Vol. 2

Use Cases of Public AI Service

March 2026

NIA NATIONAL INFORMATION SOCIETY AGENCY

TABLE OF CONTENTS

User Convenience

[Improving the Way of Working]

Administration

OVERVIEW

Objective and Approach

The report aimed to address a total of 110 major AI service demonstration cases promoted by the Korean government since 2022 to help each administrative and public institution as a reference case when planning AI tasks and developing services.

The three strategic goals of public AI are **'Solving Social Problems', 'Innovating Public Service', 'Improving the Way of Working'**, and presenting major technology elements, development contents, performance, and expected effects for AI utilization in each case.

Use Cases

The project represents projects discovered, supported, or planned through the public sector Hyper-scale AI service development project from 2022 to March 2025, serving as representative examples related to Hyper-scale AI among all projects.

Between 2022 and March 2025, a total of 110 public AI service demonstration projects were carried out across sectors below.

- **Hyper-scale AI Service:** Deploying Hyper-scale AI models to enhance government services.
- **National Dream Project:** Executing a set of high-impact initiatives designed to address major public and corporate needs.
- **GovTech Startup Support:** Fostering AI-driven GovTech solutions that strengthen public infrastructure.
- **Digital Government Infrastructure:** Building a data-driven administrative foundation for intelligent service implementation.
- **Ministry-Collaboration AI Diffusion:** Promoting AI adoption across ministries by sharing models, data, for coordinated government innovation.
- **Flagship Project:** Integrating AI into core public services and launch high-profile flagship projects to drive nationwide digital transformation.
- **Digital Government Collaboration:** Fostering cross-sector collaboration between government, private sector, and local governments.

Of these, 54 projects were implemented under the Social Problem Resolution category, focusing primarily on social care, low birth rate, welfare benefits, crime prevention, economic activities, healthcare, disaster response, traffic safety, and environment. A further 37 projects were conducted in the Public Service Innovation category, encompassing areas such as information provision, counseling support and user convenience. In addition, 19 projects were carried out in the Work Method Innovation category, targeting general administration like R&D, documentation, customs clearance, business processing etc.

Overall, these demonstration projects reflect the government’s efforts to expand the adoption of AI technologies across diverse public sectors to enhance citizen convenience and improve the quality of public services.

USE CASES OF PUBLIC AI SERVICE

Innovating Public Service

1

Information Provision (Inquiries, Search)

SOCIAL ISSUE 01

Automated Public Policy Application Service for Citizens

KEY TECHNOLOGICAL FEATURES

- **Data Linkage and Authentication System:** Linking with public data and identity verification systems to support automated applications
- **Customized Policy Guidance and Notification Features:** Providing users with real-time access to eligible policies and notification capabilities

OVERVIEW

Currently, citizens are required to search for relevant information across multiple government agencies and navigate complex application procedures in order to access public policies. As a result, many eligible individuals fail to receive the support benefits available to them.

This service leverages AI to analyze individual circumstances and automatically recommend relevant public policies tailored to each citizen. In addition, it streamlines and simplifies the application process, thereby improving accessibility to government support and enhancing the overall efficiency of policy delivery.

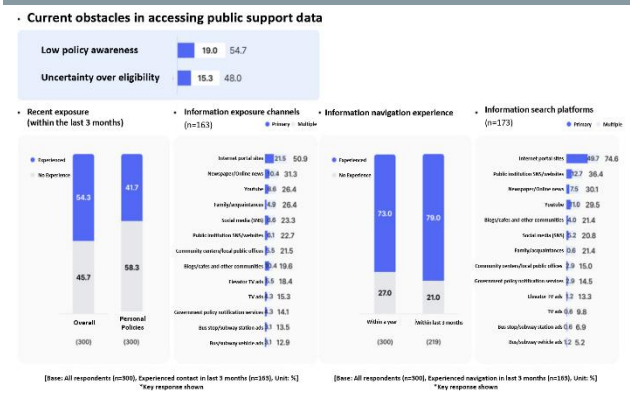
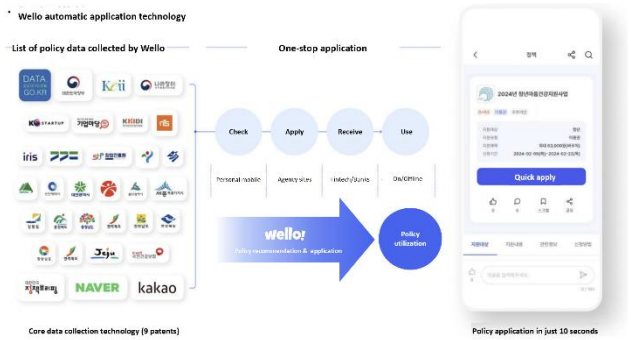
Table 1. Project Detail

Project Category: GovTech Startup Support (2024)
Implementing Agency/Partner Organization: Wello Co., Ltd.
Project Description The system utilizes AI to analyze individual profiles—including income, age, residential area, and family composition—to recommend optimized public policies. This ensures users do not miss out on benefits for which they are eligible. By integrating with public data portals and identity verification systems, the platform automatically identifies eligible policies and, with user consent, facilitates the application process. This eliminates complex manual documentation and automates the entire journey, significantly enhancing citizen convenience. Furthermore, the AI monitors newly introduced policies or updates to existing support programs in real-time. It provides personalized notifications and advance alerts for approaching deadlines, ensuring citizens can apply within the required timeframe.

EXPECTED OUTCOMES

This service enables citizens to more easily apply for public policies they are eligible to receive, while reducing complex administrative procedures to enhance the government's administrative efficiency.

Figure 1.1. Before and after implementation (AS-IS vs TO-BE)

Before (AS-IS)	After (TO-BE)
	
• Limited access caused by information silos and inefficient application processes	• Enhanced accessibility through intelligent search and one-stop application

SOCIAL ISSUE 02

Time-Slot News Summarization and Daily Reporting Service

KEY TECHNOLOGICAL FEATURES

- **LLM Training Dataset:** Building approximately 1,000 training data points using generative models
- **Fine-tuning:** Developing a model capable of generating issue keywords through fine-tuning
- **NewsDigest:** Developing a module that analyzes major sector news daily/hourly, providing daily/hourly NewsDigest results via API

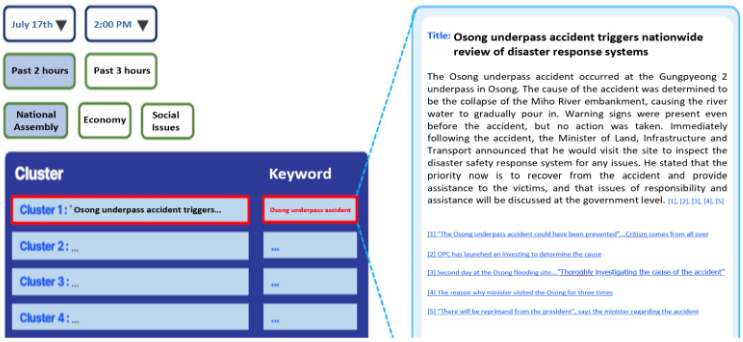
OVERVIEW

To help viewers quickly navigate the news landscape, there's a growing demand for tools that extract core keywords and organize information by timeline. The National Assembly Library's new pilot service meets this by transforming vast news data into concise reports. By categorizing news by topic and time, it provides an efficient briefing experience for rapid decision-making.

Table 2. Project Detail

Project Category: Hyper-scale AI service (2023)
Implementing Agency/Partner Organization: National Assembly Library
Project Description To enhance news accessibility, the project developed a specialized generative model that distills complex news into keywords of four words or less. Based on 10,000 training data points, these keywords allow users to anticipate event details at a glance. By connecting to the NewsDigest API, the service categorizes and summarizes key news items by time and sector, ultimately offering automated daily reporting for selected topics.

Figure 2.1. NewsDigest demo service screen



EXPECTED OUTCOMES

The service met keyword extraction requirements and proved effective in sector-specific news categorization. It concludes that future integration into actual services is highly feasible, with significant potential for enhancing news briefing efficiency. Continuous data optimization will be pursued to handle diversifying news content.

SOCIAL ISSUE 03

Q&A Chatbot for Library Archives and Services

KEY TECHNOLOGICAL FEATURES

- **LLM Training Dataset:** Collecting and preprocessing training data by extracting and converting Q&A sets using the PDF Annotation Tool
- **Retrieval Engine Implementation:** Completed implementation of a search engine using documents from the National Knowledge Resource Sharing Service and frequently asked questions from websites as QA data
- Integration of the Retrieval Engine with the MaumGPT model (13B parameters)

OVERVIEW

The diversification of library services and expansion of data have made it difficult for users to obtain desired information immediately, and people have faced limitations in addressing this. Therefore, this project aims to utilize generative AI to learn library data and provides a service that allows users to easily find the information they seek.

Table 3. Project Detail

Project Category: Hyper-scale AI service (2023)
Implementing Agency/Partner Organization: The National Library of Korea
Project Description The project utilized a PDF Annotation Tool to extract Q&A sets, followed by data preprocessing and a rigorous review of labeled data for model training. This training incorporated 286 items from the National Library of Korea, including FAQs, user guides, procedure manuals, and promotional materials. To optimize performance, the system trained a retrieval engine on MaumGPT (13B parameters) and implemented two conventional search engines for comparative analysis. Ultimately, the service developed a GPT-based chatbot designed to provide precise answers by leveraging website FAQs and documents from the National Knowledge Resource Sharing Service.

EXPECTED OUTCOMES

The project developed a high-precision GPT model using a curated dataset of 286 Q&A pairs. Comparative analysis results are as follows:

- Retrieval + MaumGPT: 89% accuracy (97/108 queries).
- General MaumGPT: 12% accuracy (6/50 queries).

The system enhances user convenience by effectively translating general inquiries into structured library service information.

SOCIAL ISSUE 04

Specialized AI Chatbot for Science and Technology Information

KEY TECHNOLOGICAL FEATURES

- **Data Preprocessing:** Preprocessing approximately 1,839 passages of internally held data
- **Training and Indexing**: Training and indexing research report data into SearchGPT

OVERVIEW

The project recognizes a critical need for expanded training data due to the current shortage of AI models tailored for science and technology. By promoting intelligent and efficient task operations for all relevant personnel, the service seeks to improve the practical usability of the organization's vast data resources. Consequently, the project is dedicated to constructing a domain-specific ultra-large language model and deploying diverse generative AI solutions customized to meet professional requirements.

Table 4. Project Detail

Project Category: Hyper-scale AI service (2023)
Implementing Agency/Partner Organization: Korea Institute of Science & Technology Evaluation and Planning (KISTEP)
Project Description The project leveraged extensive datasets from the 5th Science and Technology Basic Plan managed by the Korea Institute of Science and Technology Planning and Evaluation—comprising 10,000 meeting agendas, 5,000 trend documents, and 20,000 internal reports. By developing an in-house conversational AI and a fine-tuned pilot model, the service delivers tailored science and technology insights to a wide range of stakeholders. Specifically, it assists businesses with optimal partner matching, supports researchers in navigating complex regulations and drafting proposals, enables government ministries to execute data-driven policy planning, and provides the general public with science-based solutions for social issues. Furthermore, the establishment of a document-search and Q&A-responsive GPT allows users to access precise information with significantly enhanced speed and accuracy.

EXPECTED OUTCOMES

By developing a suite of user-customized generative AI services, the project establishes a robust utilization framework for the proprietary data held by the Korea Institute of Science and Technology Planning and Evaluation (KISTEP). The service is expected to maximize the intelligence and operational efficiency of tasks performed by diverse stakeholders across the science and technology sector.

SOCIAL ISSUE 05

PoC: Bus Route Guidance Chatbot for Gwangju Metropolitan City

KEY TECHNOLOGICAL FEATURES

- Data retrieval via Gwangju City Bus Open API
- Application of natural language consultation dialogue methods for customer support, data entry, and document processing using machine learning on an LLM-based platform

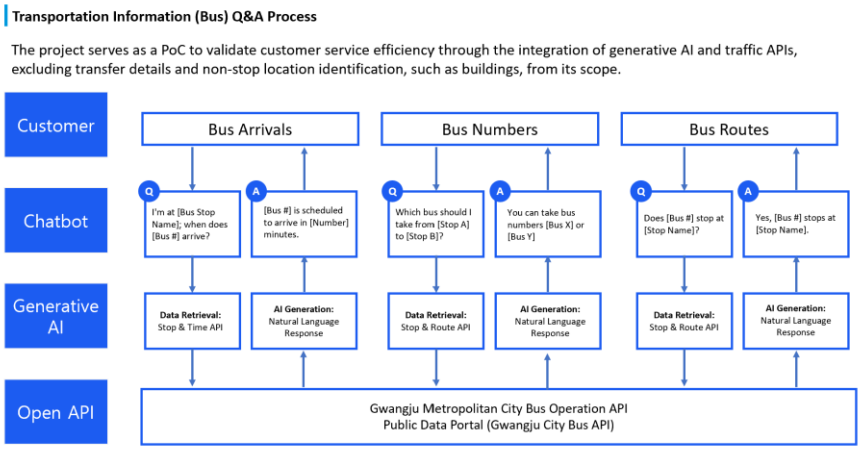
OVERVIEW

The Gwangju Metropolitan City 120 Call Center operates with 14 staff members working from 8 a.m. to 10 p.m. As of 2023, it receives approximately 20,000 consultation inquiries per month. Approximately 32% of all public inquiries, or 6,000 cases, involve simple traffic information, predominantly bus numbers and routes. As these basic bus number inquiries are now routed through the call center alongside existing public inquiries, the workload burden on the call center staff has increased, necessitating service improvements.

Table 5. Project Detail

Project Category: Hyper-scale AI Service (2024)
Implementing Agency/Partner Organization: Gwangju Metropolitan City
Project Description The project conducted in-depth Q&A sessions with operational consultants based on a pre-PoC questionnaire to establish a comprehensive classification of customer consultation tasks. Using this foundation, the system leverages an LLM to categorize user intent, which then triggers queries to the appropriate traffic information APIs. The service generates natural language responses from the retrieved data and incorporates a verification layer within the generative AI chatbot PoC to ensure the accuracy and relevance of every answer.

Figure 5.1. Q&A process



The service provides real-time responses to inquiries regarding bus arrivals, stops, and first/last bus schedules by integrating traffic information Open APIs with a generative Chatbot.

EXPECTED OUTCOMES

By automating the processing of public transit complaints—a task that previously accounted for over 30% of an annual consultant's workload—the service has significantly reduced operational burdens and improved overall job satisfaction. Furthermore, the adoption of cutting-edge generative AI technology not only enhances administrative efficiency but also serves as a scalable model for implementing innovative technological solutions across various other public services in the future.

Figure 5.2. PoC results

SOCIAL ISSUE 06

Intelligent Personalized Information Service for Citizens

KEY TECHNOLOGICAL FEATURES

- Intelligent integrating information inquiry service through deep learning model training
- **Developed APIs:** Integrated information inquiry for affiliated institutions, event/festival notifications, event/festival participation applications, passport application waiting information, etc.

OVERVIEW

Government and local authority websites serve as primary channels* for policy information dissemination, yet user satisfaction remains low due to poor accessibility and search accuracy. Furthermore, dozens of separate websites exist for each ministry and local authority's affiliated organizations, requiring citizens to visit and search individual sites to find information—a cumbersome process. Consequently, an API providing integrated information has been developed and opened for Hwaseong City.

* Hwaseong City website averages 6,000 daily visitors, with approximately 20,000 integrated search queries

The local government plans to establish an AI-based prediction model by analyzing citizen participation status and satisfaction levels for each event and festival. This model will be used to predict citizen participation rates and satisfaction levels for future events and festivals.

Table 6. Project Detail

Project Category: Digital Government Infrastructure (2023)
Implementing Agency/Partner Organization: Hwaseong City, URP Co., Ltd.
Project Description - **Intelligent Integrated Information Search API:** The service provides localized information on nearby hospitals, public health centers, and postpartum care centers based on the user's residence, while offering guidance on regional support programs and welfare services to ensure systematic pregnancy and childbirth management. - **Event Notification and Application API:** The project consolidated disparate event, festival, and educational data from various subsidiary organizations into a single system, developing integrated functions within private messenger platforms to streamline information delivery and application processes. - **Passport Application Wait Time Information API:** The system provides real-time updates on passport application ticket issuance and estimated wait times through private messenger channels, significantly improving user convenience and administrative transparency.

EXPECTED OUTCOMES

The system is expected to enhance citizens' satisfaction with the public website by making daily life information, which was previously difficult for citizens to access, easy and convenient to use.

Figure 6.1. Business Overview

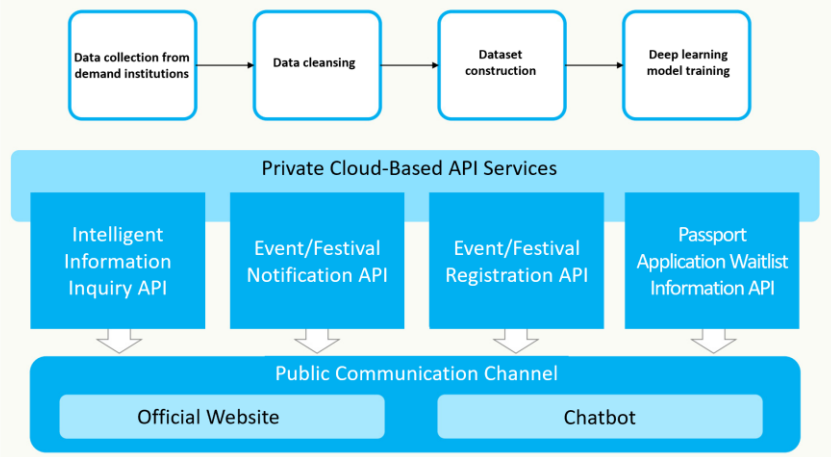


Figure 6.2. Event notification before and after implementation (AS-IS VS TO-BE)

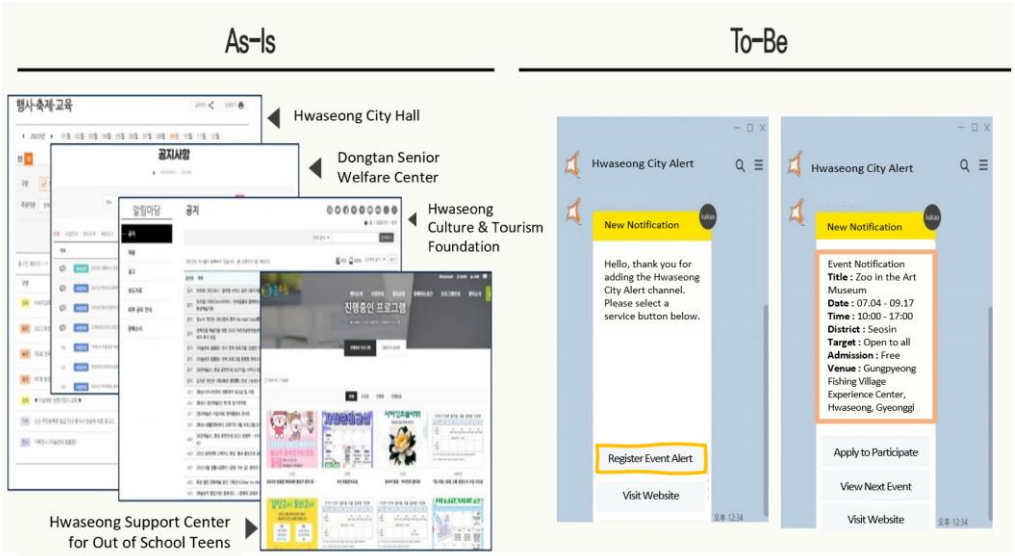
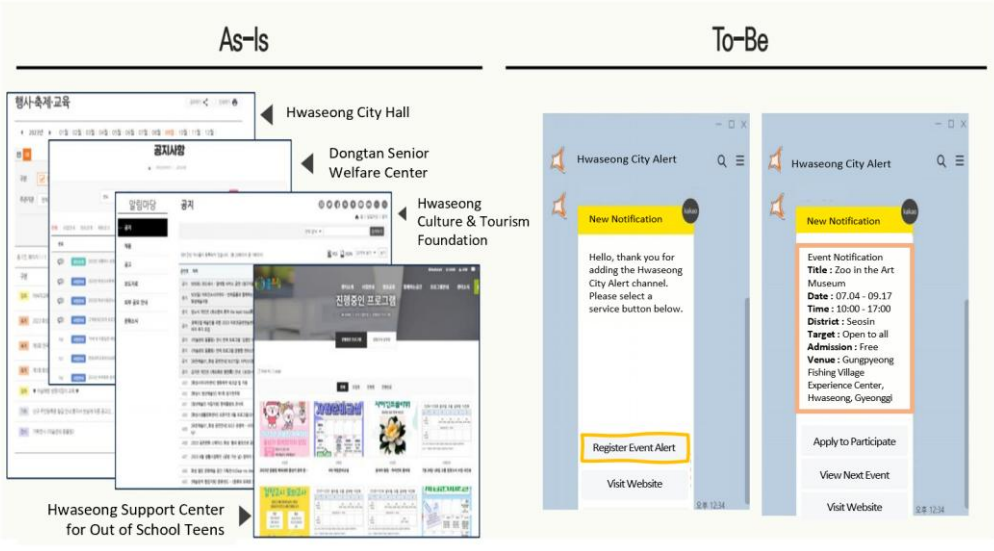


Figure 6.3. Passport application before and after implementation (AS-IS VS TO-BE)



SOCIAL ISSUE 07

RAG-Based Criminal Justice Analysis and Civil Complaint Response System

KEY TECHNOLOGICAL FEATURES

- Improving search summarization performance by enhancing response accuracy and consistency using G-eval
- Implementing lightweight training using the LoRA-based PEFT technique and selecting the optimal Adapter by testing various model versions

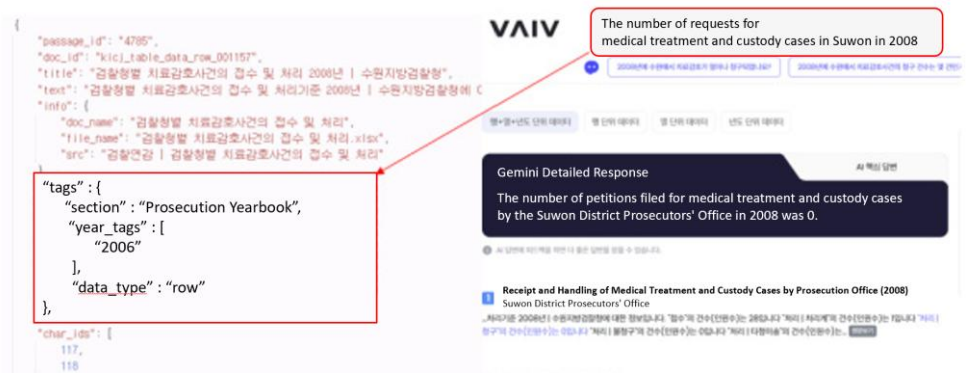
OVERVIEW

The project develops a RAG-based conversational AI that summarizes crime statistics into natural language, utilizing G-eval for automated quality refinement and LoRA-based PEFT for efficient LLM optimization. By balancing training efficiency with high performance, the service streamlines public data exploration, significantly enhancing citizens' accessibility to and comprehension of vital crime information.

Table 7. Project Detail

Project Category: Hyper-scale AI service (2024)
Implementing Agency/Partner Organization: Korean Institute of Criminology and Justice (KICJ)
Project Description The system is constructed by converting crime statistics into a searchable text structure and integrating it with a RAG solution, while performance is optimized by implementing filtering for key crime-related keywords.

Figure 7.1. System overview



EXPECTED OUTCOMES

By automating civil complaint responses, the service enables employees to focus more on their core tasks, thereby enhancing overall organizational productivity and significantly increasing public satisfaction with response quality.

SOCIAL ISSUE 08

Smart Search and Chatbot for Real Estate and Construction Data

KEY TECHNOLOGICAL FEATURES

- **LLM Training Dataset:** Preprocessing of Civil Complaints and Legal Statute Data Held by the Korea Real Estate Board
- **SearchGPT/SmartChat-Based Service:** Training and Indexing Preprocessed Data on SearchGPT, and Generating Queries Based on SmartChat

OVERVIEW

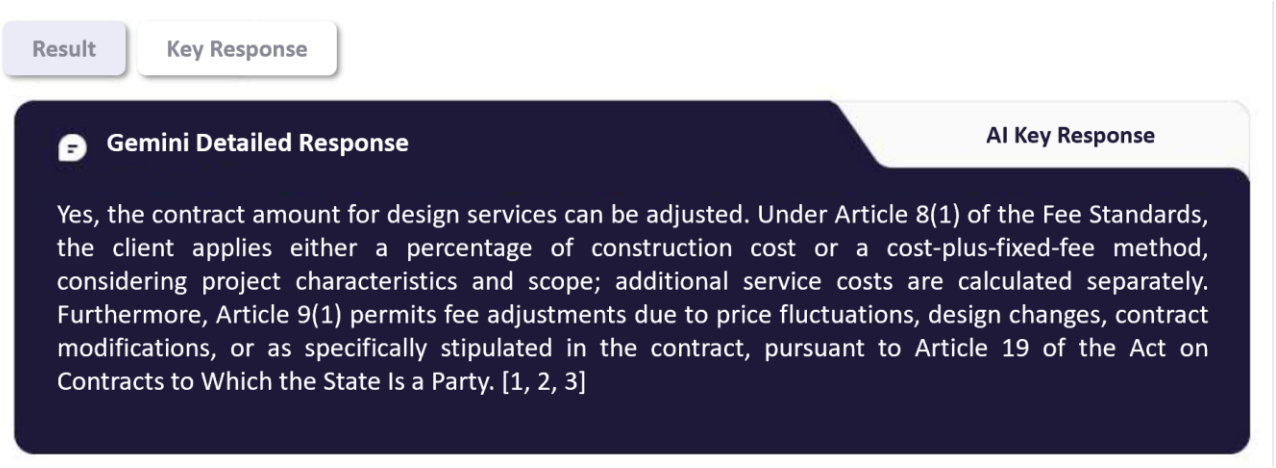
To resolve the inefficiency of processing 10,000+ annual building-related complaints, the Korea Real Estate Board developed a SearchGPT/SmartChat-based pilot service to reduce staff workload.

Table 8. Project Detail

Project Category: Hyper-scale AI service (2023)
Implementing Agency/Partner Organization: Korea Real Estate Board (REB)
Project Description For the service pilot, approximately 900 complaint cases, 100 pages of the Building Act, and 400 pages of daily life regulations were used as training data. For model training, the Korea Real Estate Board preprocessed its complaint and legal data and reviewed the query generation results from the SmartChat system. Subsequently, the preprocessed data was trained and indexed in SearchGPT/SmartChat, ultimately leading to the development of a testable demo page. Key services include a SmartChat-based chatbot service for complaints and construction-related information, and a SearchGPT-based information retrieval function for various laws and regulations or complaint responses.

Figure 8.1. SmartChat-based chatbot service screen

Figure 8.2. SearchGPT-based chatbot service screen



EXPECTED OUTCOMES

The service reduces the burden of civil petitions by providing easy access to historical data through integrated search and chatbot functions. Following the successful PoC, we plan to implement a full-scale service featuring building regulation updates and an advanced similar-petition search function for the upcoming year.

SOCIAL ISSUE 09

Q&A Support Chatbot for Broadcast Advertising Documents

KEY TECHNOLOGICAL FEATURES

- **Data Collection:** Collected internal data from KOBACO (Korea Broadcasting Advertising Corp.), along with external laws and legal precedents via APIs and public web pages
- **OCR Implementation:** Applied Optical Character Recognition (OCR) models to process scanned documents and image data, enabling text extraction from existing archives.
- **RAG Utilization:** Employed Document Layout Analysis (DLA) models to extract content and structure internal documents for efficient RAG (Retrieval-Augmented Generation) learning.

OVERVIEW

The project was initiated to improve the civil complaint handling process, as staff members faced difficulties in providing rapid responses due to repetitive and time-consuming manual searches. Korea Broadcast Advertising Corporation (KOBACO) developed a RAG-based Q&A service targeting relevant documents and integrated it with a chatbot system to enhance operational efficiency.

