



The ICSCoE Report is a public relations newsletter on ICSCoE's activities.



## 9th Core Human Resource Development Program Introducing Activities of Our Final Projects Part2

### Improving Security Across Parent and Group Companies

#### Backgrounds and Issues

A growing number of incidents have seen parent companies suffer cyber attacks that originated through their group companies, making it an urgent priority to raise security levels across the entire corporate group. However, group companies typically have far fewer human, physical, and financial resources than their parent companies, which limits how much they can strengthen security on their own. At the same time, parent companies themselves have often been unable to specify the required security level for their group companies or ensure that it is met, and in practice this has sometimes left them without effective, enforceable control.

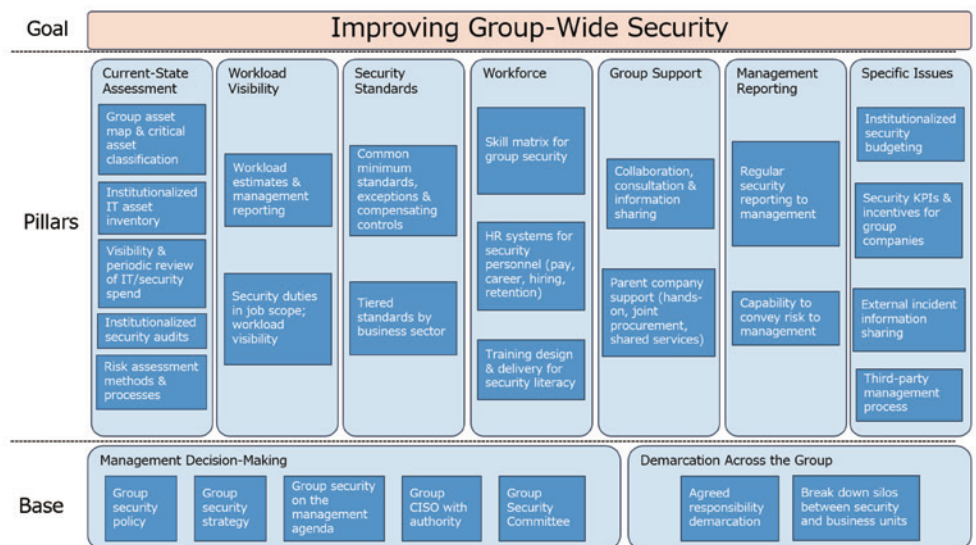
Against this backdrop, our project brought together eight members, one per enterprise, each drawing on their own group's experience, to fundamentally rethink and discuss the question: "Why do security efforts at group companies fail to advance?" Our root-cause analysis revealed that most factors stemmed not from a lack of execution capability on the front lines, but from the governance and design layer on the parent-company side — namely, undecided policy direction, unassigned authority, and the absence of supporting mechanisms.

#### Issue-Solving and Outcomes

Building on these findings, we identified the actions that parent-company security departments should take, and structured them into a "Group Security Issues Model."

As the highest-priority issues, we focused particularly on two pillars establishing a common group-wide security policy and developing a risk assessment methodology and process. The project team compiled practical implementation steps and guidance for assessing current conditions in supplementary documents.

The final deliverable consists of an eight-chapter main document, along with a list of root causes and issues (Appendix A), a glossary (Appendix B), and the process used to derive the root causes were derived (Appendix C). The deliverable has also been published online as the "Group Security Guide" for internal and external practitioners to reference. We encourage you to read the guide.



Group Security Issues Model



### What is your utmost benefit from this project?....

My greatest benefit was the opportunity to put into practice a process, using root-cause analysis, to break down the vague sense that "security measures at group companies are not progressing" into its underlying factors and pinpointing the root causes. If we focus only on the surface-level symptom — the slow progress of security countermeasures at each company — we may tend to place the blame on the front lines. As we went deeper into the analysis, a clear pattern emerged: in many cases, the root cause lay in the absence of policy, authority, and mechanisms on the parent-company side. Discussing and analyzing this together with eight members from other industries gave me perspectives I could never have gained by looking only at the group I was seconded from. It also showed me that this is a structural challenge common across industries and gave me comrades-in-arms who are working together toward a solution. This was also a major benefit of the project.



