hotline

Utilizing the large model's strong intent recognition and multimodal interaction capabilities, initiatives are underway to explore business upgrades in voice scenarios such as home network troubleshooting. A pilot program for "voice + semantics" end-to-end intent understanding is set to launch by the end of October 2024. This includes the rollout of a nighttime voice portal and the activation of a smart hotline for after-hours customer service. These developments aim to meet customers' urgent needs, enable self-service information retrieval, and optimize the allocation of nighttime customer service resources.

Enhanced agent assistance - smart filling

By employing the large model's strong intent recognition and automatic script generation capabilities, this service offers various types of support to customer service agents. During call interactions, it transitions from "manual input/search" to "smart judgment with proactive notification/generation" for transaction recommendation scripts, consultation service scripts, and complaint work order creation, thereby equipping agents with superior tools. With over 700 online agents utilizing the system, the usage rate of instant work order promotion has reached 62.87%, and the time taken for agents to fill in the work orders has been reduced by 20 seconds.

Enhanced smart coaching

To tackle the high training frequency and broad coverage, and high skill requirements for phone operators, the large model enhances smart generation, automatic supervision, and error correction capabilities throughout the training and quality inspection processes. By 2024, it has automated the mining of training scripts and generated comprehensive training materials. This initiative has been piloted across provincial enterprises, creating 20 customized courses that support online training for about 17,000 participants.

03 Key innovations

Integration of specialization and collaboration

Built on China Mobile's in-house "Jiutian Large Model" as a foundational capability, the model harnesses advanced deep learning algorithms to enable efficient and smart interactions across various fields. By incorporating reinforcement learning from human feedback (RLHF) and utilizing a "large-small model" synergy, it significantly enhances performance in customer service applications, delivering more accurate and effective Q&A interactions.

Integrated training & inference and technological innovation

Driven by the development and training of the Jiutian Large Model, the Customer Service Large Model boasts industry-leading parameters in tens of billions. This advancement has spurred independent innovation in an end-to-end, one-stack core technology based on the Transformer algorithm. It encompasses data construction, pre-training, fine-tuning, and inference, exemplifying excellence in converting technology into productivity.

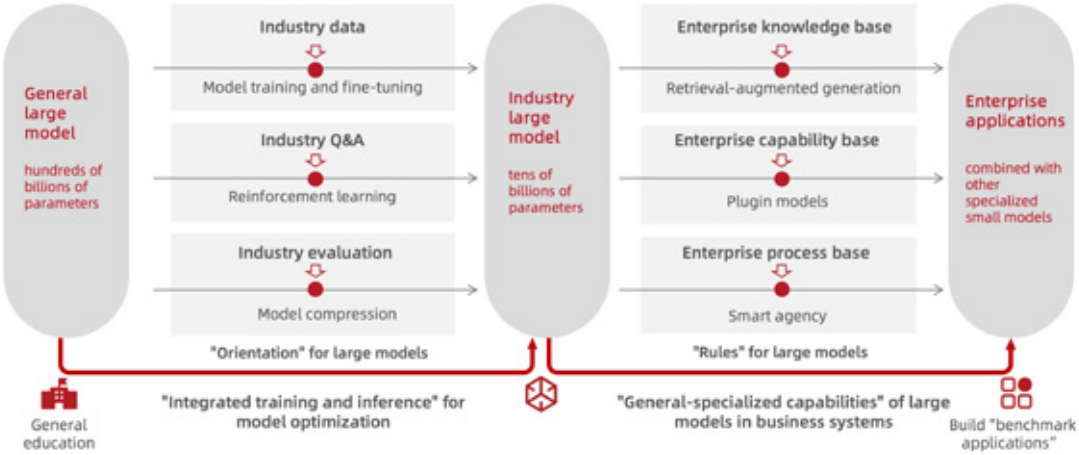


Figure 3: Innovative Approach to the Development of the Customer Service Large Model

Flexible customization and emotion empowerment

By flexibly applying natural language processing (NLP) and deep learning algorithms, the model captures user emotions in real-time through emotion recognition and computation technologies. While answering customer inquiries, it also provides emotional support, care, and precise recommendations, delivering a thoughtful, humanized, and personalized service experience.

Professional training for safe and reliable performance

The training process emphasizes value handling and employs cutting-edge AI-generated content (AIGC) risk control solutions to efficiently identify risks in real-time AIGC. This ensures responses are safe and controllable, demonstrating excellent ethical and safety control capabilities.

04 Business models

Through commercial exploration, China Mobile's Customer Service Large Model has formed two business models: standardized products and customized implementation. Standardized products include large model management platforms and large model scenario applications, allowing customers to "order" based on their needs. These come with standardized charges depending on the selected functional modules, enabling quick deployment. Customized implementation is available to meet specific needs. This includes customized development services for large model training, fine-tuning, and scenario applications, with charges based on specific workload. This approach provides maximum flexibility and the highest stability requirements.

05 Core values

In-house Customer Service Large Model to upgrade customer service experience

China Mobile's Customer Service Large Model, built on China Mobile's proprietary "Jiutian General Large Model", is specifically trained with data tailored to the customer service sector. It features interactive capabilities of a large model with tens of billions of parameters and professional customer service capabilities covering all of China Mobile's business areas. To date, over 80 million users have been invited to test this model, leading to widespread application and a significant enhancement in customer service experience.

Innovative development model for a pioneering development pathway

China Mobile's Customer Service Large Model is built on the innovative AI + X systematic AI development model, seamlessly integrating both general and specialized capabilities to promote model training and inference. This approach has created a development pathway that transitions from "shelves" to "showcases". By starting with lightweight scenarios, the model fully leverages its strengths in intent understanding and knowledge generation, thereby upgrading existing service intelligence and management models.

REFLECTION

Reflections

The implementation of the Customer Service Large Model in enhancing smart customer service scenarios has been a gradual journey. It has drawn strength from the independently developed robust smart interaction foundation, a rich professional dataset, and continuous iterations of the reinforcement learning technology. Through multiple version iterations and extensive internal testing, the model has continually improved its accuracy, completeness, and friendliness. This progress has driven diverse value in technological innovation, flexible customization, and emotional empowerment, providing customers with thoughtful, humanized, and personalized service experience.

Follow-up plans

Exploring the productization and ecosystem development of service capabilities empowered by the Customer Service Large Model to enhance service value

Efforts will focus on packaging the mature applications of large customer service models into monetizable service capabilities. Through standardized prefabrication and personalized customization, products such as smart interaction, smart agent assistant, and smart coaching will be developed to meet the needs of various service industries, including aviation, banking, and consumer services. Additionally, expanding collaboration ecosystems will facilitate technological cooperation related to large models, promote joint efforts in smart computing infrastructure, and drive standardization efforts. This will speed up the productization and visibility of service value, accelerating the formation of new quality productive forces in the service sector.

Facilitating structural transformation of personnel

Plans will continue to support the structural transformation of frontline service personnel, gradually transitioning certain operational and production personnel into smart service support engineers, while some existing phone operators will be shifted to model fine-tuning and data training roles. Additionally, a job cluster and certification system for digital service support engineers will be developed to aid in personnel transformation and enhance capability management and assessment.

"Diting" — AI Agent for Customer Service

As the AI era unfolds, China Telecom has expanded its business from the conventional communications field to cloud, network, data, intelligence, security, platform, and other fields. However, conventional natural language understanding (NLU) cannot meet the requirements of complex service scenarios. China Telecom faces a series of problems, such as increasing stress on agents, uneven quality of service, and high training costs. To address the pain points of its customer service hotline 10000, China Telecom Hubei Branch partnered with China Telecom Artificial Intelligence Technology Co., Ltd to build "Diting", an AI agent for customer service. Large language models (LLM) are leveraged to optimize intelligent customer service processes, delivering a heartwarming and professional service experience and boosting user experience with lower operating costs.

Zhang Min
General Manager, China Telecom Hubei Branch

SOLUTION PARTNERS

As the telecommunications industry keeps transforming and evolving and the business is continuously enriched and iterated, the conventional NLU no longer meets the telecommunications customer service needs in complex service scenarios. Numerous human customer service requests need to be processed. In one-to-one service scenarios with customers, Hotline 10000 agents face increasing stress and higher demands for remote services. The uneven quality of their responses causes the training costs to keep rising. These pain points have greatly deteriorated the user experience of China Telecom and driven up its operating costs.

To figure out a problem- and customer-oriented solution, China Telecom Hubei Branch established a deep partnership with China Telecom Artificial Intelligence Technology Co., Ltd. To tackle the weak links of the Hotline 10000 service, they explored the AI+ concept to rebuild the customer service process and create the AI-powered customer service agent "Diting". They integrated the capabilities of LLMs into the intelligent customer response process to enhance human-like service capabilities, guiding customer service personnel to provide customers with a heartwarming and professional service experience.

CHALLENGES

Challenges faced by operators' customer services:

1. As business operations grow more varied and complex, AI is urgently needed to boost efficiency.

With the arrival of the intelligent era represented by large models, the products powered by China Telecom Hubei Branch have become increasingly complex. In addition to basic services such as landline, broadband, and mobile services, they also include Chinatelecom Cloud, two-line services, big data applications, security, satellite, AI large models, and industry applications. These products, combined with customer characteristic tags and personalized needs, have resulted in a richer array of sales products and marketing packages. All of these require support for complaints, obstacles, consultations, and business opportunities through the 10000 hotline system. A brief statistic shows that the services provided by the 10000 hotline have exceeded 200 scenarios and are continuously expanding, necessitating the introduction of AI tools for smart knowledge compilation and searching to improve operational efficiency.

2. Customer service agents lack proficiency, leading to repeated calls and a poor user experience.

The rapid changes in new products and services have led to a proliferation of policy documents for products, sales items, and scenarios. This situation is compounded by a high volume of recorded hotline calls and inefficient knowledge base management. Consequently, new employees typically need 3 months of training before they can start working, and frequent personnel turnover results in a low first contact resolution rate. To address these challenges, there is an urgent need for introducing smart assistant services in a co-pilot model to guide agents in improving service efficiency.

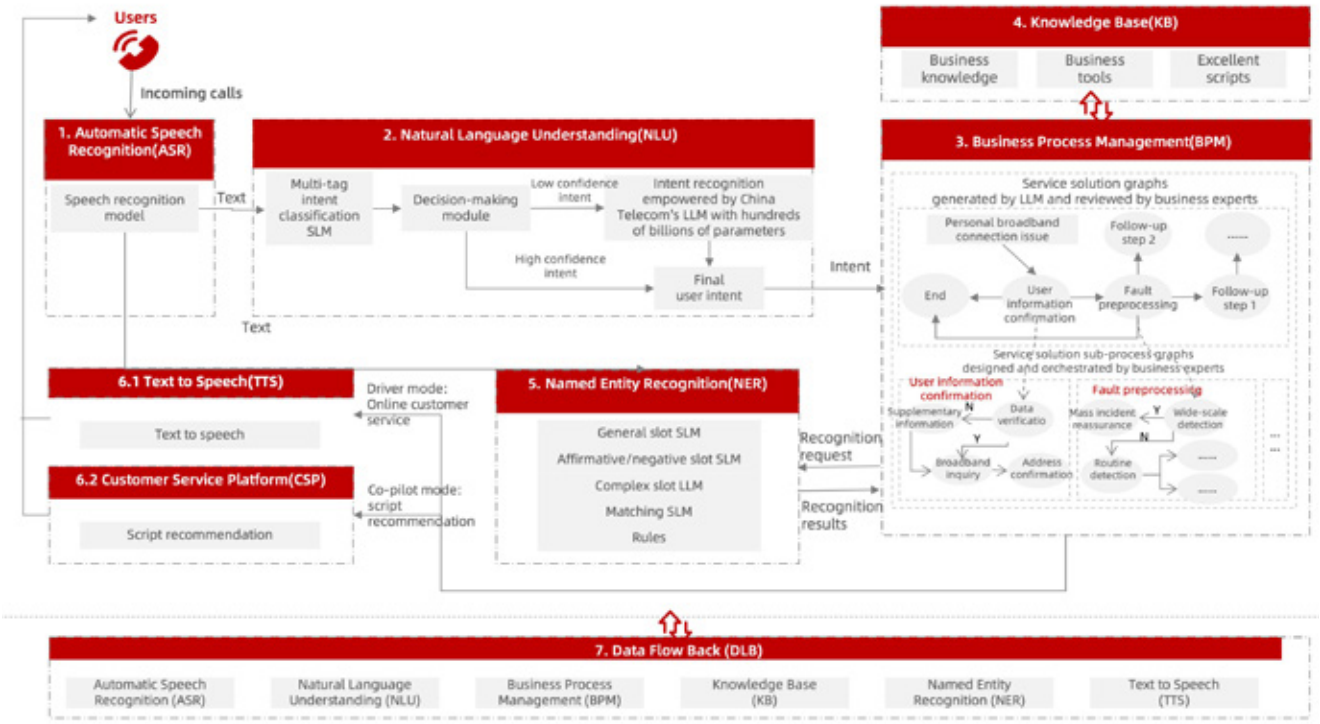
3. Customer service requires high investments in human capital.

China Telecom Hubei Branch currently has over 300 agents handling voice hotline services and nearly 50 agents dedicated to instant messaging (IM) services over the Internet. Investing heavily in human resources not only raises operational costs but also complicates risk management during pandemics and emergencies. There is an urgent need to introduce AI-powered human-machine dialog agents in a driver model to divert incoming calls, reducing costs and increasing efficiency.

SOLUTION AND VALUE

01 Overall structure/solution

The AI-powered customer service agent "Diting" consists of eight major capabilities: automatic speech recognition (ASR), natural language understanding (NLU), named entity recognition (NER), business process management (BPM), knowledge base (KB), text to speech (TTS), customer service platform (CSP), and data flow back (DLB). The following figure shows the overall system architecture.



After a user's call is connected, the AI agent converts the call content between the user and the agent to text using China Telecom's proprietary automatic voice recognition capabilities for multiple dialects, as shown in the first point in the figure.

The AI agent performs NLU on the dialog text between the user and the agent, as shown in the second point in the figure. First, China Telecom's small language model (SLM) for intent classification performs pre-recognition and sends the results to the decision-making module. If a high-confidence intent is identified, the intent is directly output as the final intent. If no high-confidence intents are identified, China Telecom's LLM starts to perform intent recognition by using the TeleChat2-115B model that support hundreds of billions of parameters. This LLM will be continuously upgraded and evolved.

The AI agent provides the user intent to the process management platform. The platform searches for a matching service solution generated by models and automatically executes it, as shown in the third point in the figure. The business knowledge, tools, and scripts required by the service solution can be obtained from the KB, as shown in the fourth point in the figure. The user number and certificate information required by the service solution can be recognized through NER, as shown in the fifth point in the figure.

The BPM will output targeted service scripts based on the execution results of the service solution. Depending on the scenarios, the targeted service scripts can be synthesized into speech to provide online customer services in the form of human-machine dialogs with the driver. This way, the agents can instantly respond to and serve customers around the clock, as shown in the 6.1 in the figure. On the other hand, the targeted service scripts can be pushed to the customer service platform as a reference for the agents, as shown in the 6.2 in the figure.

The DLB module can evaluate all the preceding service processes online. The data that is flowed back can be used for evaluation and training,

continuously optimizing the model, service solutions, intelligent knowledge base, and text-to-speech results, as shown in the seventh point in the figure.

To enhance reuse and sharing, we can also decouple the capabilities from business processes by layer. At the product layer, the ASR, TTS, and LLMs, SLMs, agent capability engine, and KB can be decoupled as atomic capabilities to adapt to models and modules from different manufacturers. As applications, products support personalized configuration and low-code visualization and can connect to the call center and the call platform of customers across different industries. At the PaaS layer, the MRCP call engine, the AI-powered data annotation and evaluation middle platform, the LLM training and inference engine, the data flow-back, and the monitoring and statistics engine are provided. They have complete capabilities that feature high scalability and availability and support grayscale release and sustainable operations. This allows the product to be implemented with all-in-one capabilities in vertical industries.

02 Application scenarios

Based on China Telecom's leading AI technology and intelligent interaction technology, the AI-powered customer service agent "Diting" provides solutions that include products, applications, and PaaS and assist ecological partners with open capabilities.

The AI agent applies to customer service scenarios, for example, agent assistance (co-pilot), interactive voice response (IVR, driver), and smart hotline. In addition, the value of this product has proved in all industries. Technology upgrades and marketing have broadened its applications across various industries. The new product and its applications have boosted efficiency and value.

Online customer service — "Diting" is on call at any time. It can provide online customer services in the form of human-machine dialogs like a driver. This way, the AI agent can instantly respond to and serve customers around the clock. Its features include online smart dialog, online service result notification, the ability for users to engage in casual chat and interrupt AI agent's dialogs, mute detection, and automatic transfer to human support for complex scenarios as per user requests.

Agent assistance — The AI-powered customer service agent "Diting" is immune to the complexity of telecommunications services and user expressions. It recommends scripts in dialog to agents like a co-pilot, guiding them to accurately understand customer intents. The agent enables them to efficiently resolve customer complaints and accessibility issues, provide consultations, and capture business opportunities.

Government hotline — Many local governments have used "Diting" as agent assistants to provide voice interaction, reducing the workload of agents and improving their work efficiency. "Diting" can quickly and accurately identify citizens' intentions through RAG knowledge-enhanced Q&A and multi-turn dialog based on large models. It guides the public through service processes and extracts work order elements to assist in form filling. Currently, "Diting" has expanded beyond the telecommunications industry to the general public service and government sectors.

Use cases

Livelihood appeal project in Shenzhen

To implement the "Shenzhen Artificial Intelligence Action Plan", Shenzhen has released a total of 73 application scenarios in 3 batches. As the first scenario, livelihood appeals are included in the pilot. The livelihood appeal platform is upgraded from "Process + IT" to "Process + AI", reengineering the entire process of appeal services. The plan includes 28 smart application scenarios based on 8 roles and 17 service processes of the platform. An open competition mechanism is adopted to select the best 8+12 large models for trials. The pilot scenarios will be implemented by September 2024. In the project, China Telecom's multiple features are integrated into the intelligent agent assistant scenario. They include the multi-dialect ASR for real-time translation, AI-powered customer service agent "Diting", intelligent transaction processing assistant, intelligent knowledge base, and content security review. The platform intelligently analyzes and recognizes citizens' intents, recommends items and knowledge, assists citizens in filling in forms, and provides security warnings and guidance. The call processing time decreases from an average of 51 seconds per case to 30 seconds, improving the efficiency by 41%.

Dongguan municipal government service hotline 12345

The hotline adopts the "Xingchen" government service LLM and the AI agent "Diting". These capabilities are deeply embedded into the production and operation processes of the hotline to provide a wide range of features for phone operators. These features include real-time transcription, intelligent knowledge recommendation, intelligent process navigation, intelligent script recommendation, intelligent call work order summary and conversation element extraction, LLM-based intelligent form filling and dispatching, and intelligent knowledge base Q&A. They help managers analyze data operations analysis and monitor the number of calls and work orders, and automatically identify concentrated hotspots. "Diting" can assist the agents and provide them with intelligent training, reducing the operator training time by 50% and the training cost by 70%. The platform helps polish the skills of the agents, improving their service efficiency. The average call duration decreases from 65 seconds to 42 seconds, improving overall customer service efficiency by 35%. This greatly improves the call receiving capacity of the government service hotline 12345 and significantly enhances the perception and satisfaction of citizens. This greatly shortens the wait time and call duration, improving public satisfaction.

03 Key innovations

In terms of algorithm engineering

the AI-powered customer service agent "Diting" adopts a system architecture where SLMs and LLMs collaborate to adapt to different scenarios with their complementary advantages. You can manage the entire life cycle and pipeline of AI models, including model training optimization, fine-tuning, inference acceleration, model evaluation, and data flow back. The China Telecom LLMs are used for complex intent recognition, knowledge collection and editing, and complex entity information extraction. The China Telecom SLMs are used for ASR, simple intent recognition, knowledge matching, and simple entity information extraction. The decision-making module allows you to set thresholds and coordinates the working models. By collaborating with SLMs and LLMs, "Diting" can accurately and quickly identify user intents and automatically extract key information such as the ID card number, numbers, addresses, and packages mentioned by users. In complex scenarios, it can effectively interact with customers in multiple rounds in the driver or co-pilot mode as required by specific industries and solve tough problems for them by serving as online agents or recommending scripts.

In terms of business processes

the Hotline 10000 system of China Telecom Hubei Branch provides services in 200 service scenarios and manages nearly 800 business APIs. "Diting" introduces a two-layer knowledge graph to accumulate the experiences of service experts and accurately adapt to different intents. This paves the way for offering personalized service solutions to customers. The Layer-1 knowledge graph focuses on service solutions, ensuring that accurate service actions are taken to meet varying customer needs. The Layer-2 knowledge graph focuses on service components and defines the capabilities required for service actions, ensuring that service actions are efficiently taken. The two-layer service graph helps quickly generate and optimize personalized service solutions for customers and improves the automation and intelligence of service processes. It makes customer service more efficient, accurate, and personalized.

In terms of autonomous learning,

"Diting" adopts the reinforcement learning mechanism. By leveraging data flow back, the AI agent simulates the interaction with customers in customer service environments. This way, the models can continuously learn and evolve while dealing with practical problems. In addition, the AI agent can dynamically adjust its service policies based on user feedback and actual results. As the AI agent is used more frequently, the models will gradually adapt to different user groups and service scenarios, providing more personalized and targeted services. This interactive process is similar to how humans accumulate experience and improve their behavior while interacting with customers in service environments. It keeps learning and growing, getting smarter as you use it more frequently. This way, the system continuously adapts to ever-changing customer needs, delivering higher-quality customer service more efficiently.

04 Business models

End-to-end solution

"Diting" is not restricted to the telecommunications industry. With powerful intelligent processing capabilities, it can adapt to and serve multiple vertical industries such as governmental affairs, education, tourism, catering, retail, e-commerce, medical care, and finance. It can provide customized end-to-end service solutions tailored to industry-specific service needs and pain points. These solutions encompass the entire process, from customer consultation and problem diagnosis to service provision, after-sales support, and data analysis and feedback. "Diting" optimizes customer experience and enhances the competitive edge of enterprises, boosting their service quality and customer satisfaction.

Modular service solutions

In addition to comprehensive solutions, "Diting" provides highly modular model services. If you need to recognize technical terms in the medical industry, consult on courses in the education sector, and recommend products in the retail industry, you can use this AI agent. It can provide distinct professional models, such as those for intent recognition, process generation, and script optimization, tailored to industry-specific needs, offering intelligent solutions for various services. These models can be seamlessly integrated into a company's existing systems to enhance service intelligence.

Modular service tools

To meet the requirements for automated services in different industries, "Diting" provides a wide range of general modular service tools. Solution orchestration tools enable enterprises to customize service solutions based on their business processes with a few steps. Knowledge management tools help enterprises efficiently organize and use knowledge resources. Capability integration tools enable enterprises to seamlessly connect the intelligent capabilities of "Diting" and other systems to implement global intelligence.

Cross-industry data insights and business optimization

With the intelligent robot and platform services of "Diting", enterprises can enhance customer service quality and gather extensive user interaction data. "Diting" conducts in-depth data mining using advanced data analysis technology and outputs analysis reports for related industries. These reports provide valuable market insights and business optimization suggestions for enterprises, helping them capture market opportunities and design targeted marketing strategies.

05 Core values

Improve competitive edges

According to user feedback, most enterprises that use "Diting" have improved their customer satisfaction by more than 10%. After receiving timely and professional services, customers are more likely to make repeat purchases or purchase higher-value products and services. In addition, these customers are more likely to recommend this company to other people, forming word-of-mouth marketing influence. Ultimately, more potential customers are attracted and engaged.

Improve operational efficiency

"Diting" plays a crucial role in the digital transformation of an enterprise. By automatically processing customer queries, "Diting" significantly shortens the average response time of customer service personnel by 30% compared to traditional systems.

Reduce operating costs

"Diting" significantly cuts human capital expenses by reducing the number and working hours of customer service personnel. In addition, the modular reuse design of "Diting" reduces system construction and maintenance costs, allowing enterprises to deliver higher quality services at a lower cost.

REFLECTION

Reflections

- When recognizing complex intents and key information, the AI agent collaborates with SLMs and LLMs and uses multi-agent orchestration technology. The understanding capability of the LLMs improves the recognition accuracy, while the computing power consumption is greatly reduced, preventing LLM hallucinations.
- When executing service solutions, the AI agent employs visual service orchestration and large-scale parallel processing technology. A service solution graph is modeled to capture the expertise of customer service experts.
- With data flow back, the AI agent can evaluate results and autonomously learn the data, continuously optimize the models and service solutions, getting smarter with every action it takes.