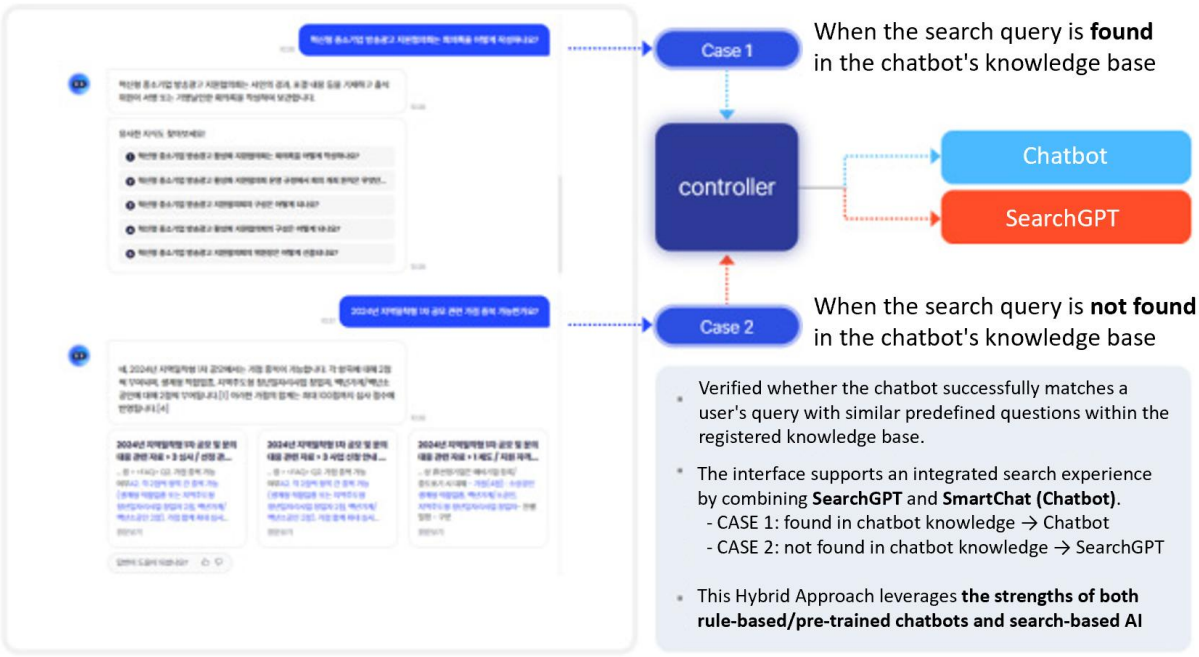
Table 9. Project Detail

Project Category: Hyper-scale AI service (2024)
Implementing Agency/Partner Organization: Korea Broadcast Advertising Corporation (KOBACO)
Project Description To support the broadcast advertising chatbot, relevant documents were restructured into search-optimized formats. Generative AI was used to automatically generate chatbot queries, leading to the development of a specialized chatbot service that combines RAG technology with KOBACO's advertising data. The service, which supports staff in responding efficiently to complaints, is classified into two types: - **Chatbot-based responses:** Provides answers based on pre-trained knowledge from existing documents (internal data, laws, precedents, etc.). - **RAG-based SearchGPT:** For queries outside the training set, it provides highly relevant content using RAG-powered search capabilities.

EXPECTED OUTCOMES

By replacing repetitive and exhaustive tasks with AI, the service is expected to create a work environment where employees can focus on more professional and high-value tasks.

Figure 9.1. Query Handling Logic



SOCIAL ISSUE 10

Advanced Chatbot Development for National Digital Transformation

KEY TECHNOLOGICAL FEATURES

- **SearchGPT:** Completed preprocessing of 700 data passages, followed by training and indexing; integrated with a demo page to provide search services.
- **SmartChat:** Completed generation and verification of 1,800 queries; applied to the chatbot solution and integrated with a demo page for automated query generation.

OVERVIEW

While guidelines, templates, and reference materials are currently distributed via official documents, there are limitations in service response due to a small workforce and an Excel-based data collection system. To overcome this, we built an automated and efficient monitoring and analysis service for national digital projects using Hyper-scale AI technology.

Table 10. Project Detail

Project Category: Hyper-scale AI service (2023)
Implementing Agency/Partner Organization: NIA (National Information Society Agency)
Project Description For AI model training, data was first extracted from 10 initial Q&A sets, internal budget documents, and laws. The development followed these procedures: ①Preprocessing of training data (700 passages) ②SmartChat-based query generation and verification ③Training and indexing of preprocessed data into SearchGPT/SmartChat ④Development of a demo page for functional testing Through this advanced chatbot, we provided a consultation support service to respond to inquiries and data requests regarding national digital projects and budgets from central ministries and public institutions.

EXPECTED OUTCOMES

As a result of the chatbot advancement, we reduced processing time for identifying accurate information and follow-up tasks through rapid and consistent responses. We successfully provided an intelligent search service capable of answering various inquiries based on internal data. Furthermore, we expect to enhance performance through future improvements, such as converting tables and numerical data into text or structured formats for additional training.

SOCIAL ISSUE 11

Generative AI Chatbot for Public Information and Web Content Guidance

KEY TECHNOLOGICAL FEATURES

- **Data Indexing Optimization**: Focused on enhancing the relevance and accuracy of results through optimized data indexing and embedding
- **RAG Utilization:** Optimized the retrieval phase, flexibly applied various query strategies, and utilized embedding similarity
- **Implementation of Advanced RAG**: To overcome the limitations of Naive RAG (which simply combines retrieval and generation), Advanced RAG technology was applied to implement sophisticated search functions and context-based response generation.

OVERVIEW

Through an analysis of the official and sub-websites of the Korea Institute for Gender Equality Promotion and Education (KIGEPE) and its customer consultation operations, this service provides customized gender equality content.

Public demand for information on preventing the "Five Major Types of Violence" (power-based sexual violence, digital sexual violence, domestic violence, dating violence, and stalking) is increasing annually. Consequently, it was necessary to develop a system that integrates prevention education and victim support to provide real-time information to the public.

Table 11. Project Detail

Project Category: Hyper-scale AI service (2024)
Implementing Agency/Partner Organization: Korea Institute for Gender Equality Promotion and Education (KIGEPE)
Project Description The AI Chatbot Service provides the following features: - **Content Search:** Prevention education, violence prevention, and crime types - **Civil Complaint Filing:** Reporting damage and victim support - **Service Requests:** Modification or deletion of information and data - **History Management:** Customized support based on past user history

Figure 11.1. Current status of data provision

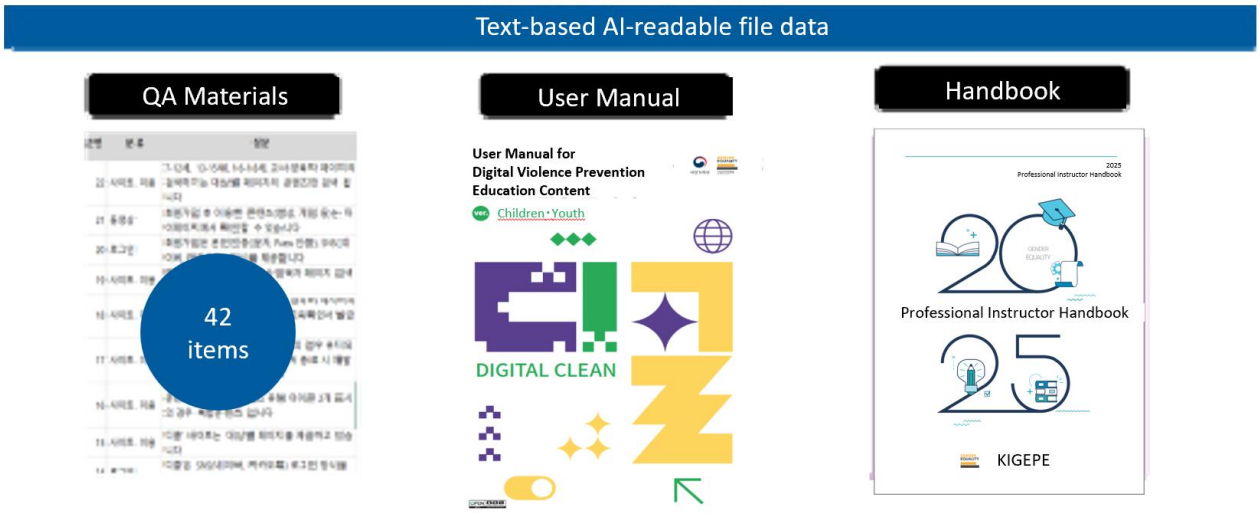
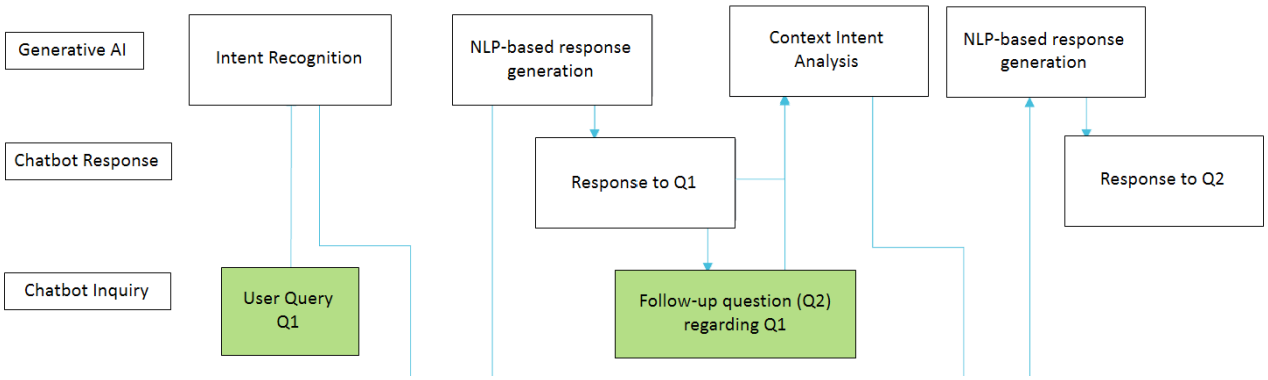


Figure 11.2. Chatbot Q&A Process



The project involved crawling notices and educational announcements from the official website and its sub-sites (Education Center, E-learning Center, Gender-On, and D-Clean). By training on a diverse dataset of video, image, and text-based educational content, this service was configured as a chatbot that provides guidance on KIGEPE (Korea Institute for Gender Equality Promotion and Education) and gender equality training, enabling users to locate specific content immediately.

EXPECTED OUTCOMES

This AI chatbot service has streamlined consultation procedures through automated data collection and analysis. It is expected to significantly reduce the overall workload required for handling consultations and inquiries.

SOCIAL ISSUE 12

Personalized Student Information Support via Web Chatbot

KEY TECHNOLOGICAL FEATURES

- **Hyper-scale AI Integration**: Implementing customized Q&A services by adopting Hyper-scale AI chatbot technology to ensure personalized and human-like interactions
- **Advanced RAG Architecture**: Building a Retrieval-Augmented Generation (RAG) framework based on Hyper-CLOVA X (LLM), structured Q&A, and rule-based scenarios to maximize response accuracy

OVERVIEW

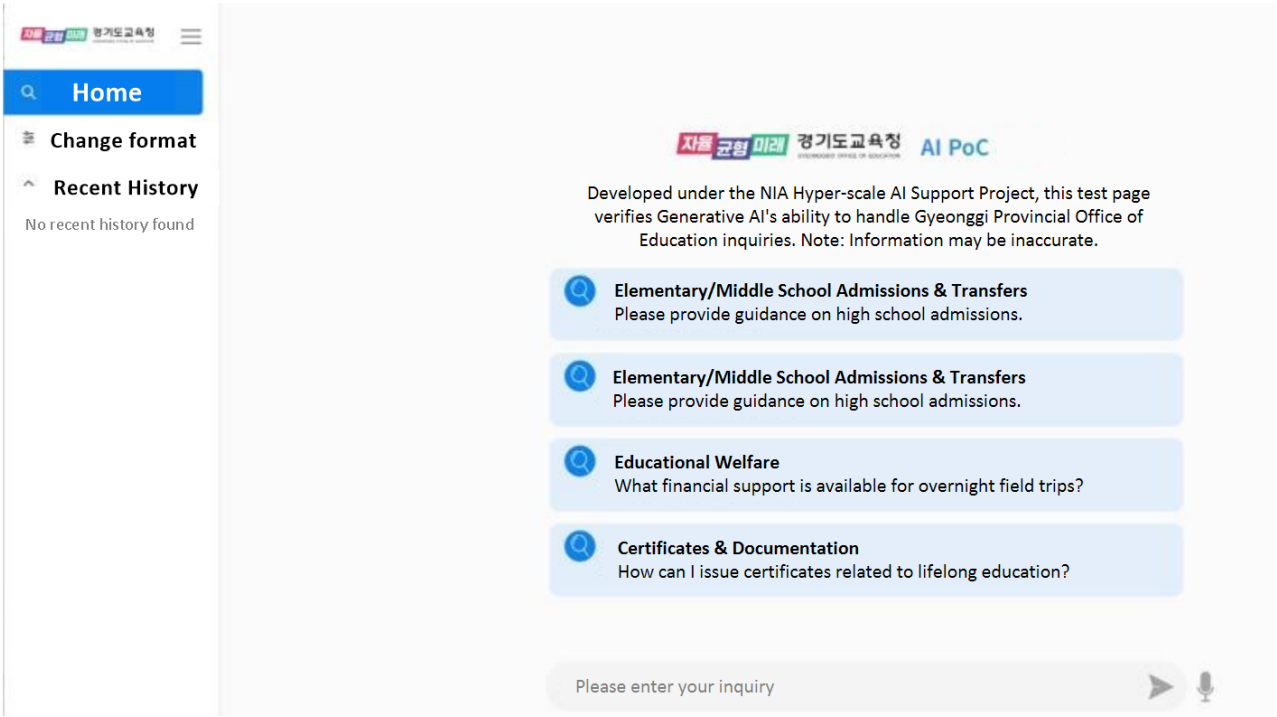
School environments face challenges where faculty members struggle to focus on core educational activities due to a high volume of simple, repetitive inquiries* from students and parents.
*Academic/event schedules, attendance/field trip application procedures, facility usage, school newsletters, etc.

Currently, education-related information is fragmented across multiple channels, forcing users to navigate various services to find information. This has increased the administrative burden of responding to repetitive complaints. To address this, the project aimed to build a customized guidance system at the school level using Hyper-scale AI technology.

Table 12. Project Detail

Project Category: Hyper-scale AI service (2024)
Implementing Agency/Partner Organization: Gyeonggido Office of Education
Project Description The project collected data from educational service materials and website announcements for students and parents. Requirements were analyzed through user interviews to derive improvements for service enhancement. Additionally, this project explored cloud-based system configurations and AI model optimization to develop a strategic plan for a "Educational Community Portal." By collecting rule-based chatbot Q&A data (including scenarios) and institutional website data from the Gyeonggi Provincial Office of Education, the project implemented data preprocessing, embedding, and Vector DB construction. Finally, this service established a Generative AI chatbot that responds to various simple inquiries by integrating Hyper-CLOVA X APIs with optimized prompting.

Figure 12.1. PoC implementation case



EXPECTED OUTCOMES

Through this Generative AI chatbot pilot project, the service provides an integrated and optimized Q&A system for a wide range of education-related inquiries. By offering 24/7 consultation and customized information in a single location, the project enables students and parents to receive immediate responses, thereby significantly enhancing user satisfaction. Furthermore, this service is expected to create an environment where faculty members can focus more effectively on core educational activities by automating simple, repetitive inquiries and reducing their overall administrative workload.

2

Counseling Support

SOCIAL ISSUE 13

Public Institution Customer Service Assistant Service Powered by Hyper-scale AI

KEY TECHNOLOGICAL FEATURES

- **NLP & API Integration**: Processing natural language queries from consultants by performing keyword decomposition and synonym expansion, followed by seamless integration with the Hyper-CLOVA X API
- **Accuracy Ranking Optimization**: Tuning search result rankings to ensure high relevance and precise alignment with the user's intent
- **Fine-tuning Data Construction**: Building specialized datasets for model fine-tuning to accurately capture and reflect the specific information used by consultants
- **Kubernetes Rolling Deployment**: Establishing a CI/CD pipeline where development updates are continuously containerized and deployed to the cluster via Kubernetes rolling updates

OVERVIEW

As the demand for non-face-to-face services grows, the workload and labor intensity of public call centers have significantly increased. To address this and improve public satisfaction, the project introduced an AI-powered Civil Complaint Assistant service utilizing Hyper-scale AI technology. By providing precise answers to inquiries along with relevant source information, this AI assistant service has successfully enhanced both citizen satisfaction and the operational productivity of consultants.

Table 13. Project Detail

Project Category: Hyper-scale AI service (2023)
Implementing Agency/Partner Organization: Hwaseong City Hall
Project Description By integrating Hyper-CLOVA X with cloud-based search capabilities, this system enables the retrieval of publicly available civil complaint data and the identification of relevant search results on a large scale. The system provides consultants with automated summaries and source references via Generative AI, enabling them to verify answers quickly and significantly reduce response times. To enhance response quality, the project trained the model on a vast dataset of unstructured administrative manuals and forms—including 834 cases from Suncheon City, 444 from Uijeongbu City, and 422 from Hwaseong City. By applying RAG (Retrieval-Augmented Generation), the service successfully mitigated hallucinations and improved the overall accuracy of the AI assistant. For continuous service optimization, an Agile methodology was adopted to consistently refine answer relevance and user convenience. Furthermore, to objectively verify the improved accuracy of civil complaint responses, "Generative Q&A Model Accuracy" was established as a quantitative performance metric, with formal testing conducted by an accredited certification testing agency.

Figure 13.1. Service flow

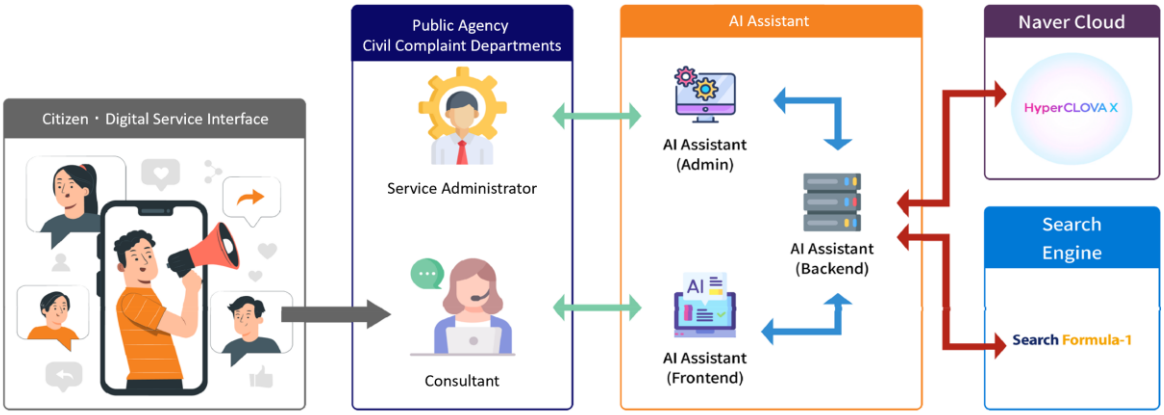
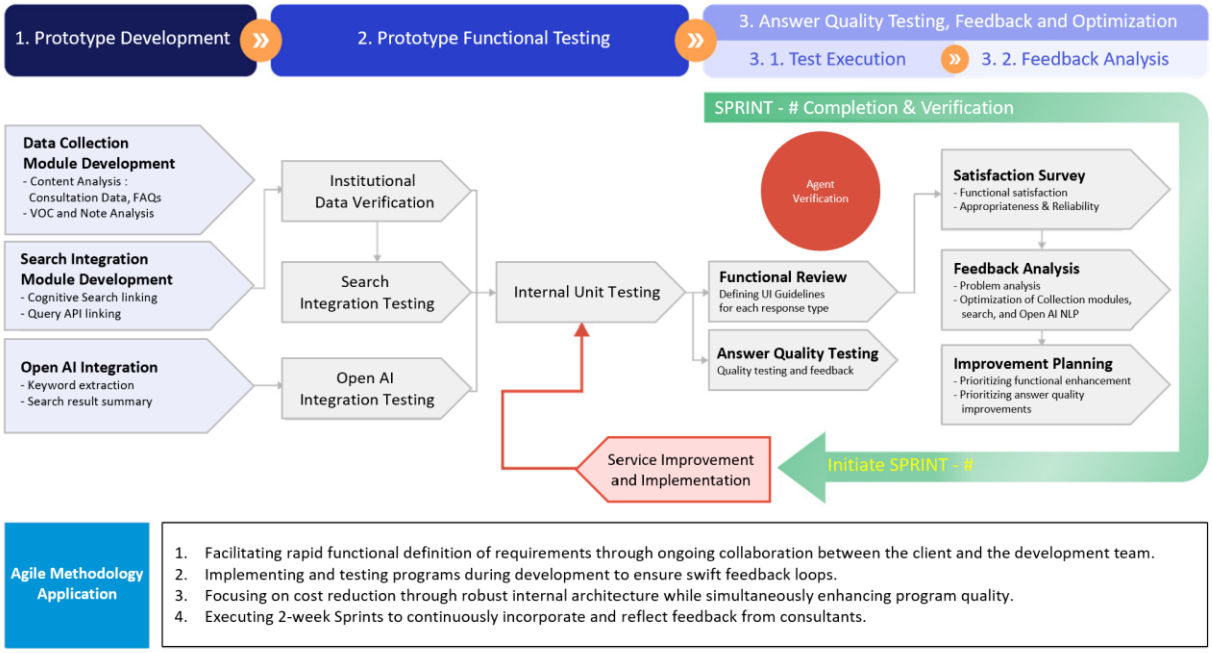


Figure 13.2. Process flow



EXPECTED OUTCOMES

The use of the AI Civil Complaint Assistant is expected to reduce the data search time for each inquiry by over 30 seconds, leading to a more than 10% increase in total call handling capacity. Additionally, it is anticipated that the daily workload of call center staff will be reduced by approximately 42 minutes, allowing for a 20% increase in their rest periods. Enhanced operational productivity for call center workers also signifies an improvement in citizen satisfaction. High-quality consultation is expected through faster processing and responses that consistently apply a professional tone and manner.

Based on this project, there are plans for an initial expansion to the three pilot call centers in Hwaseong, Suncheon, and Uijeongbu. Subsequently, the service will be continuously scaled to include nationwide public inquiry agencies, as well as private sectors such as finance, insurance, and logistics.

SOCIAL ISSUE 14

AI-based Analysis and Classification Enhancement for e-People Civil Complaint Data

KEY TECHNOLOGICAL FEATURES

- **AI Training Data Construction**: Established a comprehensive AI training dataset leveraging 14,000 cases of civil complaint records to ensure model accuracy and reliability
- **Semantic Reasoning & Auto-Classification**: Developed an advanced complaint analysis and automatic recommendation model by utilizing AI-driven semantic inference and automated classification technologies
- **Module Integration & Automation**: Built a seamless module integration framework and an automated system architecture to optimize overall operational workflow

OVERVIEW

The number of civil complaints filed through e-People (People's Online Petition and Discussion Portal) is continuously rising, with a high volume of simple and repetitive inquiries adding to the administrative burden. Furthermore, the low accuracy of the current system's agency recommendation feature has led to an increased workload for personnel in charge.
To address these issues, the project aims to establish an AI-driven complaint analysis and automation system that enhances complaint summarization, time-series clustering, and agency recommendations by utilizing semantic reasoning and automatic classification technologies. With the Anti-Corruption and Civil Rights Commission (ACRC) as the primary demand organization and VAIV Company as the AI technology provider, this project seeks to improve response efficiency through systematic data analysis and automated classification models.

Table 14. Project Detail

Project Category: Hyper-scale AI service (2024)
Implementing Agency/Partner Organization: Ministry of Health and Welfare, Korea Social Security Information Service, etc.
Project Description This project focuses on developing summarization technology to automatically extract key information from e-People (People's Online Petition and Discussion Portal) data and establishing a monitoring system by applying time-series clustering. Furthermore, it aims to develop a text summarization and agency classification model that provides significantly higher accuracy compared to existing systems. The specific procedures are as follows: - **Redefining Data Analysis & Classification:** Improving the complaint classification system to overcome existing limitations and ensure more accurate recommendations for agencies and departments. - **Training Data Construction & Model Development:** Building an AI training dataset based on 14,000 complaint records from specific regions and periods to train semantic reasoning and auto-classification models. - **Summarization & Agency Recommendation Model:** Utilizing AI-driven summarization models to extract core content and developing a model that automatically recommends the appropriate handling organization.

- **Module Integration & Automation Framework:** Integrating the summarization and recommendation models into the system to automatically deliver summarized results and recommended agency information to the personnel in charge.

EXPECTED OUTCOMES

By introducing an AI-based complaint analysis and automation system, it is expected to increase the efficiency of complaint processing and alleviate the workload of personnel. Enhanced response speed through AI summarization and clustering allows the system to automatically extract key information and group similar complaints, enabling more rapid responses. The goal is to improve the agency recommendation accuracy to over 90%, ensuring that complaints are promptly assigned to the correct organization. This automated system optimized by AI technology will ultimately accelerate digital innovation in administrative services.

Based on the results of this pilot project, there are plans to upgrade the complaint handling systems of e-People and related organizations using AI. This initiative aims to create an environment where citizens can receive faster and more accurate responses to their inquiries.

SOCIAL ISSUE 15

AI Chatbot Service for Handling Public Inquiries on Airport Facilities

KEY TECHNOLOGICAL FEATURES

- **Customer Response GPT**: Conducted Fine-tuning using MaumGPT after preprocessing the corporation's internal data to optimize it for Large Language Model (LLM) training
- **Retrieval Engine Implementation**: Mitigated hallucination effects by implementing a search engine based on a Q&A dataset of 835 internal customer response records

OVERVIEW

Visitors using the airport require various information regarding convenience and commercial facilities. However, current response methods rely heavily on phone calls or on-site information desks, often making it difficult to provide rapid responses. In response, the Korea Airports Corporation (KAC) aims to develop a chatbot service model that aggregates airport facility information to effectively handle simple and repetitive customer inquiries. The goal of this project is to enhance customer satisfaction, reduce the workload of counseling personnel, and maximize the convenience of airport use.

Table 15. Project Detail

Project Category: Hyper-scale AI Service (2023)
Implementing Agency/Partner Organization: Korea Airports Corporation (KAC)
Project Description To develop the chatbot service, Korea Airports Corporation is building a learning model based on internal documents and data. The specific details and procedures are as follows: - **Data Aggregation**: Compiled facility information from 14 airports, utilizing a total of 835 records, including 423 cases of airport facility Q&A and 412 cases of commercial facility Q&A. - **Step-by-Step Development:** Executed a comprehensive process ranging from requirements analysis and data collection to preprocessing, training data generation/refinement, search engine training, and final chatbot production. - **Targeted Design:** Designed to provide swift and accurate answers to frequently asked questions about convenience and commercial facilities. - **Operational Efficiency:** Enables visitors to easily obtain necessary information without wait times while alleviating the repetitive task burden on human agents.

EXPECTED OUTCOMES

The introduction of the chatbot service will improve the response speed to airport visitor inquiries and maximize the efficiency of complaint processing. Furthermore, Korea Airports Corporation has established an automated system for resolving complaints by refining and preprocessing internal data through AI. This system guides users through representative questions and resolves inquiries regarding airport facilities in real-time. Consequently, this chatbot service is expected to increase visitor satisfaction while reducing the workload of airport operations staff, contributing to the establishment of a more efficient information delivery system.

SOCIAL ISSUE 16

Demonstration of Civil Petition Counseling Information Service for Yangsan Residents

KEY TECHNOLOGICAL FEATURES

- **SearchGPT Integration**: Utilized SearchGPT to enable rapid retrieval and lookup of internal administrative data
- **SmartChat Solutions**: Implemented a chatbot solution via SmartChat to alleviate the workload of consultation services and enhance overall service quality

OVERVIEW

While the demand for administrative services in Yangsan City has been increasing, the city faced a shortage of personnel to handle the load. In this environment, it became necessary to introduce a new service to promptly process the diverse civil complaints and requests of residents.

Accordingly, the project aims to develop an AI-based consultation service that allows even non-experts to quickly search for internal information and respond to inquiries. This initiative is designed to reduce the burden on officials and improve the quality of public services.

Table 16. Project Detail

Project Category: Hyper-scale AI Service (2023)
Implementing Agency/Partner Organization: Yangsan City
Project Description An AI-powered information search and consultation system was established using past civil complaint and administrative data held by Yangsan City Hall. Approximately 70 cases of administrative and complaint-related records were utilized to perform data training and indexing within the SearchGPT framework. Additionally, a chatbot solution was implemented by adding a query generation function based on SmartChat. This system is designed to integrate diverse data scattered across various departments in Yangsan City, making it easier for citizens to search for the information they need. Through this, citizens can obtain civil complaint information more quickly and accurately, while officials can respond to inquiries with greater efficiency.

EXPECTED OUTCOMES

By providing a search service based on civil complaint and administrative data, information from various departments within Yangsan City Hall—including architecture, informatics, health, and population—can be utilized efficiently. This is expected to reduce the burden on personnel and improve the quality of resident services by increasing response speeds. Furthermore, the system will be continuously improved by learning from newly accumulated data. By linking search and chatbot services, a more sophisticated response system can be established. This will create a framework for responding rapidly to resident demands and maximizing the efficiency of administrative processing.

SOCIAL ISSUE 17

Expert Chatbot Service for Telecommunication Services

KEY TECHNOLOGICAL FEATURES

- **Data Preprocessing**: Preprocessed 310 data passages related to "Smart Choice" to optimize them for AI utilization
- **Training & Indexing**: Applied SearchGPT to the preprocessed data and implemented a chatbot solution featuring query generation based on SmartChat
- **Hybrid Service Model**: A hybrid system combining SearchGPT-based search/summarization with SmartChat-based chatbot services

OVERVIEW

Providing telecommunications information and services in a way that users can find conveniently and understand easily will establish the platform as a go-to service for everyone. To this end, the Korea Association of ICT Promotion (KAIT) has developed services such as real-time civil complaint consultation and personalized rate plan recommendations.