### Methods for utilizing project outcomes .....

After returning to my company, I will start by comparing the list of root causes and issues compiled through this project against my own group's current situation to identify where the gaps lie. In particular, I intend to focus on the two highest-priority areas : establishing a group security policy and developing a risk assessment methodology and process. I will leverage supplementary documents accompanying our deliverables to assess the actual conditions across our group companies and to help prioritize countermeasures accordingly. Going forward, I would also like to share our deliverables with other companies in our industry, with the aim of advancing security across the industry.



Mr.TSUKISHIMA Kazuma



### This is unique to the ICSCoE! .....

Our instructors, with their vast knowledge, gave us new perspectives and insights we could not have reached on our own, day after day. It was an inspiring year that flew by. I also built relationships with fellow trainees that crossed industry lines, forming a network of dependable colleagues I know I can turn to for support going forward. As for the curriculum, classroom instruction brought everyone's foundational knowledge of both IT and OT up to a common level. Although I initially felt my OT knowledge was lacking, I was able to keep pace with the coursework. Later, through active learning in a simulated environment, I was able to gain knowledge that went well beyond what the classroom sessions alone. That hands-on approach is something I feel is unique to this program.



Project members

# Secure Deployment and Operational Validation of a Domain-Specific Local LLM in Confidential Environments

## Backgrounds and Issues

In recent years, technological advancements in large language models (LLMs) have significantly transformed approaches to operational efficiency and decision-making support. Expectations for the use of generative AI are also growing in highly confidential environments, such as the maintenance and operation of mission-critical systems in the electric power, finance, and telecommunications industries, as well as industrial control systems (ICS) isolated from external networks.

In these environments, demand for local LLMs is rapidly increasing, as they can be operated within an organization's own infrastructure without transmitting data to external services. However, the assumption that an offline environment is inherently secure is not necessarily correct. Even LLMs deployed locally have AI-specific attack surfaces.

Examples include prompt injection attacks that embed malicious instructions within user inputs or reference data, attacks targeting data sources used in Retrieval-Augmented Generation (RAG), and model poisoning attacks that attempt to compromise the model itself. In addition, establishing effective LLMOps practices, including continuous updates to model and software, patch management, evaluation, and monitoring, remains a critical challenge.

This project focused on building and validating a domain-specific local LLM environment that enables the secure and effective use of proprietary data in confidential environments, without relying on the assumption that “offline equals secure.”

## Issue-Solving and Outcomes

This project went beyond theoretical research and desktop analysis by building an actual local LLM environment and conducting end-to-end verification, including attack simulations and evaluation of mitigation measures.

Project members actively participated in every phase of the initiative, including hardware selection and procurement of GPU-equipped systems, software environment setup, selection of pretrained models, validation for business use cases, risk analysis, and assessment of risk mitigation effectiveness.

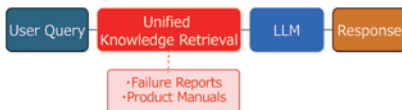
For the business-use validation, we selected incident response operations as a representative use case and incorporated actual incident resolution records and product manuals from the companies represented by project members into a RAG environment. Rather than simply ingesting data, we implemented semantic chunking, separated knowledge by information category, and designed and optimized system prompts reflecting specialized personas, such as a “power transmission and distribution maintenance engineer.”

These efforts demonstrated improved response accuracy for the target operational tasks and clarified key considerations for data preparation and prompt engineering when applying local LLMs in real-world business environments.

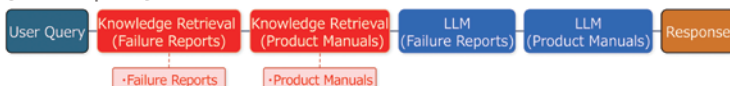
### Validation Using Real Operational Data

- When failure reports and product manuals are stored in a single knowledge repository, retrieval may return information with low relevance.  
⇒ This can lead to inaccurate or unreliable AI-generated responses.
- By separating knowledge retrieval and LLM processing according to data type, we reproduced the generation of evidence-based responses required by transmission and distribution maintenance engineers.