Follow-up plans

- Cross-industry applications: "Diting" has shown strong adaptability across various industries and is included in the key AI-powered product catalog of China Telecom Hubei Branch. In the future, the AI agent will be applied to multiple vertical industries in depth, such as medical consultation, e-commerce services, and financial services. Industry-specific intelligent customer service solutions will be jointly customized in partnership with leading enterprises in each industry, helping enterprises improve service quality and customer satisfaction.
- Technology upgrade and iteration: "Diting" will continuously undergo technology upgrades and iterations. User feedback and interaction data will be continuously collected to iteratively optimize SLMs and LLMs, improving the recognition accuracy and shortening the response time. Attention will be attached to the latest research results in the LLM field and advanced algorithms and technologies will be introduced to improve the intelligence of the system.
- Ecological construction and cooperation: "Diting" will build a cross-industry partner ecosystem. A close partnership will be established with leading companies across various industries to jointly explore the application of intelligent customer service in more scenarios. In addition, the APIs and development documents of "Diting" will be open-sourced to encourage developers to conduct secondary development and drive application innovations accordingly. By establishing a developer community, we will promote technical exchange and collaboration to jointly advance intelligent customer service technology and foster a thriving ecosystem.

In summary, the AI-powered customer service agent "Diting" will continue to expand its application scope and market influence and provide intelligent and automated customer service solutions for more industries. It will help enterprises in terms of digital transformation and upgrade.

Digital Human Product Assistant Empowered by Large Models

AI has become a key driving force behind a new round of technological revolution and industrial transformation, profoundly impacting global economic and social development as well as the progress of human civilization. Virtual digital humans, as a product of advancing AI technology, have gradually transitioned from concept to reality. The market for virtual digital humans is expected to exceed RMB 300 billion, with annual growth rates surpassing 50%. They are increasingly being applied in various scenarios such as exhibition halls, cultural tourism, education, live streaming, and gaming. Virtual digital human technology fueled by large model Q&A can reduce costs in service industries, enhance AI interaction experiences, and provide core interactive mediation for the metaverse.

Jiang Yongqiang
Deputy Director of the Digitalization Center, China Telecom Shandong Branch

SOLUTION PARTNERS

China Telecom interacts with customers using digital human technology in scenarios like smart exhibition halls, visual customer service, dual-teacher classrooms, and live commerce, enhancing user engagement. Applicable scenarios include operator exhibition halls, historical sites, museums, theaters, sports venues, patriotic education bases, film and television, internet video sites, and short video media platforms. Digital humans are empowered by AI in logical reasoning, visual capture, voice recognition, and language expression. Key technologies include motion capture, speech synthesis, emotional simulation, facial expression generation, VR and AR presentation, and real-time rendering. By leveraging the language capabilities of large models, digital humans can obtain language expression and emotional capabilities, and emotion types and key elements can be extracted for coding. Using EEG signals and eye movement signals and conveying emotions through natural language and voice tone, these digital humans interact with users more like real humans. The project has successfully been implemented in various locations, including Shandong Telecom's smart exhibition hall, Zibo Telecom's video customer service, China Life Insurance in Shandong, Zaozhuang Vocational College of Science and Technology, and Linyi Shunhe E-commerce Logistics Park, bringing direct economic benefits exceeding tens of millions of RMB and generating significant economic and social value.

CHALLENGES

Technology and interaction experience

Digital humans are widely used in smart customer service and marketing, but these fields often feature relatively simple service content. Intense industry competition has led to isolated R&D among leading players, making it difficult to share AI technology frameworks and corpora. For the time being, digital humans are mostly deployed in simpler environments and for basic issues. They may provide irrelevant responses or fall in an infinite loop during scenario switches or multi-turn dialogs, which hinders user experience.

Content creation and operation

The core value of virtual digital humans lies in their ability to provide diverse content and services. However, the quality of interactive content varies significantly, and digital humans struggle to meet the growing personalized demands of users. Ongoing content updates and operations also pose challenges, as virtual digital humans must continuously introduce new content and features to maintain user interest and engagement.

Law and ethics

Digital humans raise various legal and ethical issues. Legally, the protection of intellectual property, data privacy, and the legal status of virtual assets related to virtual digital humans remain unclear, potentially leading to disputes. Issues such as whether the design and voice characteristics of virtual digital humans are protected by intellectual property rights and whether they hold copyright in generated content need legal clarification. Furthermore, questions about whether virtual digital humans should possess consciousness and rights, and how to prevent them from having a negative psychological impact on humans, require careful consideration to ensure their development aligns with ethical standards.

SOLUTION AND VALUE

01 Overall structure/solution

Utilizing the extensive corpus of the Xingchen large model, virtual digital humans can replace human explainers. With voice interaction and emotional exchange, these virtual digital humans can provide personalized and professional explanations to enhance visitors' learning and interaction with relevant services.

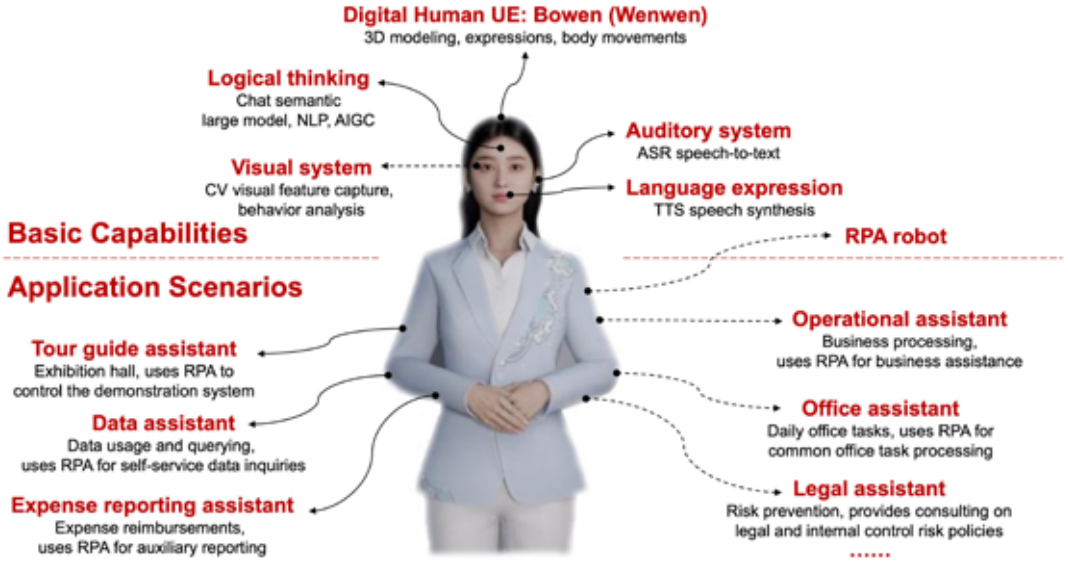


Figure 1: Technical Architecture of Virtual Digital Human

02 Application scenarios

China Telecom's digital human applications are primarily in the cloud-network and government-enterprise sectors. They facilitate interactive automation in exhibition halls (such as internal networks, data centers, cloud platforms, and equipment operating conditions) through two-way interaction with clients. They also empower employees in assistant scenarios (like visual agent and smart customer service), engage in cultural tourism (interacting with visitors at historical sites, museums, theaters, sports venues, and patriotic education bases), support educational environments (dual-teacher classrooms and holographic classrooms, where the teaching instructor is replicated 1:1 at other teaching locations, with some courses led by virtual digital teachers), and serve e-commerce streaming (innovative streaming scenario with virtuality and reality combination and virtual human assistants to improve consumer experience).

Exhibition assistant

The economic benefits have been significant, replacing much of the reception, consultation, and explanation work, thus reducing labor costs. The Shandong Telecom exhibition hall spans 1,750 square meters and features eleven major exhibition areas and thirty-seven booths. Digital human assistants take over tasks like one-click exhibition setup, smart interaction, and automated tours. The AI-enabled platform underpins digital assistants, providing a one-stop solution for capabilities such as voice processing, large model operations, knowledge bases, and Robotic Process Automation (RPA). This has resulted in annual savings of at least RMB 600,000, with an expected indirect revenue increase of over RMB 10 million.

Education and training

Zaozhuang Vocational College of Science and Technology, located in the ZTE Production-Education Integration Pilot Zone in Zaozhuang's Shizhong District, covers an area of 575 acres and is planned to accommodate 15,000 full-time junior college students. Due to the large student population and space limitations, there is an urgent need for dual-teacher classrooms and holographic projections. AI digital human enable easy replication of teachers' images, voices, and actions, assisting primary instructors in dialog and interaction with students, providing feedback

and suggestions. By generating digital human teaching videos, the efficiency of content creation is significantly enhanced, effectively meeting students' needs for relevant courses.

Government and enterprise services

For citizens visiting government service halls for the first time, the "AI Digital Human" can provide detailed guidance and process introductions, helping them quickly understand the necessary steps and materials for the services they require. The "AI Digital Human" offers 24-hour unmanned smart support, assisting the public with basic information verification, answering questions, and offering timely help.

Cultural and tourism services

In Jinan, the cultural tourism attractions span large areas and lack professional tour guides. The digital human can offer personalized guided tours, tailoring itineraries and content based on visitors' behavior patterns and interests, greatly enhancing their touring experience. The AI digital human can also combine real scenic views for live broadcasts, interact with visitors in real-time, lead them on virtual tours, showcase local specialties, and explain regional features and cultural heritage. This approach helps to uncover potential visitors, expand the reach of cultural tourism promotion, and enhance the visibility of attractions, driving traffic to physical sites.

Livestreaming hosts

Digital humans serve over 300 high-definition streaming merchants and thousands of hosts at the Linyi Lanshan Shunhe E-commerce Logistics Park. Through the integration of real-time livestreaming and virtual human co-hosting, innovative streaming scenarios are created, improving the consumer experience.

Figure 2: Digital Humans in China Telecom Exhibition Halls

Figure 3: Digital Humans in Livestream Shopping

03 Key innovations

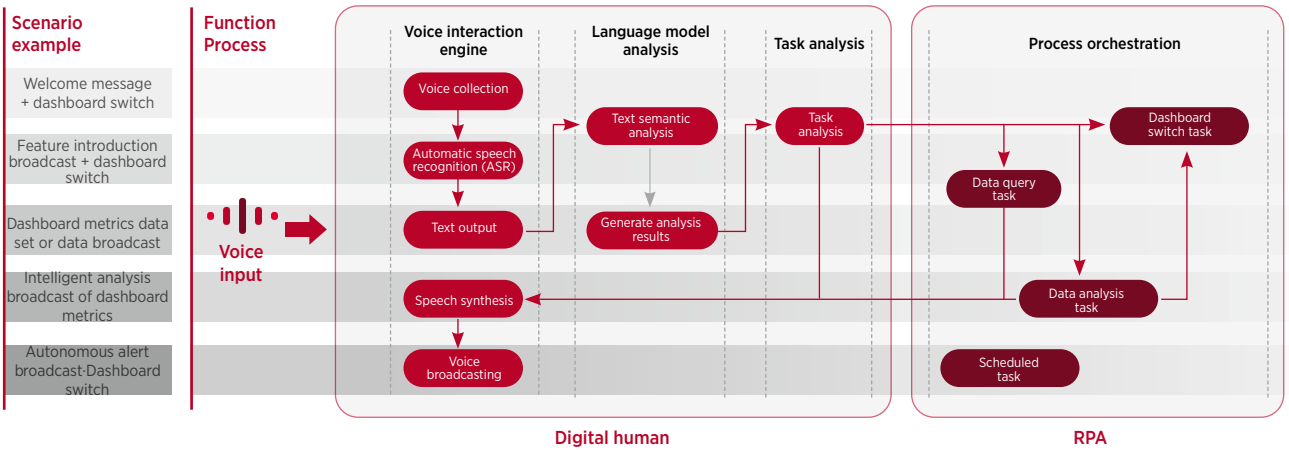


Figure 4: Functions of Digital Human Assistant Function Workflow

Scenario Implementation Highlights

- Natural language interaction: intuitive, real-time intelligent responses that are convenient and efficient;
- Intelligent decision support: data-driven analysis providing accurate and timely decision support, reducing decision risks;
- Real-time data monitoring and alerts: alerts set up to notify when thresholds are exceeded.

Scenario Implementation Deficiencies

- Supported scenarios are limited: data querying, analysis, and alerts only support predefined analytical algorithms and models, as well as alert indicator data.

Digital human automation capabilities

Virtual image
Digital humans replace human presenters for voice interaction. They provide explanations, answers, and guided services through their virtual images, offering personalized and customized visitor experiences that enhance user perception and technological engagement in exhibition halls.

Behavioral capabilities:
Research focuses on integrating Robotic Process Automation (RPA) with virtual digital human images, which greatly expands the capabilities of virtual digital human representations and enables them to "explain while operating."

Logical thinking
Using large models, digital humans can engage in smart Q&A, addressing both specialized questions and open-ended interactions.

Speech synthesis
By utilizing trained acoustic models and vocoders, speech synthesis can be performed. The main process involves converting input text into outputs from the acoustic model, which are then passed to the vocoder to generate sound waves.

Emotion recognition and augmented reality

Digital humans can understand and simulate human emotions, offering a more natural and personalized interaction experience. This includes emotion modeling, generation, recognition, and encoding. Through self-learning, digital humans continuously improve their behavior and interaction methods to predict audience actions and optimize the interaction experience. In virtual and augmented reality, digital humans create immersive interaction experiences, including virtual scene construction and augmented reality interactions.

Innovation in work modes

Innovation in digital human user experience and semantic large model applications: Digital human images replace human presenters, using large models and voice interaction technology for intention analysis, enabling interactive explanations, answers, and guided services.

Innovation in exhibition hall work modes: An integrated support system is established for exhibition scenarios, from preparation and warm-up interactions to automated sequential tours, reducing costs and enhancing intelligence.

04 Business models

Production services
Digital human producers offer digital human customization services

Feature services
Digital human developers or providers offer customized services like voice recognition and emotion analysis

Technical support
Digital human technology developers offer technical support or product services such as technology, algorithms, and applications

Subscription services
Digital human developers or providers offer subscription services for digital human features and services, charging monthly or annually

Training and consulting
Digital human developers or providers offer training, guidance, and consulting services in areas such as development, usage, operation, and deployment

Commercial pathways for virtual digital human applications

Platform services
Digital human developers or providers offer platform services that enable the creation, customization, and management of digital humans

Agency services
Digital human agents manage commercial collaborations and provide related services

Marketing services
Digital human service providers and MCN agencies offer personalized marketing and planning services

Operation services
Digital human service providers, MCN agencies, or management firms offer content production, live streaming operations, and other services

Copyright/Licensing
Digital human developers or providers offer usage rights or licenses for digital humans to generate commercial revenue

Figure 5: Commercial Pathways for Digital Humans

Digital humans mainly focus on interactive sectors like cultural tourism, education, and e-commerce. Revenue sources include customized platform construction, service subscriptions, and content operations.

Platform contracting

Developers of digital humans offer self-creation and custom platform development services, with one-time fees that cover custom development, upgrades, and maintenance. This model suits organizations lacking technical implementation capabilities that require comprehensive solutions.

Service subscriptions

Ideal for service contracts, this model allows industries to subscribe to digital human capabilities, choosing types, rendering methods, and background displays, with costs lower than platform development. Fees are charged based on specific subscription features.

Content operations

Leveraging the Xingcheng large model, businesses can quickly generate marketing materials, such as product detail images in seconds, resulting in significant cost reductions and efficiency improvements. The system supports context-aware multi-turn dialog, enhancing communication to be smoother and more natural. Even in complex dialog scenarios, AI digital humans maintain coherence, progressively deepening their understanding of user needs to provide more precise solutions.

05 Core values

The precise alignment of digital humans with specific scenarios enhances application functionality and supports intelligent service solutions. This has led to the production of five patents related to intelligent digital human control, voice extraction, and process model handling, as well as two software copyrights for the iTelBot process automation robot. Additionally, there is one evaluation of digital human exhibition applications based on large models. As exhibition assistants, digital humans have undertaken most reception, consultation, and explanation tasks, reducing labor input and saving RMB 600,000 in labor costs annually. The large model agents and retrieval enhancements have infused digital humans with powerful language understanding capabilities. In the daily operations of government service and cultural tourism services, these technologies have reduced labor costs and resource consumption by over 50%. Furthermore, the incorporation of external knowledge bases has improved the system's answer accuracy to over 96%, addressing common issues of misunderstanding and hallucination found in traditional customer service robots. Digital humans can understand and simulate human emotions, offering a more natural and personalized interaction experience. This includes emotion modeling, generation, recognition, and encoding. Through self-learning, digital humans have entered reality to offer emotional companionship. Digital humans have generated tens of millions of RMB in direct revenue and have indirectly driven even greater business income in related application scenarios.

REFLECTION

Digital humans have immense potential to promote the smart transformation of industries. They have successfully been applied in various fields, including smart exhibition halls, education, cultural tourism, and e-commerce. Through technological innovation and optimization of business models, they have brought significant economic and social benefits to partnering organizations. The automation capabilities of digital humans, along with their emotion recognition and augmented reality technologies, as well as innovative work models, have greatly enhanced user experience and service efficiency.

Looking ahead, there is still considerable growth room in areas such as digital human technology research and development, application demonstration, and industry clustering.

Further enhance the intelligence to make digital humans super-smart agents capable of handling complex tasks

Digital humans possess strong learning and reasoning abilities, able to quickly and accurately tackle intricate tasks. By utilizing technologies such as machine learning, deep learning, natural language processing, computer vision, and reinforcement learning, they can learn from vast amounts of data, identify patterns, and understand and process natural language. They can also make decisions based on observations and environmental perceptions. Next, there is a need to pay more attention to issues such as privacy protection, data security, and human-machine relationships. Appropriate mechanisms should be introduced to monitor and constrain their behaviors to avoid potential risks and misuse.

Engage in more creative work to accelerate the transformation of creative generation

Digital humans can produce unique ideas and concepts, including in fields like art, film, and music creation by learning and mimicking human creative processes.

Foster cross-industry collaboration by providing more innovative ideas

Digital humans can provide innovative solutions and ideas across different industries by leveraging their knowledge and skills, for example, introducing successful practices and strategies from one field into another.

Understand and simulate human emotions to build deeper social relationships with human being

Digital humans can analyze human speech, tone, body language, and facial expressions to understand emotional states and respond appropriately. This includes engaging in deep conversations, listening, offering understanding and encouragement, and providing emotional support to help alleviate stress and anxiety.

Yuanjing Office Large Model Smart Assistant

The Yuanjing Office Large Model Smart Assistant aims to provide governments, enterprises, and individuals with a relatively standardized large model product suitable for various scenarios. It aims to enhance office efficiency, optimize workflow, and create a platform for selecting, modifying, and applying models, along with a family of standardized general application products, while offering personalized solutions.

Ding Ding
Deputy Director of China Unicom AI Innovation Center
Solution partners

SOLUTION PARTNERS

The Yuanjing Office Large Model Smart Assistant is a smart office and life assistance platform based on China Unicom's in-house Yuanjing Large Model. It aims to meet the modern demands of individuals and enterprises for efficient and intelligent office and life support. The platform provides automated and intelligent tools for the entire office process, to be specific, collecting, sorting, handling, obtaining, learning, applying, and demonstrating various modalities of structured and unstructured data, including documents, speech, images, and videos. Features include corporate encyclopedia, creative writing, text interpretation, work notes, listening and recitation, digital business card, AI image interpretation, audio transcription, voice cloning, and more. It forms a personalized knowledge base and intelligent assistant application that encompasses all aspects of corporate decision-making, management, and execution, as well as the daily office lives of individual users. This puts an end to the limitations of knowledge collection, organization, accumulation, application, and inheritance in traditional office settings, while reducing cost and improving efficiency in government and enterprise management and fueling new growth. It allows individual users to "play" and "utilize" in both office and entertainment contexts, assisting them in transitioning from manual work to smart assistance, and ultimately to fully smart office processes, significantly enhancing personal effectiveness.

CHALLENGES

The daily operations of telecommunication companies usually face such challenges as poor information communication and cumbersome approval processes, impeding operational efficiency and undermining their capabilities to innovate and compete in the market. To improve operational efficiency and business collaboration, the Yuanjing Office Large Model was developed to reshape office processes through intelligent means. This model leverages three major intelligent applications—Yuanjing Vision, Yuanjing Hearing, and Yuanjing Language—to finish the last mile of user office scenarios.

Difficult information acquisition

Individuals frequently struggle to sift through vast amounts of information, making it difficult to quickly filter and find what they need.

Inefficient document reading

Manual document interpretation and processing are time-consuming and exhausting.

Difficult image interpretation

Managing and interpreting image content require specialized skills and tools, posing difficulties for the average users.

Challenging audio processing

Transcribing and managing audio content from meetings or lectures require significant manual effort, resulting in inefficiency.

Creative block in writing

Individuals may experience creative block and struggle with self-expression during writing and content creation.

Challenges in daily work recording

Individuals often forget or overlook details when documenting tasks, leading to lost or disorganized information.

The Yuanjing Office Large Model Smart Assistant effectively addresses these challenges, improving knowledge management and utilization while boosting overall efficiency.

SOLUTION AND VALUE

01 Overall structure/solution

The Yuanjing Office Large Model Smart Assistant is a smart office support platform based on China Unicom's in-house Yuanjing Large Model. Designed for government, enterprise, and individual office and life scenarios, it utilizes an AI paradigm that allows for model selection, modification, and usage in a low-code format. This platform facilitates the creation of personalized knowledge bases and smart assistant AI agents that encompass all aspects of government and enterprise decision-making, management, and execution, as well as the office activities and daily life of individual users. As models, applications, and solutions evolve through practical use, the platform is developing a product system of "platform + standardized product family + rapid solutions", as shown in Figure 1.

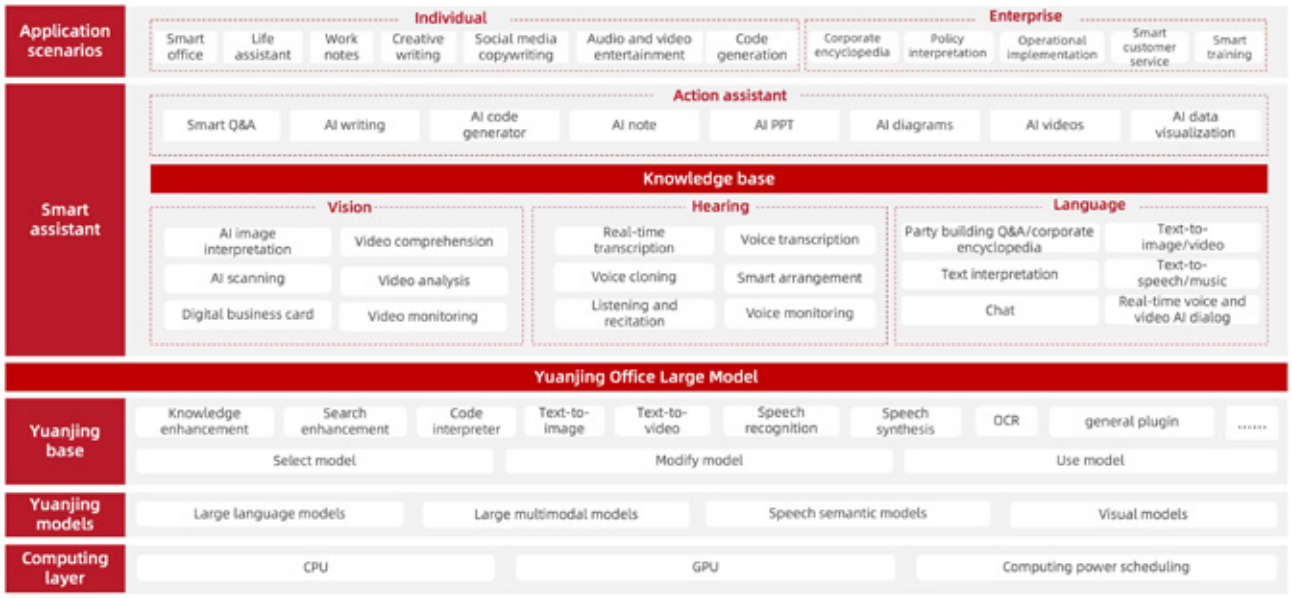


Figure 1: Architecture of the Yuanjing Office Large Model

02 Application scenarios

The Yuanjing Office Large Model Smart Assistant empowers office intelligence through three key applications: Yuanjing Vision, Yuanjing Hearing, and Yuanjing Language. These technologies have been implemented in various internal companies, including Unicom Digital Tech.