Currently, complaint handling is only available during weekday daylight hours and is limited to agents responding directly to user questions. This structure faces challenges in recommending necessary precautions or customized products for users selecting telecom services. Therefore, the project aims to introduce an interactive AI chatbot service to overcome these limitations.

Table 17. Project Detail

Project Category: Hyper-scale AI service (2023)
Implementing Agency/Partner Organization: Korea Telecommunications Operators Association(KTOA)
Project Description The project developed a real-time specialized chatbot service for telecommunications by training on data such as legal regulations, terms of use, subscriber information (rate plans, usage patterns, early termination fees), and telecom service details. By learning from civil complaints, terms of service, and legal data, the chatbot is capable of responding to inquiries in real-time at a professional level equivalent to a human agent. It analyzes words and sentences entered by users to provide natural answers and recommendations. Furthermore, it communicates with users in a conversational manner—following rational telecom consumption procedures—to recommend the most suitable rate plans tailored to their specific needs.

EXPECTED OUTCOMES

Through the interactive AI chatbot consultation service, necessary information is provided to users at the right time and place, contributing to the operational efficiency of service personnel. Additionally, it is possible to develop a "telecom lifestyle check-up" service by analyzing the history of user usage patterns.

Furthermore, the service aims to contribute to reducing national telecommunications costs by promoting the refund of unclaimed telecommunications credits and providing guidance on discount systems. By implementing Hyper-scale AI in public complaint response services, the project will continue to drive improvements in overall service quality.

SOCIAL ISSUE 18

Improving Military Manpower Administration Public Services using Hyper-scale AI

KEY TECHNOLOGICAL FEATURES

- **Data Preprocessing**: Type conversion, unstructured data refinement, and JSON transformation (includes agency-specific processing such as URL handling, synonym dictionary application, and subtitle separation)
- **Model Selection & Configuration**: Selected customized models including an Embedding model (in-house model based on E5, fine-tuned by SDS), a Reranking model (in-house model based on E5, fine-tuned by SDS), and an optimized Samsung LLM
- **System Optimization**: Achieved performance optimization through meticulous tuning of Parsing and Embedding parameters.

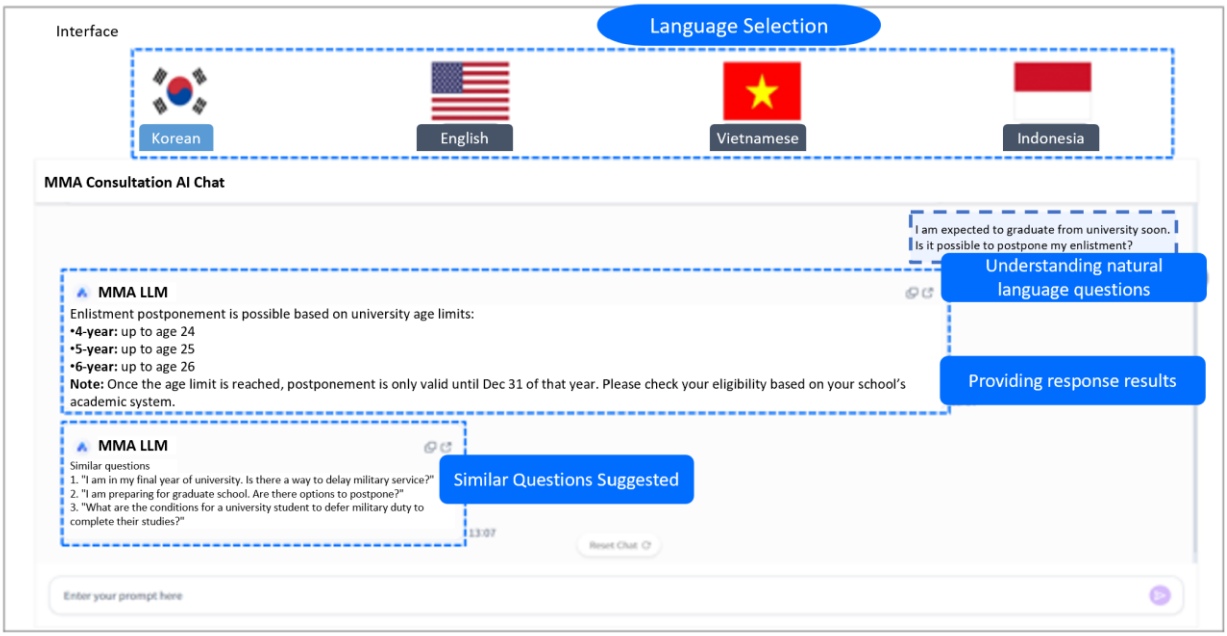
OVERVIEW

The existing chatbot provided for military service consultations followed a rule-based structure, which lacked flexibility in responses and presented difficulties in data management and maintenance. To address these challenges, the Military Manpower Administration (MMA) aims to enhance the efficiency of its consultation services by introducing Hyper-scale AI, overcoming the technical and managerial limitations of the previous system.

Table 18. Project Detail

Project Category: Hyper-scale AI service (2024)
Implementing Agency/Partner Organization: Military Manpower Administration (MMA)
Project Description This service integrates Hyper-scale AI into the existing rule-based civil complaint (text-based) chatbot of the Military Manpower Administration. First, a Vector DB was constructed using military service consultation manuals and chatbot glossaries to provide the foundation for the service. As a result of optimizing data through specialized preprocessing, the system achieved a response accuracy of 94% and offers multi-language consultation, including English, Vietnamese, and Indonesian. Considering that overseas citizens are also eligible for enlistment; the multi-language service allows users to select their preferred language for inquiries. The AI understands natural language questions and provides answers based on official manuals and glossaries. The responses are personalized, presenting direct answers to questions alongside relevant similar case studies.

Figure 18.1. AI Consultation Scenario



EXPECTED OUTCOMES

By providing a context-based conversational service rather than a simple menu-selection method, this project enhances user satisfaction. Furthermore, the automation of data preprocessing and DB conversion is expected to deliver significant achievements in terms of data management and maintenance efficiency.

SOCIAL ISSUE 19

Enhancement of Chatbot Services based on Power Supply Terms, Regulations, and Civil Petition History

KEY TECHNOLOGICAL FEATURES

- **SearchGPT**: Completed preprocessing of internal data (604 passages), followed by training and indexing to provide a search service integrated with a demo page
- **SmartChat**: Generated and verified 1,949 Q-A pairs, then applied them to the chatbot solution to provide an automated query generation service via the demo page

OVERVIEW

The existing chatbot system at Korea Electric Power Corporation (KEPCO) had low utilization due to its overly simplified responses compared to direct human consultations. To overcome this, the project aimed to upgrade the chatbot to improve the automated response rate and provide consistent, standardized response drafts through AI.

Table 19. Project Detail

Project Category: Hyper-scale AI service (2023)
Implementing Agency/Partner Organization: Korea Electric Power Corporation (KEPCO)
Project Description To train the AI model, 1,599 cases of KEPCO's operational FAQ data were extracted. The project followed a structured procedure: 1. Preprocessing training data (604 passages), 2. Generating and verifying queries based on SmartChat, 3. Training and indexing the preprocessed data in SearchGPT/SmartChat, and 4. Developing a demo page for testing. The main services include an "Intelligent Search Service" for data such as sales operational guidelines and a "Q&A Service" through the AI chatbot to facilitate conversation-based complaint handling.

EXPECTED OUTCOMES

By integrating the intelligent search module with the existing chatbot, the project saved on the budget required for service quality upgrades. Compared to previous services, the query range has expanded, enabling the system to answer questions that previously had short responses or contained special characters. Furthermore, collecting a wider variety of real user questions and including them in the training dataset is expected to continuously enhance the performance of the chatbot model.

SOCIAL ISSUE 20

Demonstration of AI Human for Handling Civil Petitions in National Pension Lump-sum Payments

KEY TECHNOLOGICAL FEATURES

- **SmartChat-based Query Generation**: Conducted data preprocessing, indexing, and generated Q-A pairs based on SmartChat
- **SmartChat**: Implemented a hybrid approach combining the strengths of SearchGPT and SmartChat for both search and chatbot solutions

OVERVIEW

As public interest in the National Pension grows, the volume of related civil complaints has surged, leading to increased workloads and a shortage of response personnel. In particular, the high frequency of inquiries regarding lump-sum pension payments requires rapid and accurate responses. To address this, the project aims to demonstrate a complaint response system by introducing an AI Human-based concierge service capable of responding in a manner similar to actual staff.

Table 20. Project Detail

Project Category: Hyper-scale AI service (2023)
Implementing Agency/Partner Organization: National Pension Service (NPS)
Project Description To build the AI Human-based response service for lump-sum pension inquiries, various data were collected and training models were developed. Over 1,000 sets of data, including the National Pension Service's (NPS) guidelines, official documents, guides, and Q&A records, were utilized. The model development process followed a structured sequence: 1. Data preprocessing and dataset construction, 2. AI model integration and fine-tuning, 3. Development of chatbot functions for lump-sum payment inquiries, and 4. Service testing in connection with the AI Human interface.

EXPECTED OUTCOMES

The introduction of the AI Human-based system will enable more accurate and prompt responses to inquiries regarding lump-sum pension payments. This is expected to reduce the burden on personnel and increase citizen satisfaction. Furthermore, by supporting keyword and query-based searches, the system allows users to find desired information quickly, providing an environment where pension-related complaints are handled effectively through intelligent search and chatbot functions. Based on the Proof of Concept (PoC) results, future plans include additional model training to expand the service beyond pension inquiries to various other administrative services.

Figure 20.1. AI-powered intelligent consultation system

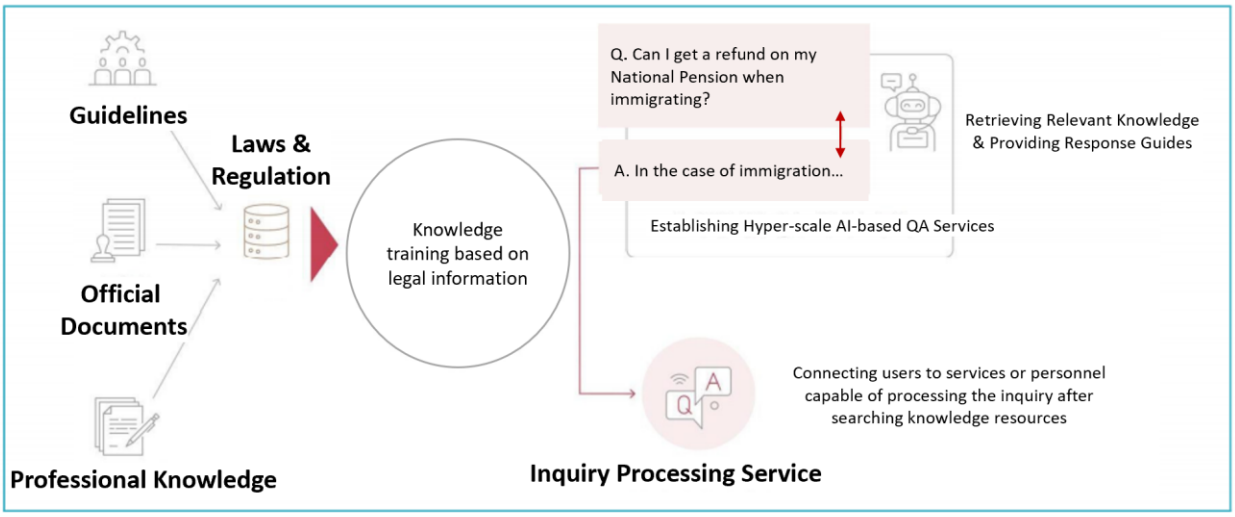
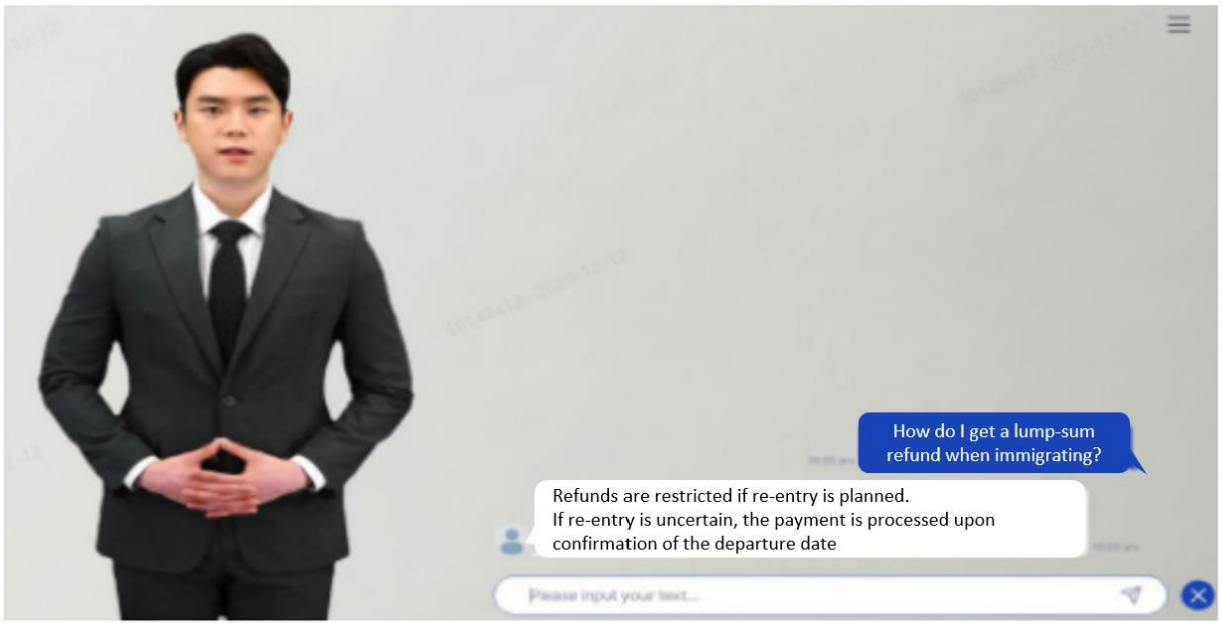


Figure 20.2. AI Human-based lump-sum inquiry response test



SOCIAL ISSUE 21

Hyper-scale AI-based Civil Petition Counseling Support Service for Foreigners

KEY TECHNOLOGICAL FEATURES

- **Service Availability**: Provides 24/7/365 multilingual civil complaint services based on natural language query processing
- **Knowledge Base**: Trained on administrative guides from the Ministry of Justice, Ministry of Employment and Labor, and Korea Immigration Service
- **Model Training Pipeline**: Follows a structured workflow: Data Collection/Integration → Document Parsing → Data Structuring → Chunking Optimization (breaking data into meaningful semantic units for efficient AI processing)

OVERVIEW

Ansan City is exploring ways to enhance the efficiency and quality of civil complaint services for its foreign residents. The current system, which relies heavily on phone calls and in-person visits, faces challenges such as high dependency on human consultants and the excessive time required for staff to retrieve information. To overcome these limitations, the city aims to introduce and validate a Hyper-scale AI-powered chatbot service for complaint response.

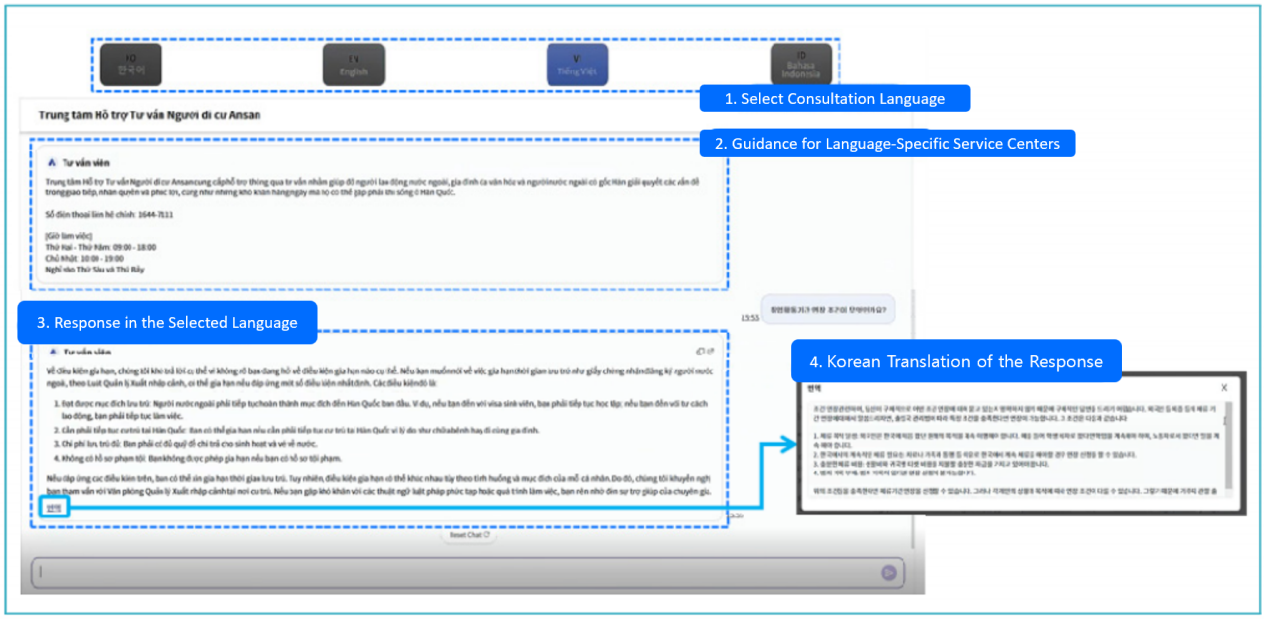
Table 21. Project Detail

Project Category: Hyper-scale AI service (2024)
Implementing Agency/Partner Organization: Ansan City
Project Description This project establishes a Vector DB based on existing consultation casebooks and Q&A datasets to enable Hyper-scale AI-driven inquiry and response processing. This allows residents to receive faster and more accurate consultations, while human staff can focus on more complex cases. Additionally, the system is designed to support multiple languages—including Vietnamese, Indonesian, Korean, and English—improving accessibility for foreign residents and minimizing inconveniences caused by language barriers.

EXPECTED OUTCOMES

The introduction of an AI-based automated consultation system is expected to significantly boost the operational efficiency of human consultants. As the AI handles basic and repetitive inquiries, the burden on staff will be reduced, creating an environment where they can concentrate on high-priority, complex problem-solving. Furthermore, decreased waiting times for residents will improve service quality and reduce overall operational costs, providing a faster and more convenient consultation experience for the foreign community.

Figure 21.1. Multilingual Consultation Scenario



SOCIAL ISSUE 22

Enhancing Welfare Services for Suwon City Residents

KEY TECHNOLOGICAL FEATURES

- **RAG Architecture**: Constructed a Retrieval-Augmented Generation (RAG) architecture utilizing an Embedding Model to enhance response accuracy
- **Hybrid Search & Re-ranking**: Implemented advanced document retrieval and re-ranking through a combination of Dense and Sparse Embedding vectors

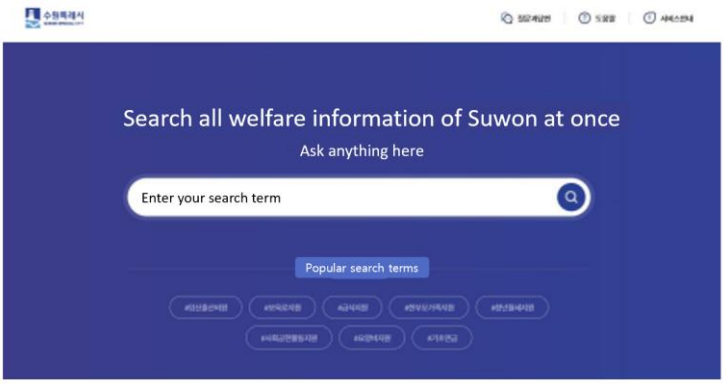
OVERVIEW

Previously, welfare information for Suwon City was fragmented across multiple systems, making it difficult for citizens to find necessary information quickly. This project was initiated to improve welfare-related civil services by consolidating them into a single, integrated service platform.

Table 22. Project Detail

Project Category: Digital Government Infrastructure (2024)
Implementing Agency/Partner Organization: Suwon City Hall, Lifelog Co., Ltd.
Project Description The project established a system capable of sentence-based searches by integrating information on welfare policies, services, and regulations provided by both the central and local governments. Raw welfare policy data was reorganized, and sentence-based search functions were strengthened by analyzing user intent. Additionally, a specialized welfare task support service for government officials was developed to facilitate inter-departmental collaboration and allow for easy verification of legal information.

Figure 22.1. Welfare Service Intelligent Search System



EXPECTED OUTCOMES

By increasing accessibility to welfare services, the project will enhance citizen satisfaction and improve the operational efficiency of government officials, enabling the delivery of rapid welfare administrative services.

3

User Convenience

SOCIAL ISSUE 23

National AI Coaching Solution Pilot Program

KEY TECHNOLOGICAL FEATURES

- **AI-Driven Customized Training**: Utilizes computer vision and motion capture technology for real-time match analysis and motion tracking, supported by machine learning-based predictive models
- **Real-time Data Management**: Manages training data and provides instant feedback through cloud-based mobile platforms

OVERVIEW

With challenges such as the aging of national athletes and a lack of scientific specialization in para-sports leading to a decline in performance, there is a critical need to optimize athlete management and coaching environments. Since 2023, the Ministry of Culture, Sports, and Tourism has supported a consortium of AI specialists to develop and validate AI solutions that provide personalized training programs and data-driven strategic support.

Table 23. Project Detail

Project Category: Ministry Collaboration-Based (2023-2025)
Implementing Agency/Partner Organization: Ministry of Culture, Sports and Tourism (MCST), Simplatform Co., Ltd.
Project Description The project involves collecting and processing match and training data across 11 sports (including badminton, athletics, and taekwondo) to build AI training datasets. Based on this data, the AI models are trained to analyze athlete movements and predict performance. Specifically for badminton, the solution provides: • Analysis of actual practice and match footage. • Personalized training plan establishment. • Strategic support through video analysis of competitors.

Figure 23.1. Before and after implementation (AS-IS vs TO-BE)

Before (AS-IS)	After (TO-BE)
• **Simple Video Storage**: Match and training footage were merely recorded and stored without analytical processing. • **Reliance on Subjective Experience**: Coaching relied heavily on the personal experience and intuition of individual coaches. • **Visual-Based Analysis**: Posture and technical analysis were limited to the naked eye, leading to subjective feedback and inconsistent results.	• **Strategic Pattern Analysis**: AI analyzes shuttlecock trajectories and landing positions to identify competitor patterns and provide customized strategic coaching. • **Skeleton-Based Motion Analysis**: High-speed cameras and skeleton recognition technology analyze swing poses in detail. • **Performance Optimization**: Optimizes athlete movements by reflecting individual physical characteristics and injury risks for maximum efficiency.

EXPECTED OUTCOMES

Over two years of continuous development, the solution achieved a technical milestone of 81.6% accuracy in human detection and tracking. Notably, this AI solution was put into practical use during the 2024 Paris Paralympics for the badminton team.

Looking forward, the introduction of systematic, scientific training methods is expected to boost international competitiveness. This AI-based system will contribute to the overall advancement of para-sports and holds high potential for expansion into general sports training beyond elite athletics.

SOCIAL ISSUE 24

AI Tutor for 'Alpha Campus' S&T Learning Platform

KEY TECHNOLOGICAL FEATURES

- **Knowledge Base Construction**: Trained the chatbot using internal institutional data to build a specialized knowledge domain
- **sLLM-based Hybrid System**: Utilizes a Small Large Language Model (sLLM) integrated with a custom RAG (Retrieval-Augmented Generation) solution. This hybrid approach provides predefined answers for standard FAQs while generating search-based responses for complex queries

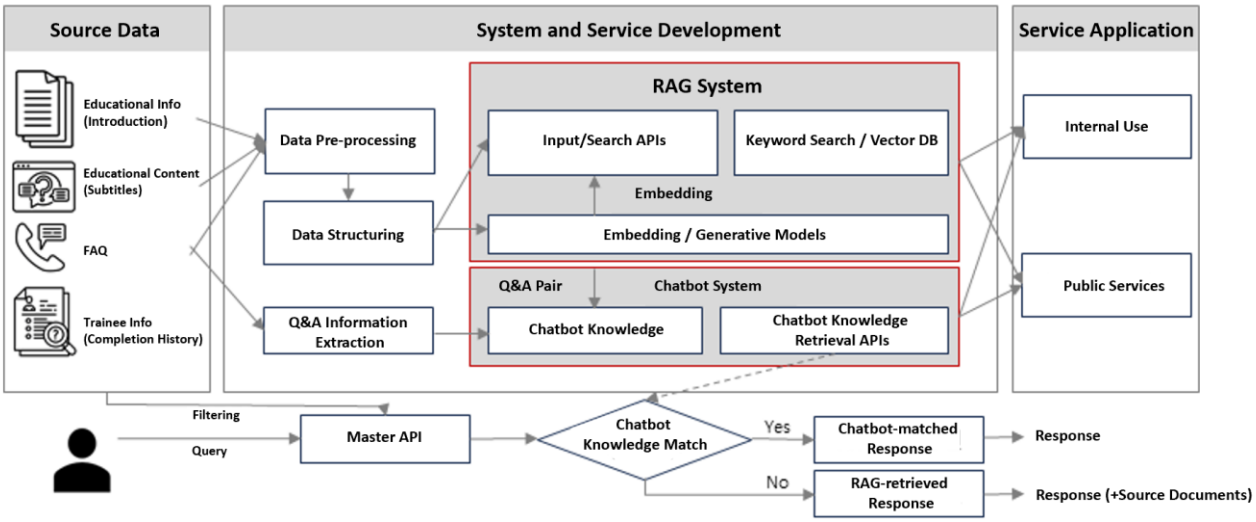
OVERVIEW

KIRD operates 'Alpha Campus,' an online lifelong learning platform for science and technology professionals. Currently, their learner support center faces challenges in real-time processing due to repetitive inquiries and limited ability to handle in-depth questions regarding learning materials. To address this, KIRD conducted a Proof of Concept (PoC) for a generative AI-powered 'AI Tutor' to provide instant solutions, enhance learning immersion, and improve user convenience.

Table 24. Project Detail

Project Category: Hyper-scale AI Service (2024)
Implementing Agency/Partner Organization: Korea Institute of Human Resources Development in Science and Technology (KIRD)
Project Description The project involved digitizing approximately 2,000 pages of data—including education plans, curricula, FAQ lists, and video subtitles for research ethics courses—after de-identifying personal information. This resulted in 5,472 registered knowledge entries. Hybrid Response Logic: - **Predefined Chatbot Knowledge**: Handles standard inquiries such as website usage and completion procedures. - **RAG-based Search**: Handles complex topics like research ethics laws and regulations by retrieving information directly from the database when a direct match is unavailable.

Figure 24.1. Flow chart



EXPECTED OUTCOMES

The AI Tutor makes it easier for professionals to explore courses and enhances learning effectiveness. A key strength is its ability to provide immediate information on research ethics theories and cases by learning from video subtitles. Moving forward, the system will expand to include the latest laws and guidelines. This "One-stop Learning Support Service" acts as both a customer center and a private tutor, supporting self-directed learning and improving the overall usability of the Alpha Campus platform.

SOCIAL ISSUE 25

AI-Integrated Platform for Terms and Conditions Review

KEY TECHNOLOGICAL FEATURES

- **Training Data Construction**: Digitized diverse types of raw data, including civil complaints, terms and conditions (T&Cs), and past review materials
- **LLM-based AI Solution**: Developed an automated platform featuring classification of T&Cs, detection of unfair clauses, and generation of improved sentence drafts to streamline information processing

OVERVIEW

Due to limited personnel and manual-oriented workflows, the Fair-Trade Commission (FTC) has faced challenges such as declining specialization and processing delays in T&C reviews. To address this, the FTC is supporting an AI specialist consortium (since 2024) to develop an AI T&C Review Platform. The goal is to enhance operational efficiency and accuracy, ultimately contributing to the protection of consumer rights.

Table 25. Project Detail

Project Category: Ministry Collaboration-Based (2024-2026)
Implementing Agency/Partner Organization: Korea Fair Trade Commission (KFTC)
Project Description The project leverages the FTC's internal data to train an LLM-based solution capable of categorizing terms and detecting unfair provisions. - For Complainants: Provides guidance to ensure complaints are drafted according to formal requirements. - For Reviewers (Admins): Acts as a decision-support tool by classifying types, flagging unfair clauses, and suggesting revised wording. - For Businesses: Offers a self-check service where companies can cross-reference their T&Cs with similar past cases and complaints to rectify unfair terms voluntarily.

EXPECTED OUTCOMES

By applying AI across the entire review lifecycle—from automatic classification to clause identification and revision suggestions—the platform is expected to significantly boost work efficiency. Ultimately, this will play a vital role in protecting consumer interests and establishing a fair-trade order.