[Before Data Separation]



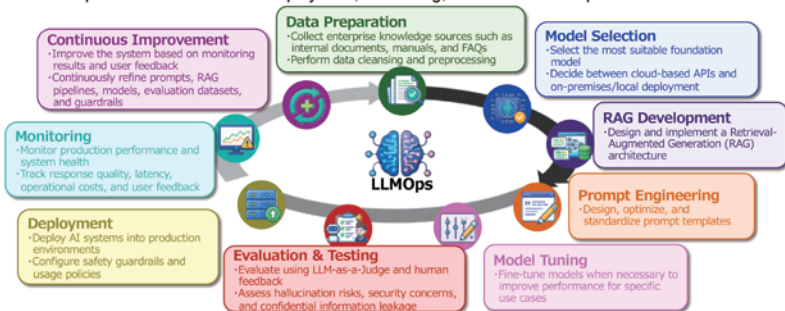
[After Data Separation]



Validation using members' real-world business data

### LLMOps (Large Language Model Operations)

- An operational framework for continuously managing the lifecycle of LLM-based applications, from development and evaluation to deployment, monitoring, and continuous improvement.



Because AI systems inherently involve uncertainty, such as response accuracy, inappropriate outputs, latency, and operational costs, **LLMOps is not a one-time activity but an ongoing continuous process.**

LLMOps (Large Language Model Operations)

We also identified potential attack surfaces within the validation environment, analyzed risks that could significantly impact business continuity, and conducted hands-on verification of prioritized mitigation measures to confirm their effectiveness.

Knowledge gained through these iterative cycles of build, attack, detect, and defend was organized into comprehensive documentation to reduce dependency on individual expertise. The resulting deliverables provide practical guidance for professionals evaluating local LLM adoption and secure operation in confidential environments, supporting activities such as risk assessment, system design, and operational planning.



### What is your utmost benefit from this project?....

The greatest benefit is the opportunity to experience the entire lifecycle of a cutting-edge local LLM firsthand, from environment construction and cyberattack validation to evaluation of defensive measures.

At the start of the project, there were significant differences among members in terms of LLM and cybersecurity knowledge. However, through approximately three months of collaborative activities, all members developed a common understanding of key concepts and terminology, enabling productive discussions on equal footing.

In particular, we gained extensive practical experience across a wide range of areas, from infrastructure-related topics such as hardware selection and software environment configuration to system prompt optimization for improving RAG performance, and even demonstrations of advanced model poisoning attacks.

Being able to acquire practical, real-world knowledge rather than merely theoretical understanding has become a valuable asset that we expect to leverage in our future work.



### Methods for utilizing project outcomes .....

The validation reports and the deployment and operational documentation produced by this project are intended to serve as practical guidelines when project members return to their organizations and begin formally evaluating and promoting local LLM adoption.

Today, many organizations remain cautious about the business use of generative AI because of concerns regarding confidential information leakage. By leveraging the outcomes of this project, organizations can gain insights into secure on-premises LLM architecture design and establish evaluation criteria for AI-specific security risks such as prompt injection attacks.

The deliverables can also serve as foundational materials for discussions among operational teams, related

departments, and management regarding compliance with internal security policies. Furthermore, they provide valuable guidance for conducting pre-deployment risk assessments and designing post-deployment LLMops processes.



### This is unique to the ICSCoE! .....

One of the most distinctive aspects of the ICSCoE is its environment, where professionals from diverse industries, technical domains, and backgrounds come together from their respective organizations, engage in in-depth discussion, and work as one team toward a common goal.

The ICSCoE is far more than a lecture-based training program. Trainees can gain hands-on experience using actual equipment, including high-performance GPU servers, and can validate realistic threat scenarios such as supply-chain risks and attack detection in practical environments.

In addition, whenever we encountered challenges during the project, we received direct, expert feedback from instructors who are actively engaged in the cybersecurity field. Learning how to evaluate security risks from multiple perspectives proved to be an invaluable experience.

The knowledge gained through this project, together with the professional network built with fellow participants who shared both challenges and successes throughout the program, represents one of the most significant benefits uniquely offered by the ICSCoE.



Mr.TAKAHASHI Shinji (front row, center) and project members