Yuanjing Vision

As one of the core highlights of the Yuanjing Office Large Model, it endows machines with insight capabilities that transcend traditional visual limits. By seamlessly integrating visual information with data insights, it ushers in a new era of smart analysis and efficient utilization of visual information. Moving beyond simple recognition and extraction; it focuses on deep understanding and creative application, offering unprecedented convenience and insights to individuals, enterprises, and various industries. Key features include AI image interpretation, AI scanning, and digital business cards, as shown in Figure 2.

Figure 2: Yuanjing Vision

AI image interpretation For image content, Yuanjing Vision provides powerful image analysis and interpretation. Whether it's recognizing objects or scenes in images, or analyzing emotional tones and compositional features, it can quickly deliver accurate results. This feature has widespread applications in advertising analysis, market research, and security monitoring, helping businesses quickly extract valuable information from vast amounts of image data.

AI scanning It supports high-precision scanning and recognition of various certificates, documents, and files. Yuanjing Vision can swiftly capture and extract key information such as text, numbers, and QR codes in files like ID cards, passports, invoices, and contracts and then convert them into editable electronic formats. This reduces manual input errors and improves work efficiency.

Digital business card Yuanjing features easy generation of personalized electronic business cards, covering comprehensive information in a professional and convenient manner. It offers various templates to create a unique image and supports dynamic updates to ensure information accuracy. Users can easily save locally and share without hassle. Additionally, it offers voice cloning for introductions, enhancing personal interactions and making business communication more engaging and effective.

Yuanjing Hearing

This integrates advanced audio transcription technology to quickly and accurately convert speech to text, enhancing work efficiency and optimizing information management. It introduces voice cloning technology for personalized voices, improving user experience. The application also supports the transcription of various audio files and automatically processes uploads, streamlining workflows. Yuanjing Hearing makes audio transcription more efficient and flexible, meeting diverse office needs. Its capabilities include real-time transcription, smart meeting, and listening and recitation, as shown in Figure 3.

Figure 3: Yuanjing Hearing

Real-time transcription In meetings or daily office settings, Yuanjing Hearing supports real-time audio recording and transcription into text simultaneously. Users can focus on discussions or work exchanges without the distraction of note-taking. It also allows users to link recordings with transcribed text, easily tagging them as work notes for later reference and organization.

Smart meeting Yuanjing Hearing provides a one-stop solution for meeting scenarios. It not only transcribes meeting audio in real-time but also automatically generates meeting minutes and intelligently extracts key information based on discussion content. At the same time, the system supports speech recognition and differentiation of meeting participants, creating individual speaking records for each speaker and facilitating subsequent analysis and review.

Listening and recitation For lengthy articles or complex information, Yuanjing Hearing can quickly summarize core points and present them to users through voice output. This feature helps users quickly grasp key information, significantly enhancing the convenience and efficiency of information retrieval. Users can receive information through listening during busy work intervals, effectively alleviating visual fatigue.

Yuanjing Language

This application provides precise and efficient smart services by integrating party-building guidance, corporate insights, creative inspiration, work assistance, and casual chats, all powered by deep learning and natural language processing technologies. Its seven capabilities include party building Q&A, corporate encyclopedia, text interpretation, creative writing, work notes, text-to-image, and chat, as shown in Figure 4.

Figure 4: Yuanjing Language

Party building Q&A With a rich database of party-building knowledge, Yuanjing Language can respond quickly to policy interpretations, theoretical studies, and practical guidance, assisting in deepening party-building efforts and advancing organizational development.

Corporate encyclopedia Yuanjing accesses comprehensive information across various industries, from market dynamics to industry trends, and from corporate management to technological innovation, providing authoritative and accurate answers to help businesses gain a competitive advantage in a fierce market.

Text interpretation With powerful natural language processing capabilities, Yuanjing can accurately understand text needs. Be it complex academic literature or everyday communication, Yuanjing can quickly distill key points and conduct in-depth analysis, helping users grasp the essence of information.

Creative writing Inspire creativity and support content creation. Whether for article writing, advertising, or brand storytelling, Yuanjing generates creative and personalized text based on user needs.

Work notes A convenient note-taking feature allows users to record work ideas, meeting highlights, or daily insights at any time. By intelligently organizing note content, it forms a clear work outline to improve work efficiency.

Text-to-image Transform text into vivid images. From design sketches to mind maps and creative posters, Yuanjing can generate images that meet needs based on descriptive text, visualizing creativity.

Chat In addition to formal work scenarios, Yuanjing Language offers a smart conversational experience. Users can engage in natural, smooth chats with the system, sharing life moments and discussing hot topics, enjoying a delightful communication experience brought by intelligent technology.

03 Key innovations

The first key innovation is the multi-task speech generation model architecture, as shown in Figure 5. This model, trained on over 100,000 hours of high-quality speech data, excels in zero-shot voice cloning of human voices, human-like naturalness, and multi-language and multi-dialect generation capabilities. Yuanjing Hearing is built on this model architecture.

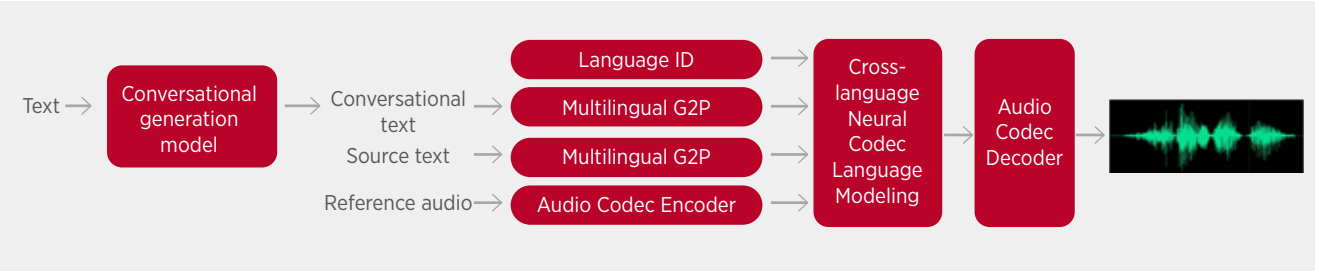


Figure 5: Architecture of the Multi-Task Speech Generation Model

The second key innovation is the Yuanjing large multimodal model (LMM) with 204 billion parameters, as shown in Figure 6. It features an innovative composite visual coding module by combining mainstream ViT visual encoders with CNN architectures that have been proven effective in various traditional visual tasks. This allows for visual feature extraction from different dimensions, achieving fine-grained perception and understanding of image content. By accurately aligning the visual and language modules, the model supports complex logical reasoning and visual mathematical calculations, effectively mitigating the hallucination issues in large models.

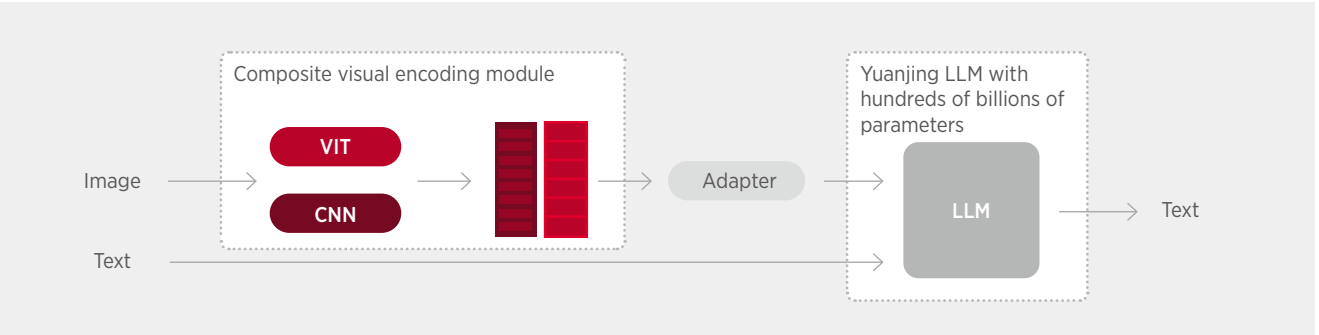


Figure 6: Yuanjing LMM with 204 Billion Parameters

The third key innovation involves the core components of the AI agent added to the Yuanjing Office Large Model Smart Assistant, as shown in Figure 7. These enhancements enable the large model to better discern user intent, autonomously invoke multiple tools, and offer customizable tool extensions.

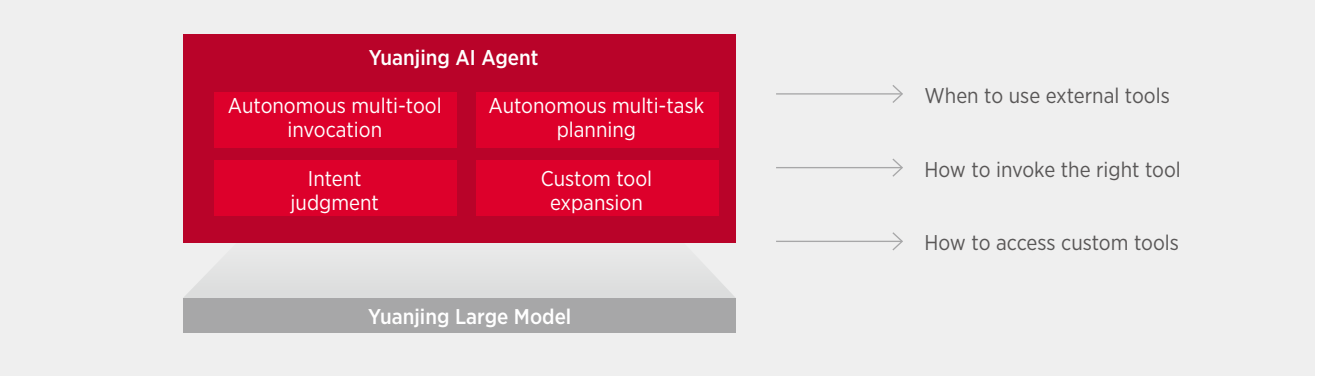


Figure 7: Core Components of the AI Agent

04 Business models

	Individual-focused products	Enterprise-focused products
Target users	Customers who require cross-platform and software capabilities for the collection, storage, and utilization of smart multimodal information and a personalized smart life assistant based on personal knowledge bases for work, social, and entertainment purposes.	State-owned and central enterprises seeking smart office solutions and secure party-government operations, especially those with strong ties to China Unicom.
Promotion mode	Leverage China Unicom's mobile packages and channels to conduct online and offline marketing activities; enhance advertising placements and social media presence, featuring keywords like "China Unicom's Yuanjing Smart Assistant" and "Yuanjing Smart Assistant - Your Efficiency and Entertainment Companion".	Utilize China Unicom's resources and its in-house large models that highlight independent and controllable management to provide office services to enterprises and institutions with high data/supplier security demands. These services will be gradually introduced to relevant government bodies and enterprises.
Product forms	Cloud services, PC and mobile apps, mini programs, and applications on large platforms.	Cloud services, private services, integrated hardware-software products, office large model API services, as well as mini programs and applications on large platforms.
Profit mode	Basic capabilities are free for individual users, with a token limit. Additional usage incurs a pay-as-you-go fee. Additional value-added services include cross-platform and software capabilities for the collection, storage, and utilization of smart multimodal information, personalized knowledge bases, and smart assistants. Payment options include subscriptions based on membership, annual or monthly fees, and usage volume.	For enterprise users, provide basic services through cloud services, private services, and integrated hardware-software solutions, along with customized and value-added services.

05 Core values

The launch of the Yuanjing Office Large Model Smart Assistant represents a significant milestone in the smart and digital transformation of the telecommunications industry. By effectively tackling industry challenges and bottlenecks, it boosts operational efficiency and customer experience, paving the way for continuous innovation and growth in the sector. AI, as the core driving force behind this transformation, will continue to play a pivotal role in future business development.

REFLECTION

The introduction and application of the Yuanjing Office Large Model Smart Assistant have yielded positive outcomes in technology and product development and industry impact.

Large user base The current user base exceeds 100,000, mainly consisting of enterprise users. Looking ahead, services will be expanded to both enterprise and individual customers, with an expected user base of over 400,000, providing a solid market foundation for ongoing development and promotion.

Rich coverage of scenarios With over 20 office components, it covers smart Q&A, smart writing, voice explanations, and various office scenarios, meeting diverse business needs and enhancing product applicability and user engagement through extensive coverage.

Significant efficiency improvement Smart methods greatly enhance office efficiency, achieving a qualitative leap in productivity and providing enterprises with a competitive edge. Data-driven decision-making reduces the time and efforts in the decision-making process.

Enhanced customer experience Smart customer service and personalized offerings improve customer satisfaction and loyalty, fostering positive interactions between enterprises and users. The smart transformation of customer services enables businesses to better understand customer needs and provide precise services.

Driving business innovation Data-driven decision-making mechanisms allow enterprises to respond promptly to market changes and innovate products and services. By analyzing market trends, businesses can maintain a competitive edge in a fierce market landscape.

Looking forward, the Yuanjing Office Large Model Smart Assistant will deepen its collaboration across diverse industries, exploring new application scenarios to meet the ever-changing market demands. We are committed to providing high-quality services by continually optimizing product features and enhancing user experiences. Through technological innovation and practical applications, we anticipate ushering in a new chapter in the telecommunications industry.

Chunghwa Telecom Applications of AI-powered Smart Customer Service

With a commitment to being a "leader in smart life and enabler in the digital economy", Chunghwa Telecom is dedicated to becoming an international benchmark enterprise, a digital ecosystem co-creator, and a leading technology group. Emphasizing the core values of "integrity, trust, innovation, and accountability", Chunghwa Telecom has developed omnichannel smart customer service solutions. By leveraging AI, it has implemented a human-robot collaboration mode that seamlessly integrates digital employee robots with customer service agents. This approach enhances the service speed, accuracy, and user-friendliness, providing a superior AI-powered Smart Customer Service experience that leads users toward a sustainable digital life.

Chang Pen-Yuang
General Manager of Chunghwa Telecom Enterprise Business Group

SOLUTION PARTNERS

Chunghwa Telecom provides an omnichannel AI-powered smart customer service solution to support customers using all kinds of communication media, online and offline. Customers can connect to the all-around AI-powered customer service through various channels, such as telephones, mobile phones, apps, websites, smart counters, interactive self-service machines, and physical robots.

Chunghwa Telecom's AI-powered Smart Customer Service offers smart voice customer service, smart text customer service, and customer voice analysis. It aims to deliver faster and more convenient services, enhancing the overall customer experience. By utilizing diverse communication channels, Chunghwa Telecom gains a deeper understanding of customer needs, enabling them to offer more targeted solutions.

Chunghwa Telecom's AI-powered smart customer service solution boosts the service efficiency of the customer service center, reduces the wait time, and improves the overall service quality. Leveraging cutting-edge technologies and extensive service experience and data, Chunghwa Telecom is committed to continuously refining and innovating its smart customer service products to provide an exceptional service experience for its customers and meet their ever-evolving needs.

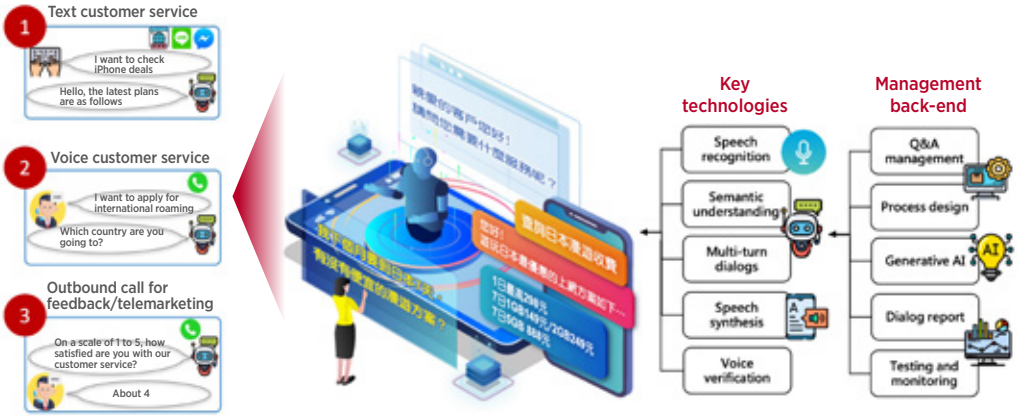


Figure 1: Overview of the AI-powered Smart Customer Service

- Universal knowledge base that serves calls from multiple channels, such as webpages, social media, and phones
- Over 15 Q&A dialog modes
- Comprehensive AI-powered knowledge management back-end with more than 20 features
- Generative AI-based knowledge base, reducing the burden of customer service agents by 70%

CHALLENGES

Dilemma faced by traditional voice customer service

Modern consumers have high expectations for customer service, seeking quick responses and efficient assistance. Long wait times often lead to dissatisfaction. Traditional customer service faces several challenges that hinder agents from promptly meeting various customer needs:

- Lengthy voice notes: Conventional voice customer services based on the interactive voice response (IVR) system play lengthy voice notes when a call is connected. The long wait time easily leads customers to hang up.
- Inconvenient operation: Conventional IVR systems require keypad inputs, which are inconvenient for mobile users who must switch between listening and input modes during a call.
- Poor call routing: Poorly designed keypad provides excessive or insufficient options, causing customer frustration and leading them to request direct transfer to the switchboard. This inefficiency increases the workload on customer service agents.

These issues contribute to extended wait times and customer dissatisfaction. Frustrated customers can cause emotional stress for agents, leading to decreased service quality and higher turnover rates. Moreover, a significant portion of customer inquiries involves repetitive and routine questions. When skilled human resources are tied up handling these inquiries, they miss opportunities to innovate and improve service processes.

The necessity of a text customer service system

- Limitations of FAQs: FAQs on websites often fall short in addressing all customer inquiries. Text customer service agents are required.
- Long wait time of voice customer service: The voice customer service is limited to one-on-one interactions. When call volumes spike, customers often face long wait time.
- Repetitive customer service tasks: Certain events can trigger a flood of similar questions, overwhelming agents with repetitive tasks.

Difficulties in assessing the customer service quality and standards

- Difficult to understand customer needs: Customer service centers must delve into customer concerns and intents and dig deeper into their potential needs. However, the vast amount of data and call recordings in medium to large centers makes it challenging to effectively analyze and respond to customer needs.
- Lack of compliance with service specifications: Compliance with the standardized service specifications can significantly enhance the overall service quality of agents.
- Low efficiency in conventional methods: Conventional methods such as manually obtaining and listening to voice recordings or writing case descriptions are inefficient, time-consuming, and labor-intensive. These methods often lead to incomplete information, making it difficult to accurately assess service quality.

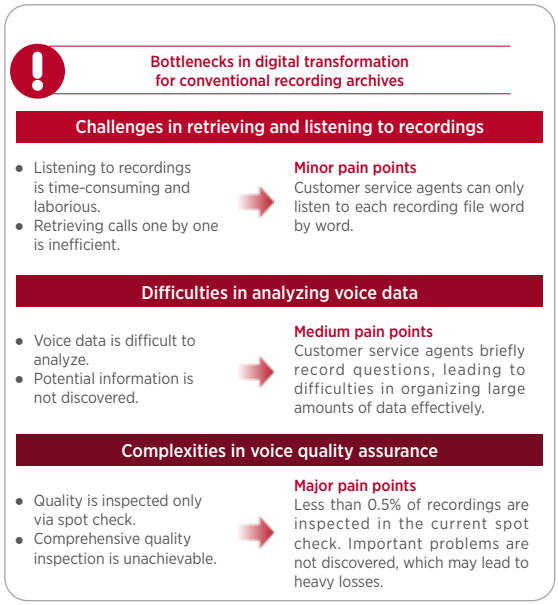


Figure 2: Difficulties in subsequent services and analysis of customer service

SOLUTION AND VALUE

01 Overall structure/solution

Chunghwa Telecom's AI-powered Smart Customer Service offers comprehensive omnichannel support for customer service applications. By leveraging the company's strengths in communications, the technologies developed by Chunghwa Telecom Laboratories, and the knowledge base of enterprise customer services, the service supports tailored configurations and data importing across various on-site environments. In addition, it seamlessly integrates with enterprise information systems, such as ERP, CRM, and HR systems, and connects to multiple channels such as voice calls, corporate webpages, brand apps, and social media.

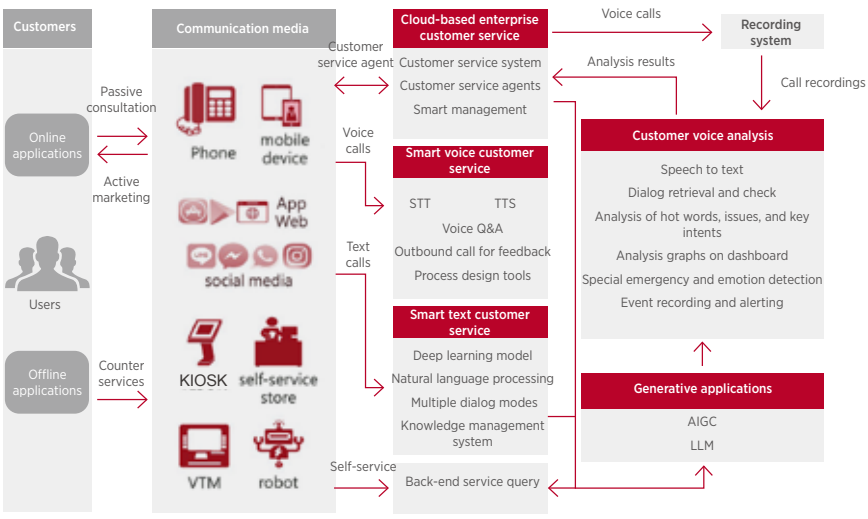


Figure 3: Architecture of Chunghwa Telecom's AI-powered Smart Customer Service Solution

Chunghwa Telecom's AI-powered smart customer service solution includes cloud-based enterprise customer service, smart voice customer service, smart text customer service, and customer voice analysis.

Cloud-based enterprise customer service

Cloud-based enterprise customer service supports multi-tenant applications. It provides the voice customer service platform, the computer telephony integration (CTI) service, the online recording system, and smart customer service module integration. It is equipped with the most advanced carrier-grade customer service platform architecture and services and a complete security mechanism that helps reduce security costs. It can assist customers in building a multi-site customer service center, facilitating centralized management and call dispatching.

Smart voice customer service

The smart voice customer service adopts the AI-powered voice recognition and semantic understanding to deliver near-human quality speech using a hybrid speech synthesis method. Through human-machine collaboration, it provides sentence feedback and error cause analysis, enabling continuous optimization of the voice customer service. Compared with conventional IVR, the smart voice customer service improves the efficiency of customer service agents and shortens their response time.

The smart voice customer service can automatically make calls to request customer feedback and conduct voice interaction with customers. This improves the survey response rate and the number of customers reached, maximizing the efficiency of customer service agents and reducing the labor costs associated with manual feedback calls. The smart voice customer service also performs post-call analysis on the survey results to refine its features.

Smart text customer service

The smart text customer service provides multiple dialog modes and is capable of semantic understanding. The service supports omnichannel integration, allowing customers to ask and answer questions through various channels. They can update and optimize the knowledge base of the service in real time via the back-end management system. The service provides more than 15 dialog interaction modes, including the default welcome menu and customized welcome menus. The Q&A modes include basic Q&A, personalized query, chat Q&A, typo detection, sensitive word detection, guidance Q&A, graphic Q&A, multi-turn dialogs, recommended Q&A, fuzzy Q&A, smart word prompts, and repeated sentence Q&A. In special Q&A situations, customer service agents can take over to provide services, allowing customers to access information easily and promptly.

The smart text customer service features semantic understanding capabilities. AI technologies adopted include sentence normalization, sentence segmentation, intent recognition, dialog management, knowledge base retrieval, and sentence generation. The back-end management system enables continuous optimization through feedback, allowing for rapid learning and increased service efficiency.

The back-end management system provides a feedback-based learning mechanism, allowing users to edit the knowledge base, manage vocabulary and intents, edit intents, and set and test parameters. It also allows for observing dialogs and importing sentences. A complete data circulation mechanism is in place to ensure the real-time updating of the knowledge base for the smart text customer service, enabling effective handling of emergencies.

Customer voice analysis

The customer voice analysis platform (DeepVoice) is jointly developed by the Development Department and Enterprise Business Group of Chunghwa Telecom. It is the first solution in Taiwan that deploys Chunghwa Telecom's proprietary voice recognition and AI-powered text analysis technologies in all customer service agents at telephone customer service centers of the government and public institutions.