SOCIAL ISSUE 26

Generative AI Agent for Legal Case Analysis Support

KEY TECHNOLOGICAL FEATURES

- **LLM: Solar-Legal-LNC**: Based on Upstage Solar-mini (10.7B), optimized for the Korean language and legal domain
- **Model Training Pipeline**:
  - **Preprocessing:** Text Cleaning (NLP) and Optical Character Recognition (OCR)
  - **Domain Adaptation:** Continuous Pre-training (Further Pre-training)
  - **Post-training:** Supervised Fine-tuning (SFT) and Instruction-Finetuning
  - **Preference Alignment**: Applied DPO (Direct Preference Optimization) for model alignment
- **Advanced RAG**: Integration of Fine-tuning and Graph RAG (Entity-Relationship Extraction & Knowledge Graph Construction)

OVERVIEW

Domestic legal professionals face excessive workloads due to a surge in cases, leading to trial delays and inefficient legal services. By introducing AI-based case analysis and natural language Q&A, repetitive administrative tasks can be automated, allowing legal experts and courts to focus on critical resolution. A generative AI agent-based service is essential to improve the speed and accessibility of legal services.

Table 26. Project Detail

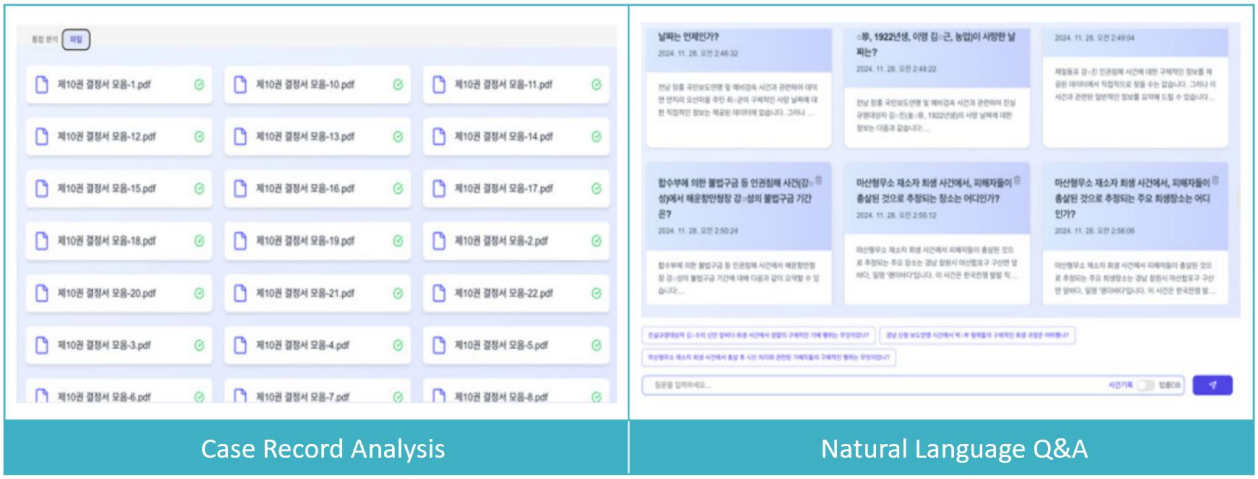
Project Category: Flagship Project (2024-2025)
Implementing Agency/Partner Organization: Law & Company Consortium
Project Description The project established an AI dataset by refining over 4.54 million legal records, including precedents, statutes, and commentaries, to design specialized language models for legal documents and Q&A. These core functions enable the automatic analysis of 8,000 pages per hour and provide natural language Q&A based on case records. By utilizing large-scale legal data and AI technology, the system serves as a decision-support assistant for legal professionals, drastically streamlining the heavy task of analyzing massive case files.

EXPECTED OUTCOMES

The service significantly boosts the operational efficiency of legal experts by reducing work time by over 20% and increasing the number of cases processed by more than 15%. Beyond these productivity gains, it enhances the overall quality of legal services and improves accessibility for the public, thereby contributing to the promotion of national legal welfare.

The developer, Law & Company Consortium, successfully validated the legal case analysis service through a pilot program with the law firm Hwawoo in 2024. Building on this achievement, the consortium plans to officially expand the service to various law firms and legal entities in 2025 by enhancing the system's performance and developing additional specialized features.

Figure 26.1. Service overview



EXPECTED OUTCOMES

The service significantly boosts the operational efficiency of legal experts by reducing work time by over 20% and increasing the number of cases processed by more than 15%. Beyond these productivity gains, it enhances the overall quality of legal services and improves accessibility for the public, thereby contributing to the promotion of national legal welfare.

The developer, Law & Company Consortium, successfully validated the legal case analysis service through a pilot program with the law firm Hwawoo in 2024. Building on this achievement, the consortium plans to officially expand the service to various law firms and legal entities in 2025 by enhancing the system's performance and developing additional specialized features.

SOCIAL ISSUE 27

LLM-Based Regulatory Monitoring and Interpretation Service

KEY TECHNOLOGICAL FEATURES

- **LLM-based Natural Language Processing:** Analyzes vast amounts of legal and regulatory data to provide core summaries and interpretations
- **Real-time Regulatory Monitoring System**: Automatically collects and analyzes official notices of legislative amendments and regulatory changes from government agencies in real time
- **Customized Regulatory Information Delivery**: Provides tailored regulatory summaries and actionable execution guides based on the specific needs of businesses and policy-makers

OVERVIEW

Enterprises and policy-makers must constantly adapt to a shifting regulatory environment, yet interpreting massive legal documents requires significant time and human resources. Existing monitoring methods rely heavily on manual expert reviews, making immediate responses difficult. This service leverages AI-driven automated analysis to provide real-time updates on the latest regulatory changes and offers simplified summaries and interpretations to support efficient policy response.

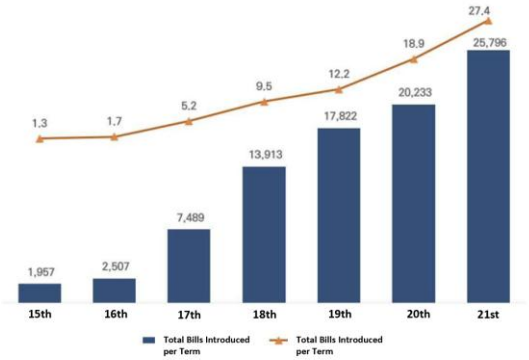
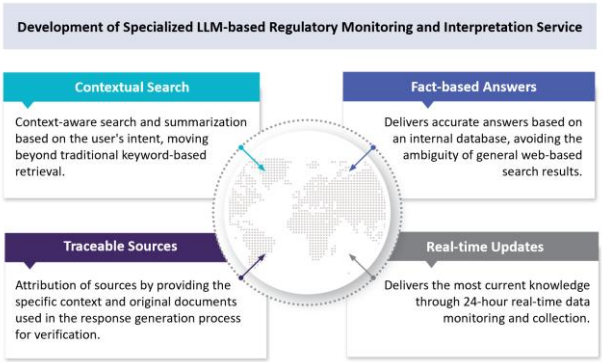
Table 27. Project Detail

Project Category: GovTech Startup Support (2024)
Implementing Agency/Partner Organization: CG INSIDE CO., LTD.
Project Description The system serves as a bridge by automatically collecting and analyzing the latest regulatory data to present it in a user-friendly manner. Using Large Language Models, it breaks down complex legal texts into clear compliance requirements and summaries, enabling stakeholders to respond swiftly to legislative amendments. The platform monitors government portals, public data hubs, and legal information sites to detect changes the moment they occur. Furthermore, it delivers personalized information filtered by industry and specific areas of interest, providing step-by-step execution guides to minimize legal risks and support effective policy establishment.

EXPECTED OUTCOMES

Through this system, businesses and policy-makers can track regulatory shifts in real time and ensure rapid, accurate responses to new requirements. The AI-powered interpretations help users easily grasp complex legal contents, significantly reducing potential legal risks and operational overhead.

Figure 27.1. Before and after implementation (AS-IS vs TO-BE)

Before (AS-IS)	After (TO-BE)
 <Figure> Trends in Bill Introductions by National Assembly Term	
Over 100 daily legislative updates from various authorities, exceeding manual management capacity.	Driving legal market innovation and LegalTech growth through affordable, AI-driven regulatory services.

SOCIAL ISSUE 28

Optimizing the Regulatory Information Search System

KEY TECHNOLOGICAL FEATURES

- Optimization and Reinforcement Learning based on User Feedback
- Korean LLM-based Fine-tuning
- Utilization of X2BEE Base Models: AI Chatbot, AI Search, AI Customer Service Analytics, AI Marketing Copywriter, etc.

OVERVIEW

As the business environment becomes more sophisticated and government regulations grow increasingly complex and diverse, there is a rising demand for a systematic search infrastructure that allows corporations and the public to access regulatory information with ease. This project was planned to enhance the information accessibility and utility of the existing Regulatory Reform Portal (www.better.go.kr) and the National Law Information System.

Table 28. Project Detail

Project Category: Digital Government Infrastructure (2024)
Implementing Agency/Partner Organization: Korea Institute of Public Administration (KIPA), etc.
Project Description The GBC export support platform 'gbcprime' was implemented to provide export consultation chatbots, overseas market information, customized buyer recommendations, best practice guidance, and information on related GBC programs and support initiatives. Considering its Proof of Concept (PoC) status, a representative test utilizing the GBC Prime LA Office database (2022–2024) demonstrated suitability with a 0% error rate across 183 cases involving 50 concurrent users.

Figure 28.1. Contents of the service

EXPECTED OUTCOMES

Through Agile-based LLM development and fine-tuning, prediction uncertainty was reduced by 21.9%, and the user follow-up question rate decreased by 24.6%. By enabling corporations and the public to access regulatory information quickly and accurately, this project has improved regulatory compliance rates and contributed significantly to enhancing administrative efficiency.

SOCIAL ISSUE 29

Personalized Information Platform for Functional Health Foods

KEY TECHNOLOGICAL FEATURES

- **NLP-based Health Functional Food Data Analysis**: Structuring and standardizing health functional food data using Natural Language Processing (NLP), a field of AI that enables computers to understand and process human language
- **Public-Private Data Convergence System**: Integration of Public Data (Open API) and Corporate Databases (Dtree)
- **Health Functional Food Recommendation and Interaction Analysis Algorithm**: Providing customized recommendations based on health status and medication information

OVERVIEW

Despite the rapid growth of the health functional food market, consumers are not provided with sufficient information regarding potential interactions between these products and medicines, nor the proper methods of consumption. Furthermore, the lack of recommendation systems tailored to individual health conditions increases the risk of side effects from indiscriminate intake.

This project utilizes Natural Language Processing (NLP) to standardize health functional food data. By providing personalized healthcare solutions, we aim to support the safe and informed use of health functional foods.

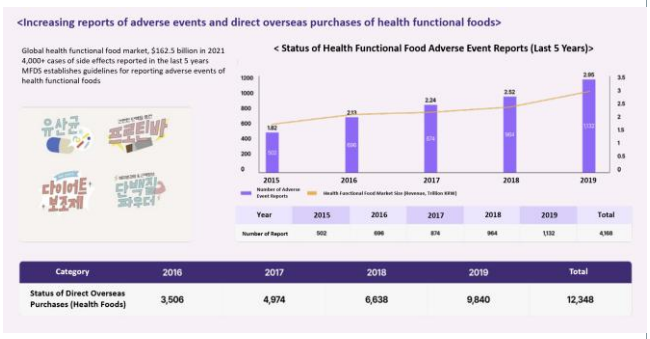
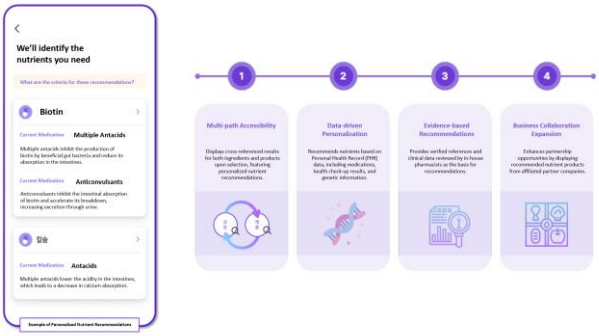
Table 29. Project Detail

Project Category: GovTech Startup Support (2024)
Implementing Agency/Partner Organization: Ones Global Co., Ltd.
Project Description This platform provides health functional food data analysis and personalized health management features, supporting consumers in making safer and more effective choices. By utilizing Natural Language Processing (NLP)—an AI technology that enables computers to understand and process human language—the platform collects and analyzes public and private data to establish a structured database. This system automatically classifies ingredients, functionality, and intake methods while significantly enhancing search capabilities. In particular, the integration of the Ministry of Food and Drug Safety (MFDS) data with the pharmaceutical and disease databases from the developer, Ones Global, allows for a systematic analysis of interactions between health functional foods and medicines, providing highly precise and customized health information. Furthermore, the platform develops personalized recommendation algorithms that reflect a user's health status, current medications, and lifestyle habits, proactively preventing potential side effects by analyzing interaction risks.

EXPECTED OUTCOMES

Through this platform, consumers can make safer and more effective selections of health functional foods, while the provision of interaction data between these products and medicines helps prevent adverse side effects. Furthermore, by offering precise health management solutions that leverage the convergence of public and private data, the platform will significantly improve the quality of personalized healthcare services.

Figure 29.1. Before and after implementation (AS-IS vs TO-BE)

Before (AS-IS)	After (TO-BE)
	
The rise in adverse event reports and overseas direct purchases of health functional foods necessitates personalized recommendation services.	Diversification of access paths, nutrient recommendations based on user variable data, provision of evidence for nutrient recommendations, and provision of product information.

SOCIAL ISSUE 30

Advanced AI Search & Summarization for Smoking Cessation Info

KEY TECHNOLOGICAL FEATURES

- **Establishment of Training Data and Tag Rules**: Constructing Tag Rules based on existing tags to enable document retrieval utilizing tag information
- **Application of RAG Solution (VAIV SearchGPT)**: Transitioning from simple keyword-based search to semantic search (VAIV Search) to improve search accuracy and support natural language inquiries

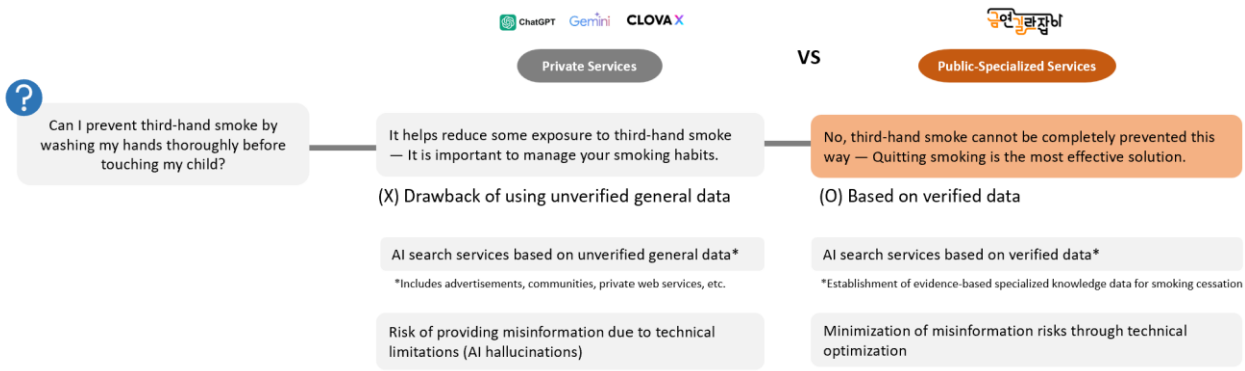
OVERVIEW

The No-smoking Guide (nosmokeguide.go.kr), operated by the Ministry of Health and Welfare and the Korea Health Promotion Institute, is South Korea's representative smoking cessation portal, providing various information and action programs. To meet the rapidly changing digital environment and rising public expectations, this project aims to implement an "Evidence-based Smoking Cessation AI Search Service" powered by Hyper-scale AI technology utilizing accumulated smoking cessation databases. This initiative is designed to enhance user convenience and public satisfaction with online smoking cessation services. Furthermore, it serves as a primary communication channel for national tobacco control policies, providing accurate health information and countering misinformation from the tobacco industry.

Table 30. Project Detail

Project Category: Hyper-scale AI (2024)
Implementing Agency/Partner Organization: Korea Health Promotion Institute (KHPI)
Project Description To train the model for this service, we collected and refined smoking cessation data accumulated since 2001, internal databases, and tobacco harm reports from the Korea Disease Control and Prevention Agency (KDCA). Through data preprocessing, we established 1,093 sets of training data and 1,039 sets of TagRules. Following procedures for smoking cessation data training and generative AI algorithm development, we built a Retrieval-Augmented Generation (RAG) based search service capable of providing customized responses to user inquiries. Starting with a Proof of Concept (PoC) in 2023, we performed optimization of specialized search and generative AI models in 2024, successfully increasing AI search accuracy from 83% to 95%. This service provides intuitive and efficient information retrieval for citizens seeking smoking prevention and cessation information. Based on the latest internal data, the system understands user intent and summarizes core content. When users enter questions regarding tobacco control policies, laws, the risks of smoking, or effective cessation methods, the system provides immediate summaries and an "Original Document View" feature for in-depth exploration.

Figure 30.1. Key advantages compared to existing services



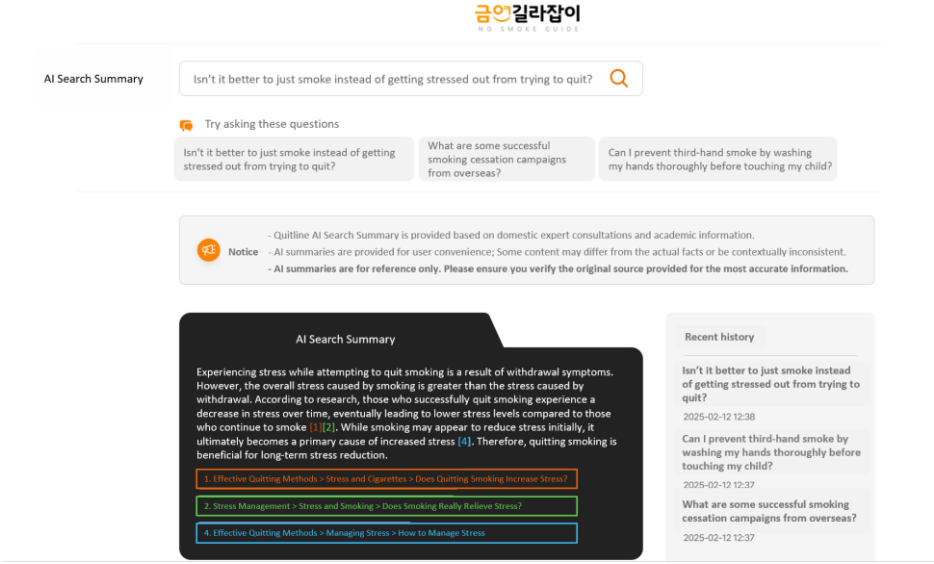
EXPECTED OUTCOMES

For service commercialization, we have completed the development of both user and administrator interfaces. Since late November 2024, the "No-smoking Guide AI Search & Summary" service has been in pilot operation on the No-smoking Guide website. During its first month of operation in 2024, the service recorded 5,430 uses, achieving a high overall satisfaction score of 4.6 out of 5.0.

By moving beyond traditional keyword-based searches and applying generative AI technology, we can provide users with the smoking cessation information they need in an intuitive and efficient manner. This allows us to offer highly accessible smoking cessation support services to smokers motivated to quit, while effectively delivering reliable information to the general public interested in the subject.

Furthermore, by revitalizing outdated services and enhancing their appeal, we expect to encourage more citizens to use the No-smoking Guide service. This will, in turn, increase the effectiveness of national smoking cessation policies and contribute significantly to the promotion of public health.

Figure 30.2. Service interface of Quitline AI-specialized search service



SOCIAL ISSUE 31

Edutech-Based Fitness Measurement & Health Management

KEY TECHNOLOGICAL FEATURES

- **AI-based Physical Fitness Analysis Model**: Analyzing individual fitness measurement data to provide customized health management plans
- **Establishment of IoT Sensor Networks**: Real-time data collection utilizing smart wearable devices and fitness measurement equipment
- **Customized Health Management and Feedback System**: Providing exercise and dietary management solutions tailored to the health status of individual users

OVERVIEW

As physical decline and health-related issues become more prevalent in modern society, the need for personalized health management is steadily increasing. Existing fitness measurement methods often rely on periodic examinations, which have limitations in providing real-time analysis and immediate feedback. This service utilizes AI and IoT technologies to analyze fitness measurements and health management data in real time, aiming to support more effective health maintenance by providing personalized health management solutions.

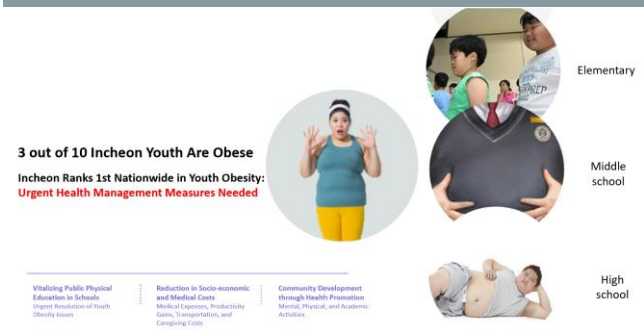
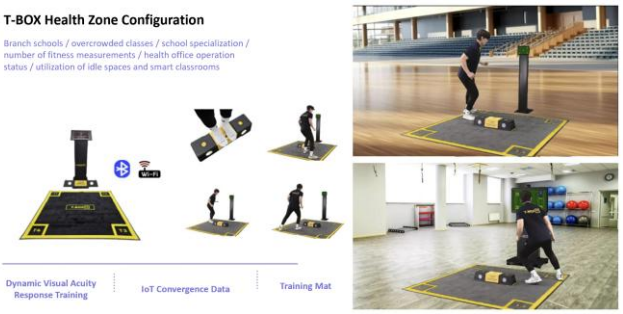
Table 31. Project Detail

Project Category: GovTech Startup Support (2024)
Implementing Agency/Partner Organization: Incheon Facilities Corporation, The Better Exercise Co., Ltd.
Project Description This system collects fitness measurement data and analyzes it to provide optimal health management strategies. By utilizing AI algorithms, the system evaluates the user's fitness data, exercise capacity, and overall health status. Based on these insights, it establishes personalized exercise and health management plans to encourage continuous health improvement. The system also leverages smart wearable devices, such as smartbands and smartwatches, along with fitness measurement equipment to gather real-time data. This allows for the analysis of the user's exercise patterns, heart rate, blood pressure, and oxygen saturation, enabling the early detection of potential health abnormalities. Based on the AI analysis results, the platform offers customized solutions including exercise recommendations, nutritional management, and health goal setting. Furthermore, it effectively supports health promotion by providing continuous feedback tailored to changes in the user's health condition.

EXPECTED OUTCOMES

Through this system, individuals can manage their physical fitness and health more effectively, which is expected to contribute significantly to the improvement of user health and the prevention of chronic diseases.

Figure 31.1. Before and after implementation (AS-IS vs TO-BE)

Before (AS-IS)	After (TO-BE)
	
Socio-economic and medical losses caused by the social issue of youth obesity, such as 3 out of 10 Incheon youth being obese, and the need for transformation in physical education.	Vitalizing public physical education, reducing social, economic, and medical costs, and contributing to the health promotion of youth in the local community.

SOCIAL ISSUE 32

AI Persona-Based Cultural Content Matching

KEY TECHNOLOGICAL FEATURES

- **Personalized Recommendation Service via GIS Visual Map Based on AI Algorithms**: Providing tailored recommendations through an interactive GIS visual map powered by AI recommendation engines
- **Development APIs**: Developing and providing APIs for cultural facility visitor demographics, basic facility information, local accommodations/restaurants/menus, credit card spending patterns, exhibition schedules, and integrated processed data for recommendations

OVERVIEW

Currently, enjoying cultural and artistic services requires individuals to manually search for information on museums, art galleries, and performance halls or rely on promotional materials from blogs and social media. However, personalized recommendation services remain insufficient, and small-scale cultural service providers often face difficulties in actively disseminating information through promotional activities. To address these challenges, we developed a nationwide platform that provides information on cultural facilities, exhibitions, and performances using an AI persona-based visual map. This project involved developing personalized cultural content matching services and releasing various APIs related to cultural facilities to the public.

Table 32. Project Detail

Project Category: Digital Government Infrastructure (2024)
Implementing Agency/Partner Organization: Human & Forest Co., Ltd., etc.
Project Description To establish a comprehensive database, we verified the validity of 9,839 cultural facilities nationwide, ultimately selecting 7,178 target sites and defining effective zones for extracting visitor population data. We generated master data for location information and constructed analytical datasets focusing on visitor trends and surrounding commercial districts. This included data on local accommodations, restaurant menus, credit card affiliate information, and consumer spending interests. Additionally, to power the recommendation service, we integrated performance data from the Korea Culture Information Service, the Korea Performing Arts Box Office Information System (KOPIS), and other cultural content sources. To develop the AI persona-based matching service, we defined specific personas by analyzing cultural consumer and content data, leading to the creation of a customized AI recommendation service (cultureplus.kr) based on integrated processed data.

Figure 32.1. Home menu detailed features

▶ **Home Menu Detailed Features**

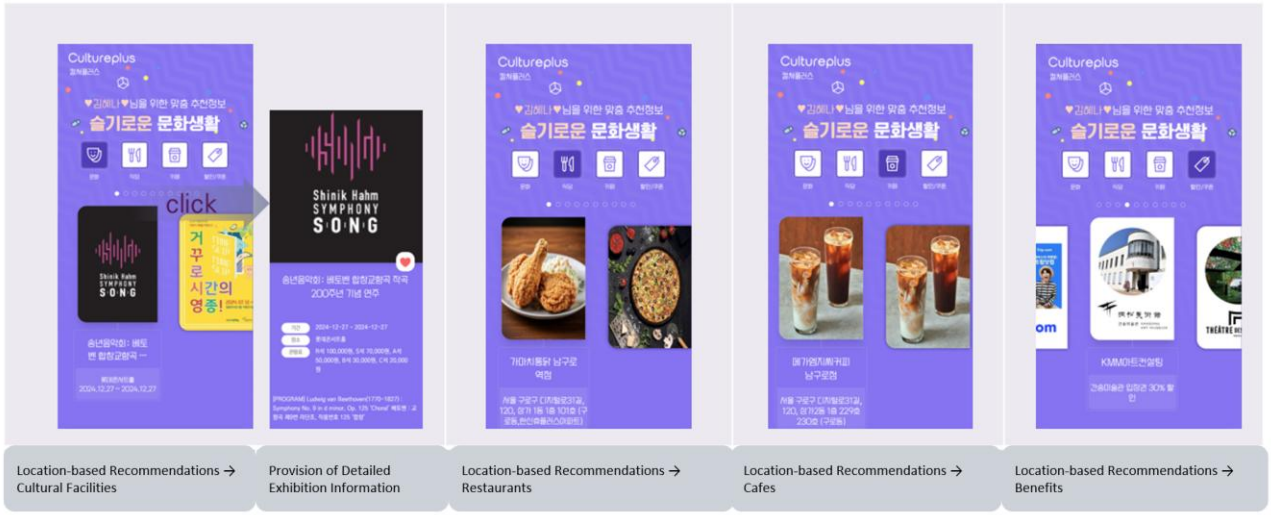
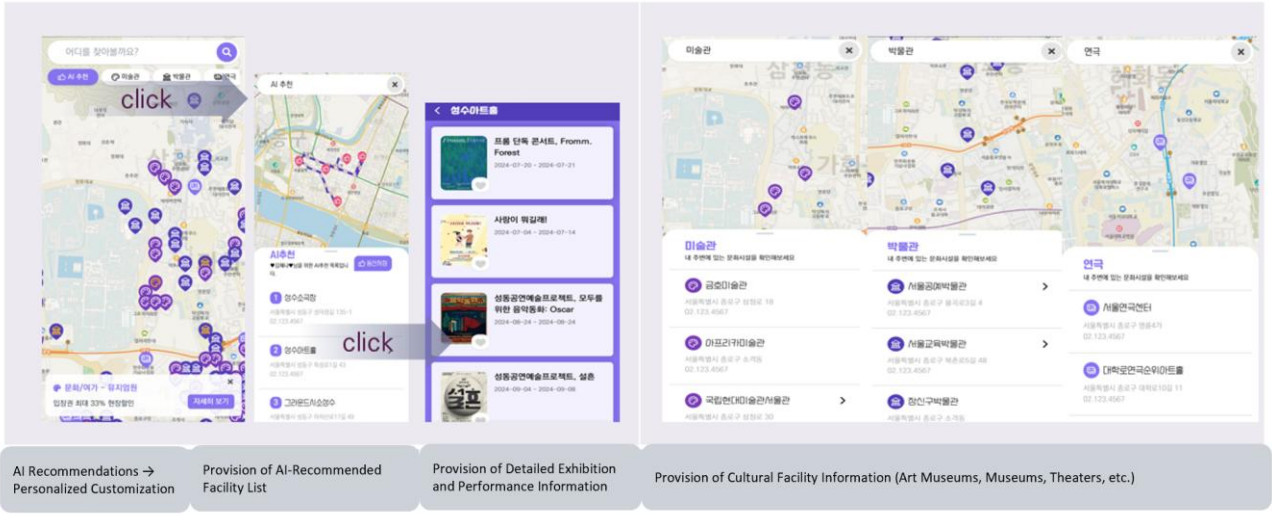


Figure 32.2. Map menu detailed features

▶ **Map Menu Detailed Features**



EXPECTED OUTCOMES

The AI persona-based matching service strengthens the connection between cultural consumers and providers while enhancing accessibility to cultural information, thereby increasing cultural diversity and broadening the options available to the public. Furthermore, by linking tourism elements such as sightseeing, dining, and entertainment, the service is expected to increase regional visits and generate significant economic impact.

SOCIAL ISSUE 33

Customized Exhibition Environment for the National Museum

KEY TECHNOLOGICAL FEATURES

- **Personalized Exhibition Recommendation Service based on AI Algorithms**: Providing tailored exhibition recommendations by analyzing user preferences and behavior
- **Development APIs**: Developing and releasing APIs for visitor trends, household trends, demographics (gender/age), visits by day and time, integrated visitor data, and customized exhibition recommendation data.

OVERVIEW

While the demand for leisure activities has increased with the emphasis on work-life balance, the digital transformation of art museum operations—a core pillar of civic leisure and culture—remains insufficient. Identifying visitor characteristics is essential for providing customized exhibition and educational content; however, even representative institutions like the National Museum of Modern and Contemporary Art (MMCA) lack sophisticated customer management systems that reflect granular visitor traits. This situation is even more challenging for smaller public and private museums. To address these gaps, this project aimed to develop comprehensive APIs that can scientifically track and utilize visitor data.

Table 33. Project Detail

Project Category: Digital Government Infrastructure (2023)
Implementing Agency/Partner Organization: Human & Forest Co., Ltd., etc.
Project Description The project focused on the development and release of three core API categories. First, we established the Visitor Statistics API per Exhibition, which provides detailed statistics for each exhibition at MMCA by utilizing telecommunications data (LG Uplus) and CCTV analysis data. Second, the Visitor Trend Statistics API was developed in conjunction with telecommunications data to aggregate statistics based on detailed visitor profiles, including gender, age, household type, lifestyle-based movement patterns, and preferred media content types. Finally, we developed the Personalized Exhibition Recommendation API, which generates tailored exhibition lists by considering the specific tastes and living radius of the users. These APIs are designed to be open for wider institutional use.

Figure 33.1. API Development list

No.	API (Korean)	API (English)	Detailed Features (English)
1	Cultural Facility Visitor Trends	getVisitorPastYear	Visitor counts for the past 1 year
2	Cultural Facility Visiting Household Trends	getHouseholdVisitsPastYear	Trend and share of visiting households for the past 1 year
3	Visitor Gender/Age Trends	getVisitorAgeGender	Visitor counts by gender and age for the past 6 months
4	Visitor Trends by Day of the Week	getVisitorAgeDayOfWeek	Visitor counts by day of the week for the past 6 months
5	Visitor Trends by Hour	getVisitorHourly	Visitor counts by the hour for the past 6 months
6	Visitor Aggregated Data	getVisitorDataAggregation	Aggregated data of cultural facility visitors
7	Personalized Exhibition Recommendation Data	getVisitorrecommendation DataAggregation	Aggregated data for personalized exhibition recommendations

Figure 33.2. Web Interface for Personalized Exhibition Recommendations at the National Museum of Modern and Contemporary Art (MMCA)

EXPECTED OUTCOMES

By scientifically collecting and analyzing data on viewing patterns and exhibition environments, this service fosters a personalized cultural experience, which is expected to significantly enhance public cultural welfare and attract new audiences. This data-driven approach will drive service innovation to satisfy the public's cultural needs. Furthermore, by supporting evidence-based decision-making for museum policies, the system will help secure the validity of national museum operations, revitalize museum management, and contribute to the sustainable qualitative and quantitative growth of national cultural institutions in the long term.

SOCIAL ISSUE 34

Sentiment Analysis Module for User-Friendly Tourist Information

KEY TECHNOLOGICAL FEATURES

- **Development of AI-based Sentiment Analysis Modules**: Selecting sentiment-specific keywords and establishing a sentiment analysis API
- **Application of SearchGPT-based Search Services**: Supporting sentiment analysis of social data related to tourist attractions using SearchGPT and providing search results integrated with social data analytics

OVERVIEW

General integrated search services typically provide results sorted only by recency or accuracy, which highlights the growing need for more user-friendly service models. There is a demand for sorting methods based on sentiment analysis—such as ranking results by positive or negative sentiment—and ensuring that documents of high public interest are displayed at the top. To address this, the Korea Tourism Organization (KTO) aims to demonstrate a user-friendly tourist information service by applying sentiment analysis modules, thereby providing an environment where these technologies can be effectively applied to social data.

Table 34. Project Detail

Project Category: Hyper-scale AI service (2023)
Implementing Agency/Partner Organization: Korea Tourism Organization (KTO)
Project Description The system utilizes information about tourist attractions from blogs, lifestyle, and cultural news, classifying data into eight specific categories: Satisfaction, Taste, Price, Scenery, Service, Convenience, Transportation, and Cleanliness. Each category is analyzed for positive and negative keywords; for example, if a keyword like "atmospheric" or "scenic" is detected in the Scenery category, it is classified as positive. To implement this, we selected sentiment-based keywords to develop an analysis module and created a demo page where these modules can be tested in connection with existing social data.

EXPECTED OUTCOMES

By selecting sentiment keywords for key tourist information and applying analysis modules, we have provided a search service that is both user-friendly and highly relatable to the public. This initiative is expected to contribute to a virtuous cycle of increasing tourist demand and fostering cultural development at various attractions. Looking ahead, the service has high potential for expansion into various fields, such as providing translation services for foreign tourists or designing personalized regional travel planners.

SOCIAL ISSUE 35

Immigration Solution with On-Device AI Document Reading

KEY TECHNOLOGICAL FEATURES

- **Multispectral Visual Intelligence Technology**: Detecting forgeries in public documents by utilizing Ultraviolet (UV) and Infrared (IR) imaging
- **AI-based Automated Inspection System**: Authenticating public documents and detecting forgeries in passports and identification cards
- **Blockchain-based Data Security**: Selecting sentiment-specific keywords and establishing a sentiment analysis API
  Application of SearchGPT-based Search Services: Securely storing decoded public document information on a blockchain to enhance data integrity and reliability

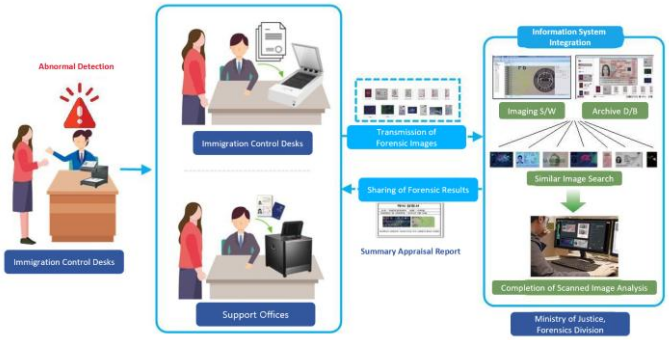
OVERVIEW

Rapidly and accurately identifying forged passports and identification cards is critical during the immigration clearance process. Traditional visual inspection methods face limitations due to increasingly sophisticated forgery techniques, and their heavy reliance on manual labor hinders the speed of the inspection process. This project aims to develop an innovative solution that combines AI-based automated inspection systems with multispectral imaging technology to swiftly analyze public documents for forgeries and revolutionize the immigration clearance procedure.

Table 35. Project Detail

Project Category: GovTech Startup Support (2024)
Implementing Agency/Partner Organization: Logichain Inc.
Project Description This system offers a range of features designed to improve the reliability of document reading and enhance the efficiency of immigration clearance. First, by applying multispectral visual intelligence technology, the system analyzes security features in passports and IDs using UV and IR imaging to detect forgeries. It is engineered to ensure stable reading performance under various lighting conditions. Second, an AI-based automated inspection system has been implemented, utilizing deep learning-based image analysis models to precisely inspect security elements within public documents and automatically identify forged patterns. Furthermore, the system integrates facial recognition technology to verify the identity of the person against their passport photo. Finally, a blockchain-based data security system has been established where document data read during immigration is stored on a blockchain, fundamentally preventing any possibility of data tampering. This ensures the construction of a highly credible inspection system.

Figure 35.1. Before and after implementation (AS-IS vs TO-BE)

Before (AS-IS)	After (TO-BE)
	
Current products and services are limited to detecting forgery at specific points or determining authenticity through the insight of external experts, such as by visual inspection or scanning holograms and hidden text.	Using AI-based pattern extraction and anomaly detection technology, it automatically verifies the authenticity of documents and reduces associated costs.

EXPECTED OUTCOMES

This system is expected to accelerate the speed of immigration clearance and improve the accuracy of forged document detection. Moreover, by introducing blockchain-based data security technology, the project will contribute to establishing an internationally trusted immigration system.

SOCIAL ISSUE 36

Integrated Testbed Based on Private Cloud

KEY TECHNOLOGICAL FEATURES

- **Establishment of Private Cloud-based Infrastructure**: Building a robust foundational infrastructure leveraging private cloud resources
- **Support for MLOps, Data Analytics, and Hyper-scale AI**: Facilitating AI model development through integrated MLOps environments, advanced data analysis, and Hyper-scale AI utilization
- **Efficient Development via SaaS-based AI Platforms and LLM Coding Assistants**: Enhancing development efficiency by employing SaaS-type AI platforms and LLM-powered coding assistance tools

OVERVIEW

With the acceleration of digital transformation and the integration of AI into daily life, there is a growing necessity for everyone to develop innovative technologies and services based on private-sector capabilities. However, small and medium-sized enterprises (SMEs) and ventures, which often face more challenging ICT environments compared to large corporations, struggle to secure the infrastructure needed to develop, test, and verify advanced technologies. Furthermore, individual citizens and startups with creative ideas often face barriers in creating innovative services due to a lack of familiarity with planning, coding, and testing. Consequently, there is a vital need for a foundational environment where SMEs, startups, citizen developers, and public institutions can freely develop, test, and verify Digital Government Infrastructure advanced technologies and services.

Table 36. Project Detail

Project Category: Digital Government Infrastructure (2024)
Implementing Agency/Partner Organization: Ministry of Science and ICT (MSIT), CloIT Co., Ltd., etc.
Project Description The Integrated Testbed consists of three core pillars designed to facilitate the free development and validation of innovative services: Digital Resources, Development Support Tools, and Professional Technical Support. First, Digital Resources provide various APIs, including public and private data, and support multi-cloud environments (CPU, GPU, storage, etc.). Second, Development Support Tools include software development suites (DevOps, No-Code, Low-Code, etc.) and platforms for data analysis and AI training. Third, Professional Technical Support offers educational opportunities and technical consulting in areas such as AI/ML, Big Data, Hyper-scale AI, and Cloud-Native technologies. Together, these elements create an environment where anyone can develop and verify innovative services.

Figure 36.1. Service development support list

Digital Resources	Development Support Tools	Expert Technical Support
♦ Digital resources including private/public data and service functions (APIs) ♦ Support for private multi-cloud services such as CPU/GPU and storage	♦ Data storage, processing, and analysis, as well as AI model training and development tools (MLOps) ♦ Support for LLM-based coding and private hyperscale AI development tools	♦ Technical support for the development, testing, verification, and deployment of innovative technologies and services ♦ Support for specialized commercialization consulting, Agile coaching, and technical training

Figure 36.2. Definition of integrated testbed & landscape

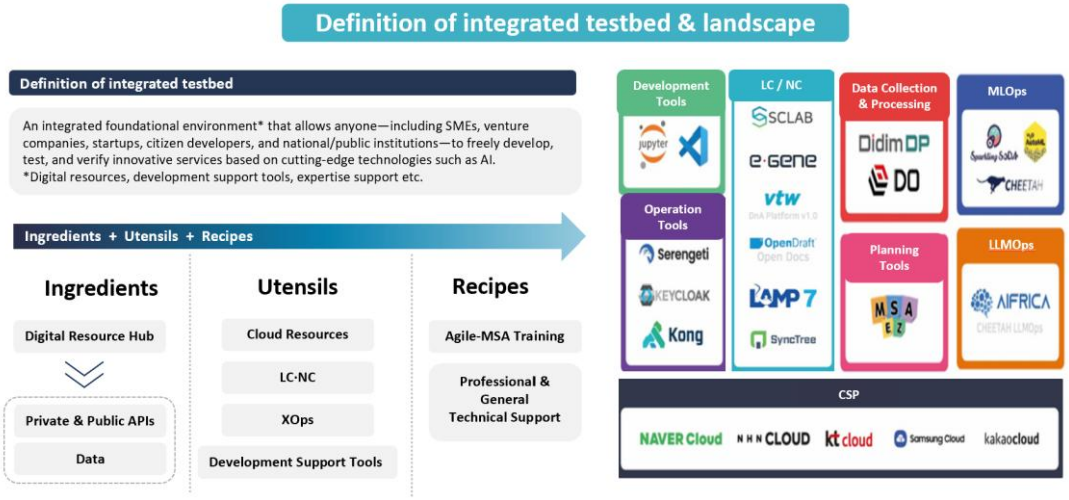


Figure 36.3. Contest (Challenge · Hackathon) Awards Ceremony

EXPECTED OUTCOMES

To solve national social issues and promote public innovation by combining digital resources with AI, the "Challenge 2024" and "Hackathon 2024" were held using the Integrated Testbed. Moving away from traditional contest formats, these events provided coaching and mentoring from experts at AI and ICT companies to ensure that identified ideas could be developed into optimal AI services. As a result, 10 business models, such as a tailored AI real estate recommendation service, and 11 innovative services, including a customized job search service for seniors, were developed. Additionally, a total of 233 technical support cases were completed, encompassing professional/general technical assistance, coaching, education, and business matching. While SMEs, startups, and citizen developers previously lacked the environments to develop and verify innovative services, the Integrated Testbed is expected to empower anyone to create innovative, tangible services for the public.

SOCIAL ISSUE 37

Intelligent Water Administration

KEY TECHNOLOGICAL FEATURES

- **Transmission of Self-Metering Data via AI-powered Optical Character Recognition (OCR)**: Enabling the transmission of self-metering data by recognizing digits on water meters through AI OCR technology
- **Implementation of Water Usage Prediction and Anomaly Detection**: Applying a localized model trained on water meter data based on the Azure Vision model to predict usage and detect abnormal patterns

OVERVIEW

As water metering data has traditionally been managed manually, difficulties have arisen in billing and handling civil petitions. To address this, an intelligent water administration service was planned to integrate water bill inquiry, payment, and usage monitoring. This service implements an optimal algorithm by merging the core Azure Vision AI model with specialized local datasets to modernize administrative processes.

Table 37. Project Detail

Project Category: Digital Government Infrastructure (2024)
Implementing Agency/Partner Organization: Yangpyeong County Office, NICE Zinidata Co., Ltd.
Project Description Yangpyeong County introduced a mobile-based "Self-Metering Function," allowing residents to verify their own water consumption and pay bills directly. Furthermore, prediction models and anomaly detection functions were developed to enable public officials to easily monitor water usage status. These tools allow for the early identification of potential leaks or meter malfunctions by analyzing usage patterns against predicted data.

Figure 37.1. Service overview & key development status

EXPECTED OUTCOMES

By improving the water metering and billing management processes, the system is expected to reduce administrative costs and significantly enhance convenience for both water users and public officials. The transition to an automated, AI-driven system ensures higher data accuracy and faster response times for water-related administrative service

USE CASES OF PUBLIC AI SERVICE

Improving the Way of Working

1

Administration

SOCIAL ISSUE 1

Hyper-scale AI-Driven Integrated Research Support Services

KEY TECHNOLOGICAL FEATURES

- **Hyper-CLOVA X-based Platform**: Utilization of CLOVA Studio to develop specialized LLM (Large Language Model) services for research assistance
- **Data Infrastructure**: Construction of a dedicated Data Lake for research data and implementation of a RAG (Retrieval-Augmented Generation) based search system

OVERVIEW

The need to improve research quality has emerged by creating a Generative AI-based environment that minimizes the significant human and time resources typically spent on literature review and data analysis. An integrated research support service based on a Hyper-scale AI platform will provide the foundation for an efficient research environment. This is expected to contribute to national development and social value by enabling timely responses to major national tasks and social issues.

Table 1. Project Detail

Project Category: Hyper-scale AI service (2024)
Implementing Agency/Partner Organization: National Research Council for Economics, Humanities and Social Sciences (NRC), NAVER Cloud Corporation
Project Description Researchers can engage in dialogues with the AI regarding new research topics and methodologies. They can utilize auxiliary services such as document review, translation, and the generation of meeting minutes or presentations throughout the lifecycle of research planning, execution, and performance dissemination. • **Proposal**: Suggests new themes through research exploration and topic validation. • **Efficiency**: Enhances focus and quality via report generation and brief summarization, enabling better data analysis and resource forecasting. • **Collection**: Gathers diverse materials through similarity analysis with related studies, assisting in the sharing of research outcomes. The service aims to innovate the research environment by reducing the time researchers spend on policy studies, which currently averages over a year. Regarding Research Planning, the platform supports decision-making by suggesting research topics and methodologies and generating preliminary research proposals. For Research Execution, it provides services such as research proposal analysis, report summarization, and similarity mapping with related studies. Additionally, researchers can access Auxiliary Support features like automated meeting minutes and grammar/context error detection. Finally, to facilitate the Dissemination of results, the system offers verification of research regulations and the automated generation of presentation materials.

Figure 1.1. Service overview

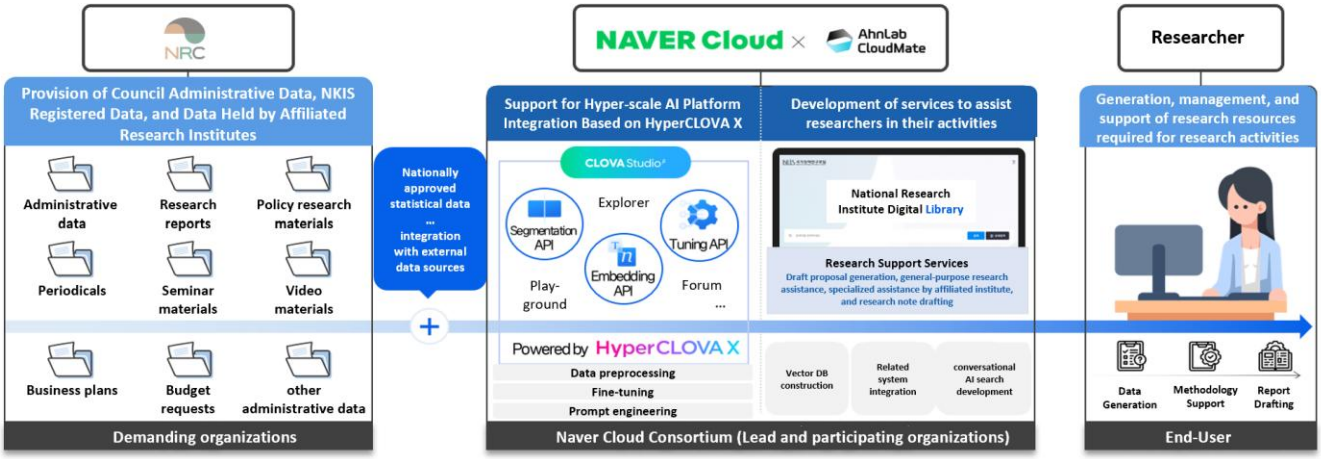
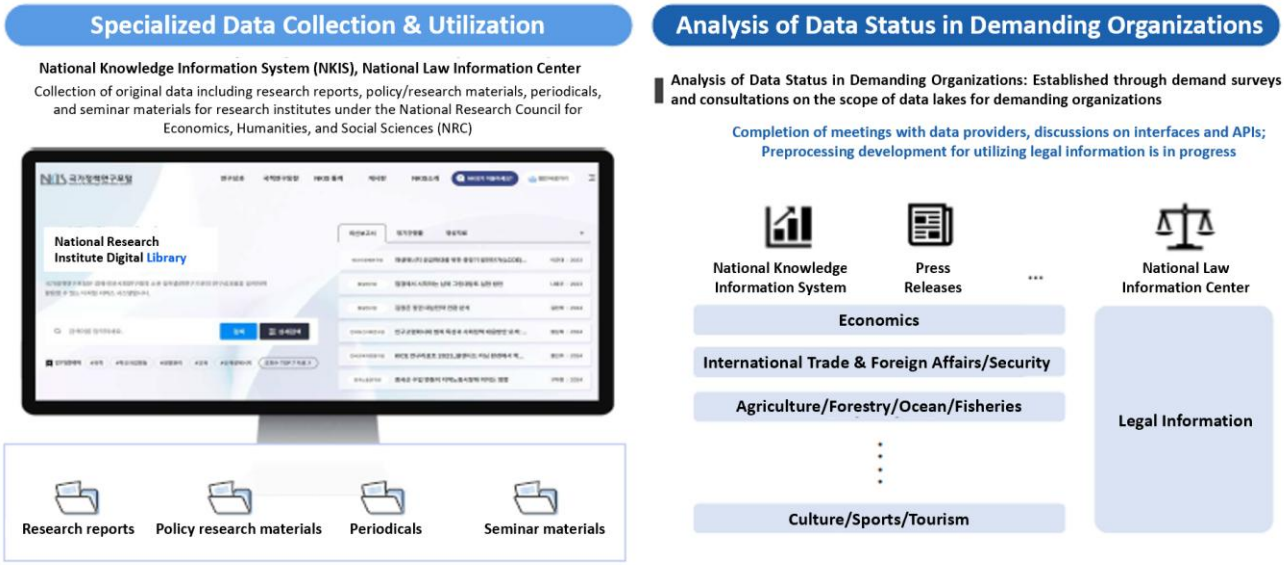
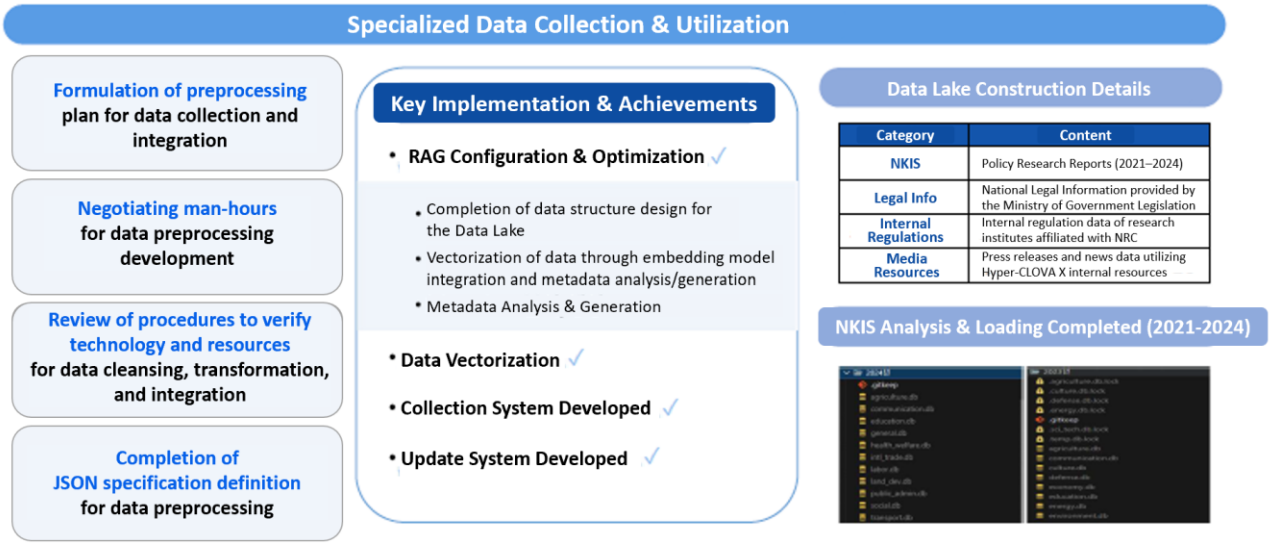


Figure 1.2. Data integration & AI service architecture



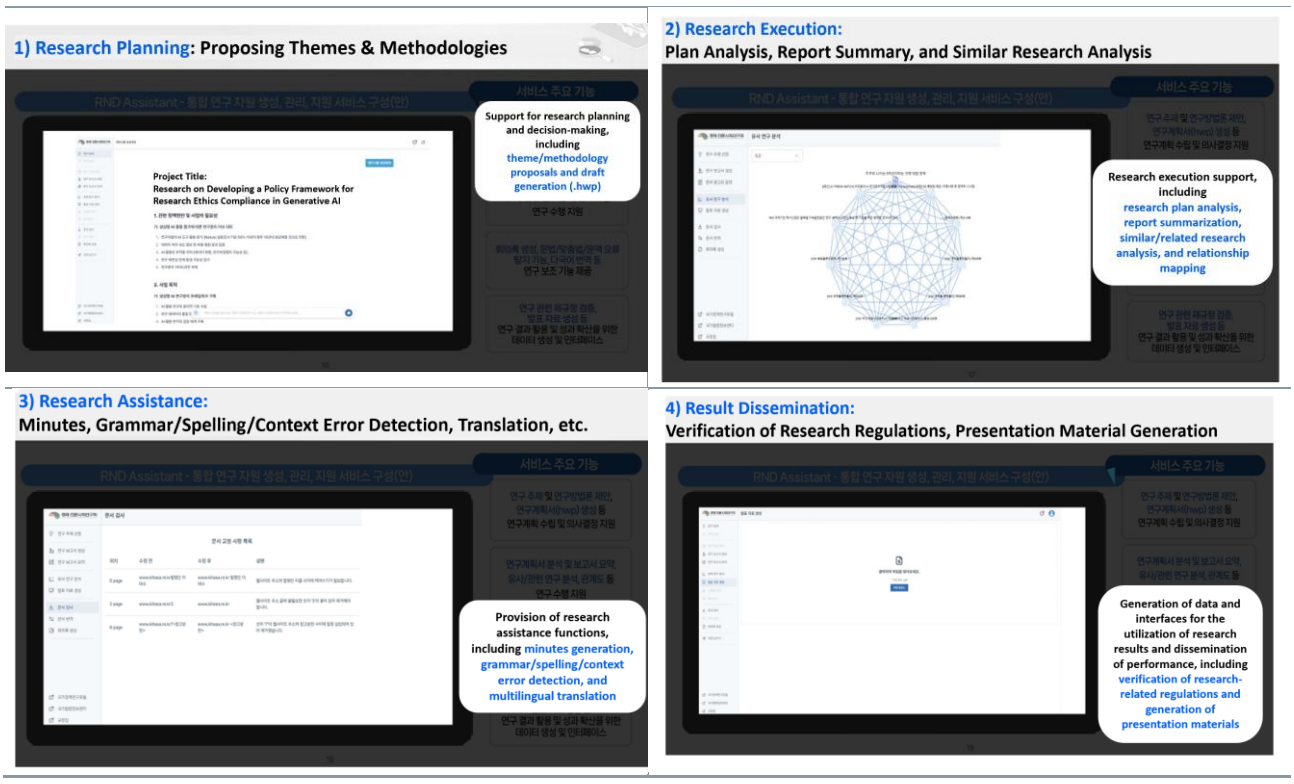
The project involved developing a conversational service interface and establishing cloud infrastructure. A Data Lake was constructed and used for training/retrieval, incorporating data from the National Knowledge Information System (NKIS), the National Law Information Center, internal regulations of 11 affiliated research institutes, and news articles.

Figure 1.3. Specialized data collection and utilization



- **Data Collection & Pre-processing**: Original data—including research reports, policy materials, periodicals, and seminar papers—is collected and processed. Through a "Data Pre-processing Pipeline," the team completed RAG design/optimization testing, data vectorization, and the development of systems for data collection and real-time updates.

Figure 1.4. Full-cycle AI research assistance

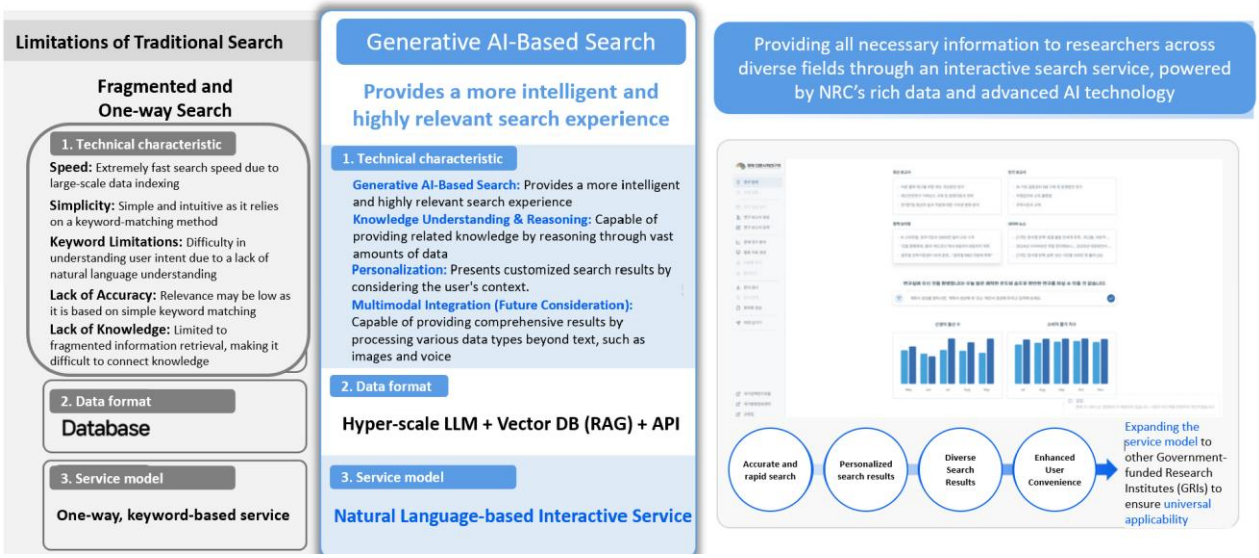


EXPECTED OUTCOMES

The provision of an integrated research support service using Hyper-scale AI significantly improves research efficiency. By automating data analysis and processing, researchers can save time and focus more on core analysis and strategic thinking. Building an environment where vast amounts of research data can be learned through AI improves the overall quality of research results and contributes to better policy decision-making. Furthermore, the digital platform makes collaboration between researchers from different fields easier, and convergence research becomes possible through AI trained on multifaceted data.