After a call with the customer service agents ends, the service automatically converts the recording file into a translated text file. In addition, semantic analysis and processing are automatically performed on each text file, including intent sentence recognition,

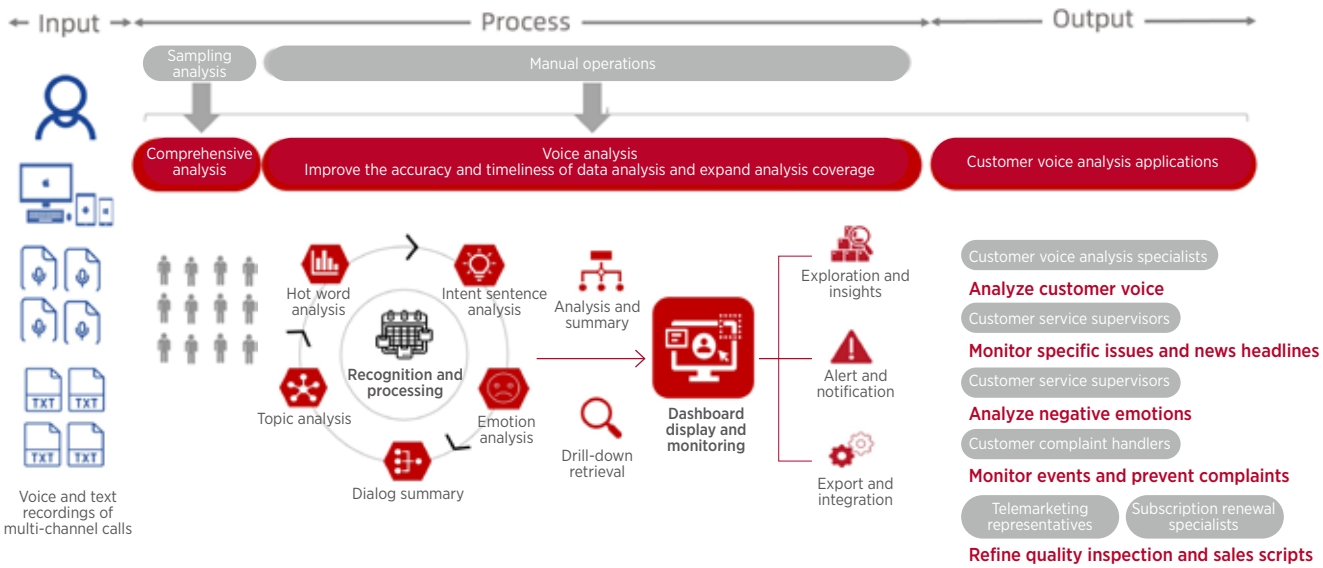


Figure 4: Customer voice analysis process of Chunghwa Telecom

emotion recognition, and conversation summary. The analysis results are stored in an efficient and stable system and presented on a streamlined visual UI, meeting the requirements that the data can be searched, retrieved, and analyzed.

The DeepVoice platform analyzes data by using a deep learning model, delivering 96% accuracy in intent sentence recognition and 90% accuracy in emotion recognition. It also provides a user-friendly UI, allowing users to retrieve and filter data and present the analysis results in visual graphs. At the application level, the platform takes into account multiple application scenarios of customer service centers and assists users in primary work tasks, including issue analysis, routine observation, case collection, and automatic event notification. The platform reduces 90% of call analysis time and helps enterprises capture the voices of customers and improve customer service satisfaction.

02 Application scenarios

The application scenarios of Chunghwa Telecom's AI-powered smart customer service can be described in chronological order: before, during, and after human service.

Before human service: Automatically helps customer service agents receive customers and answer questions, reducing their workload

The service supports customer calls from multiple channels, such as voice channels (local fixed-line calls and mobile calls) and online channels (apps, web pages, and SMS messages). When a call is connected, the service automatically takes over the frontline customer service agents by using IVR-based voice navigation and an online customer service robot. It can greet customers, identify customer needs, directly answer their questions by using text or voice, and lead them to the target feature, helping them complete the task on their own. If human customer service is required, the service can transfer the customer to the customer service agent with the best-matched skills. If an outbound call is required, the outbound dialing robot can assist the customer service agent in making an external call, increasing their productivity. This design enables customers to receive immediate assistance when they contact government or enterprise customer service. Additionally, it prioritizes handling a large number of common and simple requests, reducing the workload of customer service agents.

During human service: Assists customer service agents in quickly and accurately resolving customer issues

When a customer service agent starts a conversation with a customer, the online response assistant can generate a list of recommended responses in real time, allowing the operators to click one to answer the customer. This reduces the workload of manual input and minimize the customer's wait time. The online response assistant autonomously trains the large language model (LLM) by using previous conversations. By leveraging AI, the assistant can understand the context and automatically generate recommended answers, without requiring human efforts in data annotation. This greatly improves the efficiency of operators.

Additionally, the AI-powered smart customer service provides the KM copilot feature that integrates the generative AI model and prompt engineering technology to assist AI trainers in quickly building or optimizing the robot's knowledge base. It includes a wealth of modes, such as knowledge classification, Q&A generation, paraphrasing, reference answer generation, answer summary, tone changing, synonym generation, and key point summary. These modes further shorten the manual configuration time by 70%, improve the work efficiency of AI trainers, and accelerate the generation of up-to-date business information to customers. With the online response assistant and KM Copilot, customer service agents can accurately and quickly respond to customer questions, boosting both service quality and efficiency.

After human service: Helps the government and enterprises improve public satisfaction through AI-powered voice analysis

The independently developed customer voice analysis platform DeepVoice can analyze customer voices to evaluate the performance of customer service agents, prevent customer complaints, retain at-risk customers, explore business opportunities, and improve customer experience and satisfaction. By leveraging speech-to-text (STT) and generative AI technology, the platform can automatically summarize key points, extract intents, analyze customer emotions, divide topics into groups, and predict trends based on conversations with customer service agents. It helps enterprises understand the impact of new services or new policies on customers. In addition, to ensure data confidentiality and integrity, the platform adopts security measures such as web page credential encryption, graphic verification codes, and authentication.

The DeepVoice platform takes the lead in adopting Chunghwa Telecom's customer service center. It is the first solution in Taiwan that deploys voice recognition and generative AI in all customer service agents at the telephone customer service centers of the government and public institutions. The platform won the 2024 Taiwan Excellence Award, and the 2024 FCA Innovative Business Award sponsored by Business Next. These awards are a recognition of the prospect of this product. This product is applicable to government institutions, state-owned enterprises, medium and large customer service centers, and organizations that need to analyze the behavior and intents of a vast number of customers.

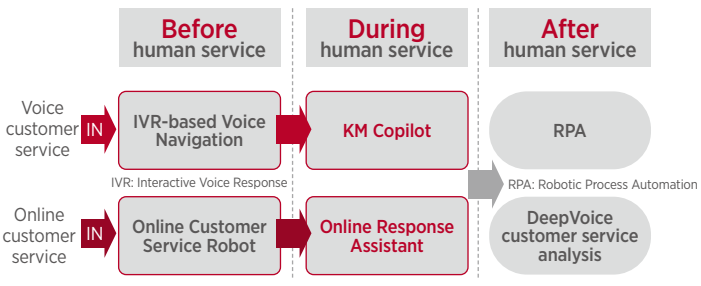


Figure 5: Application scenarios of Chunghwa Telecom's AI-powered customer service

03 Key innovations

The AI-powered customer service solution brings the following key innovations to enterprise customers:

Generative AI application	Generative AI is introduced to enhance dialog summary, emotion analysis, issue analysis, and other features. LLM-based generative AI technology improves the intelligence of the customer service system, enabling it to more accurately understand and respond to customer needs. In addition, more advanced speech recognition and synthesis technologies are developed and deployed to deliver a more natural and seamless customer experience.
Hybrid of cloud and on-premise inferences	Customers can flexibly choose the cloud or on-premise LLM inference service for generative AI-based dialog analysis.
Automatic data labeling module	Users can classify and label the response sentences first, such as sentences used for greetings or confirming the name and gender of the customer. After the labels are imported into the model for training, the system automatically labels dialogs. This feature can be used for quality inspection, telemarketing, and customer voice analysis.
AI-powered automated voice quality inspection module	Currently, the voice quality inspection service of most customers is still semi-automatic. Chunghwa Telecom is planning to provide AI-powered automated voice quality inspection and automatic scoring features.
Improved system integration	An IP private branch exchange (IP-PBX) system and a computer telephony integration (CTI)-based customer service system can be integrated with the AI-powered customer service to ensure seamless connection between the systems. This improves the efficiency and reliability of the overall solution. Flexible APIs will enable enterprise customers to perform customized development and integration as needed.

04 Business models

The AI-powered customer service solution is provided in the form of project deployment. The following models are provided for customers to select:

One-time purchase of on-premise deployment

- The one-time software licensing fee is charged. Customers can use the solution permanently, but may need to pay additional maintenance and upgrade fees.
- Cloud service leasing
- Monthly or annual fee: A fixed fee is charged based on the number of users or the feature tier.
- Tiered subscription: Different features and services are provided based on tiers, such as the basic edition, advanced edition, or enterprise edition as needed.

Cloud-on-premise hybrid model

- Cloud and on-premise deployment: Customized service development is provided for on-premise deployment, which can be combined with a cloud subscription for value-added features.

Customers can select a model based on their business model, the specific situation of the project, and customer needs. The flexible project deployment models help meet the special needs of different customers. In addition, a partner ecosystem can be built by partnering with professional AI technology companies and solution providers to jointly develop and promote innovative AI-powered customer service applications. The open technology ecosystem will engage more third-party developers and partners to jointly enrich and optimize the customer service solution.

05 Core values

Leveraging generative AI to boost the service efficiency of customer service agents

The AI-powered KM Copilot feature is introduced to understand and organize data and generate content that is different from but has similar characteristics with the input data. This feature can be used to assist in building the knowledge base and quickly creating an intelligent customer service robot. It helps AI trainers reduce the editing time of the robot's knowledge base by 70%. Combined with prompt engineering, the accuracy of generated content can reach more than 90%.

Improving the automation rate by integrating the AI-powered customer service robot and applications

Chunghwa Telecom's AI-powered customer service robot Qbao has answered a growing number of calls, which is more than 10 times the number answered three years ago. The IVR-based voice verification feature of the intelligent voice customer service deals with 7.13 million calls per year, the outbound dialing robot serves 4.21 million calls per year, and the IVR-based self-service processes 1.77 million calls per year. In 2023, the intelligent voice service has served 16.2 million calls. The AI-poweredautomation rate of Chunghwa Telecom's customer service center is greatly improved.

Enabling AI-powered innovation and enterprise transformation with Chunghwa Telecom's proprietary speech recognition engine

The massive language data database receives more than 1 million cases per day. Extensive experiences of 2,000 customer service agents are continuously accumulated. Based on them, Chunghwa Telecom can customize the AI-powered customer service system for enterprises in different fields. It can deliver 90% to 95% accuracy in speech recognition and provide a hybrid recognition engine to meet requirements for multilingual situations, including Mandarin, English, and Taiwanese. It helps enterprises automate customer service through AI and enable them to achieve AI-powered innovation and transformation.

Reducing the data collection time by 33.3% with omnichannel speech-to-text

The omnichannel STT feature automatically generates conversation summaries based on the transcribed text. Customer service agents can quickly grasp the key points of calls without listening to the recordings word for word. This feature significantly improves data collection efficiency. In addition, it greatly shortens the time required for data processing and improves the accuracy of problem solving.

Improving the data value of customer service recording files

The customer voice analysis platform automatically transcribes conversation recordings and performs full-text searches. It transforms recording files into searchable data, greatly improving the value of customer service recording files. Users can perform quick searches and filter the data by keyword, emotion, and call time to find the required information. When data usability is improved, the value of customer service recording files is also maximized.

Allowing customer service agents to focus on meeting customer needs with an automated customer service process

The service automatically generates conversation summaries and performs text analysis. Customer service agents no longer need to manually record the call process, and can focus more on satisfying customer needs. This feature reduces the workload of customer service agents and improves their work efficiency and concentration. Automated emotion analysis and question classification enable customer service agents to understand customer situations more quickly and provide targeted responses, enhancing both service quality and customer satisfaction.

Increasing customer engagement with the automated feedback tracking feature of the customer voice analysis platform

The customer voice analysis platform provides automated emotion analysis and question classification to help enterprises track and receive customer feedback. This feature enables enterprises to understand customer needs and make adjustments and improvements accordingly. The real-time feedback tracking mechanism improves customer satisfaction and enhances customer trust and loyalty, boosting customer engagement and stabilizing customer relationships.

REFLECTION

The era of generative AI has arrived. Chunghwa Telecom can use AI to process text, images, videos, audio, programs, synthetic data, workflows, and object models. Chunghwa Telecom is committed to establishing a customer service application ecosystem to integrate upstream and downstream partners. It serves institutions and enterprises whose customer service centers need AI transformation, to speed up the intelligence development of customer services in Taiwan.

Looking ahead, from 2025 to 2030, Chunghwa Telecom will continuously evolve AI technology and expand the applications of generative AI to provide 24/7 uninterrupted intelligent in-depth services. Chunghwa Telecom aims to resolve the current issues, such as the shortage of customer service agents, difficulties in onboarding training, and the lack of 24/7 human customer services due to labor regulations. In addition, it will leverage automation and AI to explore customer needs and analyze customer voices, improving service quality and user experience of enterprise customers. Chunghwa Telecom is committed to providing heart-touching customer services and better customer service solutions for enterprises through continuous innovation.

1 Anchoring in the industry

- Shorten VOC analysis time through human-machine interaction
- Improve customer satisfaction and minimize complaints

2 Expanding the market

- Forge alliances with the Call Center ecosystem, both upstream and downstream
- Target institutions/enterprises that require AI-powered transformation
- Advance intelligent development of customer services in Taiwan

3 Further evolution of AI

- Continuously enhance trustworthy generative technologies
- Expand the diverse intelligent application scenarios for AIGC
- Develop training materials for customer service agents
- Analyze and establish strategic business insights for enterprises

Future plans

2022

- Pilot POC

2023

- Initiate implementation plan

Drive continuous innovation to deliver superior customer service solutions

2024

- Event-driven VOC analysis
- Automated customer call analysis: Uncover customer needs & predict and address issues
- Implement AIGC to boost analysis efficiency
- Analyze successful sales pitches and optimize response scripts: Empower industry growth

2025-2030

- Prevent customer complaints
- Manage subscription retention
- Manage customer churn
- Convert opportunities into sales
- Manage brand reputation
- Customer persona analysis: Committed to delivering an emotionally engaging intelligent customer service experience

Figure 6: Future plans

Sichuan Educational Examination Authority Launched AI Proctoring Project

> The National College Entrance Examination, or Gaokao has a direct bearing on the future of many students and their families. We must rigorously enforce exam discipline and crack down on cheating. For this purpose, we should make full use of AI and other new technologies and measures to timely identify violations, raise awareness, and enhance deterrence effects, thereby ensuring fairness and justice in the exam.
>
> Huang Qiang
> deputy secretary of the Communist Party of China Sichuan Provincial Committee and governor of Sichuan province

SOLUTION PARTNERS

Gaokao, a major national examination and an important mechanism for screening talents for the country, has a direct bearing on the future of many students and their families. In response to the Ministry of Education's initiative to implement real-time inspection of Gaokao, Sichuan Province has utilized intelligent technologies such as machine vision, target tracking, and multimodal for exam paper confidentiality, standardized examination management, and violation identification. This is how the province ensures exam security and fairness. Sichuan has been one of the first provinces in the country to play back exam videos with AI, providing technical reference standards for the Ministry of Education to push forward intelligent proctoring. By 2024, Sichuan has realized real-time proctoring in Gaokao through "cloud deployment", creating a "six-in-one" exam security system (Figure 1). This includes: 1) working with China Mobile to establish dedicated lines and deploy cloud computing resources for 357 exam sites and 196 rooms to store exam papers and confidentiality materials (Figure 2), enabling unified analysis and cost optimization via AI models at the provincial level; 2) developing the "Guidelines on AI Recognition and Analysis of Abnormal Behaviors in Sichuan Educational Examinations" which clarifies 44 criteria for intelligent analysis and improving technical schemes; 3) enhancing emergency response and handling mechanisms, and organizing multiple tests and drills to ensure the system operates safely and stably.

Exams in all ages are closely connected to the glory and dreams of students. China Mobile works to safeguard the original intention of education with the power of technology!

Figure 1: "Six-in-One" Exam Security System

Figure 2: AI Proctoring Cloud Resource Usage in Sichuan

CHALLENGES

At present, exams of all types are proctored by teachers manually, with the help of IT means such as video surveillance and signal shielding in standardized exam rooms. However, this has some drawbacks:

Provincial/municipal exams require many proctors

Large-scale recruitment exams at national/provincial/municipal levels require a large number of proctors who are mostly teachers recruited from various schools.

Visual blind spots

Teachers cannot stay aware of the whole room all the time. The current model of one video proctor for every nine rooms not only requires lots of manpower but rests with the physiological limits of proctors. Optimal results are hard to achieve in terms of attention, focus, energy, and efficiency.

Uncontrollable personnel quality

As teachers of all types are recruited as proctors, it is hard to guarantee that each proctor will follow examination procedures, leading to inconsistent proctoring standards.

Poor generalization and inaccurate recognition in traditional small models

Small models based on supervised learning use classification recognition algorithms. Due to the lack of enough parameters and structurally complex networks to capture the complex relationships and patterns in the data, small models tend to memorize training samples. They are overly sensitive to noise and details. When there are major changes in the exam environment or when cheating behaviors are inconsistent with the training samples, there is a significant decrease in the accuracy of cheating detection.

SOLUTION AND VALUE

By combining multiple models like a small model, a large visual model, and a multimodal large language model, the AI proctoring system analyzes exam scenes in real-time to ensure optimal balance in computational cost, recognition accuracy, and coverage, thus greatly reducing manpower, energy, and time input. The system can ensure fair evaluation based on unified standards. By introducing AI-assisted video proctoring rule and embedding integrated devices in exam centers or providing cloud-based proctoring services, the product offers real-time video monitoring with intelligent analysis and alerting to ensure comprehensive surveillance of the examination process (see Figure 3). The product uses technical means to address the difficulties of detecting, tracing, and identifying violations on an integrated system through unified platform management.

Low costs

A unified system is built on the provincial level to avoid high costs of separate construction for each examination site. Local computing resources are concentrated in the cloud through a cloud-based solution, where computing power is leased, reducing costs to 1/10 of previous costs.

Fast recognition

The product can realize high-concurrency processing at millisecond speed. The transition from manual recognition in minutes to machine recognition in seconds has increased the proctoring efficiency by over 30 times.

Unified standards:

The product runs as a program on the computer to ensure that each examinee in the video is analyzed by the same standard.

Full coverage

All scenes in a standardized examination site are covered, examinees are monitored in different zones, and real-time videos are thoroughly analyzed.

Accurate recognition

High-precision algorithms trained on a large scale, with two levels of manual verification, ensure accuracy in recognition.

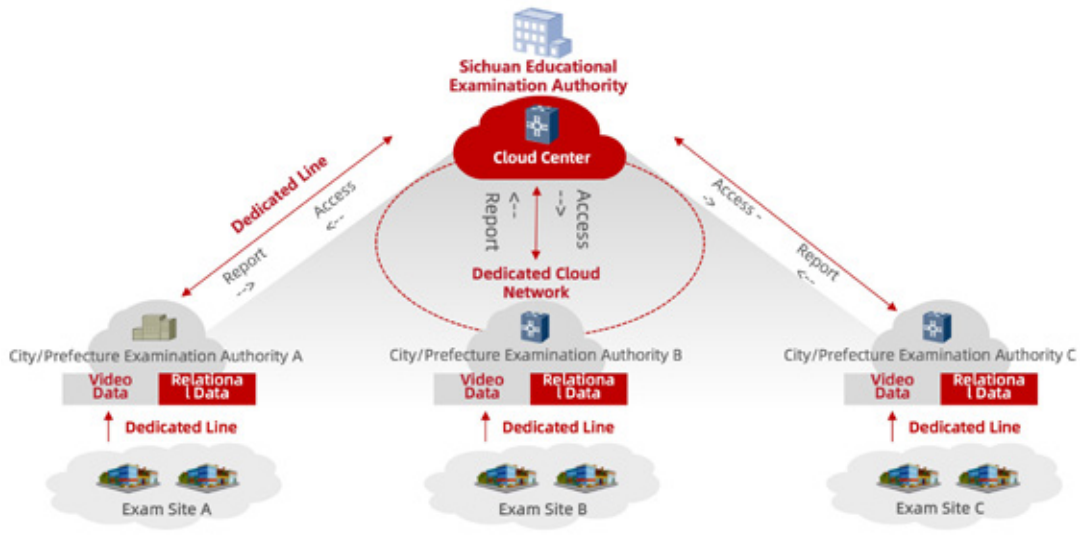


Figure 3: Examination Site Architecture in Real-Time Proctoring Analysis Scenario

01 Overall structure/solution

The sytem uses a cloud-based architecture (see Figure 4) that relies on cloud computing and works with a computing center and a proctoring management and proctoring business system.

The computing center pulls video streams from the proctoring system through a dedicated public cloud network to extract frames, conduct algorithm analysis, identify cheating behaviors, and generate suspected cheating tasks.

The proctoring management and proctoring business system is also deployed on the cloud so that examination institutions and sites in schools can access the system through a dedicated public cloud network for proctoring purposes.

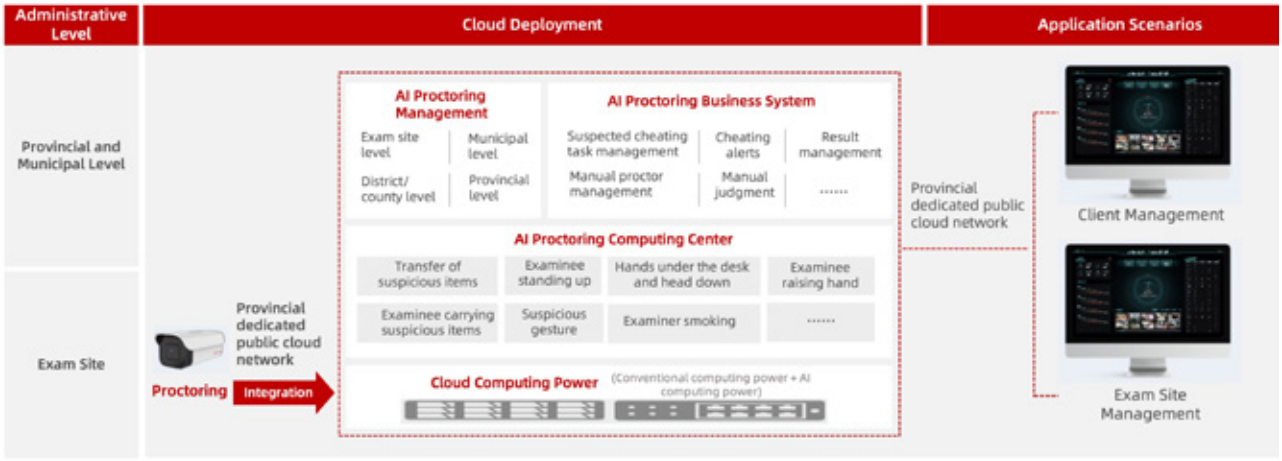


Figure 4: AI Proctoring Cloud-Based Deployment Architecture

02 Application scenarios

As a real-time behavior analysis system, the AI proctoring system is primarily used in national examinations such as the Gaokao and graduate entrance exams. It analyzes and alerts on abnormal behaviors by individual/group, and deviations from examination procedures (see Figure 5), including the analysis of abnormal behaviors of examinees, proctors, and relevant staff.

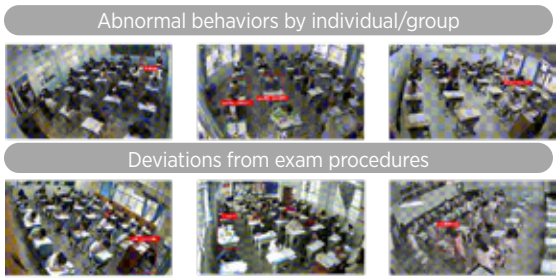


Figure 5: Classification of Abnormal Behaviors in Real-Time Proctoring Scenarios

03 Key innovations

One-stage object detection algorithm

A one-stage object detection algorithm is used to increase the speed by 70% as compared to two-stage object detection algorithms while ensuring accuracy. A robust "network expansion" method is used to enhance the performance of small models while balancing computational complexity and memory usage. Simple and effective strategies are employed to expand large object detectors based on optimal partitioning. The algorithm includes the adaptive anchor box calculation that effectively adjusts anchor box sizes based on the dataset to match targets in images, enhancing training efficiency and significantly reducing convergence time. Adaptive image scaling reduces unnecessary black border/padding during prediction and increases detection speed. Relevant loss function calculation methods contribute to marked improvements in detection rates for overlapping targets.