In addition, resources such as manpower and time can be utilized more effectively. Hyper-scale AI allows for the preemptive filtering of similar or redundant research, thereby shortening the overall research period. Through a standardized learning data and service structure, the resources required for repetitive data training and service construction are minimized, ensuring a sustainable and cost-effective research ecosystem.

Figure 1.5. As-is vs To-be



SOCIAL ISSUE 2

Risk Management Solution for National R&D Project Spending

KEY TECHNOLOGICAL FEATURES

- **AI-Based Expense Management Chatbot**: Provides automated responses and guidance by learning R&D laws and ministry-specific guidelines
- **Automated Data Collection & Document Generation**: Collects R&D expenditure data and automatically generates required administrative forms

OVERVIEW

Traditional management of national R&D project expenses has faced challenges, including a high administrative burden on researchers and institutions, and operational inefficiencies caused by complex regulations and evidentiary requirements. This solution aims to alleviate the administrative workload of researchers and enhance the accuracy and efficiency of budget execution by utilizing AI and automation technologies.

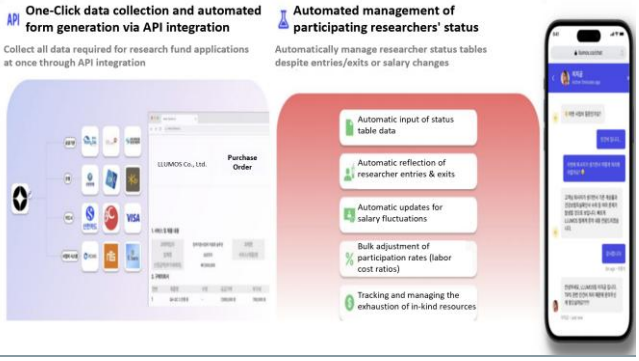
Table 2. Project Detail

Project Category: GovTech Startup Support (2024)
Implementing Agency/Partner Organization: LLUMOS Co., Ltd.
Project Description This solution includes comprehensive functions to efficiently manage the execution of national R&D project expenses while minimizing the administrative burden on researchers. Through the AI-powered expenditure support chatbot, researchers can easily verify laws, guidelines, and evidentiary requirements related to budget execution. The system provides immediate responses to user inquiries and simplifies administrative procedures through automated guidelines. Furthermore, the solution automatically collects data necessary for expense execution and generates required documentation. This integrated approach reduces human error and maximizes the overall efficiency of administrative workflows.

EXPECTED OUTCOMES

The implementation of this solution is expected to significantly enhance the efficiency of national R&D project expense management and greatly reduce the administrative burden on researchers. By streamlining these processes, researchers can focus more on their primary research and development activities, while the transparency and accuracy of budget execution are improved. Ultimately, the system provides a foundation for more accountable and productive research environments by eliminating the friction of complex financial administration.

Figure 2.1. Before and after implementation (AS-IS vs TO-BE)

Before (AS-IS)	After (TO-BE)
	
Existing commercial accounting ERPs cannot provide the necessary evidence and forms required for National R&D project administration	Reduction in rejections due to human error, decrease in fund execution time, and saving of research administrative costs

SOCIAL ISSUE 3

Data Analytics for Policy Proposal Support

KEY TECHNOLOGICAL FEATURES

- **LLM Training Dataset**: Collection and pre-processing of training data, including the conversion of 10 sets of Q&A pairs using a PDF Annotation Tool
- **Retrieval Engine Implementation**: Implementation of a search engine using research reports within the NKIS (National Knowledge Information System) as metadata
  GPT Fine-tuning: Fine-tuning of MaumGPT (13B parameters) using research reports from NKIS focused on disaster and safety themes

OVERVIEW

Researchers unfamiliar with programming languages often face difficulties in performing in-depth analysis for future forecasting or policy support. To improve the utility and efficiency of analyzing vast amounts of raw research reports, there is a need for a service that can autonomously analyze and learn from data to generate optimized results, such as statistics and visualizations. This project aims to demonstrate a service by implementing a GPT-based chatbot capable of providing natural language answers based on the research outputs stored in the NKIS.

Table 3. Project Detail

Project Category: Hyper-scale AI service (2023)
Implementing Agency/Partner Organization: National Research Council for Economics, Humanities, and Social Sciences (NRC)
Project Description For model training, raw research report data was sourced from the NRC Data Information System (NDIS), the National Knowledge Information System (NKIS), and the Policy Research Information Service Management (PRISM). The technical procedure involved establishing a retrieval engine based on training data and performing fine-tuning on the MaumGPT model. Subsequently, the retrieval engine and the fine-tuned MaumGPT model were integrated using proprietary middleware (Orchestra). This service is specifically designed to support researchers in future forecasting and policy-making by enabling the large language model to autonomously process vast datasets and present optimized analysis results, including statistical data and visual materials.

EXPECTED OUTCOMES

This project successfully demonstrated the creation of a GPT-based chatbot capable of generating answers directly from the original texts of research reports. By utilizing the developer's specialized tools to support workflow creation and UI integration, the project proved the feasibility of a full commercial service launch. Furthermore, the implementation of a retrieval engine effectively addresses the "hallucination" issues common in standard chatbots. By overcoming these existing limitations, the service is expected to become a highly practical and reliable tool for researchers, providing accurate information based on verified policy and research data.

SOCIAL ISSUE 4

Institutional Research Data Utilization for Policy Support

KEY TECHNOLOGICAL FEATURES

- **SearchGPT/SmartChat-Based Service**: Learning and indexing pre-processed data within SearchGPT and generating queries based on SmartChat
- **Hybrid Integration**: A service approach for the chat interface that combines the respective strengths of search engines and chatbot solutions through a hybrid integration method

OVERVIEW

The Korea Federation of SMEs (KBIZ) sought to introduce AI technology to discover policy indicators through internal data searches and to provide similar reference materials when responding to requests from institutions and members. Accordingly, this project aimed to verify the feasibility of a SearchGPT/SmartChat-based service to support policy-related tasks using key internal documents such as reports, news articles, and legal data.

Table 4. Project Detail

Project Category: Hyper-scale AI service (2023)
Implementing Agency/Partner Organization: Korea Federation of Small and Medium Business (KBIZ)
Project Description To train this service model, 859 queries were generated using a dataset consisting of 1,963 documents, 10,000 news articles, and 1,963 statutes. The model training procedure began with the pre-processing of internal data, followed by the generation and review of queries based on SmartChat and the indexing of the processed data. Subsequently, a demo page was developed to verify the integration of the trained model with the service interface. This system is designed as a search and chatbot service utilizing major internal document data, specifically aimed at supporting staff members during the policy proposal process.

EXPECTED OUTCOMES

This service supports the discovery of policy indicators for the Korea Federation of SMEs, the proposal of new policies, and the drafting of press releases by learning from key research materials. By converting existing statistical information from table formats into structured text, the system allows staff to easily secure various quantitative data points through simple searches. Furthermore, instead of reviewing the entire content of reference materials like research reports, staff can engage in Q&A with the chatbot to find specific evidence for policy formulation. This is expected to significantly reduce the time required for data processing and improve overall administrative efficiency.

SOCIAL ISSUE 5

AI-Powered Patent Examination Support System

KEY TECHNOLOGICAL FEATURES

- **Embedding Search**: Development of Context-aware Chunking capable of understanding and segmenting context for domain-specific indexing optimization
- **RAG (Retrieval-Augmented Generation)**: LLM training to mitigate hallucinations
  - **Pre-training:** Additional learning of domain-specific expressions and vocabulary to understand specialized technical knowledge
  - **Instruction-Tuning:** Fine-tuning based on instructions expected by professional domain users
  - **Alignment-Tuning:** Optimization based on human preferences using RLHF (Reinforcement Learning from Human Feedback) and DPO (Direct Preference Optimization)

OVERVIEW

The Korean Intellectual Property Office (KIPO) currently ranks fourth globally in patent application volume. With applications in key 4th Industrial Revolution sectors growing at an average annual rate of 13.5%, the increasing complexity of examinations necessitates an improved workflow environment. To overcome the high costs of infrastructure and initial investment required for Large AI development, KIPO collaborated with domestic companies leading the Generative AI sector to build a specialized support system.

Table 5. Project Detail

Project Category: Hyper-scale AI service (2024)
Implementing Agency/Partner Organization: Ministry of Intellectual Property (MOIP), Saltlux, Inc.
Project Description Due to the highly specialized nature of intellectual property, a consortium was formed with Saltlux and GenIP. Saltlux managed data collection, pre-processing, model training, and UI development, while GenIP provided legal counsel on IP law, ensured regulatory compliance of refined data, and validated the model's accuracy. The collaboration resulted in four core functions: Chatbot-based Examination Q&A, Automated Search Query Generation, Similar Trial/Judicial Precedent Search, and Written Opinion Summarization. The Examination Q&A function uses RAG to provide conversational answers within a defined scope of patent laws, examination guidelines, and KIPO publications. It prevents hallucinations by providing the original source text alongside the answer. The Search Query Generation function analyzes existing examination patterns to extract keywords and expand them through LLM-driven synonym/antonym mapping. This standardizes the "prior art search" process, which previously relied on individual expertise and apprenticeship-style learning. For the Similar Precedent Search, a dual structure of keyword and vector search (utilizing cosine similarity and Nearest Neighbor algorithms) allows examiners to find relevant trial records accurately. Finally, the Written Opinion Summarization tool allows examiners to upload lengthy submissions from applicants to receive concise summaries and engage in a follow-up Q&A to ensure no critical points are overlooked.

EXPECTED OUTCOMES

As application volumes and technical complexity rise, there is increasing social pressure for higher examination quality. Current dissatisfaction among applicants often stems from inconsistent standards regarding clarity or inventive steps across different examiners. The tools developed in this project will enhance the efficiency and quality of work for examiners and trial judges by providing fast, accurate information. Specifically, the search query generation tool will serve as a standardized guideline for new examiners, reducing the learning curve.

Furthermore, AI support will minimize subjective errors in judgment, thereby increasing the consistency and reliability of the entire examination and trial process. This is expected to lead to higher satisfaction among the public and patent applicants. Through continuous system refinement and advancement, KIPO aims to build a more efficient patent examination environment, strengthening national intellectual property competitiveness and solidifying South Korea's status as a global leader in the IP sector.

Figure 5.1. Examination Q&A

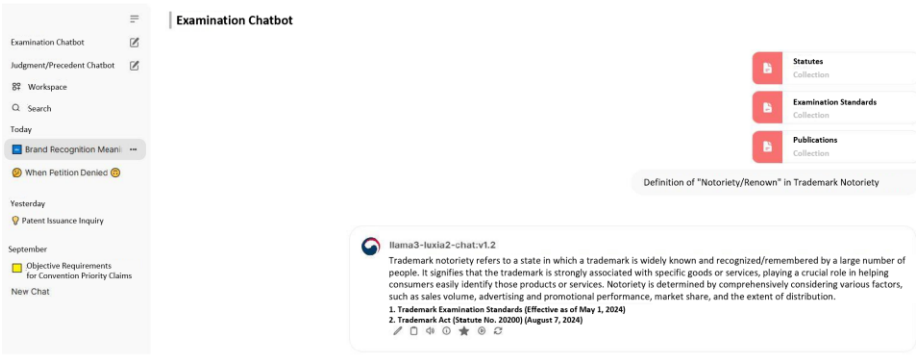


Figure 5.2. Similar prior art patent search screen

Figure 5.3. Search screen for similar trial decisions and judgments

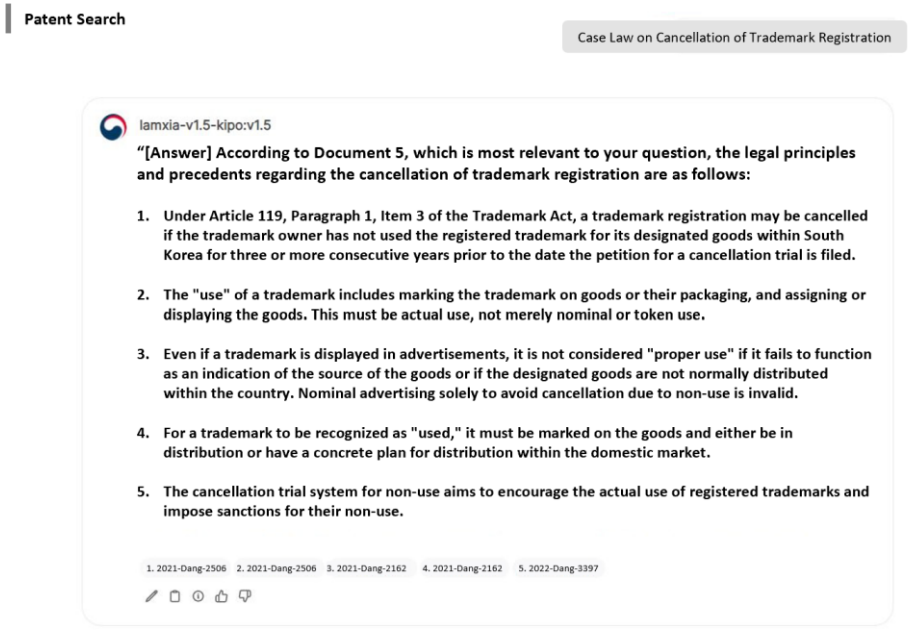


Figure 5.4. Opinion summary screen

SOCIAL ISSUE 6

Hyper-scale AI-Based Smart Firefighting Administration

KEY TECHNOLOGICAL FEATURES

- **LangChain-based RAG Model**: Implementation of a Retrieval-Augmented Generation model using LangChain, featuring data refinement, embedding, vector store integration, and search optimization based on Cosine Similarity
- **High-Performance Vector DB**: Construction of a Vector Database using advanced indexing techniques to enhance search performance and handle large-scale, high-dimensional data
  Blueprint Analysis Model: Utilization of the YOLOv8 computer vision engine for object detection, with image pre-processing and data transformation performed via the Roboflow tool

OVERVIEW

The field of fire safety and building permits requires a complex, expert-level review of building blueprints, including area, floor height, and usage, based on numerous laws and technical manuals. However, maintaining expertise is challenging due to personnel rotations, and citizens often struggle with the intricate permit procedures. To address this, the Seoul Metropolitan Fire & Disaster Headquarters developed a specialized Hyper-scale AI model for the fire and construction sectors. By providing fire administration assistants and blueprint image analysis services, the system ensures professional expertise, resolves difficulties in mastering complex manuals, and creates an environment where fire safety facilities can be inspected conveniently.

Table 6. Project Detail

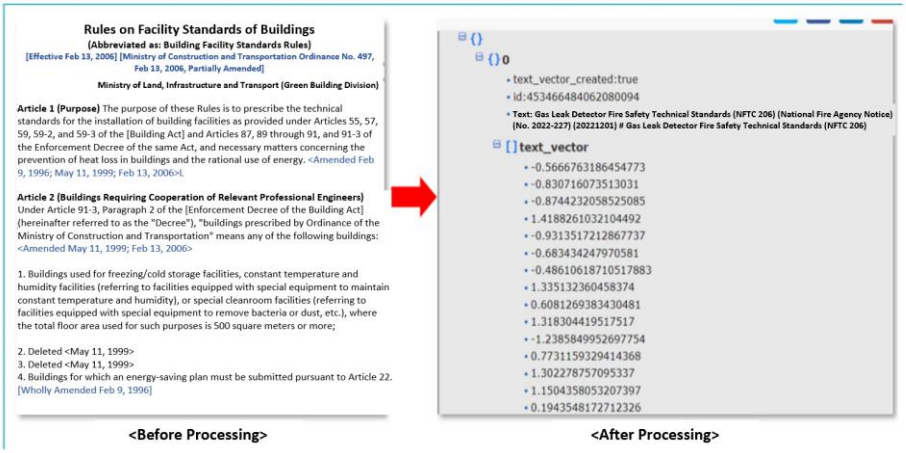
Project Category: Hyper-scale AI service (2024)
Implementing Agency/Partner Organization: Seoul Metropolitan Fire & Disaster Headquarters, Ailis Frontier Co., Ltd.
Project Description To train the specialized model, the project secured a vast dataset including 40 statutes, 164 fire safety standards, 17,990 civil complaint records, 31 manuals, and 9,559 fire safety blueprints. Specifically, 536,390 data points from laws and manuals were processed for RAG training, while 10,092 fire safety symbols from 1,026 blueprints were precisely labeled for image analysis. The Fire Safety RAG Chatbot generates answers based on 42 types of fire and building regulations to eliminate hallucinations, enabling officials to provide professional responses to civil inquiries. It also includes a checklist of fire safety facilities categorized by area, usage, and number of floors to prevent omissions in the permit process. The Blueprint Analysis Model employs a two-track strategy: a "Symbol Extraction Model" for 20 standardized symbols and a "Symbol Selection Model" for 320 non-standardized symbols, ensuring practical utility across diverse blueprint styles.

EXPECTED OUTCOMES

The implementation of this service has revolutionized the review process—which previously required examining up to five types of reference materials (blueprints, installation schedules, laws, manuals, etc.)—shortening the average processing time per case from 6 hours to just 5 minutes. The system achieved a user satisfaction score of over 80 points among officials in 25 fire-related departments, contributing to easing the hurdles and avoidance associated with building permit tasks.

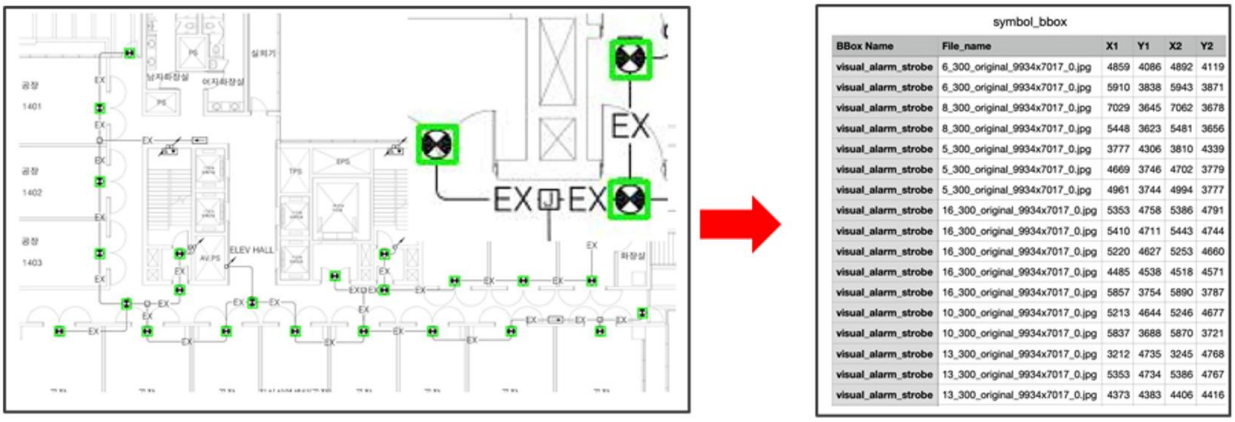
Since this service is applicable to fire agencies nationwide, KIPO plans to expand its reach to improve administrative efficiency across the country and minimize confusion by providing consistent legal standards and information. Furthermore, the blueprint analysis function is expected to be extended to private construction companies, local government departments, and facility maintenance firms, fostering a more robust safety management ecosystem.

Figure 6.1. Comparison of legal data before and after processing



<Before Processing> <After Processing>

Figure 6.2. Drawing Symbol Data Labeling



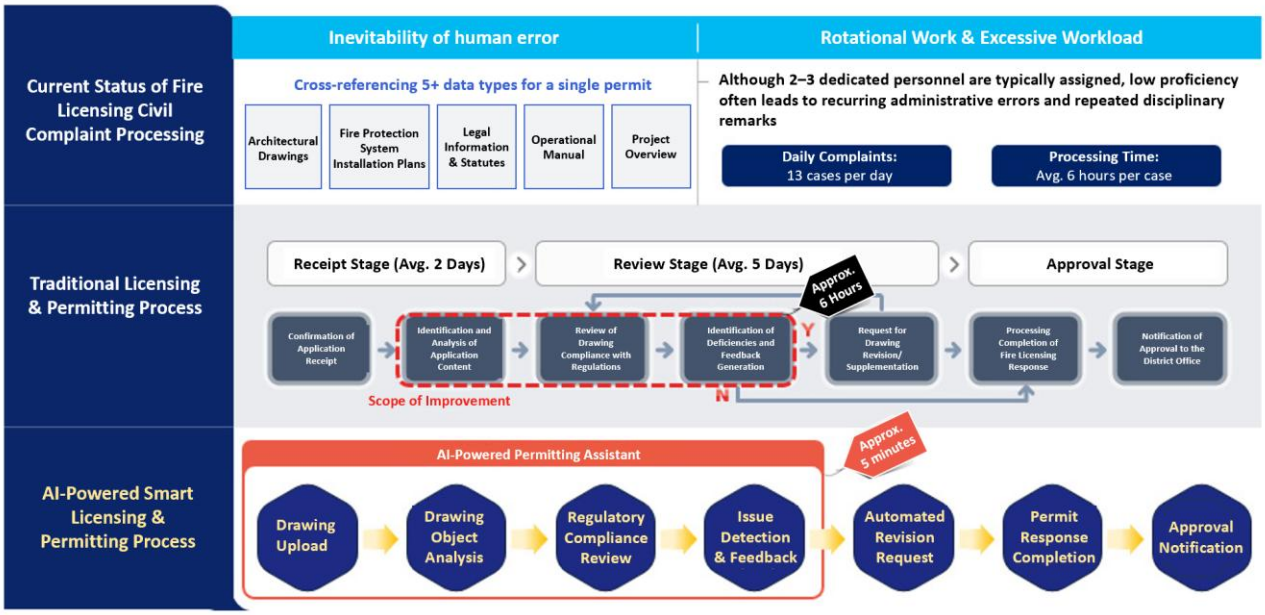
symbol_bbox

BBox Name	File_name	X1	Y1	X2	Y2
visual_alarm_strobe	6_300_original_9934x7017_0.jpg	4859	4086	4892	4119
visual_alarm_strobe	6_300_original_9934x7017_0.jpg	5910	3838	5943	3871
visual_alarm_strobe	8_300_original_9934x7017_0.jpg	7029	3645	7062	3678
visual_alarm_strobe	8_300_original_9934x7017_0.jpg	5448	3623	5481	3656
visual_alarm_strobe	5_300_original_9934x7017_0.jpg	3777	4306	3810	4339
visual_alarm_strobe	5_300_original_9934x7017_0.jpg	4669	3746	4702	3779
visual_alarm_strobe	5_300_original_9934x7017_0.jpg	4961	3744	4994	3777
visual_alarm_strobe	16_300_original_9934x7017_0.jpg	5353	4758	5386	4791
visual_alarm_strobe	16_300_original_9934x7017_0.jpg	5410	4711	5443	4744
visual_alarm_strobe	16_300_original_9934x7017_0.jpg	5220	4627	5253	4660
visual_alarm_strobe	16_300_original_9934x7017_0.jpg	4485	4538	4518	4571
visual_alarm_strobe	16_300_original_9934x7017_0.jpg	5857	3754	5890	3787
visual_alarm_strobe	10_300_original_9934x7017_0.jpg	5213	4644	5246	4677
visual_alarm_strobe	10_300_original_9934x7017_0.jpg	5837	3688	5870	3721
visual_alarm_strobe	13_300_original_9934x7017_0.jpg	3212	4735	3245	4768
visual_alarm_strobe	13_300_original_9934x7017_0.jpg	5353	4734	5386	4767
visual_alarm_strobe	13_300_original_9934x7017_0.jpg	4373	4383	4406	4416

In-drawing symbol labeling

Matching symbols, file names, and coordinate values

Figure 6.3. Improving the fire & construction licensing and permitting process



SOCIAL ISSUE 7

Generative AI for Military Facility Construction Administration

KEY TECHNOLOGICAL FEATURES

- **Specialized Infrastructure**: Construction of a Smart Database for architectural regulations and a domain-specific sLLM (small Large Language Model)
- **Advanced RAG Integration**: Implementation of a dedicated RAG (Retrieval-Augmented Generation) framework to seamlessly combine the legal database with the sLLM

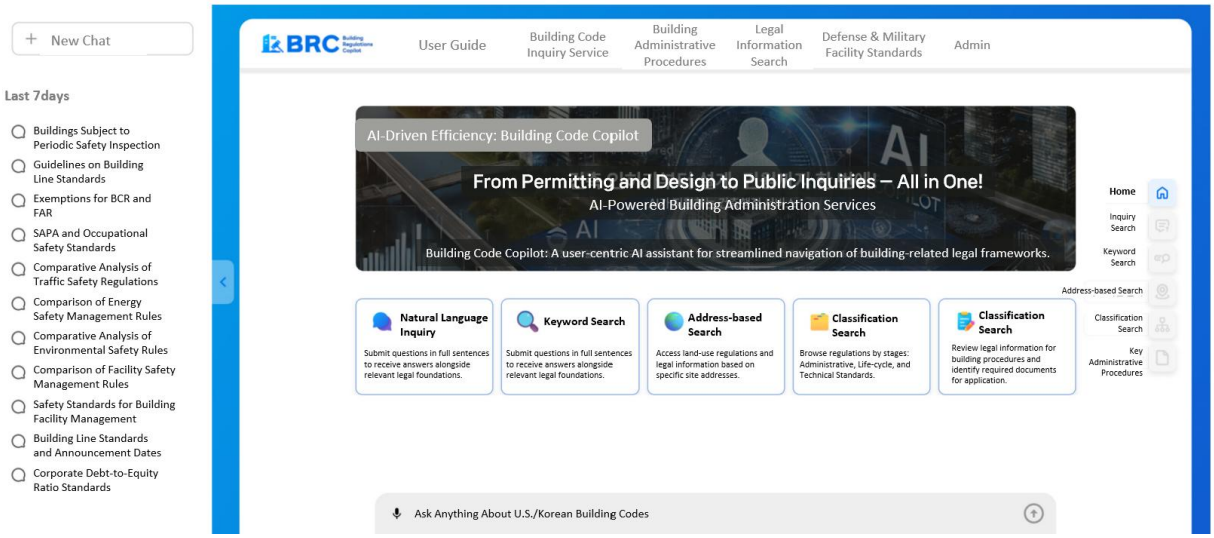
OVERVIEW

The Ministry of National Defense (MND) has faced frequent legal disputes and accident risks due to unauthorized or unregistered buildings within military bases. Coupled with the professional limitations of administrative staff and a surge in civil complaints, a systematic solution became essential. To resolve this, the MND developed the "Architectural Regulation Copilot," which integrates scattered laws, administrative rules, ordinances, and design standards. By providing user-centric, easy-to-understand legal guidance to various stakeholders, this service aims to eliminate confusion during the permit and construction process while enhancing the efficiency of architectural administration.

Table 7. Project Detail

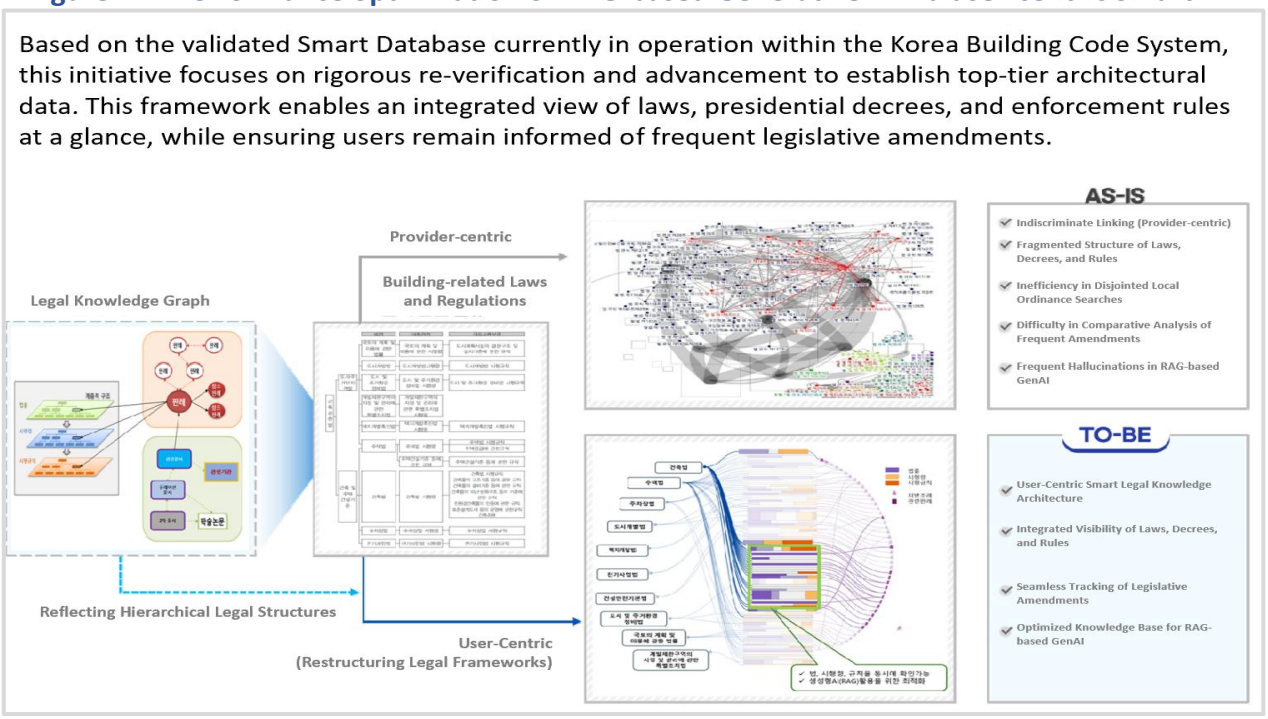
Project Category: Hyper-scale AI service (2024)
Implementing Agency/Partner Organization: Ministry of National Defense (MND), SQISOFT Co., Ltd.
Project Description To optimize AI model training, the current legal-centric architectural regulations were restructured into a user-centric classification system—categorized by administrative rules, technical standards, and building lifecycle stages. Data is managed at the specific level of articles and clauses to enhance embedding accuracy. Furthermore, unstructured civil complaint data, including records from the e-People (Kookmin Shinmun-go) portal, was processed for supplementary training. The core development includes three specialized modules: "Architectural Admin AI" for permit support, "Architectural Design AI" for design and construction tasks, and "Architectural Complaint AI" for public assistance.

Figure 7.1. Service overview



Through this Copilot service, permit officers receive administrative support, practitioners obtain assistance in design, construction, and supervision, and general users benefit from streamlined civil complaint responses.

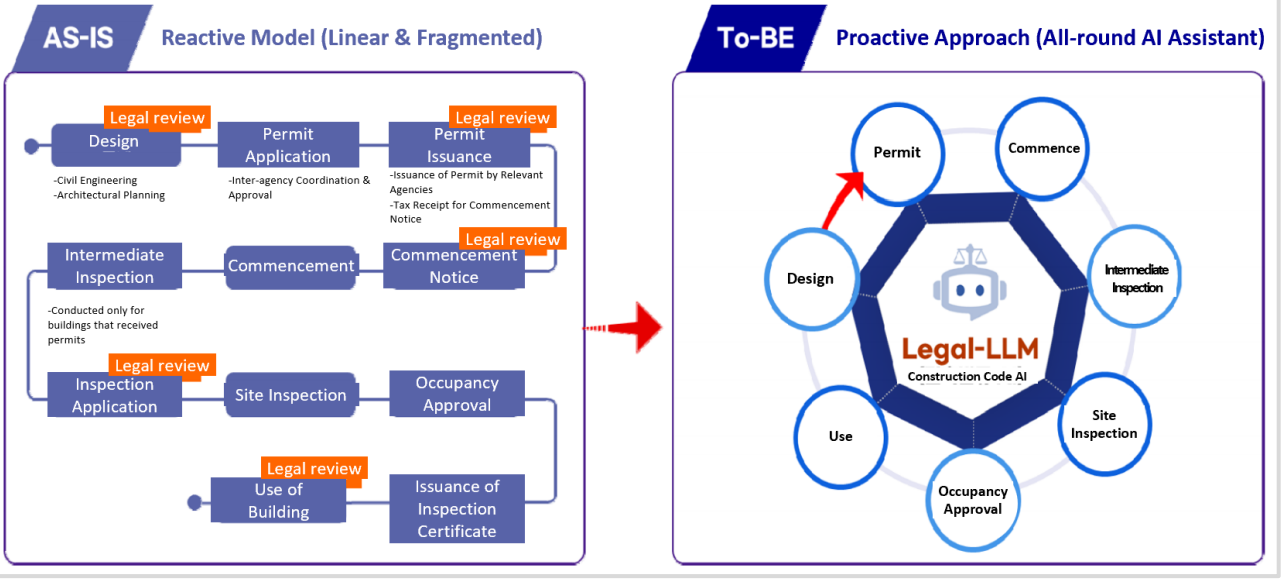
Figure 7.2. Performance optimization of RAG-based Generative AI via user-centric Smart DB



The system is built upon a verified "Smart DB" from the Korea Building Regulation system, allowing users to view laws, enforcement decrees, and regulations at a glance and stay updated on frequent revisions.

Figure 7.3. From Reactive Penalties to Proactive Compliance

Shifting from a reactive approach—where identifying complex legal errors often leads to post-construction penalties (fines, enforcement levies)—to a proactive prevention model. This service provides instantaneous legal verification throughout the entire lifecycle, assisting both government agencies and citizens in real-time.



Unlike traditional post-processing methods—which often resulted in fines or enforcement levies due to overlooked regulations—this service provides a preventative legal review at every stage of the construction process, enabling both institutions and citizens to verify compliance in real-time.

EXPECTED OUTCOMES

The development of this Generative AI-based service efficiently organizes complex architectural laws and provides regulation services through a user-friendly classification system. This allows for easier identification of restrictions and improved management systems, preventing operational confusion even when non-experts are involved in administrative procedures. In particular, the ability to verify regulatory compliance across the entire lifecycle—from design to permit approval—is expected to drastically reduce administrative lead times.

The MND anticipates significant quantitative improvements, including reducing the administrative cancellation rate from over 10% to below 5%, increasing budget execution rates by more than 3%, and shortening processing times by over 30%. Local governments are also expected to see administrative burdens lightened, with processing and complaint response times reduced by 20% and 30%, respectively. By providing diverse services including legal interpretations, precedents, and Q&A, the system has bolstered trust between public institutions and private users. Aligned with the Defense Innovation 4.0 plan, this Copilot service will support highly specialized new projects, ensuring a safe environment through strict regulatory compliance and enhancing public trust in military facilities.

Regarding the expansion plan for this service, a directive has been issued to the Army, Navy, Air Force, Marine Corps, and units directly under the Ministry of National Defense (MND)—which are

responsible for facility project requirements exceeding 5 trillion KRW annually—as well as to intermediate review authorities. These entities are instructed to utilize the system and submit improvement requests. The consolidation of these key issues and feedback is scheduled to be completed within the first half of 2025.

To further advance the service, the MND is currently securing budget and establishing operational strategies in connection with the "Next-Generation Integrated Defense Facility Information System Project," which is set to launch in 2025. This integration will ensure the long-term sustainability and technological sophistication of the Architectural Regulation Copilot within the broader defense infrastructure.

SOCIAL ISSUE 8

AI-Integrated Customs Video Management Solution

KEY TECHNOLOGICAL FEATURES

- **Image Separation & Storage Service**: AI automatically separates overlapping items within X-ray images and visually organizes them by matching each item with its declared cargo information
- **Image Tracking & Merging Service**: Provides a real-time integrated view on a single screen, combining separated image data with AI analysis results
- **Image DB Search Service**: High-speed AI-driven search to identify similar cargo among tens of millions of X-ray images stored in the existing database
  Real-time Alert Service: AI trained on hazardous item imagery identifies similar shapes in real-time and issues warnings to users upon detection

OVERVIEW

The explosive growth of overseas direct purchases has led to a surge in X-ray image data at customs. This has raised concerns regarding decreased accuracy in customs clearance due to the accumulated fatigue of human inspectors using traditional visual methods. Furthermore, missed images have caused persistent data inconsistency issues. To ensure the reliability and speed of the clearance process, the Korea Customs Service has been supporting an AI specialist consortium since 2023 to develop and demonstrate AI solutions for the automatic separation, tracking, and searching of X-ray images.

Table 8. Project Detail

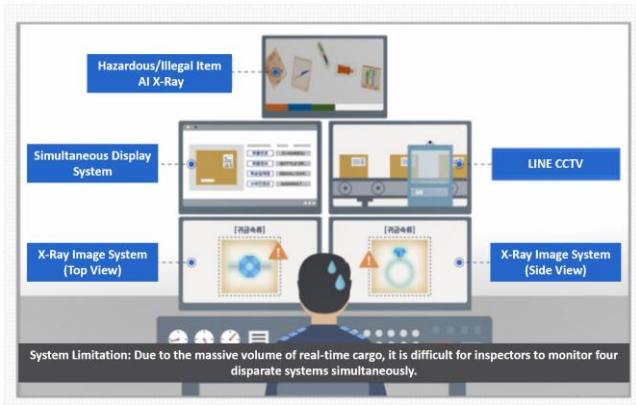
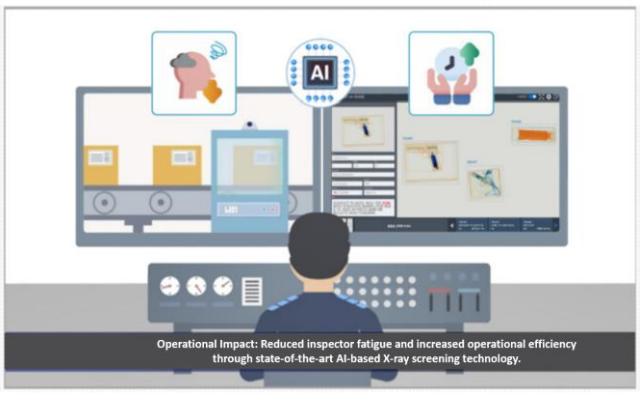
Project Category: Ministry Collaboration-Based (2023–2025)
Implementing Agency/Partner Organization: Korea Customs Service
Project Description The AI-integrated customs image management solution was developed based on large-scale X-ray imagery and cargo declaration data held by the Korea Customs Service. Over 11 million X-ray data points were processed and utilized to develop and validate AI solutions focused on three core functions: image separation/storage, tracking/merging, and DB search. This AI solution encompasses a variety of sophisticated features. The Image Separation & Storage Service automatically disentangles overlapping items in X-ray views, matching them with declaration info for visual clarity. The Image Tracking & Merging Service integrates these separated images with AI diagnostic results into a single, real-time interface. Additionally, the Image DB Search Service enables rapid identification of similar cargo from tens of millions of archived records. Finally, the Alert Service utilizes AI trained on hazardous materials to detect similar shapes in real-time, immediately warning inspectors of potential threats.

EXPECTED OUTCOMES

The AI-based customs image management solution has undergone continuous refinement through field demonstrations at major customs sites, including the Incheon International Airport Customs Express Cargo Logistics Center and Pyeongtaek Customs. As a result, customs inspection accuracy has improved to over 95% compared to manual visual inspections, with image similarity searches achieving an average accuracy of 96.7%.

Notably, the time required for cargo matching and image searching has been reduced from several minutes to under 1 and 2 seconds, respectively. This has led to a 26-fold increase in customs inspection speed (from 5 cases/min to 130 cases/min). This performance boost is expected to lower the fatigue of customs personnel, significantly reduce human error, and facilitate safer and faster clearance. Furthermore, by proactively blocking the entry of hazardous items such as narcotics, the system is anticipated to secure both national security and public safety simultaneously.

Figure 8.1. Before and after implementation (AS-IS vs TO-BE)

Before (AS-IS)	After (TO-BE)
	
• Customs inspectors manually verify information across individual systems to process clearances. - Key Challenges: Cumulative fatigue due to surging volume (e.g., cross-border e-commerce), resulting in clearance delays and compromised detection accuracy.	• Automated detection of hazardous/illegal items using AI, providing a Unified Dashboard by synchronizing source data, including X-ray images and cargo declaration info. - Strategic Outcome: Effectively blocking the entry of hazardous/illegal goods and enhancing the reliability of customs operations.

SOCIAL ISSUE 9

Lightweight Digital Twin for Public Work Environments

KEY TECHNOLOGICAL FEATURES

- **Lightweight Digital Twin Modeling**: Provides efficient and intuitive 3D simulations by optimizing public datasets for lower hardware overhead
- **Real-time Data Integration & Analysis**: Reflects real-time conditions by linking IoT sensors and public data platforms, utilizing AI for advanced analysis
- **Visualization Technology for Public Services**: Offers data visualization and simulation functions to support administrative decision-making based on AI insights

OVERVIEW

Existing Digital Twin technologies often require massive data processing and high-performance hardware, creating barriers for immediate adoption by public institutions. This project applies a more agile, Lightweight Digital Twin technology to enhance administrative efficiency and support rapid policy-making. This approach accelerates digital transformation across various public service sectors, including spatial management, facility maintenance, and disaster response.

Table 9. Project Detail

Project Category: GovTech Startup Support (2024)
Implementing Agency/Partner Organization: d.AGIT Co., Ltd.
Project Description The system establishes a lightweight Digital Twin environment using public data to support the administrative tasks of government agencies. First, it converts spatial data, facility information, and environmental data into optimized 3D models. By providing a more intuitive interface than traditional, heavy-duty digital twins, the system remains accessible for diverse public service applications. For real-time analysis, the platform integrates with IoT sensors and public data systems to collect and reflect data instantly within the digital twin environment. This enables dynamic management of facilities, traffic flow analysis, and disaster response. Furthermore, the system supports Policy Simulation and Decision Support. Administrators can run "what-if" scenarios for improving traffic flow, analyzing disaster response strategies, or planning infrastructure maintenance, ensuring the most effective outcomes for public projects.

EXPECTED OUTCOMES

The implementation of this system significantly improves the work efficiency of public institutions and enables rapid, data-driven policy decisions. It also enhances the quality of administrative services by promoting the digital transformation of public facilities and infrastructure.

Figure 9.1. Before and after implementation (AS-IS vs TO-BE)

Before (AS-IS)	After (TO-BE)
	
• Currently, commercial portal road views often show outdated imagery or are unavailable in areas where vehicles cannot enter.	• Under the new system, using affordable 360-degree cameras, specific spaces and times can be virtualized. This allows officials to check the "Road View" of a specific lot or inaccessible area remotely, eliminating the need for frequent on-site visits and providing the most current visual data available.

SOCIAL ISSUE 10

Logistics Prediction for Discontinued Military Supplies & Procurement

KEY TECHNOLOGICAL FEATURES

- **ETL/ELT Process Design**: Automation of data Extraction, Transformation, and Loading (ETL/ELT) processes
- **Data Standardization**: Ensuring consistency and compatibility across disparate data sources for seamless integration
- **AI Model Development & Optimization**: Development of predictive models for discontinued items and procurement quantities, followed by optimization through hyperparameter tuning and rigorous model validation.

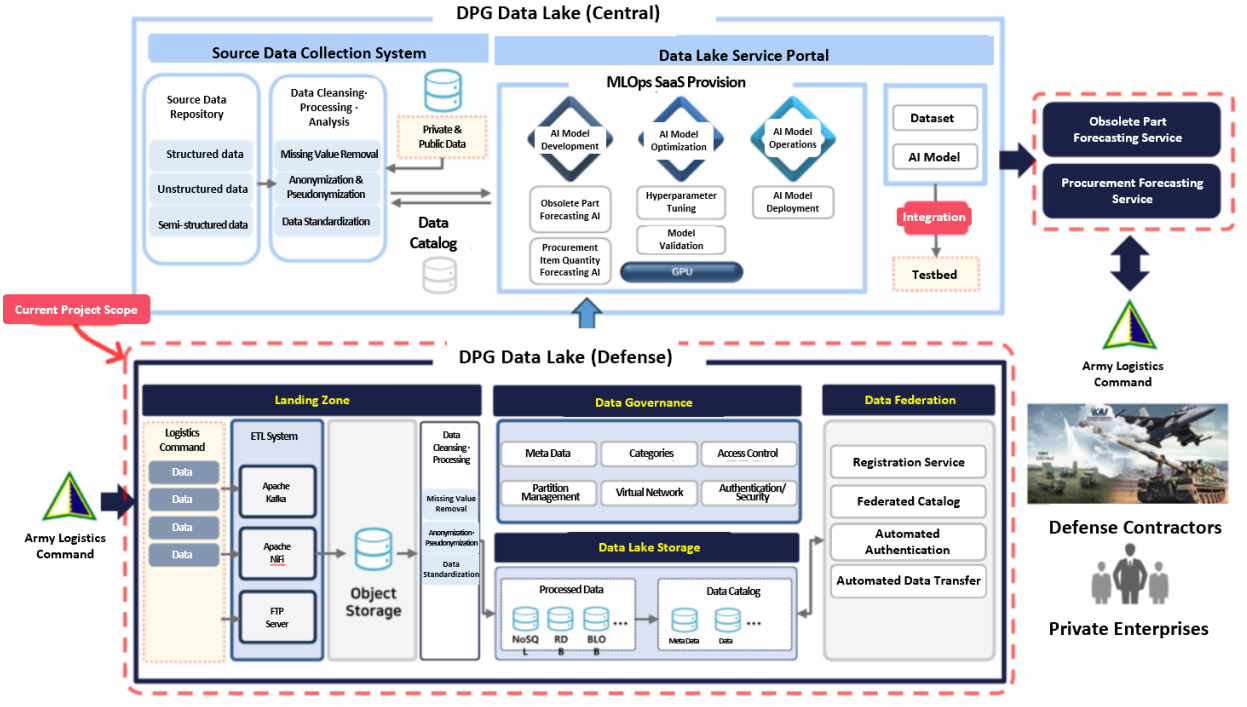
OVERVIEW

Uncertainties in the discontinuation and procurement volume of military supplies can lead to lower equipment readiness, increased costs, and threats to national security. Currently, logistics data managed individually by the Army Logistics Command, the Ministry of National Defense (MND), and the Defense Acquisition Program Administration (DAPA) is fragmented, making integrated analysis difficult. This project aims to build a system that centrally manages military logistics data and utilizes AI-based predictive models to preemptively address procurement delays and item obsolescence.

Table 10. Project Detail

Project Category: Digital Government Infrastructure (2024)
Implementing Agency/Partner Organization: Republic of Korea Army Logistics Command (ROKALC), GNEWSOFT Co.,Ltd., etc.
Project Description In this project, various public and private datasets were collected and integrated to develop AI predictive models. Specifically, we standardized procurement and discontinuation data sourced from public agencies (Logistics Command, MND, DAPA) and defense contractors to create a reliable environment for data-driven forecasting. Machine learning and data analysis techniques were then applied to optimize models for predicting discontinued parts and required procurement volumes. Considering the sensitive nature and security requirements of military data, we established a Federated Data Lake system that can be operated internally within the Logistics Command for data processing and training. By utilizing the DPG AI/Data Lake for performance evaluation, we derived the optimal predictive algorithms. The resulting "Discontinued Item & Procurement Quantity Prediction Service" is provided via API, with restricted access granted to defense contractors to support their operational planning.

Figure 10.1. Conceptual diagram: military logistics forecasting service



EXPECTED OUTCOMES

By preemptively forecasting the discontinuation of military supplies and the necessary procurement quantities, this system helps maintain high equipment readiness and prevents unnecessary cost increases. Previously, the lack of systematic management for discontinued parts led to procurement delays and operational disruptions; however, this system now enables proactive responses.

Furthermore, by providing this critical information to defense contractors, the project facilitates more efficient production and procurement planning, which is expected to revitalize the defense industry. Ultimately, this AI-driven approach enhances the stability of the military supply chain and strengthens national defense capabilities through data-based logistics management.

SOCIAL ISSUE 11

Site Selection Analytics Platform for Public Complex Facilities

KEY TECHNOLOGICAL FEATURES

- **MCLP (Maximal Covering Location Problem) Algorithm**: An optimization algorithm that determines the best placement of candidate sites to cover the maximum possible demand (population) within a limited number of facilities
- **Input Data Architecture**:
  - **Population:** Geolocation (Latitude/Longitude) and weighted population density
  - **Public Facility Locations:** Current geolocation coordinates
  - **Bus Stops:** Geolocation and passenger boarding/alighting statistics

OVERVIEW

To address the growing demand for data-driven site analysis of public infrastructure, we designed a platform that systematically supports AI-based optimal site selection for public multi-purpose facilities on a national scale. The platform utilizes MCLP and GA (Genetic Algorithm) to ensure mathematically sound and socially equitable location planning.

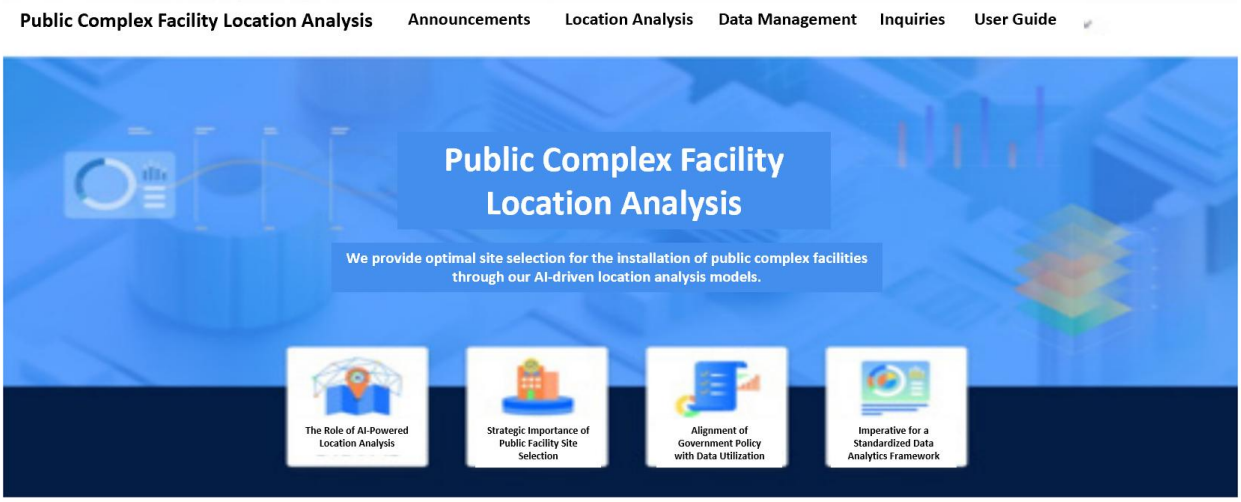
Table 11. Project Detail

Project Category: Digital Government Infrastructure (2024)
Implementing Agency/Partner Organization: Chung-Ang University Industry-University Cooperation Foundation (IACF), Sundosoft Co., Ltd.
Project Description The platform conducted site analysis for various public facilities—including libraries, public Wi-Fi hotspots, automated civil service machines, and heat-wave shelters—specifically targeting the cities of Cheonan and Gwangju. By analyzing the demand, usage status, and user characteristics of these facilities, the system enables the selection of sites that align with urban planning goals and regional characteristics. A centralized platform was established to analyze the optimal locations for "Public Multi-purpose Facilities," which provide both public and private life-essential services in a single, accessible hub. The process involves mapping demand points against existing infrastructure, such as bus stops and current facilities, to identify underserved areas and ensure that the placement of new complex facilities provides the highest level of accessibility for the greatest number of residents.

EXPECTED OUTCOMES

This platform contributes to enhancing citizen convenience and maximizing cost-efficiency by ensuring objectivity and transparency in the site selection process for public facilities. By moving away from subjective decision-making and utilizing AI-driven spatial analysis, local governments can minimize budget waste and provide essential services where they are most needed. Ultimately, this approach fosters balanced regional development and improves the overall quality of life for urban residents through data-informed public administration.

Figure 11.1. Service overview & major development status



SOCIAL ISSUE 12

AI Audit Service for Detecting Public Sector Misconduct

KEY TECHNOLOGICAL FEATURES

- **Data Collection & Pre-processing**: Removal of whitespaces and noise, standardization of formats, and construction of search indexes
- **OCR (Optical Character Recognition)**: Extraction of text data from images or PDF documents to secure usable training datasets
- **Model Tuning**: Hyperparameter optimization, including the adjustment of learning rates and batch sizes, to ensure peak model performance

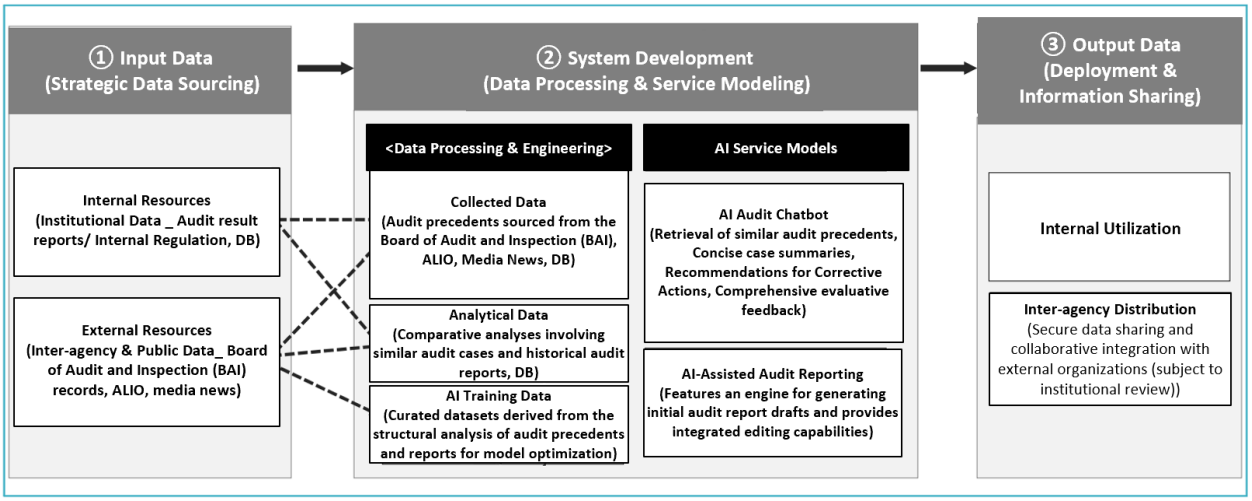
OVERVIEW

New types of employee misconduct are discovered every year in unexpected areas. To prevent such incidents within the organization, SR (Suseo High Speed Railway) aims to derive similar cases based on news articles and historical audit reports. By providing an AI-driven Audit Report Support Service, SR intends to reduce the administrative burden on auditors and support rapid decision-making regarding the analysis of past precedents and the appropriate level of disciplinary action.

Table 12. Project Detail

Project Category: Hyper-scale AI service (2024)
Implementing Agency/Partner Organization: SR Co., Ltd.
Project Description The project utilized internal and external data—including Board of Audit and Inspection (BAI) reports, public institution findings from ALIO, media/SNS content, and internal company regulations—as AI training data for the audit chatbot and report generation models. The core service features an AI Audit Chatbot that provides comprehensive responses, including summaries of similar audit cases, recommended disciplinary actions, and overall evaluations based on user queries. Additionally, the system supports the AI-powered drafting and editing of audit reports, allowing auditors to generate initial drafts automatically and refine them with AI assistance. This integration ensures that even complex misconduct cases are handled with consistency and based on factual precedents.

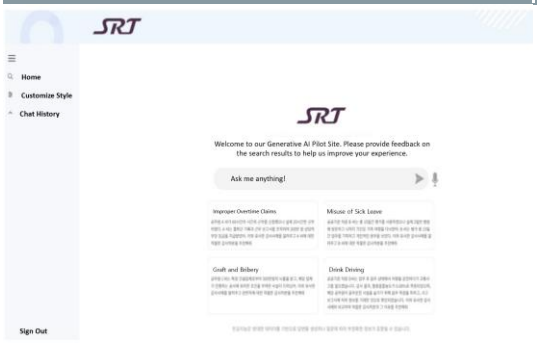
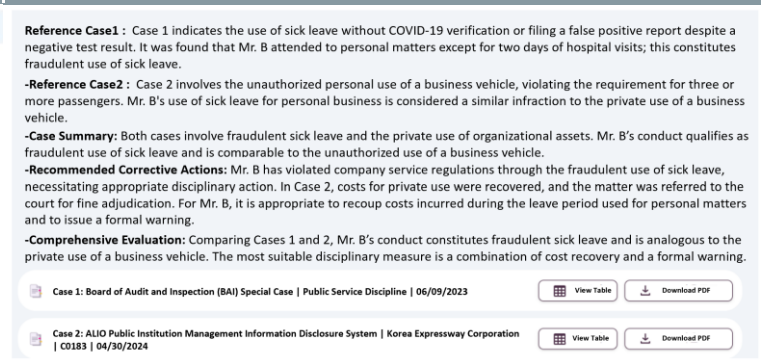
Figure 12.1. Service flow



EXPECTED OUTCOMES

Through this service, SR expects to significantly enhance operational efficiency and reduce labor costs associated with manual auditing processes. These improvements will ultimately lead to enhanced management efficiency for the public institution. By leveraging AI to analyze past cases and automate documentation, the organization can respond more proactively to misconduct risks, ensuring transparency and fairness in administrative actions while maintaining high standards of institutional integrity.