Efficient multi-object tracking

Traditional algorithms and deep learning methods are combined to extract image features for effective object tracking. Key feature areas include local/global features, templates, histograms, binary patterns, PCA, sparse PCA, SR (sparse representation), discriminative models, and generative models. The primary focus is tracking multiple objects within a scene by calculating their 2D position coordinates in each frame, thereby connecting the positions across consecutive frames to form motion trajectories.

To address issues of body occlusion or severe obstruction in examination settings, the model incorporates constraints on the activity of each component and models various body postures. This ensures stability and accuracy in detection and segmentation tasks, effectively minimizing matching errors during the detection process.

Cross-temporal contrastive learning

By leveraging the differences between pre-exam empty classroom data and during-exam examination data, cross-temporal contrastive learning enables the joint training of pre-exam partitioning algorithms and during-exam behavior analysis algorithms. This approach reduces the interference of desks on candidate recognition, enhances the model's attention to both proctors and examinees, and thereby improves model accuracy.

04 Business models

The cloud-based model disrupts the traditional business model of proctoring systems by transforming from purchasing computing power and hardware to a rental service model. Payment shifts to a comprehensive service for "cloud resources + dedicated lines + platform." A flexible payment model allows examination centers or educational management institutions to pay based on actual conditions such as the number of examination rooms and subjects.

05 Core values

Value 1

Significantly reduce the number of examination staff

A middle school, for example, has 42 exam rooms that require five video proctors to monitor the surveillance videos in all rooms during the exam. By contrast, a real-time intelligent proctoring system deployed at the exam site, along with one video reviewer and one site administrator, is enough to monitor and inspect 50 exam rooms.

Value 2

Significantly improve proctoring efficiency and quality

After implementing the real-time intelligent exam proctoring system (see Figure 6), video proctors shift from reviewing surveillance footage in real time to focusing on processing system-generated alerts of anomalies. This allows them to specifically inspect certain students when irregularities are detected, making proctoring clearer and more comprehensive without omissions. This enhances the auxiliary role of surveillance video in proctoring.

The chief and deputy examiners at the exam site can also view statistical abnormal data for the site and each room on the configured data screen, enabling timely detection and handling of abnormal situations.

Value 3

Provide a primary basis for exam decisions

Provincial, municipal, and district exam institutions can view real-time statistical alert data and details for each exam site within their jurisdiction through the main screen at the management platform. In this way, they can monitor exams in real time.

The abnormal behavior alerts generated by the real-time intelligent exam proctoring system at the exam site can be imported into the provincial big data platform for smart exam decisions in the form of video clips and images. Therefore, higher-level exam institutions can have a primary basis for exam decisions.

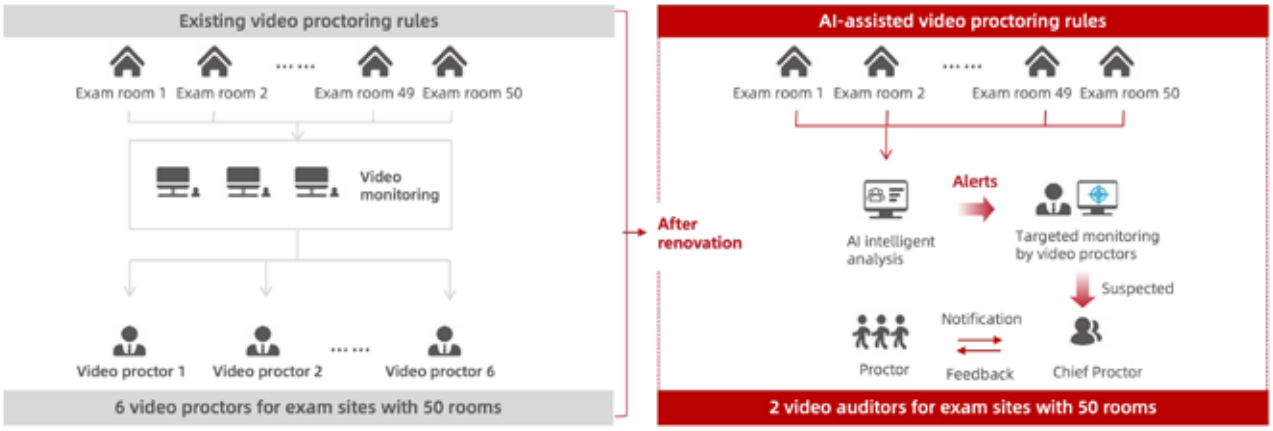


Figure 6: Comparison Before and After Implementing the Real-Time Intelligent Exam Proctoring System

REFLECTION

The AI proctoring project marks an active attempt to apply artificial intelligence in educational exam services. By marking and analyzing common violations and exam management issues through comparisons of human body movements and behaviors, it intelligently identifies suspected violations in surveillance videos. Its significance lies not only in improving the efficiency of traditional manual identification but also in the precise identification of violations and disciplinary issues in exam management.

Follow-up plans

- Improve the accuracy of abnormal behavior algorithms. Expand algorithm recognition coverage to identify 44 types of abnormal behaviors according to examination procedures. Enhance recognition accuracy through techniques like temporal analysis and large models to reduce false alarms.
- Explore handling mechanisms for alerts. Explore a new exam alert mechanism based on intelligent proctoring and inspection systems, with work responsibilities defined and personnel allocation optimized. Timely track and handle suspected cheating on the last mile from the proctoring center to the classroom to ensure exam security.

Smart First Aid Project Based on Medical Large Model

As a crucial part of the medical system, first aid and emergency care service ensures timely medical assistance to patients. By leveraging new quality productive forces and cutting AI+5G technologies, Sichuan Provincial People's Hospital works with China Mobile Chengdu Institute of Research and Development to explore innovative applications and practices in emergency care scenarios. The large model for first aid has enabled intelligent dispatch, intelligent assistance, intelligent triage, and emergency decision-making support, protecting public health against the clock.

Sun Mingwei
Director of Emergency Center, Sichuan Provincial People's Hospital

SOLUTION PARTNERS

Authorities like the National Health Commission and the Ministry of Science and Technology have introduced policies aimed at promoting the high-quality development of healthcare services through scientific and technological innovation. For example, the "14th Five-Year Plan for National Informatization" proposes to "strengthen the application of health and medical big data", and the "New Generation Artificial Intelligence Development Plan" emphasizes the "application of artificial intelligence in healthcare". First aid and emergency care are integral to the healthcare system, and the speed of response and quality of treatment have a direct bearing on the patients. To improve overall capability in emergency treatment and first aid, Sichuan Provincial People's Hospital has followed the "Healthy China 2030" strategy and joined hands with China Mobile Chengdu Institute of Research and Development (hereinafter referred to as "Chengdu Institute") to build an integrated system of first aid and emergency care by leveraging China Mobile's 5G + AI technologies.

At the 9th Asia-Pacific International Emergency Forum in May 2024, Sichuan Provincial People's Hospital launched China's first LLM for first aid and emergency care together with the Emergency Medicine Branch of the Chinese Medical Association and Chengdu Institute. Drawing wide attention in the industry, the model has been applied in many places on a trial basis. Moreover, some leading healthcare organizations offered to discuss the application of the model in healthcare as shown in Figure 1.

Thanks to such AI capabilities as intelligent dispatch, assistance, triage, and Q&A, an integrated platform for first aid and emergency care has been created, covering three scenarios: daily first aid, emergency rescue, and public health emergency. Chengdu Institute has developed some key systems like 120 video dispatch system, 5G first aid system, intelligent first aid assistance, and emergency rescue system, which have effectively driven the innovative development of first aid services across the country and made the first aid industry more digital and intelligent.

Figure 1: Launch of the Large Model for First Aid

The large model for first aid and its application have significantly improved the efficiency and quality of first aid and emergency care services. It can quickly and accurately identify the patient's condition, optimize the allocation of medical resources, and provide precise medical advice and decision-making support for medical staff, thus realizing rapid response and efficient treatment. The large model also plays an important role in emergency rescue as it improves the overall emergency management of the healthcare system, ensures reasonable and accurate decision-making at critical moments, and provides powerful support for safeguarding people's safety and health.

CHALLENGES

Dispatch efficiency requires immediate improvement

When an emergency occurs, EMS dispatchers need to handle a large number of incoming calls, assess their urgency, and effectively allocate limited emergency resources. Faced with great pressure on rapid and accurate information processing and decision-making, manual labor alone will impair efficiency. The large AI model integrates natural language processing and machine learning technology to automatically identify and classify incoming calls, quickly extract key data such as patient location, condition description, etc., and intelligently match first aid resources and optimize dispatch decisions, thus ensuring efficient allocation of first aid emergency resources.

First aid difficulty during waiting time

During the golden waiting time before the ambulance arrives, non-professionals may have difficulty in making the best first aid decisions due to the complex situation, pressure or lack of experience, and fail to provide timely and effective assistance. An advanced medical knowledge base with intelligent assistance is necessary to provide remote real-time medical guidance, including CPR procedures and hemostasis methods.

Skills gap among first aiders

Differences in the professional background, training level, and experience of first aiders, especially in remote areas, may lead to inaccurate injury assessment, inappropriate first aid treatment, and risks of omission and misdiagnosis. There is an urgent need to improve the quality and homogenization of emergency medical services.

Heavy emergency workload

Hospitals are plagued by heavy emergency workload due to a variety of factors. This affects medical service quality and patient satisfaction and creates pressure on medical professionals. In this light, AI capabilities should be combined with information technology to help healthcare workers optimize their workflow and improve efficiency.

SOLUTION AND VALUE

01 Overall structure/solution

Targeting the pain points of pre-and in-hospital emergency care as mentioned above, parties are joined together to develop a large model for first aid with multiple capabilities, such as knowledge quiz, key information extraction, grading and triage, plan classification, and assisted diagnosis and treatment, making first aid and emergency care more intelligent.

The model is trained and optimized with the textbook data of 18 medical disciplines from top universities in China and emergency clinical data of 5 million patients to ensure the data is diverse and balanced, with both theoretical and clinical data involved. A number of AI capabilities that go beyond mainstream large models in the industry have been delivered, reshaping first-aid services as shown in Figure 2.

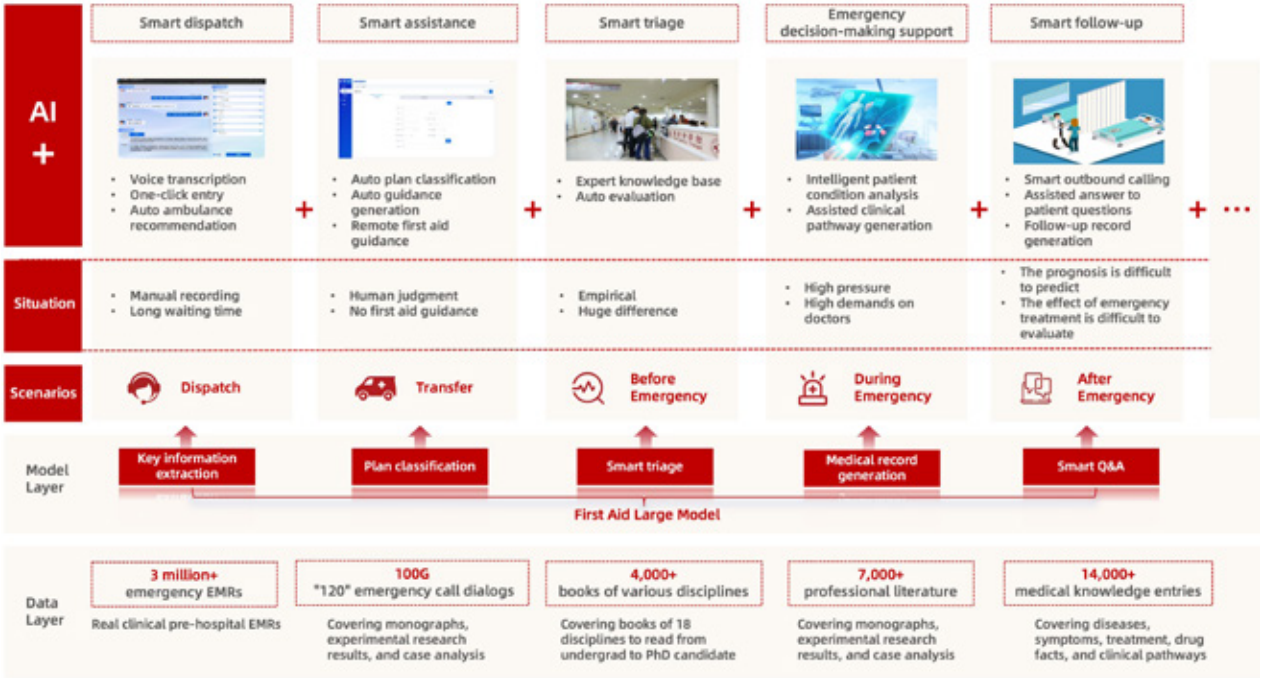


Figure 2: Empowerment of the Large Model for First Aid

02 Application scenarios

The model can be applied in diverse scenarios such as 120 emergency treatment, hospital emergency care, emergency rescue, primary health care, and CDC emergency treatment in pre-hospital, in-hospital and post-hospital care, making the whole process of first aid and emergency care intelligent.

03 Key innovations

Intelligent dispatch: With technologies like speech recognition and key information extraction, the model can intelligently extract key information from incoming calls, such as chief complaint, address, and cell phone number, and automatically recommends ambulances, thus reducing manual entry and shortening dispatch time.

Intelligent assistance: Before the ambulance arrives, the model can present the most appropriate emergency care plan and guidance measures based on its extensive knowledge base. It allows dispatchers to provide remote first aid guidance during the waiting time.

Intelligent triage: Before the patient arrives at the hospital, the system can assess the patient's severity based on vital sign data and records collected by onboard paramedics and emergency equipment. Then a pre-hospital electronic medical record is generated and sent to physicians so that they can get better prepared for rescue and treatment.

Emergency decision-making support: The system intelligently analyzes the patient's condition based on emergency data and generates a clinical path for diagnosis and treatment. Then emergency clinicians can quickly determine the patient's condition and obtain treatment recommendations. As a result, more efficient emergency treatment is realized.

04 Business models

The intelligent emergency care platform can provide a comprehensive and differentiated solution of "product+capability+service".

i. Improve emergency care efficiency with innovative information systems such as 120 video dispatch system, 5G emergency system, intelligent first aid assistance, and emergency rescue system;

ii. Provide basic service capabilities such as the large model for first aid and precise positioning, and promote new quality productive forces through joint R&D and innovation;

iii. Provide overall solutions and customized development services to meet the needs of medical institutions.

05 Core values

Faster emergency dispatch	Backed by technologies like the Attention Mechanism, the model can extract key information from incoming calls, reduce manual entry, and automatically recommend ambulances, thus shortening the dispatch time by about 1 minute.
Quality remote guidance	Once the ambulance is dispatched, the system intelligently presents the first aid plan and guidance measures so the dispatcher can provide correct and scientific remote guidance during the waiting time. In a word, the patient can receive rescue about 15 minutes earlier.
Shorter waiting time	When the patient is transported to hospital, the system can make use of its intelligent triage capability based on multi-source heterogeneous data fusion and pre-training model to intelligently assess the patient's severity. Then electronic pre-hospital medical records are generated and sent to relevant physicians, reducing record writing time by 30% and improving treatment efficiency.
Higher emergency efficiency	As for in-hospital emergency treatment, the model combines medical knowledge and expert experience to generate a chain of thought for clinical treatment based on CoT Prompting. The chain can support medical professionals in making decisions and improving treatment. Moreover, it will empower digital innovation in first aid and emergency care.
Application results	By leveraging advanced technologies of 5G, IoT and AI, Sichuan Provincial People's Hospital has provided support for regional disaster rescue operations and major events, including the "6.17 Changning earthquake" in Yibin, Sichuan, the "9.16 Lu County earthquake" in Luzhou, Sichuan, and the "9.5 Luding earthquake" in Sichuan. It is the first in the world to use 5G+AI technology in disaster medical rescue. It not only creates a new type of emergency care system at the operational level, but also simplifies the management process for relevant departments and personnel so that they can engage in medical rescue more effectively.

REFLECTION

Sichuan Provincial People's Hospital and Chengdu Institute have accumulated valuable experience in in-depth innovation and practice of the large model for first aid. By closely integrating 5G+AI technology with the pre-hospital emergency process, it not only greatly improves emergency response efficiency and treatment quality, but also optimizes the allocation of emergency care resources through data-driven decision support.

With continuous progress and innovation in AI technology, we will continue to explore the application of large AI models in emergency care. This means we will not only expand the service scope to cover a wider range of emergency care scenarios and needs, but also work to improve treatment effectiveness with greater convenience and shorter response time. In addition, data security will be our top priority. We work to protect the security and confidentiality of patient information through enhanced data encryption, access control and privacy protection measures. We want to build a smarter, safer and more efficient emergency care system, and contribute more to improving medical treatment and protecting public health.

Empowering Cultural Tourism Customer Service with Large Models

In the cultural tourism industry, customer service is a crucial aspect of enhancing user experience and brand loyalty, and it is increasingly being empowered by large model technology. Large models can efficiently and intelligently interact with users through natural language generation capabilities, significantly improving the efficiency and quality of service. In response to the customer service scenarios in the cultural tourism sector, the Software Development Center of Beijing Company under China Telecom Digital Intelligence Technology Co., Ltd. has independently developed smart Q&A based on Retrieval-Augmented Generation (RAG). The utilization of inference acceleration technologies such as TensorRT-LLM ensures accurate and fast inference by the large model. Leveraging the Telecom TeleChat-52B, it builds smart agents to address issues of hallucination, accelerating the application of large models across various business scenarios and enhancing the competitiveness of enterprise products.

Zhang Shulang
Deputy Director Expert, DICT Capability Center, Beijing Telecom

SOLUTION PARTNERS

During the 2023 Double Eleven period, the Palace Museum launched a unified visitor consultation hotline with Telecom's 400+ cloud call services. However, due to the surge in call volume, the connection rate for human agents dropped significantly during peak times. The AI R&D team at the Software Development Center of Beijing Company under China Telecom Digital Intelligence Technology Co., Ltd. utilized China Telecom's proprietary TeleChat-52B large model to extract Q&A pairs from real customer service recordings at the Palace Museum and build a training dataset. By using models like Lora for efficient supervised fine-tuning of model parameters and TensorRT-LLM and GPTQ for acceleration and quantization, it significantly reduced resource consumption and lowered the implementation costs of large models. Additionally, the model's comprehension was improved by integrating external knowledge bases for retrieval-augmented generation. Based on multi-turn dialog management and precise intent capture, the system ensures coherent dialog logic and accurate, efficient responses. Multi-turn dialog summarization and transliteration maintain information integrity in complex scenarios, while compliance matching guarantees that responses are both smart and compliant.

The self-developed smart customer service product by Beijing Telecom (see Figure 1) accurately targets customer groups in tourist attractions, hotels, and travel agencies. This effectively reduces the costs of human customer service, significantly enhances service response speed and quality, and increases customer retention and brand loyalty. The economic benefits are substantial, optimizing operational cost structures and driving word-of-mouth through exceptional customer experiences. This efficient and smart service model is setting a new trend in the field of smart customer service for the cultural tourism sector, showing vast development potential.

Figure 1: Intelligent Voice Customer Service Interface at the Palace Museum

CHALLENGES

The smart customer service voice platform integrates various technologies, including speech recognition, speech generation, natural language processing, and smart knowledge base construction. Any shortcomings in these areas can diminish user experience. The specific challenges can be divided into four categories:

First of all, data brings extremely obvious challenges. The diversity of visitor inquiries—covering bookings, transportation, sightseeing routes, historical context, and more—creates significant complexity. The variety and complexity of questions require the smart customer service system to handle a vast amount of different data types. Additionally, variations in accents among different visitors pose a major challenge to the

accuracy of speech recognition, sometimes preventing the system from correctly understanding and processing user requests. The accuracy of speech recognition is crucial for correctly matching relevant knowledge base content.

Secondly, the interactive experience is of great importance. Visitors often require multiple rounds of dialog during inquiries. If the system fails to remember and understand the prior context in a continuous Q&A scenario, it may provide incomplete or incorrect answers. A lack of contextual understanding in customer service systems severely impacts user experience and hinders satisfaction.

Thirdly, the challenges in human customer service primarily stem from response rates and costs. During peak times, such as holidays or special events, the volume of inquiries can surge, often overwhelming human agents and leading to long wait times for visitors, which in turn decreases satisfaction. Moreover, maintaining a large team of human customer service agents is costly, adding significant operational expenses. While smart customer service can alleviate this pressure, it must match the accuracy and concurrency capabilities of human agents.

Last but not least, the limitations of traditional technologies pose certain challenges. Traditional smart customer service heavily relies on localized knowledge base construction, employing template matching, intent recognition, and slot filling for Q&A. However, this approach often fails to achieve satisfactory recall rates. Due to technological limitations, traditional systems often struggle to correctly understand and respond to visitor inquiries, resulting in rigid and impersonal responses. This not only diminishes the visitor consultation experience but can also lead to the dissemination of incorrect information.

SOLUTION AND VALUE

01 Overall structure/solution

This solution is tailored for smart call center agents in the cultural tourism sector, deeply integrating various scenarios such as entry and ticket inquiries, travel guidance, and travel planning. It aims to revolutionize user experience through smart customer service. Its architecture includes four parts (see Figure 2):

1. The 400 hotline responds to incoming user calls.

2. Combining AI technologies such as speech recognition, large models, and speech synthesis, the front-end Q&A service automatically converts user inquiries into corresponding text. After filtering sensitive words, correcting text, and transcribing, it matches questions through a knowledge base to query the most relevant knowledge records. These records are then sent to the large model for responses, which are converted into speech via Text-to-Speech (TTS) and fed back to the call handling system.

3. Building on smart inquiries, the back-end management system manages Q&A data, including the discovery of new questions, classification and summarization of questions, analysis, trend prediction, and information import. This helps users quickly identify patterns in visitor inquiries.

4. Functions such as information retrieval, permission management, and Q&A management form the foundational components of the entire smart customer service system, providing corresponding capabilities for upper-level services.

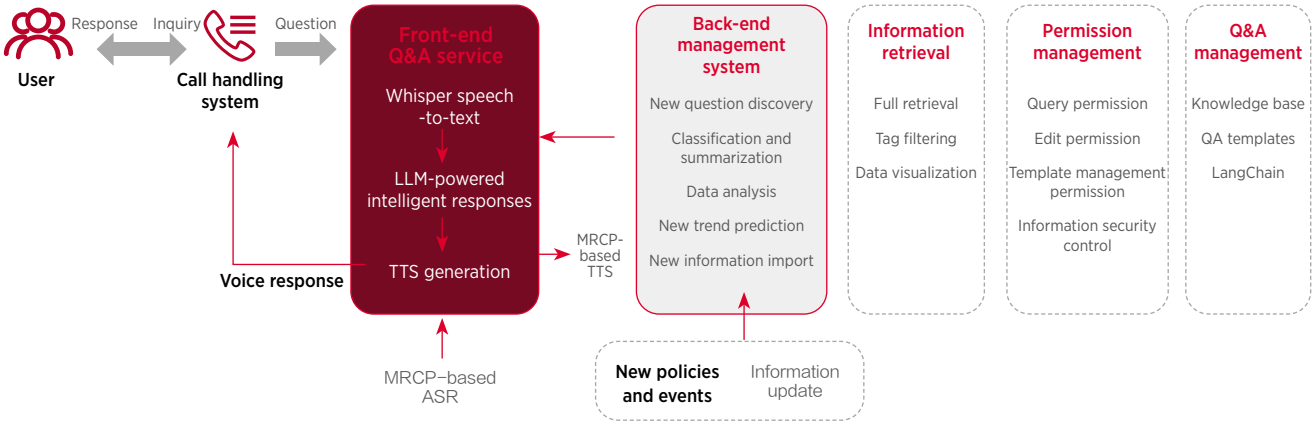


Figure 2: Intelligent Voice Customer Service Business Architecture

02 Application scenarios

This solution focuses on smart call center agents within the cultural tourism industry. The smart customer service system, powered by large models, has already been introduced in the customer service project at the Beijing Palace Museum. It delivered remarkable results and received high praise and recognition from clients. Furthermore, the capabilities developed in the smart customer service project have been transformed into three standard products: AI Intelligent Customer Service Assistant, AI Brain, and AI Cloud Call Platform.

AI Intelligent Customer Service Assistant aims to address four major pain points: inconsistent service levels, low service efficiency, difficulty in monitoring service processes, and challenges in mastering business knowledge. Based on natural language processing and large model technology, along with speech-to-text and speech synthesis, it enables understanding of customer intent, smart Q&A, decision-making support, scenario management, and smart quality inspection. This creates a smart assistant for the customer service industry, enhancing frontline service quality and efficiency, and helping enterprises reduce costs and increase effectiveness.

Targeting government, enterprise, and higher education audiences, AI Brain builds on China Telecom's self-developed large language model. It continuously integrates and learns various knowledge documents within the organization, achieving human-like interactive Q&A through smart

processing and understanding. This helps enterprises or clients quickly obtain accurate, timely, and effective knowledge from vast repositories of information.

AI Cloud Call Platform addresses common demands such as product marketing, product research, customer follow-up, and care services for enterprise clients. Powered by advanced technologies like speech recognition, text-to-speech, natural language processing, and large models, this smart call platform integrates AI calling, AI video calling, and manual outbound calls, while also providing comprehensive services such as smart quality inspection and smart assistance.

03 Key innovations

Traditional customer service in industries typically uses a matching paradigm, relying on intent recognition and slot filling to identify user inquiries and match them with standard Q&A templates. This approach lacks flexibility and fails to fully understand user semantics. This solution introduces China Telecom's self-developed Xingchen large language model to comprehend user inquiries, integrating industry knowledge through an external knowledge base. By utilizing vector matching, knowledge refinement, and the BM25 algorithm within a retrieval-augmented generation (RAG) framework, it significantly enhances the understanding of users' true intentions. The large model also enriches capabilities for multi-turn dialogs and compliant answer selection, ensuring that the final responses are accurate and compliant. Additionally, the TensorRT-LLM solution accelerates and quantizes the large model, empowering it with high-speed inference capabilities while reducing resource consumption. Multi-turn dialog summarization ensures that information remains intact in complex scenarios, while compliance matching technology guarantees that responses are both intelligent and compliant.

04 Business models

Targeting the existing clients of China Telecom's extensive cloud call products and opportunities within the cloud calling platform, the marketing team uses the smart customer service project at the Palace Museum as a benchmark case. They actively promote AI smart customer service products to users, which not only reduces the labor costs associated with traditional customer service but also enhances customer service perception. By integrating traditional communication service capabilities with large model AI capabilities, the product holds a leading position in industry competition.

Leveraging cutting-edge AI technologies such as large language models, RAG, and Agents empowers the traditional customer service call platform to provide fast, accurate, high-concurrency, and cost-effective AI voice customer service capabilities. This enhances the customer experience and core value of traditional telecom businesses while providing a platform for the commercial implementation of large models based on traditional telecom call operations.

With a focus on project execution, core algorithm capabilities of the large model are solidified and transformed into standardized products, reducing R&D costs while enabling scalable replication of R&D outcomes. This allows for rapid delivery in similar scenarios, providing quality service to more users.

05 Core values

By utilizing a foundation for smart customer service based on large model Agents and retrieval-augmented generation (RAG), we enhance AI customer service with powerful language understanding, a rich knowledge base, multi-turn Q&A comprehension, and continuous learning capabilities. This significantly lowers the cost of human customer service and reduces operational costs by over 30%, saving clients substantial human resource expenses. Moreover, the introduction of an external knowledge base has improved the accuracy of system responses to over 95%, addressing common issues in traditional customer service bots such as misunderstanding, irrelevant answers, and incomplete information. Feedback data shows that user satisfaction with the large model smart customer service has increased by over 40%. Through self-developed inference acceleration built on operator optimization and the introduction of large model post-training quantization (PTQ), the large model operates quickly and accurately while reducing deployment memory overhead by over 60%. This directly lowers the implementation costs of large models, improving connection rates during peak inquiry periods by over 50% and effectively reducing visitor wait times. The smart customer service call platform built on the large model is projected to generate direct revenue exceeding RMB 5 million in 2024, with anticipated annual revenue growth. The integration of AI large model technology with traditional business strengths enhances China Telecom's brand value.

REFLECTION

In the context of voice smart customer service, the speed of knowledge retrieval and response generation from the AI large model is also a critical factor affecting customer experience. We have significantly improved system response speeds through self-developed model adaptation and inference acceleration technology based on operator optimization, alongside post-training quantization. At the same time, we have introduced context understanding and intent capture technologies during multi-turn dialogs, allowing the system to respond more accurately to user inquiries and reducing the rate of information transfer errors. This not only strengthens customer loyalty but also enhances brand fidelity.

Next, we will continue to hone our core capabilities of the large model within projects and transform these capabilities into standardized products to improve the reusability of R&D outcomes, achieving cost reduction and efficiency enhancement. We will keep developing new functional modules for smart customer service products, regularly iterating large model versions using incrementally high-quality datasets to enhance product value and boost user experience. Furthermore, we will strengthen communication and collaboration with partners to explore future development directions in smart customer service, setting industry benchmarks in AI voice smart customer service. Finally, we will expand the application of large models, injecting their capabilities into more traditional operator business scenarios and strategically emerging business fields, empowering telecom operations with AI and making them a vital link in the product ecosystem.

Yuanjing Culture & Creativity Large Model Empowers the Dissemination of Culture

To respond to the national strategy of building a culturally strong nation and the increasing market demand for culturally enriched creative products, China Unicom, in collaboration with the Center for International Cultural Communication and China Federation of Literary and Art Circles (www.cflac.org.cn), has jointly developed the "Anique Nymph" Yuanjing large model application platform. This initiative aims to empower the anthropomorphization of cultural relics and promote the dissemination of outstanding traditional Chinese culture.

Wang Kan
Deputy Director of the Center for International Cultural Communication

SOLUTION PARTNERS

China unicom中国联通 | CICC 中国外文局文化传播中心 | 中国文艺网 | 沃橙信息

China Unicom's Yuanjing Culture & Creativity Large Model focuses on showcasing cultural relics, aiming to develop a large model and applications enriched with knowledge of these relics. The Yuanjing Culture & Creativity Large Model leverages semantic understanding and reasoning capabilities to construct high-quality textual semantic representations. It supports culture & creativity design by anthropomorphizing relics, and uses technologies like content generation and multimodal interactions, including text-to-image, text-to-video, and AI face swap, to automate the creation of explanatory texts and images for these relics. Additionally, it generates new images based on existing ones, tailored to specific prompts. The Yuanjing Culture & Creativity Large Model excels in producing personalized images in various styles while maintaining high authenticity. It addresses the complexity of cultural relic knowledge and the high costs associated with designing culture & creativity works, highlighting the potential of cultural relic data elements in the creative sector. This approach offers new impetus and direction for cultural innovation and industrial growth.

In partnership with the China Foreign Languages Publishing Administration and China Federation of Literary and Art Circles (www.cflac.org.cn), China Unicom has launched the "Anique Nymph" Yuanjing Large Model Application Platform built on the Yuanjing Culture & Creativity Large Model. This platform stands as a referenceable and replicable solution for empowering the cultural industry through AI, laying a strong foundation for industry advancement and replication, with significant potential for benchmarking.

CHALLENGES

As cultural consumption evolves, the demand for culturally enriched creative products is on the rise. and the culture & creativity industry faces several challenges:

Market demand forecasting	The industry must swiftly adapt to market shifts to align with consumer expectations. As consumer tastes and needs are constantly changing, accurately predicting market trends and customer preferences is challenging, making it difficult to ensure that products and services meet market needs.

Digital transformation	With the rise of digital media, traditional culture & creativity industry must embrace digital transformation to stay relevant in the digital age.
Pressure for design innovation	Continuous innovation is essential to captivate consumers, yet lengthy design cycles place ongoing pressure on designers to consistently deliver fresh ideas.
Cultural sensitivity	Incorporating cultural elements into designs requires a deep understanding of various cultures to avoid misunderstandings or causing offense.

SOLUTION AND VALUE

01 Overall structure/solution

The overall goal of China Unicom's Yuanjing Culture & Creativity Large Model is to use AI to anthropomorphize cultural relics, drawing inspiration from the treasures of Chinese civilization. This initiative aims to create male and female character images derived from these cultural relics, showcasing the integration of cultural innovation and technologies, and promoting the creative transformation of traditional Chinese culture while advancing the intelligent creation of culture & creativity content. The project's overall architecture, as shown in Figure 1, comprises four main layers: data, computing, model, and application layers.

The data layer serves as the foundation of the culture & creativity large multimodal model (LMM). It manages the collection, storage, and organization of multimodal data for training and real-time generation by the Yuanjing Culture & Creativity Large Model. The quality and diversity of multimodal data directly determine the model's performance and the richness of the generated results. The dataset for the Yuanjing Culture & Creativity Large Model is constructed using text and image content from various sources, including literature works, public datasets, online articles, social media, and open image libraries. It undergoes processes like cleaning, denoising, annotation, enhancement, and alignment to improve its diversity and quality.

At the core of the project, the model layer handles the integration and processing of multimodal data and intelligent content generation. The design of this layer determines the system's intelligence and creative generation capabilities. By leveraging pre-trained models like contrastive language-image pre-training (CLIP) and large language model (LLM), text data is converted into vector representations that retain semantic information, aligning text and image modalities for consistent multimodal information in semantic space. The diffusion models' stability and controllability support gradual optimization in generating target images.

The application layer is the user-facing component of the culture & creativity LMM. It provides a range of practical tools and culture & creativity services aimed directly at users, assisting them in generating and editing creative content. The design of this layer prioritizes user needs and operational habits to balance usability and functionality. Central to this layer is controllable image generation services, which produce images based on user-input text descriptions and reference images.

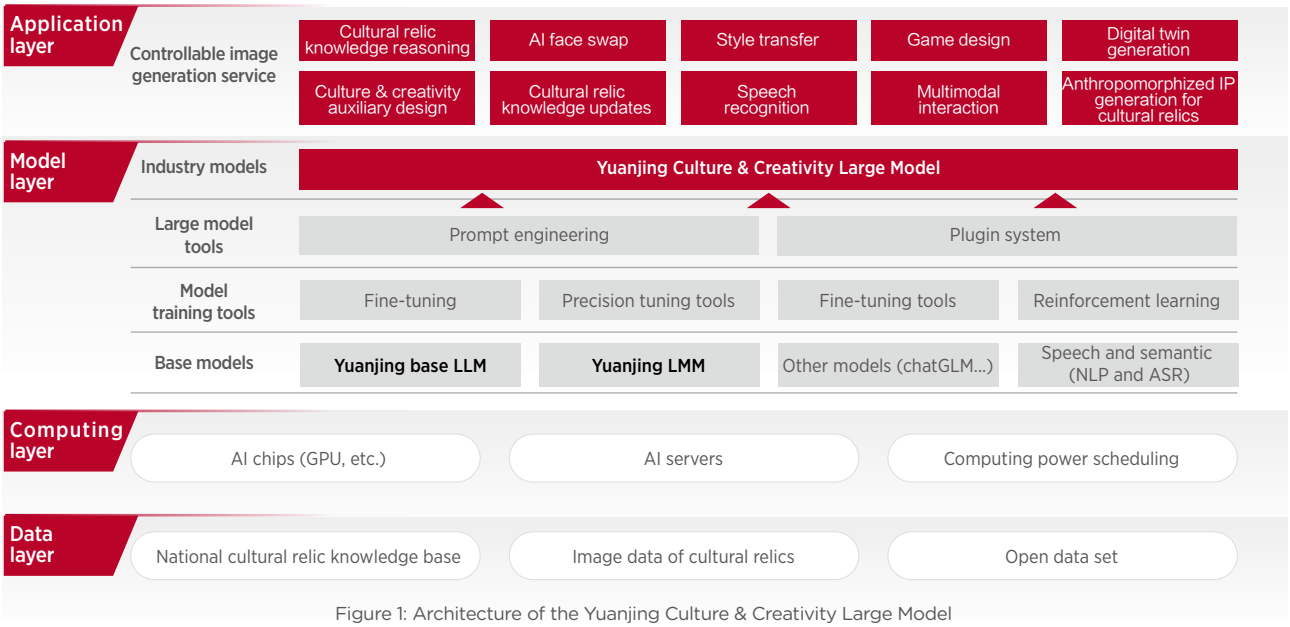


Figure 1: Architecture of the Yuanjing Culture & Creativity Large Model

02 Application scenarios

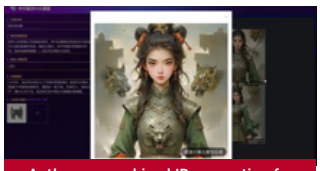
The Yuanjing Culture & Creativity Large Model builds high-quality textual semantic representations through semantic understanding and reasoning capabilities. Utilizing cultural relics from various historical dynasties as inspiration and basic materials, the model explores the elements and artistic characteristics of these relics, anthropomorphizing them through large model capabilities. Figure 2 displays the main application scenarios and distinctive features of the Yuanjing Culture & Creativity Large Model. The primary application scenarios are as follows:

Scenario 01 Based on the Yuanjing Culture & Creativity Large Model, historical relics are transformed into anthropomorphized characters with unique personalities and rich cultural connotations. This results in culturally significant character images that carry abundant historical and cultural information, providing a new medium for the dissemination of traditional culture.

Scenario 02 Inspired by the anthropomorphized images generated by the Yuanjing Culture & Creativity Large Model, a range of stories are created involving historical reenactments, cultural interpretations, and creative narratives. This allows readers to vividly experience the stories and cultural spirit behind the relics, enhancing cultural identity and emotional resonance.

Scenario 03 By combining the images of cultural relics with narrative content, digital cultural creative products are designed. This includes the creation of virtual idols, animations, games, and AR/VR experiences that possess artistic value and innovation. Through digital means, traditional culture is revitalized within modern society.

Scenarios



Anthropomorphized IP generation for cultural relics

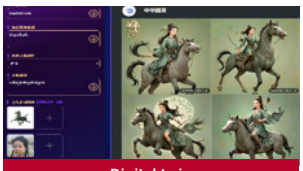
Characters are crafted using cultural relics from various historical dynasties as inspiration and basic materials, embodying the distinct features of these artifacts. Through in-depth analysis, the Culture & Creativity Large Model extracts their most representative elements, such as shape, patterns, colors, materials, and the historical stories and cultural significance behind them. This process enables the generation of IP images that preserve the essence of the relics while aligning with modern aesthetic preferences, embodying cultural depth and contemporary relevance.

Culture & creativity auxiliary design

Based on large models, digital culture & creativity products, games, cards, and other commercialized interactive content and applications are generated. The model can automatically complete certain repetitive design tasks, such as color matching, font selection, and layout adjustments, allowing designers to focus on creative ideas.

Cultural relic knowledge interaction

Through the platform, users can interact with cultural relics, which are used as clues to tell related historical stories, legends, or cultural events. This allows users to gain a deeper understanding of the cultural connotations and historical background of the relics, enhancing their awareness of the relics' historical context and cultural value. It is an innovative application that integrates education, entertainment, and interactivity, aiming to make it easier for the public to understand and learn about cultural relic knowledge through digital means.



Digital twin

Creating realistic digital twins for viewers and merging them with artifact images generates photos with personal characteristics and cultural essence. These twins are virtual characters created based on users' real images, voices, personalities, and behavioral traits. Digital twins can accurately replicate users' appearances, expressions, voices, and behaviors, enabling real-time interaction with users. They become another "self" for users in the digital world.

Features

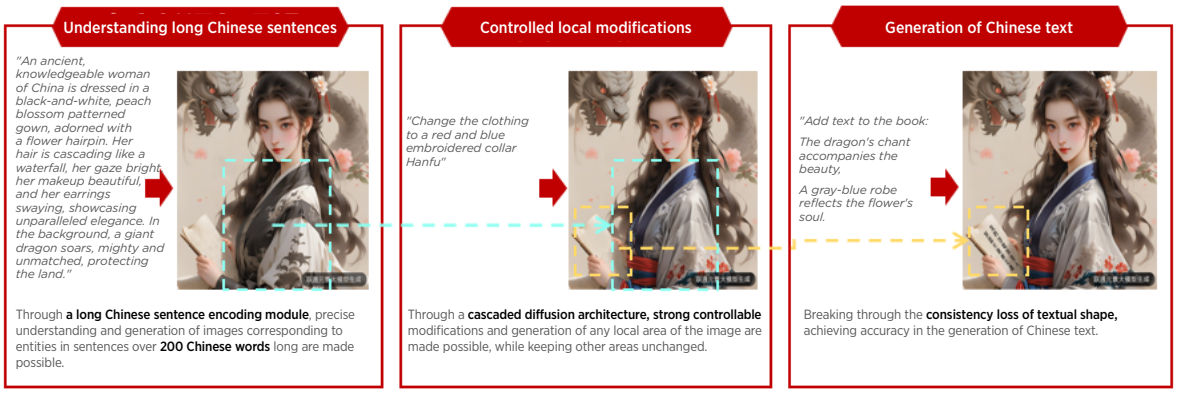
Text-to-image	Image-to-image	Style transfer	Feature alignment	Multimodal interaction
AI face swap	Speech recognition	Cultural relic knowledge reasoning	Culture & creativity auxiliary design	Cultural relic knowledge updates

Figure 2: Application Scenarios and Features of the Yuanjing Culture & Creativity Large Model

03 Key innovations

The Yuanjing Culture & Creativity Large Model is centered around the Yuanjing LLM with hundreds of billions of parameters and innovatively incorporates a composite visual encoding module, as shown in Figure 3. By integrating various visual features and contextual information, the model enhances its ability to capture image details and deepens its understanding of image content. This allows the model to possess refined content comprehension, complex logical reasoning, and precise visual computation skills.

Additionally, the Yuanjing Culture & Creativity Large Model employs an innovative long sentence encoding module and a cascaded diffusion architecture for precise entity extraction and intent understanding of long Chinese sentences. This improves the model's encoding ability for high-frequency features such as text within images, enhancing its expressiveness and generation capabilities in complex contexts. Consequently, the model has a highly controllable ability to generate images from Chinese text, accurately understanding and generating corresponding images for long sentences exceeding 200 Chinese words. It supports highly controllable modifications and generation of any local area within an image while keeping other areas unchanged, overcoming the limitations of Chinese text generation and achieving seamless integration of text and visual scenes. Additionally, the Yuanjing Culture & Creativity Large Model utilizes LoRA technology to fine-tune the large pre-trained model with industry-specific data, promoting the model's powerful generative performance while generating images that match target artistic styles.



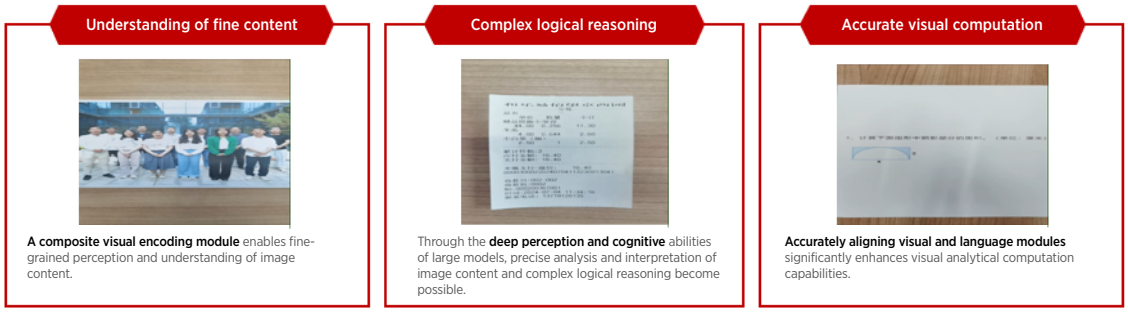


Figure 3: Innovative Highlights of the Yuanjing Culture & Creativity Large Model

04 Business models

The Yuanjing Culture & Creativity Large Model offers personalized customization services to culture & creativity enterprises and organizations to meet diverse user needs. It provides three service models: SaaS, API calls, and project customization. For independent designers or studios, the SaaS model offers a subscription-based service for standardized digital tools for cultural relics, with pricing based on the subscription period. For culture & creativity enterprises with development capabilities, API call services are available and billed based on usage frequency and data volume. For large projects with special requirements, personalized project customization services are available. Pricing for these services depends on project specifics, including needs assessment, workload, technical difficulty, and development timeline. The price covers all phases from initial consultation and requirement analysis to design, development, testing, delivery, and maintenance. The Yuanjing Culture & Creativity Large Model is expected to achieve over RMB 20 million in revenue within three years, starting next year.

05 Core values

The Yuanjing Culture & Creativity Large Model has dramatically shortened the creation cycle for anthropomorphized cultural relic character IPs by over 70%, reducing dependence on professional designers and physical materials. This advancement has significantly boosted the efficiency of culture & creativity product development, lowered the human and material costs associated with cultural design, and improved the efficiency of cultural dissemination.

Application benefits: The model has created over 200 anthropomorphized cultural relic characters and generated 210 related story series. These stories have been published on the Weixin Official Account "china-wyw" of China Federation of Literary and Art Circles, significantly enhancing the efficiency and reach of cultural dissemination.

Case 1

Based on anthropomorphized cultural relic characters, a wealth of educational and entertaining game products have been developed. These games draw on the historical periods and cultural backgrounds of the relics, featuring rich plots and tasks that encourage players to learn and explore during gameplay. Through these games, users can interact with cultural relics in a fun and engaging way, deepening their understanding and interest in traditional culture.

Case 2

Inspired by anthropomorphized cultural relic characters, a range of culture & creativity products have been launched, each conveying specific knowledge about cultural relics. This application has enriched the diversity of culture & creativity products, allowing users to learn about history and culture while collecting and exchanging cards.

Social benefits: The model has been showcased at various major exhibitions at home and abroad, drawing attention and recognition from audiences and professionals from over 20 countries and regions. It has enhanced public awareness of cultural relic knowledge and cultural literacy, disseminating images and stories of national treasures and promoting Chinese culture globally. This initiative has fostered understanding and recognition of Chinese culture around the world, thereby enhancing the global awareness and international influence of Chinese culture. Furthermore, the Yuanjing Culture & Creativity Large Model has promoted social diversity through innovative dissemination methods, providing opportunities for people from different backgrounds and ages to engage with and learn about traditional culture. This has deepened the public's sense of identification and pride in Chinese culture.

REFLECTION

The Yuanjing Culture & Creativity Large Model delivers personalized recommendations and engaging experiences through its innovative interactive approach, expanding the reach and market influence of culture & creativity products. By seamlessly integrating traditional culture with modern technology, it revitalizes heritage in the digital era and forges new pathways for the inheritance of traditional culture and the integration of modern creativity.

Looking ahead, the model plans to collaborate with museums, cultural heritage sites, and cultural relic institutions. Leveraging the content generation capabilities of the large model, it aims to transform extracted cultural elements and artistic characteristics into digital representations, bringing cultural relics to life. The model is committed to enhancing the global appeal of Chinese culture while actively supporting and promoting the unique development and prosperity of diverse cultures worldwide, thus facilitating global diversity.

Applications of Smart AI-powered Medical Assistant

In many developed regions worldwide, including Taiwan, doctors spend up to 75% of their working time on paperwork, according to some statistics. This burden extends to nurses as well, who must dedicate significant time to documenting patient records. AI-powered medical assistants can help relieve this burden by establishing services for the healthcare industry, such as speech-to-text (STT) and summarization functionalities, ultimately reducing consultation time for doctors.

YU, HSIENCHUAN
Section Chief of the Intelligent Network Division,
Chunghwa Telecom Information Technology Branch

SOLUTION PARTNERS

The smart AI-powered medical assistant applies the technologies of GPT, Claude, and Titan in the large language models (LLMs) of Microsoft Azure and AWS to process the dialogs between doctors and patients during consultations and the medical advice from doctors. The assistant provides a summary of doctors' diagnoses, allowing medical personnel to record the medical care process using voice, as shown in Figure 1. This reduces paperwork, relieving the stress and burden of frontline medical personnel. In addition, the retrieval-augmented generation (RAG) framework is used to build a health education database. This database simplifies the understanding of complex medical concepts and provides patients with treatment guidelines and personalized care instructions, empowering the public to manage their health effectively after consultations.