Figure 12.2. Overview

Service overview	Chatbot overview
	

SOCIAL ISSUE 13

e-RFP (Request for Proposal) Assistant Service

KEY TECHNOLOGICAL FEATURES

- **LoRA (Low-Rank Adaptation)**: Application of LoRA, a PEFT (Parameter-Efficient Fine-Tuning) technique, to enable efficient training and testing of multiple model versions in a lightweight manner
- **Keyword-driven Precision Enhancement**: Utilizing GPT-4 for automated keyword extraction, followed by a manual verification step by users to maximize data precision and relevance

OVERVIEW

The Public Procurement Service (PPS) aims to develop a service that processes and learns from 25,000 Request for Proposal (RFP) documents (totaling 200 million tokens) registered in its system. This service enables the automated generation of RFP drafts organized by major table of contents. Building upon the Proof of Concept (PoC) conducted in 2023, the project focuses on advancing and optimizing the draft generation model for practical administrative use.

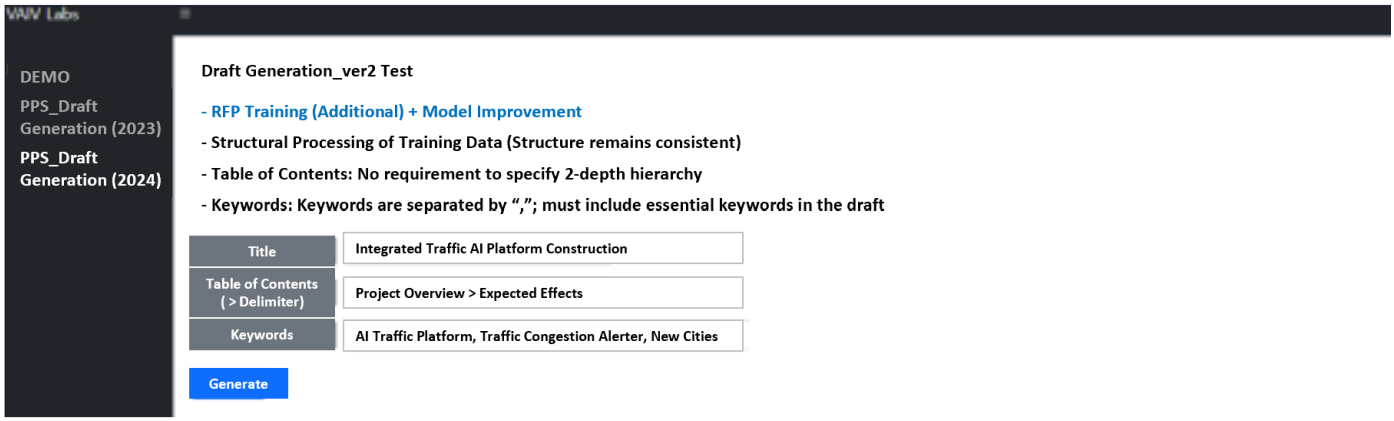
Table 13. Project Detail

Project Category: Hyper-scale AI service (2024)
Implementing Agency/Partner Organization: Public Procurement Service (PPS)
Project Description This service supports the generation of RFP drafts by collecting and processing data based on major sections and chapters, which are then used to train a Large Language Model (LLM). To respond effectively to various types of service RFPs, the team implemented LoRA to train and compare multiple model iterations efficiently. This lightweight approach ensures the system can find the optimal model structure without excessive computational costs. Furthermore, to improve the quality of training data and the accuracy of generated drafts, the system leverages GPT-4's natural language understanding capabilities to automatically extract core keywords. By incorporating a user-validation stage, the platform ensures the validity of these keywords before they are used for generation, thereby significantly enhancing data precision. As a result, users receive rapid and precise support in drafting RFPs for specific service projects, benefiting from a high-performance automated service that ensures both productivity and consistency in document preparation.

EXPECTED OUTCOMES

This service is expected to significantly assist demanding organizations in the complex task of drafting RFPs. By streamlining the initial writing phase, it will lower the barrier to document creation and contribute to the revitalization of the "e-RFP Helper System." Ultimately, the automation of specialized document drafting will reduce administrative lead times, minimize human error, and standardize the quality of public procurement documents across various government sectors.

Figure 13.1. Draft



SOCIAL ISSUE 14

Internal Welfare Data Search & Chatbot for Employees

KEY TECHNOLOGICAL FEATURES

- **RAG-based Q&A Service**: Development of a search and Q&A chatbot service based on SearchGPT (a RAG solution), targeting internal regulation manuals, relevant laws, and local ordinances
- **SmartChat Integration**: Provision of a demo-version chat interface integrated with SmartChat to facilitate interactive user engagement

OVERVIEW

The regulations of the Seoul Welfare Foundation undergo frequent revisions due to constant changes in higher-level laws and municipal ordinances. To ensure that staff can accurately grasp and reflect these latest updates, the foundation launched a data search and chat service initiative in 2024. This project aims to enhance work convenience, reduce repetitive tasks, and improve overall productivity and accuracy.

As the demand for welfare systems to address low birth rates and an aging population grows, the timely application of the latest data and legislation is critical. However, because legal and administrative documents are often scattered in unstructured formats, finding necessary information requires significant time and effort. Therefore, the development of an advanced, intelligent search system and chatbot service became essential to streamline the foundation's operations.

Table 14. Project Detail

Project Category: Hyper-scale AI service (2024)
Implementing Agency/Partner Organization: Seoul Welfare Foundation
Project Description For AI model training, the foundation collected 835 pages of internal regulations and 1,627 pages of relevant laws and ordinances (81 types). During the data pre-processing stage, non-standardized elements such as complex diagrams and images were filtered to optimize the structure for search indexing. Based on this refined dataset, a RAG-based Q&A service was developed. The core service provides regulations, laws, and ordinances ranked by relevance and accuracy in response to queries about operational procedures or standards. This allows staff to verify the latest information instantly, with the system also supporting natural language answers for data contained within tables. Additionally, a demo service was provided that allows users to test both search and chatbot functions simultaneously. This hybrid system matches user queries with pre-built chatbot knowledge first; if no direct match is found, it automatically switches to a document search to provide the most relevant information, thereby maximizing user convenience.

EXPECTED OUTCOMES

By reducing simple and repetitive tasks such as manual data searching and initial report drafting, this service allows employees to focus on core strategic tasks. The application of real-time updated information improves administrative accuracy and significantly shortens the learning curve for new employees.

Furthermore, the chatbot service ensures operational continuity by providing immediate responses even during the absence of specific personnel, thereby resolving potential work gaps. Ultimately, this intelligent system strengthens the foundation's capacity to deliver timely and accurate welfare services to the citizens of Seoul.

Figure 14.1. Implementation Model

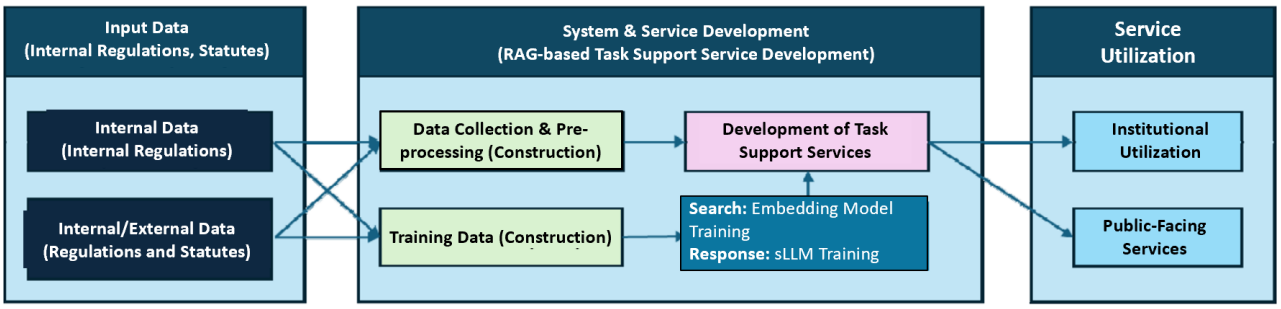
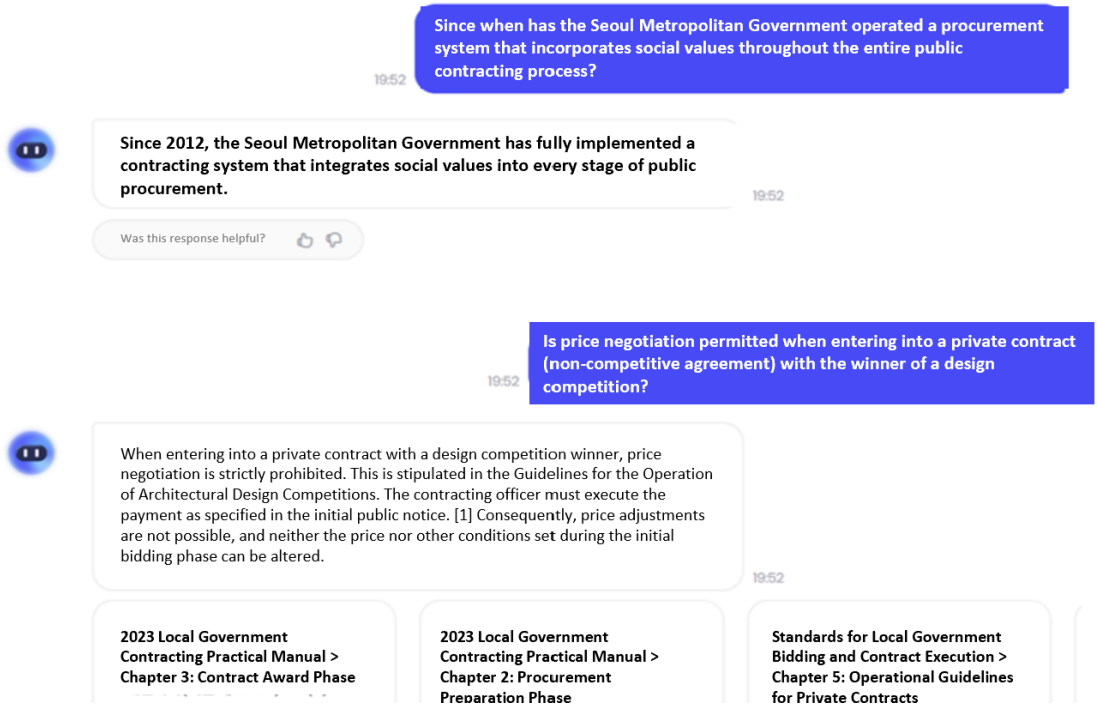


Figure 14.2. Chatbot-integrated chat service



SOCIAL ISSUE 15

PoC for Educational Administration Response via Hyper-scale AI

KEY TECHNOLOGICAL FEATURES

- **Automated Data Pipeline**: Utilization of a Hyper-scale AI platform's automated pipeline for document parsing, chunking, pre-processing, and RAG optimization
- **Accuracy Optimization & Viewer Integration**: Enhancing response accuracy through LLM parameter tuning and integrating a document viewer to provide the exact source pages within the original documents

OVERVIEW

To address the rapid surge in educational activity protection consultations* and the heavy burden of school administration, the Gyeonggi Provincial Office of Education (GOE) aims to provide an environment where faculty can focus on teaching by utilizing Hyper-scale AI for work support.

- Consultation Volume: (2021) 1,424 → (2022) 4,393 → (2023) 9,761 cases.

While many administrative tasks have been digitized, faculty members still struggle with repetitive tasks centered on searching through manuals and regulations. This pilot project focuses on School Student Records (SSR)—a task that is consistently time-consuming at the beginning and end of semesters—applying AI to streamline the process among 410 administrative tasks.

Table 15. Project Detail

Project Category: Hyper-scale AI service (2024)
Implementing Agency/Partner Organization: Gyeonggi Provincial Office of Education (GOE)
Project Description The project selected six core document types, including the "School Student Record Entry Manual" for elementary, middle, and high schools and the "Academic Achievement Management Guidelines." To ensure high accuracy, over 2,000 Q&A pairs were collected and pre-processed. A Vector Database was constructed from this refined data, and the RAG (Retrieval-Augmented Generation) system was optimized through continuous validation of answers and meticulous adjustment of LLM parameters. The service allows teachers to inquire about SSR entries through an interactive AI interface. When a teacher enters a question, the AI provides a natural language response based on the pre-trained manuals and Q&A datasets. To ensure credibility, the system lists the specific reference materials (guidelines, manuals) and presents the corresponding content, allowing teachers to verify the information instantly.

EXPECTED OUTCOMES

By shifting from keyword-based searches to context-driven natural language retrieval, the system delivers responses with over 97% accuracy. Previously, obtaining answers through the School Student Records Support Portal could take several days; now, the Generative AI service provides the necessary information instantly. This creates a robust support system for faster and more accurate record-keeping.

When applied to the broader spectrum of educational administration, this service will significantly reduce the workload of school faculty, allowing them to dedicate more time to educational activities. Looking ahead, the GOE plans to establish the "Gyeonggi Education Digital Platform"—incorporating AI and Big Data—to innovate administrative workflows and enhance communication within the educational community.

Figure 15.1. Conceptual architecture: Hyper-scale AI service

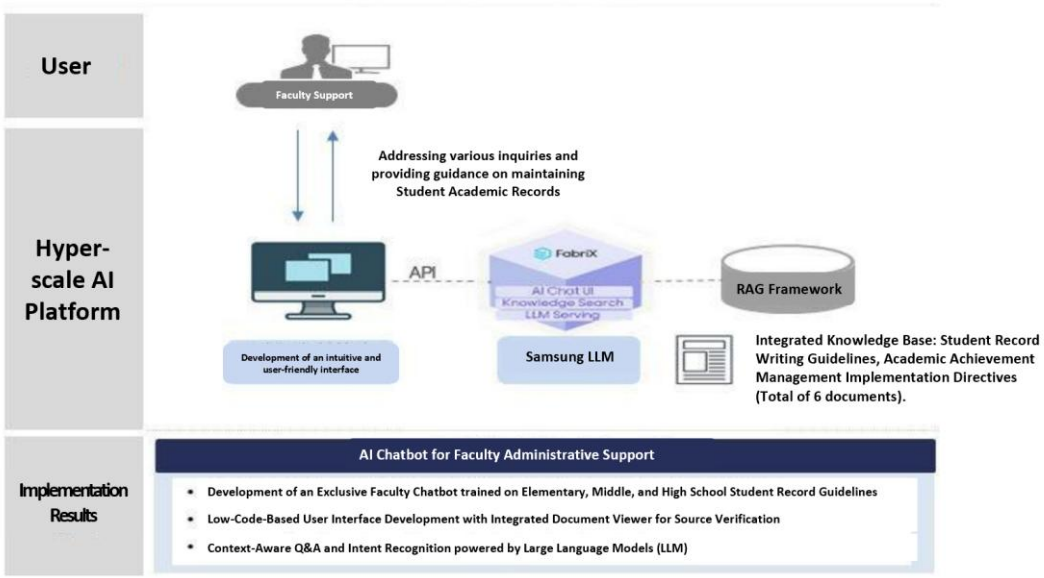
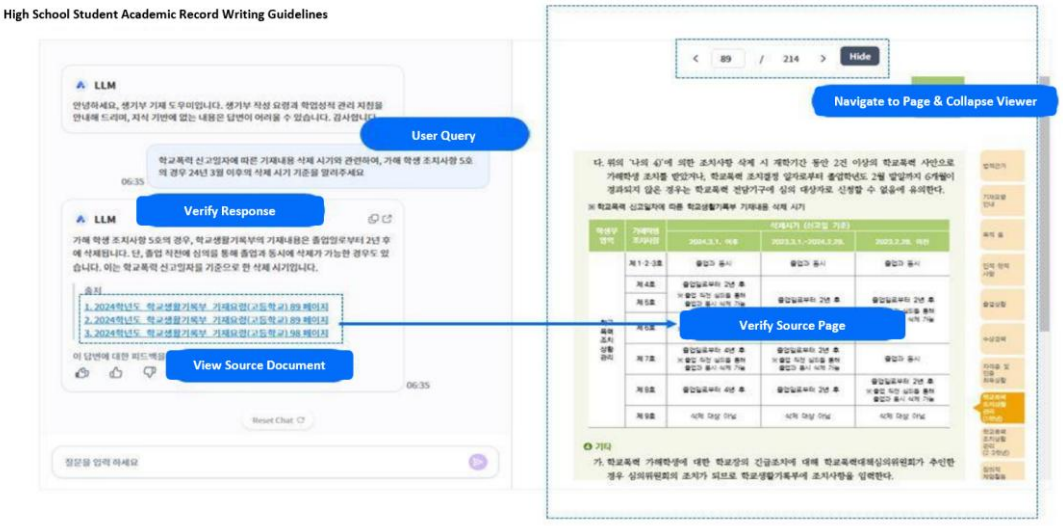


Figure 15.2. Interactive AI Chatbot Interface



SOCIAL ISSUE 16

Automated Contract Support for Schools via Adaptive Learning

KEY TECHNOLOGICAL FEATURES

- **Structured Data Indexing**: Fast structural searching based on the table of contents and specific sections of manuals or guidelines
- **Unstructured Data Indexing**: Separate indexing of core phrases and critical keywords within the data to enhance retrieval precision

OVERVIEW

Public contracts for construction, services, and goods are essential for institutional operations. However, the methods and procedures are highly complex, and various laws, guidelines, and training materials are scattered, making them difficult to search and utilize effectively. To address this, the Seoul Metropolitan Office of Education (SMOE) aims to provide a Contract Administration Support Service. By automating the document drafting and information retrieval processes, the system will save faculty members' time and allow them to focus on more specialized tasks.

Table 16. Project Detail

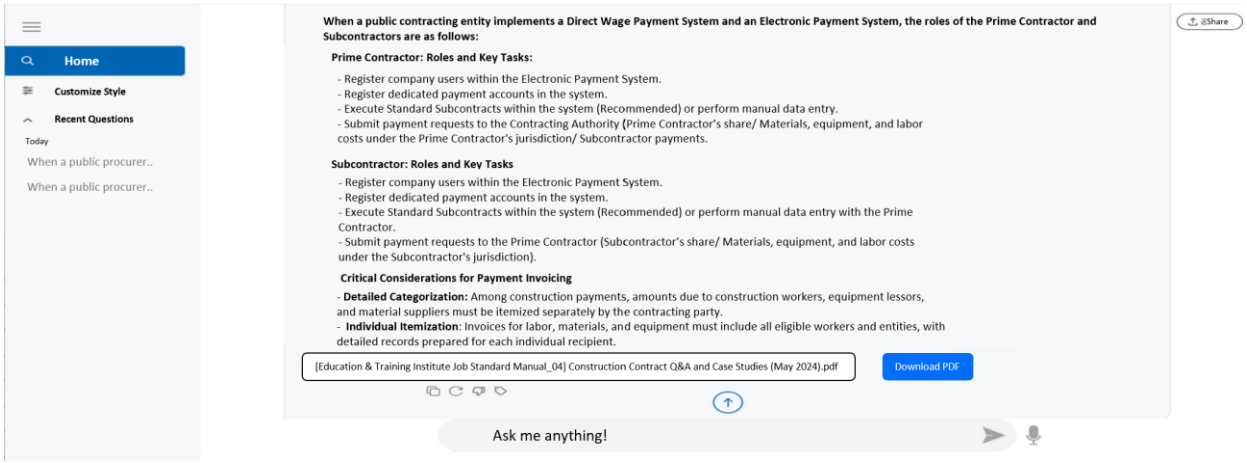
Project Category: Hyper-scale AI service (2024)
Implementing Agency/Partner Organization: Seoul Metropolitan Office of Education (SMOE)
Project Description The AI chatbot service has been trained on a comprehensive collection of contract case studies to ensure high response accuracy. When a user queries the system regarding procurement or contracting procedures, the AI provides precise guidance based on verified precedents. Furthermore, the service provides direct download links to the original source documents, allowing faculty members to verify legal grounds and official forms immediately without secondary searches. This dual-indexing approach—combining structured hierarchy with unstructured keyword extraction—ensures that even highly specific or nuanced contracting rules can be located quickly, regardless of how the user phrases their inquiry.

EXPECTED OUTCOMES

The implementation of this service creates an environment where faculty members can improve their processing speed and concentrate on higher-value professional duties. By standardizing the information retrieval process, the system strengthens the transparency and public accountability of contract administration.

Quantitatively, the SMOE expects a 30% reduction in the time required for contract-related tasks. This efficiency gain not only lowers the administrative burden but also minimizes the risk of legal errors or procedural omissions, leading to a more reliable and streamlined public procurement system.

Figure 16.1. Chatbot interface



SOCIAL ISSUE 17

Support for Heritage Record Search & Business Planning

KEY TECHNOLOGICAL FEATURES

- **Data Extraction**: Utilization of a PDF Annotation Tool to extract high-quality Q&A pairs from complex historical and administrative documents
- **Service Architecture**: Implementation of Retrieval-Augmented Generation (RAG) to support AI-driven searches for cultural heritage archives and regulations, as well as assistance in drafting project proposals

OVERVIEW

There is a growing need for efficient cultural heritage management through the application of Hyper-scale AI technologies for both preservation and utilization. Accordingly, the Korea Heritage Agency (KHA) conducted a GPT-based Proof of Concept (PoC) to deepen its understanding of AI services and to enhance operational efficiency by integrating Hyper-scale AI into daily administrative tasks.

Table 17. Project Detail

Project Category: Hyper-scale AI service (2023)
Implementing Agency/Partner Organization: Korea Cultural Heritage Foundation
Project Description To build a robust knowledge base, 1,166 Q&A data points were constructed and integrated into the model using sources such as the National Heritage Portal, publicly available web data, internal records management systems, corporate regulations, and service manuals. The development followed a systematic procedure, starting with training the RAG engine and the GPT language model on these specialized datasets, followed by integrating the models with a user interface (UI) for accessibility, and concluding with performance testing and iterative refinements to ensure reliability. The resulting GPT service supports KHA employees with functions such as Cultural Heritage Archive Search, Internal Regulation Search, Public Inquiry Response Assistance, and Project Proposal Drafting, allowing staff to handle complex inquiries and documentation with significantly greater consistency and speed.

EXPECTED OUTCOMES

According to the internal performance evaluation of the GPT model trained on KHA-specific data, the RAG model achieved an impressive accuracy rate of 94%.

By leveraging this specialized AI service, the agency can provide faster and more accurate information retrieval than traditional call center operations, significantly contributing to both citizen satisfaction and administrative efficiency. Furthermore, the automation of draft generation for project plans reduces the cognitive load on employees, allowing them to focus on the qualitative aspects of cultural heritage preservation and promotion.

SOCIAL ISSUE 18

Chatbot for Contract Management Procedures and Manuals

KEY TECHNOLOGICAL FEATURES

- **Data Pre-processing**: Segmentation of paragraphs, format conversion, Vector DB construction, and meticulous data labeling/annotation for datasets such as contract guidelines and management manuals
- **Fine-tuning & Service Implementation**: Enhancing model performance through fine-tuning to establish an AI-based contract chatbot capable of sophisticated Q&A and data retrieval support

OVERVIEW

Staff at the Small Enterprise and Market Service (SEMAS) have faced difficulties with procurement and contract management, leading to excessive workloads exacerbated by a shortage of dedicated personnel. To address this, the need arose to develop a chatbot service that leverages the organization's contract management manuals. By implementing an AI-powered contract chatbot, the project aims to provide efficient support for both new and existing staff, thereby reducing costs and improving overall operational convenience.

Table 18. Project Detail

Project Category: Hyper-scale AI service (2024)
Implementing Agency/Partner Organization: Small Enterprise & MArket Service (SEMAS)
Project Description The system is built on a configuration of Naver's Hyper-CLOVA X combined with RAG (Retrieval-Augmented Generation), allowing employees to easily access information by conducting initial Q&A sessions on procurement and contracts via the AI service. The development process involved comprehensive data pre-processing—including paragraph segmentation, format conversion, Vector DB construction, and data labeling—followed by model fine-tuning and full-scale service implementation. To enhance utility, the service features verifiable URLs for source referencing and a recommended query function. Continuous efforts are being made to increase service effectiveness through the periodic retraining of datasets to improve response accuracy. This integrated approach ensures that complex administrative guidelines are transformed into an interactive and accessible knowledge base for all employees.

EXPECTED OUTCOMES

This service reduces administrative costs and ensures high response accuracy by continuously training the model on the latest contract guides. By simplifying complex legal procedures and regulations, the chatbot improves staff comprehension and maintains operational continuity even during personnel absences. Ultimately, the system alleviates workload imbalances and fosters a more stable, efficient administrative environment within the organization.

SOCIAL ISSUE 19

Automated Document Generation via Hyper-scale AI

KEY TECHNOLOGICAL FEATURES

- **LLM-based Natural Language Processing**: Automated generation of documents tailored to specific characteristics by collecting standard and exemplary documents, followed by legal/regulatory compliance reviews
- **RAG-driven Search Functionality**: Implementation of Advanced RAG and Hybrid Search techniques to ensure high-precision data retrieval and inquiry
- **Summary Report Generation**: A feature that segments document content into chunks for summarization, followed by the generation and visual validation of the final report

OVERVIEW

The maintenance of facilities managed by the Seoul Facilities Corporation (SFC) requires the repetitive creation of essential documents. This process has historically demanded significant resources, leading to operational inefficiencies. To minimize this resource drain, a Generative AI-based service was established to automatically draft documents in standardized formats.

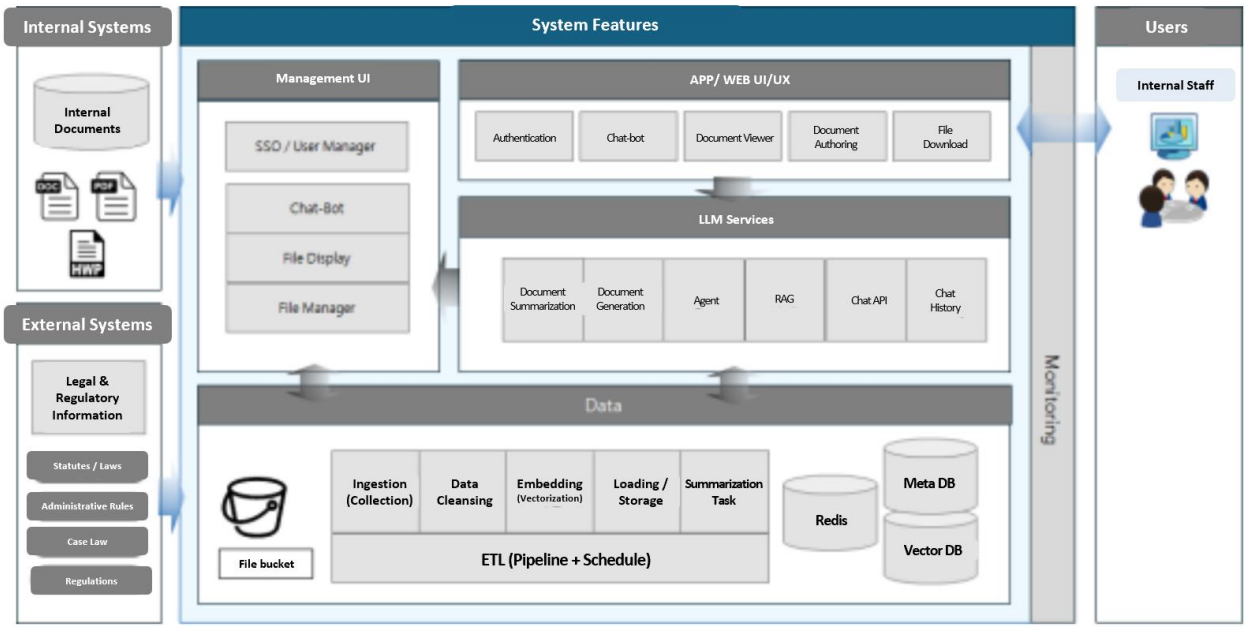
Table 19. Project Detail

Project Category: Hyper-scale AI service (2024)
Implementing Agency/Partner Organization: Seoul Facilities Corporation
Project Description The system learns from internal data and legal statutes to automatically draft documents and conduct legal reviews by leveraging a collection of standard and exemplary templates. Specifically, the Automated Standard Document Drafting Service organizes large volumes of collected data into a single, comprehensive document. Complementing this is the Exemplary Report Drafting Service, which generates optimal reports based on proven best practices even when only minimal input is provided. Additionally, the Automated Legal & Regulatory Review Service was developed to automatically reflect recent amendments in laws or regulations into existing documents. By integrating these features, the corporation has transformed its document-centric workflow into an intelligent, AI-assisted process.

EXPECTED OUTCOMES

This service has improved operational efficiency by automating document creation, allowing staff to shift their focus from administrative paperwork to their core mission of facility safety. By streamlining the drafting and review process, the organization ensures higher consistency in its documentation while proactively managing legal risks through automated regulatory updates.

Figure 19.1. Report Generation Scenario Flowchart



NIA NATIONAL INFORMATION SOCIETY AGENCY

National Information Society Agency
Use Cases of Public AI Service (Vol.2)

Published by Jong-sung Hwang, President, National Information Society Agency (NIA)

Produced by *Department of Global ICT Cooperation*

Yoon-seok Ko, Vice President
Hye-jeong Lim, Director
Joung-won Yoo, Director
Deok-won Heo, Executive Principal
Ji-young Kim, Manager

Reviewed by *Department of AI Convergence*

Byung-joo Jeong, Vice President
Won-hwak Kim, Executive Director
Min-ho Lee, Principal Manager

LLM (Large Language Model) technology was partially utilized in the translation process of this report.


NIA
NATIONAL INFORMATION
SOCIETY AGENCY