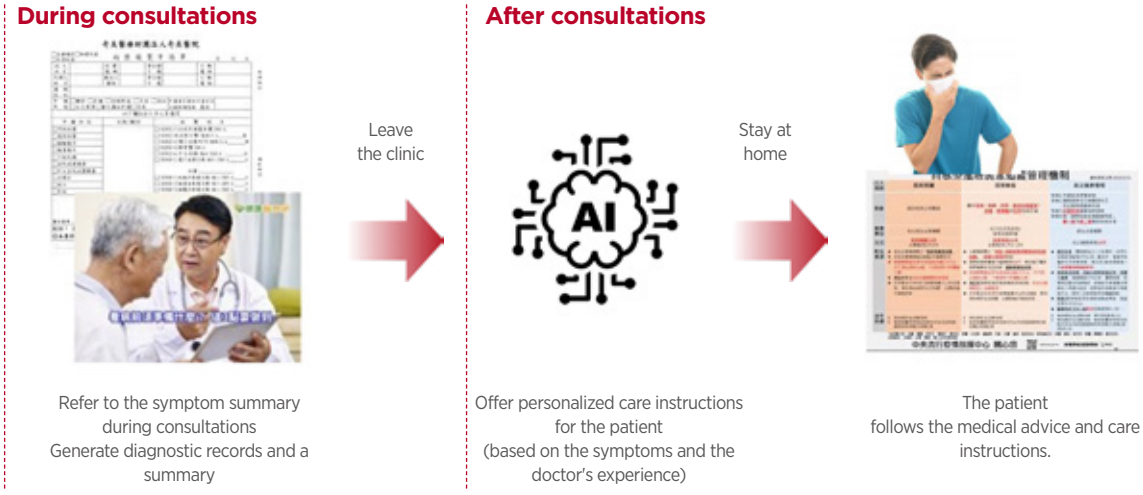


Figure 1 Flowchart of the Smart AI-powered Medical Assistant

CHALLENGES

Pain points

In medical clinics flocked with patients, doctors usually need to multitask when seeing patients because they need to build and manage medical records. This deteriorates the quality of medical care. The smart AI-powered medical assistant can summarize the dialogs between doctors and patients and the doctors' diagnoses and medical advice, relieving the documentation burden of medical personnel.

Challenges of changing the conventional mindset

Medical institutions are facilities that provide patient treatment. Due to their emphasis on ensuring patient safety, they tend to adopt a conservative approach to new technologies. To introduce the smart AI-powered medical assistant, we need to gradually change the mindset of medical institutions.

Challenges of process design and implementation

During the implementation, we need to continuously improve the assistant based on the feedback from medical institutions. If medical personnel feel that the service experience is bad, the assistant turns out to increase their workload. Therefore, how to optimize the process and mechanism and continuously make adjustments is a long-term challenge.

SOLUTION AND VALUE

01 Overall structure/solution

Apply the smart AI-powered medical assistant by integrating STT and generative AI technologies into consultations. Figure 2 shows the overall architecture of the project.

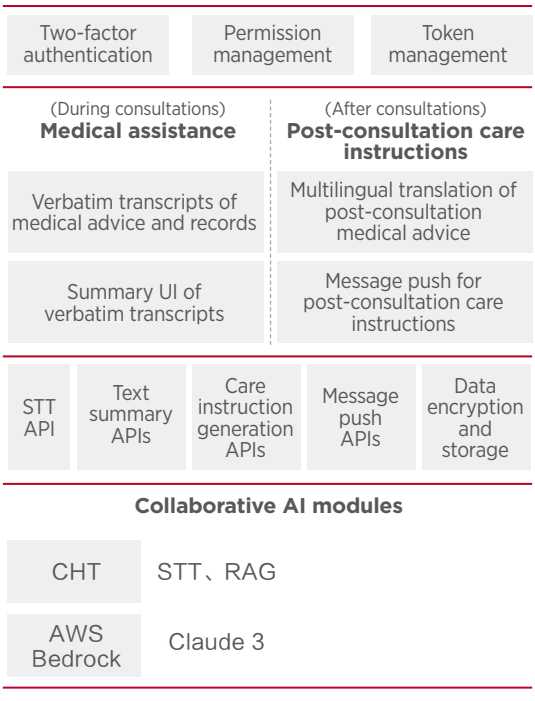


Figure 2 Overall Structure/Solution

02 Application scenarios

Summary of diagnoses and medical advice

- STT: During consultations, convert doctors' diagnoses and medical advice into a verbatim transcript using STT.
- Text summary of diagnoses and medical advice: Integrate LLMs to summarize the text transcript of diagnoses and medical advice. Provide the summary for doctors to manage medical records.

Post-consultation care instructions

- Based on the RAG framework and LLMs, provide patients with personalized care instructions, empowering the public to manage their health effectively after consultations. In addition, integrate the multilingual text translation feature to adapt to the habits of users who speak different native languages, such as new residents or foreign caregivers. This implements health equality, creating a more inclusive medical environment.

03 Key innovations

Create standard templates by converting speech to text and integrating generative AI technology. This allows medical institutions to structure medical records and efficiently analyze and summarize complex clinical records, achieving the goals of reducing data processing and analysis costs and enhancing operational efficiency.

04 Business models

Demand side

- Target audience: Major medical institutions across Taiwan.
- Customer relationship: Build partnerships by conducting a field proof of concept (POC) and jointly applying for government grant programs.

Supply side

- Key partners: Microsoft Azure and AWS.
- Key resources: Generative AI computing resources.

Value side

- Value proposition: In recent years, medical personnel have been in extremely short supply. The generative AI-based medical record summary system frees medical personnel from the complex medical paperwork, allowing them to focus more on patient treatment and care. This will further enhance patient experience and healthcare productivity.

Finance side

- Cost structure: The operation process incurs expenses and costs for system development and maintenance, cloud storage, and generative AI computing resources.
- Revenue stream: After the service is implemented, the service can be provided for medical institutions on a monthly subscription basis to obtain revenue.

05 Core values

Quantitative results

Generate verbatim transcripts of patient-doctor dialogs: Deploy the STT module to configure STT and verify the accuracy of speech recognition. Deploy the Whisper model, achieving an accuracy of 97%. This accuracy is obtained by comparing the analysis results of transcripts generated for 7 people narrating 11 text samples, each of which contains an average of 35 characters.

Qualitative impact

We collaborated with doctors from the solution partners to convert three rounds of narration to verbatim transcripts using STT and generate summaries. The accuracy of the STT-based verbatim transcripts is approximately 95%, and the summary reserves 100% of important medical information. The doctors are satisfied with the quality of the STT-based verbatim transcripts and symptom summaries.

REFLECTION

Summary

Designing a summary process for extracting and analyzing medical advice and care instructions for medical institutions is not easy, because each step of the diagnosis and treatment process involves strict regulatory requirements and specialized medical considerations. We need to take into account the characteristics of doctor-patient interaction in outpatient clinics and the actual situation of the development team. In addition, we need to keep fine-tuning the process during the implementation to effectively integrate AI LLM tools into the existing development process. We need to continuously simulate patient-doctor dialogs and convert them to text using STT, to find the optimal solution to integrate medical care with information communication service technologies in the development.

Follow-up plans

We plan to partner with hospitals according to the government's "Digital Service Innovation Grant Program" to implement the smart AI-powered medical assistant. We hope to integrate existing services with more professional medical knowledge and medication information and apply the smart AI-based assistant to more medical care fields to enhance overall satisfaction among medical personnel and the public.

5G+AI Safety Monitoring Practice in Honghe Wenyu Coal Mine in Ordos

As China vigorously promotes intelligent coal mine safety monitoring, the Honghe Wenyu Coal Mine in Ordos worked with China Mobile to deploy a 5G+AI safety monitoring system that integrates cutting-edge technologies. The coal mine has significantly improved safety monitoring capabilities, with comprehensive, round-the-clock hazard detection. Since its launch, the system has achieved an accuracy improvement of over 8%, a reduction of over 5% in missed detections, a 30% decrease in violations, and substantial savings in labor costs. This collaboration not only promotes automated coal mine safety monitoring but also sets a new industry benchmark.

Li Xiaodong
Chairman of Shandong Honghe Holding Group Co., Ltd.

SOLUTION PARTNERS

Built by Ordos Honghe Energy Technology Co., Ltd., the 5G+AI Smart Mine Project at the Wenyu Coal Mine is located in Narin Tohoi Town, Ejin Horo Banner, 50 kilometers southeast of Ordos City. The town features a superior geographical location and abundant coal resources. As a combination of the Manlailiang Coal Mine and the former Wenyu Coal Mine, the Wenyu Coal Mine covers a vast area of 9.359 square kilometers, with mining elevations of +1294m - +1135m. It has an annual production of up to 1.2 million tons.

9.359 square kilometers
vast area

+1294 m
mining elevations

120 million tons
annual production

In traditional coal mines, safety monitoring has always been a challenging task. Safety officers have to stay on duty for long periods, working with the miners at the coalface to ensure safe operations and keeping an eye on the videos in the control room to prevent accidents. To completely change that, China Mobile (Shanghai) Information and Communication Technology Co., Ltd., in collaboration with the Inner Mongolia Branch of China Mobile and its ecosystem partners, introduced an innovative 5G+AI Smart Mine solution.

The project has a total investment of over RMB 14.18 million. The Ordos Branch of China Mobile Communications Group Inner Mongolia Co., Ltd. undertook the construction of the core network and computing center, and China Mobile Shanghai Industrial Research Institute provided the industrial safety monitoring platform and key industrial software for comprehensive safety management. With private 5G networks, OnePower Smart Mine Subplatform, and AI Industrial Safety Monitoring Subplatform, the project has realized comprehensive data integration and smart scheduling management for mine production, safety, and operations.

CHALLENGES

Full coverage is not possible with costly traditional safety monitoring practice	Traditional safety monitoring practice involves establishing a CCTV monitoring system and assigning safety officers for visual inspections in control rooms. Some of the most prevalent issues include high labor costs, low efficiency of visual inspections, and the occurrence of overlooked inspections. For hard-to-reach scenarios, manual inspections are typically employed for on-site management, but safety officers only stay for a short time during this kind of inspection and effective 24-hour control is not possible.
Frequent violations and overlooked inspections due to weak safety awareness among personnel	Without safety awareness and attitudes, some employees do not take safety seriously and violations are found in scenarios such as entering and exiting mines, using cages, or wearing protective clothing. Safety officers are subjective in violation judgment because there are no consistent definitions of violations or clear violation records. As a result, many violations go unaddressed.
The system is not smart enough to forestall accidents	Traditional CCTV video surveillance only allows for live streaming and playback of recordings, making safety monitoring passive and retrospective, and it is typically used for incident review and liability determination after a safety accident occurs.

Existing monitoring systems are simple in design and limited in functions. They focus on real-time uploading and storage of images from monitoring points and fail to spot real-time occurrences of improper dressing, hazardous behaviors, and unauthorized use of equipment during work hours. In other words, they cannot proactively identify potential safety risks.

SOLUTION AND VALUE

01 Overall structure/solution

The 5G+AI safety monitoring is a comprehensive video monitoring and intelligent analysis system architecture (Figure 1) that encompasses video services and algorithm services at the SaaS layer, platform management and intelligent scheduling at the PaaS layer, and infrastructure and hardware support at the IaaS layer. By integrating various video protocols and intelligent algorithms, the system achieves real-time analysis, alerting, processing, and visualization of video data, providing a solid foundation for remote intelligent applications in business systems. The entire architecture embodies a complete process from data collection and processing to application display. It is an efficient, smart, and scalable system engineering solution.

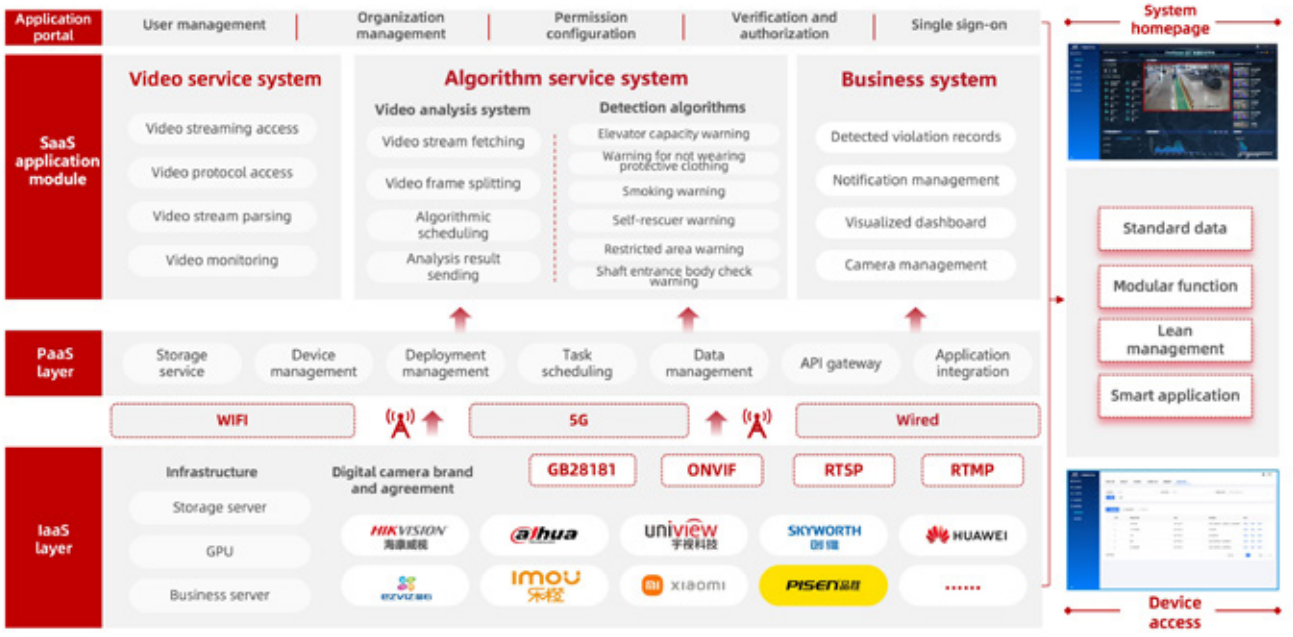


Figure 1: 5G + AI Safety Supervision System Architecture

02 Application scenarios

Visualized dashboard

The visualized dashboard (Figure 2) summarizes detected violation data and displays real-time CCTV footage, violation statistics of the day, real-time violation screenshots, violation data trends, etc.

Figure 2: Visualized Dashboard

Smart recognition

Unobtrusive monitoring and recording of behaviors in critical areas.

① Underground scenarios - unsafe behavior warning

The system will capture video data by frame from cameras in the production chambers, service chambers (Figure 3), and underground substations (Figure 4), analyze the human body entering the chambers and surrounding environment, and issue alarms on the site and on the terminal computer when dangerous dressing or behaviors, or environmental anomalies are detected. In this way, safety hazards are prevented.

② Main transportation scenarios

Main coal transportation belt - abnormal temperature monitoring (Figure 5), belt misalignment (Figure 6), belt perimeter monitoring (Figure 7), foreign object detection on the belt (Figure 8), and empty belt warnings (Figure 9)

Figure 3: Refuge Chambers - Dressing Compliance Check

Figure 4: Underground Substations - High-temperature Alarm for Equipment

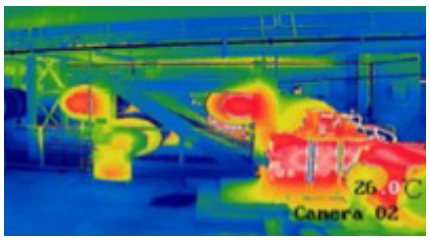


Figure 5: Temperature Monitoring

Figure 6: Belt Misalignment Detection

Figure 7: Belt Perimeter Monitoring

Figure 8: Large Foreign Object Detection

Figure 9: Empty Belt Detection

Due to factors like harsh operating environments, heavy transportation loads, and long transportation distances, belt conveyor monitoring needs to consider issues such as belt temperature, misalignment, perimeter intrusion, and foreign objects during use. The smart analysis system installs cameras above the belt based on the belt length, sets double warning lines on both sides of the belt in the video images, and extracts features from the video images with the aid of AI technologies. When the threshold is exceeded, it rapidly and accurately sends alarm messages to relevant personnel for necessary actions.

③ Shaft entrance and exit - body check warning

Safety officers will check individuals exiting the shaft with metal detectors (Figure 10) and alerts are issued on the site and on the terminal computer if checks are not conducted or conducted without metal detectors.

④ Central substations - unauthorized entrance warning

The system will capture video data by frame from cameras in the central substation (Figure 11), analyze individuals entering the central substation, and issue alerts on the site and on the terminal computer in case of unauthorized entrance to prevent safety risks.

⑤ Individual behavior - off-duty/sleeping-on-duty warning

The system will collect video data and issue alerts on the site and on the terminal computer about off-duty or sleeping-on-duty violations (Figure 12) within camera coverage.

Figure 10: Body Check Warning

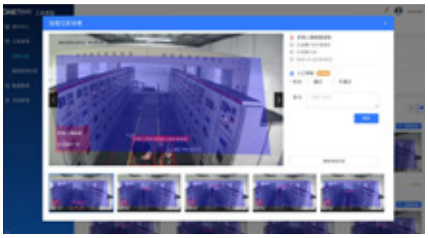


Figure 11: Unauthorized Entrance Warning

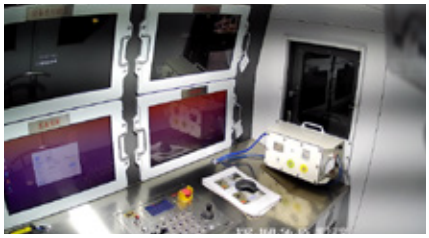


Figure 12: Off-duty/Sleeping-on-duty Warning

Alert management

The platform supports automatic alert generation based on AI-powered video content analysis results (Figure 13). Alerts can be sent via SMS or email to relevant law enforcement personnel, and optical-acoustic alarm signals can be issued as the platform is linked with alarm systems and third-party systems through APIs.

Unobtrusive recording

Automatic snapshot recording is performed when alerts occur for unobtrusive monitoring and recording without disturbance to on-site work.

Easily accessible records

Detected violations are recorded and displayed with clear timestamps for handling during the event and tracing after the event.

Prevention support

Trends in detected violations are analyzed based on data statistics and records, providing data support for proactive measures.

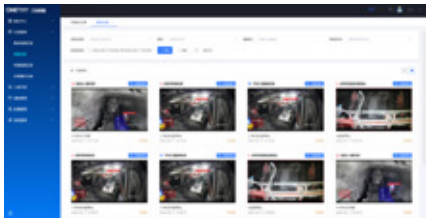


Figure 13: Warning Management

03 Key innovations

Multimodal large model integration: The project has achieved comprehensive safety supervision of mines by integrating multimodal large models. This system can analyze video data in real time and enable smart dialog and decision support through natural language processing. This integration approach enhances the system's level of intelligence and significantly reduces overlooked inspections and false alarms as compared to traditional monitoring systems.

Private 5G network and edge computing

The project introduces private 5G network and edge computing technology so that data collection and processing can be done locally in the mine. This greatly reduces data transmission latency and ensures real-time response to critical safety information. Thanks to the innovation, mine supervision relies no longer on centralized computing but on distributed intelligent processing, significantly enhancing system reliability and real-time capabilities.

Custom feature search and fast tracing

The project includes the function of custom feature search that allows for quick violation tracing within just 5 seconds when safety hazards occur. This feature makes the system easy to use and significantly improves safety management efficiency.

04 Business models

Service subscription and maintenance contracts: By providing regular maintenance and upgrade services for the smart safety monitoring system, the project adopts a one-time subscription + maintenance business model to ensure continuous and stable service delivery, thus upgrading the capabilities of large models. Furthermore, by signing maintenance contracts, customers can enjoy 24/7 technical support, reducing operational risks.

Value-added services

The project offers value-added services including data analysis report generation, trend prediction, etc. These services help businesses optimize safety management strategies and provide data-driven decision support.

Ecosystem collaboration

The project has established close ecosystem collaboration with China Mobile and other technology partners. With joint promotion and technical support from partners, the project expands its market coverage and influence.

05 Core values

Enhanced safety and cost savings

After the launch, the monitoring system has increased accuracy by 8%, reduced overlooked inspection rate by 5% and occurrence rate of violations by 30%, and halved labor costs. The significant improvement in safety and reduction in operational costs have positioned this project as a new benchmark in mine safety supervision.

Automation

The project has pushed traditional mining areas toward intelligent and automated operation, with reduced manual intervention and enhanced production continuity and stability. It has laid a solid foundation for building smart mines.

Industry demonstration

With results validated at the Wenyu Coal Mine in Ordos, the project has also provided a replicable experience for other mining areas and industrial sectors. It is a good model across a broad range of domains. In the future, this model will be seen in more industries, forming new industry standards.

REFLECTION

A total of 94 video feeds were integrated into the project, with 49 above-ground feeds covering locations such as the dispatch rooms, power distribution rooms, substations, mine entrances, and offices. Above-ground monitoring focuses on personnel surveillance, off-duty monitoring, and high-temperature hazards to enhance safety monitoring efficiency. The cameras are mostly wired.

There are 45 underground video feeds covering various locations such as coalfaces (203/204/205), central substations, belt lanes, central control rooms, conveyor belt engines, and coal spillage points. Underground monitoring focuses on belt operations, abnormal blockages, dressing, and unauthorized entrance in hazardous areas. It helps address issues of inadequate manual inspections and weak supervision. 5G base stations have been installed at underground coalfaces, and all underground videos are connected to the above-ground data center for analysis and processing through a dedicated network.

In the past, manual inspections were conducted around the mines, but risks were easily overlooked and issues could not be identified in time. The project has realized continuous 24/7 recognition based on the AI large model for safety analysis, with an expected cost reduction of over 30%. The custom feature search, in particular, allows for quick violation tracing within just 5 seconds. It has lowered the difficulty of use and improved user experience.

Plans involve exploring industrial large model benchmarks in key industries like petrochemicals, power, mining, metallurgy, and equipment manufacturing together with central enterprises and backbone companies to create a "model industrial large model" with demonstration effects. Specific measures include

(1) strengthening computing network capabilities and establishing an industrial computing network resource pool that matches edge computing power and integrates heterogeneous computing power;

(2) enhancing data capabilities by building industrial data governance systems, data sample libraries, and data security systems;

(3) enhancing platform capabilities by building platforms for industrial large model training and inference, and intelligent application generation;

(4) improving model capabilities by developing foundational capabilities in industrial visual/image analysis, text generation, knowledge QA, digital twin, and tackling challenges in large and small model fusion;

(5) enhancing ecosystem capabilities through industrial innovation bases. This means introducing partners via industry chain investments, collaborative innovation, and joint research and development to build a large model application ecosystem.

Intelligent Telecommunication Fraud Prevention

In 2022, fraud in Taiwan resulted in financial losses totaling NT$7.3 billion. Since 2023, Chunghwa Telecom has taken the lead in blocking fraud at the source. We drove innovations to develop all-around intelligent fraud prevention services for telecommunications networks, including international fraud call filtering, warning, blocking, and SMS fraud blocking. Over a year after these measures were implemented, the number of suspected fraud calls dropped by 80%. To cope with various fraud techniques, we are continuously refining the AI-powered fraud prevention technology and expanding network communication capabilities to identify and block frauds more effectively.

Ho Chung-Ding
Managing Director of Digital Innovation Laboratory, Telecommunication Laboratories, Chunghwa Telecom

SOLUTION PARTNERS

Chunghwa Telecom recognized that fraud is one of the major concerns of the public. Our analysis suggests that voice calls and SMS messages are the major fraud channels. We have established a public and private partnership with the government to block fraud at the source. In addition, we have employed AI to build a comprehensive fraud identification and blocking mechanism for the telecommunication network. The detailed case description is as follows:

International fraud call blocking

We have independently developed an international call filtering mechanism to detect international calls from spoofed caller IDs beginning with "+886" in real time. Abnormal international calls, such as spoofed fixed-line or international roaming calls, are blocked.

International call warning

We have developed the industry-first international call voice warning platform for international calls, including VoLTE and VoWiFi calls. The platform provides a 7-second voice warning to remind customers to be aware of cross-border calls and prevent fraud.

AI-powered SMS fraud blocking

We leverage AI to extract keywords, analyze fraud intents, compare similarities, and cache SMS content, to detect malicious SMS messages and block fraudulent SMS messages.

CHALLENGES

Telecommunication fraud is a critical concern that causes huge losses to individuals and society. It comes in many forms:

- Caller ID spoofing is prevalent in international calls and conventional fraud prevention methods such as post-event notification and blocklisting are ineffective. AI is needed to detect abnormal behavior and update fraud prevention technologies.
- The content of SMS messages is difficult to distinguish: SMS fraud usually impersonates the government, financial institutions, and telecommunications companies. As a result, the SMS messages sent by these institutions lost credibility, causing operational difficulties to the government and enterprises and growing property loss to the people. The current challenge is to study the SMS fraud methods and use AI to

detect whether SMS messages contain malicious URLs or fraudulent intents.

- Real-time SMS detection is difficult: For example, when a malicious URL is detected in an SMS message, most existing systems only add the sender to the blocklist. This practice cannot detect new malicious URLs. It can only add URLs to the blocklist after victims are found. If AI can be used to perform real-time detection, the number of victims can be greatly reduced.
- Cross-sector partnership in fraud prevention is difficult: Fraud activities involve multiple fields. Telecommunication operators need to partner with financial institutions, law enforcement agencies, and industries. However, cross-sector information sharing and collaboration are faced with barriers. Telecommunication operators need to expose network communication capabilities and provide open APIs to expand the fraud prevention ecosystem with the government, partners, and industries.

SOLUTION AND VALUE

International Call Blocking and Warning Service Platform: 01

01 Overall structure/solution

The system can block and warn customers of international calls from fixed lines and mobile communication networks. It imports all international calls into the international call filtering platform and blocks abnormal international calls, such as spoofed domestic-originated fixed-line or international roaming calls. In the future, we will develop an AI-powered automatic call answering system. This system can import fraudulent calls into the "generative AI-powered automatic answering" platform. In addition, it can analyze the voice to obtain fraud patterns and characteristics through machine learning. This will increase the cost of fraud and assist the government in collecting fraud intelligence.

02 Application scenarios

Block abnormal international calls

Analyze abnormal international calls and block international calls from spoofed caller IDs beginning with "+886". Then, import these calls into the generative AI-powered automatic answering system. This increases the cost of fraud and helps the government add the reported caller IDs to the blocklist.

Trigger a 7-second voice warning for an international call

We have developed the industry-first 7-second voice warnings for international calls, including 3G, VoLTE, and VoWiFi calls. The platform reminds customers to be aware of cross-border calls and prevent fraud.

03 Key innovations

Chunghwa Telecom leverages its four advantages, including customer partners, technology platforms, basic networks, and high-quality talents to lead in smart fraud prevention services, receiving industry and international recognition. It is Taiwan's leading telecommunications operator that provides comprehensive smart fraud prevention services for mobile, fixed-line, and Internet calls.

04 Business models

Chunghwa Telecom partnered with the government to prevent fraud. It helped the government prevent a potential loss of NT$8.9 billion in 2023, which exceeds the amount lost to fraud for the first time. Chunghwa Telecom is dedicated to fulfilling social responsibilities by creating a secure telecommunications environment with low fraud risks.

8.9 bn
It helped the government prevent a potential loss

05 Core values

Chunghwa Telecom is the first to launch a service for blocking and filtering international calls from spoofed numbers beginning with "+886". It has blocked all abnormal calls and sharply reduced international calls from spoofed numbers beginning with "+886" by 97%.

100%
blocked abnormal calls

AI-powered SMS Fraud Prevention: 02

01 Overall structure/solution

Chunghwa Telecom's AI-powered SMS fraud prevention technology aims to detect multiple fraud tactics in real time and share APIs. It applies to large numbers of commercial SMS messages. As shown in Figure 1, it provides the following major features:

- Multi-functional AI-powered detection: The feature uses multiple models to detect the content of SMS messages. The AI-powered malicious URL detection model checks the URLs in SMS messages for fraud risk. The AI-powered semantic detection model analyzes the text content to detect fraudulent intents. The AI-powered keyword extraction model analyzes the text content to extract keywords and identify the SMS message type.
- AI-powered detection and caching mechanism: To achieve real-time detection, the feature caches multiple analysis results determined by AI. The AI-based semantic similarity technology can find the same or semantically similar SMS messages.
- AI-powered fraud prevention APIs: Open APIs are exposed to multiple SMS services.

02 Application scenarios

When a large number of commercial SMS messages are sent, the AI-powered SMS fraud prevention platform performs real-time detection. If a suspected fraudulent SMS message is found, the platform reports the message to the competent authority for investigation and announces the risk to remind the public.

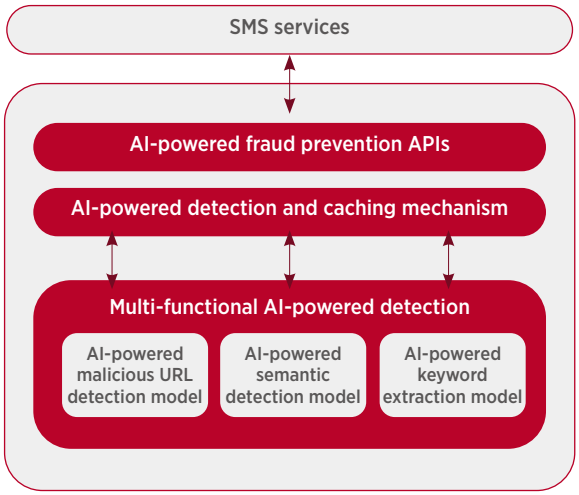


Figure 1 AI-powered SMS fraud prevention platform

03 Key innovations

Based on AI, Chunghwa Telecom has developed SMS semantic detection, a caching mechanism, and malicious URL detection to supplement SMS fraud prevention measures.

04 Business models

Open APIs are exposed to various application services to obtain revenues.

05 Core values

Chunghwa Telecom's AI-powered SMS fraud prevention technology aims to detect multiple fraud tactics in real time and share APIs. It creates the following major values:

- Fast and real-time processing: The AI-powereddetection and caching mechanism shortens the check time of each SMS message to 0.0002 seconds.
- Highly accurate AI-powered detection: AI is used for semantic analysis, malicious URL analysis, and other analysis of SMS messages, with an accuracy of over 90%.
- Shared APIs: Open APIs are exposed to multiple SMS services.

REFLECTION

Chunghwa Telecom took the lead in launching the "fixed-line international call voice warning service" on July 17, 2023. The service has effectively blocked more than 90% of fixed-line calls from spoofed local numbers that start with "+886 9". Under the guidance of relevant departments, Chunghwa Telecom once again took the lead in launching the international call voice warning service for mobile calls from spoofed numbers beginning with "+886 9". Then, it expanded the service to all international calls from spoofed local numbers and provided voice warning services for all of these calls. This prevents fraudulent calls all around. As a leading telecommunications brand in Taiwan, Chunghwa Telecom continuously drives innovations, introduces AI-powered fraud prevention technologies, and strengthens the promotion of fraud prevention. We provide complete support for the government and relevant institutions to prevent fraud and protect the public's financial security.

After AI is employed in SMS fraud prevention, AI-powered detection with high accuracy and efficiency has been available. These technologies can also be used by various application services by calling open APIs. In the future, we will speed up the promotion and application of these technologies. In addition, embedding malicious URLs in SMS messages is one of the major SMS fraud tactics. We will continuously study massive information on malicious URL tactics to better understand the occurrence frequency and methods of phishing, and seek better measures to fight phishing.

We plan to build GSMA Open Gateway APIs to expand the fraud prevention ecosystem and continuously improve fraud prevention efficiency.

Visual Intelligence Application Platform Built on the Yuanjing Port Large Model

Amid the rapid development of global trade and logistics, ports are facing increasingly complex management challenges. Artificial intelligence, especially large models, can bring new productivity and solutions. China Unicom, as a state-owned enterprise, is committed to developing core large model capabilities to address specific industry needs, such as smart port solutions. By leveraging an innovative cloud-edge-end collaborative architecture, it has created a visual intelligence application platform centered around the Yuanjing Port Large Model. This platform enhances port safety intelligence, ensuring operations are efficient, secure, and environmentally sustainable.

Wang Xiaosi
Chief Expert of Industrial AI of China Unicom AI Innovation Center

SOLUTION PARTNERS

In the development of large models and platform capabilities, China Unicom has strengthened its core technology research for large models, establishing a Yuanjing "1+1+M" large model system that includes one base model, one MaaS platform, and multiple industry models. The Yuanjing large model system features multi-parameter versions of language models to achieve industry-leading performance at the same parameter level. Additionally, it offers a large multimodal model capable of processing voice, images, text, and video, designed for rapid customization across multiple scenarios.

The visual intelligence application platform based on the Yuanjing Port large model, developed by China Unicom for the port industry, can be deeply embedded in five key operational areas of port safety management. It provides unique features such as pedestrian early warning for forklift operations, driver fatigue monitoring, and personnel uniform compliance identification, enhancing the safety and compliance of port operations, standardizing and smartening port area supervision, and significantly reducing operational oversight and safety accident risks for clients in sea ports and inland ports. China Unicom AI Innovation Center has developed the Yuanjing Port Large Model that integrates multiple tasks and modes for efficient information integration and consolidation, reducing the costs of model development and application deployment. Additionally, a cooperative mechanism for large and small models has been established to effectively tackle the challenges of small models' limited generalization ability and large models' low accuracy. The visual intelligence application platform based on the Yuanjing Port Large Model has been deployed in several ports, including Nanjing Port, enhancing daily operations such as port safety management and improving the capabilities of AI-assisted automated operations and safety management.

1+1+M

base model MaaS platform industry models

CHALLENGES

The port industry faces the following challenges during its development:

- The port environment is complicated, with various production elements such as ships, vehicles, people, and large machinery interwoven. Traditional video monitoring methods rely on human oversight, making it impossible for staff to manage the vast amounts of real-time video effectively. This often leads to security management vulnerabilities.
- Port operations are labor-intensive and carry many safety risks, making comprehensive real-time supervision difficult. There is significant safety management risk, and even having dedicated personnel monitoring core operational areas may not provide timely and effective warnings, resulting in high supervision and safety costs.
- The inability to effectively warn and reduce operational safety risks means that, after various levels of safety incidents occur, there is often a delay in notifying the appropriate personnel to execute emergency plans.
- The vast amounts of video surveillance data generated in port operations are only used for recording and post-event investigation, highlighting the urgent need for structured video data analysis to improve operational standards and efficiency.

These issues not only affect port operational efficiency but also threaten the safety of personnel.

SOLUTION AND VALUE

01 Overall structure/solution

The Visual Intelligence Application Platform based on the Yuanjing Port Large Model is an integrated smart port safety production solution that combines multiple tasks and modalities, as shown in Figure 1. The platform utilizes the Yuanjing general visual large model and large multimodal model technology, incorporating rich scene-based algorithms and an innovative collaborative mechanism for large and small models. It is deeply integrated with port operational safety supervision scenarios, providing smart perception, real-time monitoring, surveillance warnings, and analysis capabilities for port operations. This promotes safety compliance in port operations, standardizes and smartens port area supervision, and significantly enhances the efficiency and intelligence of port safety regulation.

The platform offers management functions such as model management, video convergence, visual analysis, and monitoring alerts. It is highly versatile and easy to operate, allowing even non-professionals to quickly get started and rapidly build AI analysis applications. With a rich array of scene-based algorithms, the algorithm repository provides over 100 scene-specific AI models covering production, operations, and other scenarios, with excellent and reliable model performance. Using an innovative collaborative mechanism for large and small models, the platform combines large visual and multimodal models with small CV models to leverage the high accuracy, strong generalization, and multimodal integration advantages of large models. This enhances overall analytical accuracy and improves the efficiency of rapidly launching new scene algorithms. The platform allows deployment on diverse heterogeneous gateways and can connect with cameras, edge servers, edge boxes, and other edge devices, allowing for the selection of different component specifications based on the production environment to meet various scale requirements. It features a flexible cloud-edge-end collaborative architecture and supports multiple communication protocols, enabling N:N flexible binding between devices and algorithms. The platform is adaptable to third-party algorithms and services, offering high compatibility and supporting multiple applications from a single point.

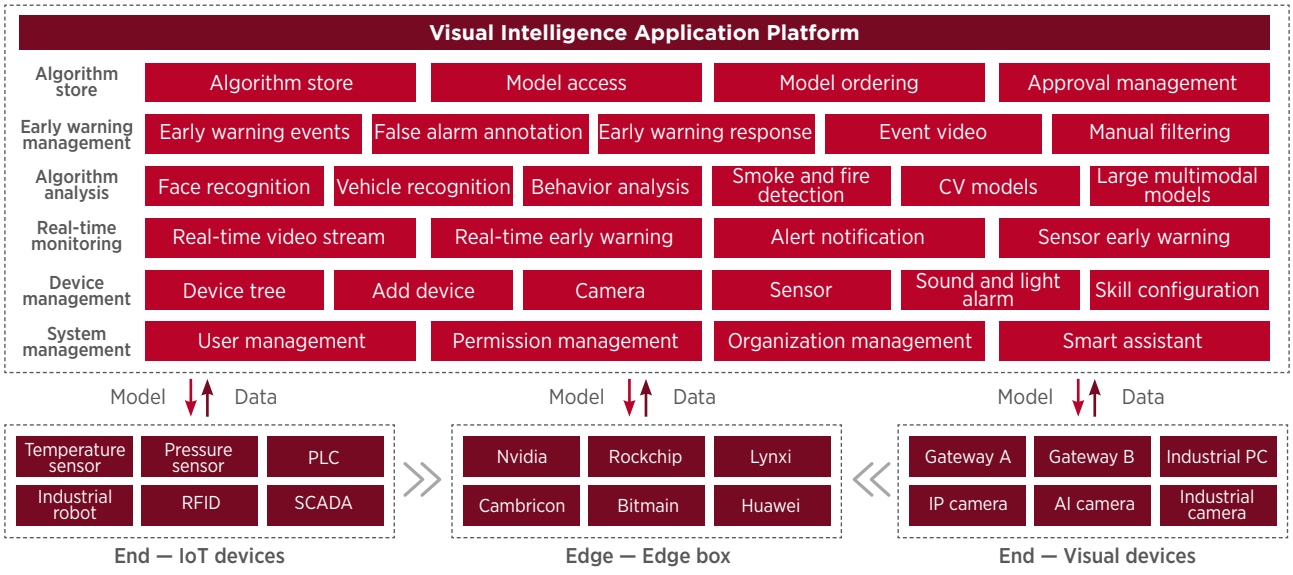


Figure 1: Architecture of the Visual Intelligence Application Platform Built on the Yuanjing Port Large Model

02 Application scenarios

The Yuanjing Port Large Model is deeply embedded in six major aspects of safety operation management at Nanjing Port, as shown in Figure 2. It addresses various operational scenarios such as container operations, bulk cargo operations, tugboat operations, mobile machinery operations, port traffic, personnel safety, and logistics management, providing unique features such as under-gate personnel warning for quay cranes, driver fatigue monitoring, and personnel uniform compliance identification.

- **Container operation monitoring:** The system can automatically recognize container numbers, check if container stacking is compliant, and monitor the loading and unloading processes to ensure operational accuracy and safety.
- **Tugboat operation monitoring:** The system monitors the tugboat operational status, including its operational trajectory and efficiency, as well as interactions with other vessels, to ensure operational safety in the port waters.
- **Mobile machinery operation monitoring:** The system monitors the operational status of mobile machinery within the port, such as forklifts and cranes, to prevent mechanical failures and operational errors, thereby improving operational efficiency.
- **Port traffic monitoring:** The system provides real-time monitoring of traffic flow and vehicle operating status within the port to optimize traffic flow, prevent traffic accidents, and ensure smooth port traffic.
- **Personnel safety monitoring:** Through facial recognition and behavior analysis technology, the system monitors the safety equipment worn by port staff and their behavioral compliance to promptly identify and warn of potential safety hazards.

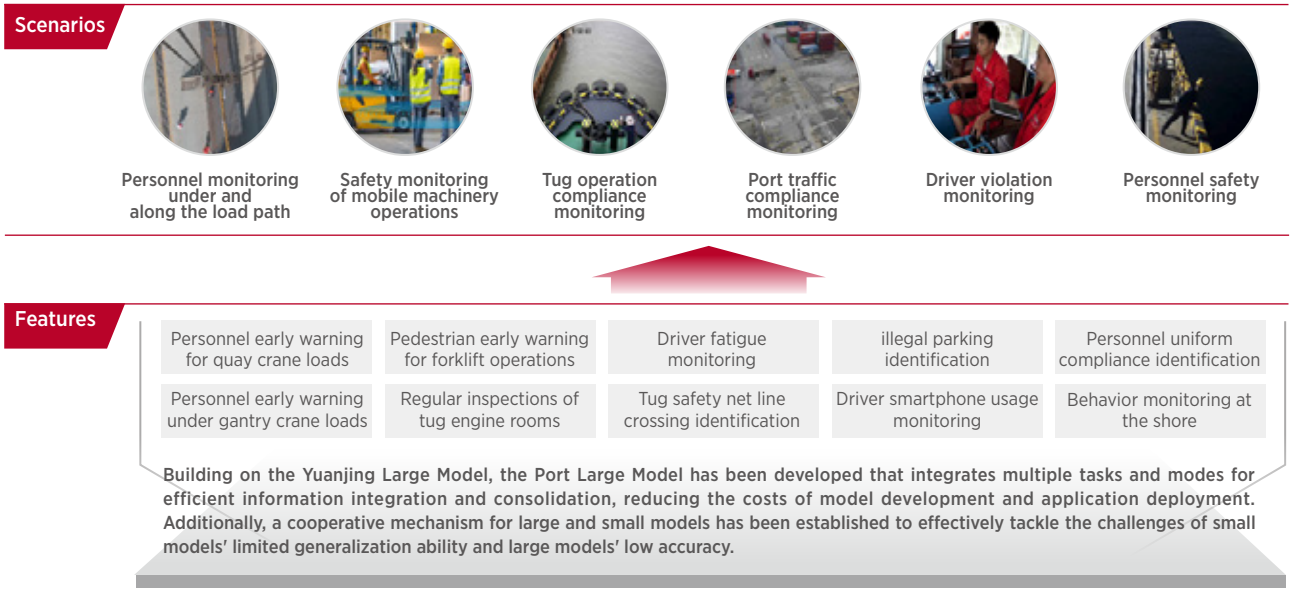


Figure 2: Application Scenarios and Features of the Yuanjing Port Large Model

03 Key innovations

Built on the Yuanjing large model, the Yuanjing Port Large Model uses a self-developed multi-task and multi-modal port safety production model to integrate efficient information and lower the costs of model development and application. The innovative collaboration mechanism for large and small models allows for precise initial screening and in-depth re-evaluation, significantly enhancing identification accuracy. By optimizing system architecture with a multitask large model, operational and maintenance costs are reduced. The open-set detection technology of the large model improves the efficiency of launching new scene algorithms. Additionally, the interactive capabilities of the large multimodal model enable flexible configuration of rules, enhancing recognition accuracy and effectively reducing false positives.

In terms of technical innovation, the Yuanjing Port Large Model has reduced false positive rates by 30% while maintaining recall rates. The large model offers second-level re-evaluation with an overall recognition accuracy of over 85%. Furthermore, model development and operational efficiency have improved by 30%, achieving zero-sample, minute-level generation capabilities, with iterative model updates completed within hours.

The Yuanjing Port Large Model provides users with innovative and flexible customization, leveraging interactive configuration of business rules based on the large model to assist in event verification and enhance platform usability.

04 Business models

The product model of the Visual Intelligence Application Platform built on the Yuanjing Port Large Model integrates overall port/industrial solutions. China Unicom AI Innovation Center provides standard platform product components, which are then assembled by Unicom's provincial subsidiaries or solution integrators into a higher-level SaaS platform. This, combined with hardware devices and networks, forms a comprehensive solution or standardized integrated product for sale.

Figure 3 illustrates the business model of the product. Currently, the target users of the product mainly include port operators and logistics service providers in sea ports and inland ports, as well as safety regulatory agencies. There are three main ways to sell the product: platform software, AI integrated machines, and algorithm models. The revenue strategy includes selling platform software licenses, model authorizations, deployment and

maintenance services, and model customization services. These are incorporated as specific charges within the overall solution or through bundled solution sales and revenue sharing with integrators.

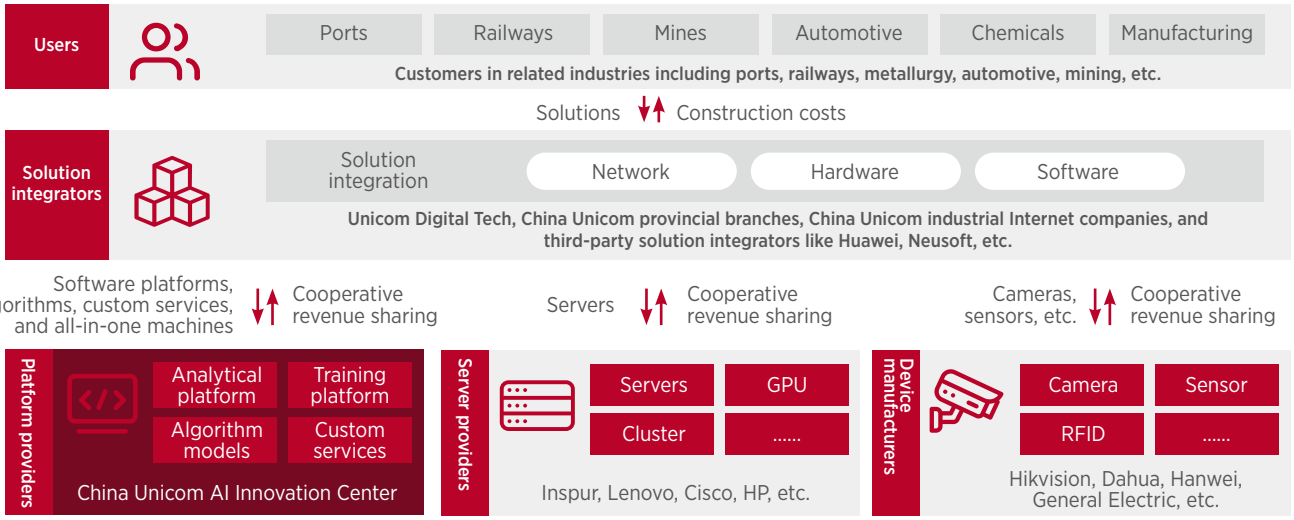


Figure 3: Business Models of the Yuanjing Port Large Model

05 Core values

The market revenue from the Yuanjing Port Large Model has exceeded 10 million yuan, and a mature business model has been established. The plan is to achieve regional scale replication by 2026, expanding business into general industrial scenarios, with a target annual revenue of over 20 million yuan and the implementation of more than 30 customer projects each year.

Application benefits: The visual intelligence application platform based on the Yuanjing Port Large Model has been deployed in several ports of different provinces and cities, enhancing daily operations such as port safety management and improving the capabilities of AI-assisted automated operations and safety management.

For example, the Nanjing Port Safety Operation Supervision Platform employs a cloud-edge collaborative architecture, covering 17 port business scenarios, including container terminal operations, water transport, and port cargo handling. The comprehensive efficiency of terminal operations has improved by nearly 20%, saving approximately 4 million yuan in regulatory labor costs annually and avoiding safety incident costs of nearly 10 million yuan.

Social benefits: The Yuanjing Port Large Model excels in improving port operational efficiency, reducing operational costs, enhancing safety in port operations, promoting industrial upgrades, and increasing international competitiveness. It drives the intelligent transformation of port businesses and sets a new benchmark for innovation and efficiency in the entire port industry.

REFLECTION

Takeaways

Technological innovation as the core driver

The self-developed multi-task and multi-modal port large model provides strong support for the intelligent transformation of ports.

Collaborative mechanism to enhance efficiency

The application of the collaborative mechanism for large and small models has significantly improved identification accuracy, saved computing resources, and reduced operational and maintenance costs.

Flexible customization to meet diverse needs

The flexible customization features of the Yuanjing Port Large Model allow configurations based on the actual needs of different ports, enhancing the availability and adaptability of the platform.

Follow-up plans

Furthering cooperation with Nanjing Port

We will continue to deepen our collaboration with Nanjing Port, further exploring intelligent needs and providing more customized solutions to assist Nanjing Port in achieving a higher level of intelligent transformation.

Expanding more port application scenarios

Building on the successful replication of intelligent technologies at Nanjing Port, we will expand into more port application scenarios to provide comprehensive intelligent solutions for ports.

Targeting international markets

We aim to target international markets, strengthening cooperation and exchanges with foreign ports to promote smart port technologies globally. By participating in international smart port projects, we will enhance the international influence of China's smart port technologies and provide efficient, safe, and green intelligent solutions for ports worldwide.

GSMA

1 Angel Lane
London
EC4R 3AB
United Kingdom

Tel:+44 (0)20 7356 0600
Fax:+44 (0)20 7356 0601