Fostering Organizational Collective Intelligence from Human-Generative AI Interaction

Shoma Hino [1], Takuya Nakata[1], Sachio Saiki[2], Sinan Chen[1], Masahide Nakamura [1,3]
[1]Kobe University, 1-1 Rokkodai-cho, Nada-ku, Kobe, 657-8501, Japan
[2]RIKEN AIP, 1-4-1 Nihonbashi, Chuo-ku, Tokyo, 103-0027, Japan
[3]Kochi University of Technology, 185 Miyanokuchi, Tosayamada, Kami, Kochi 782-8502, Japan
, Email: hino-s@es4.eedept.kobe-u.ac.jp, tnakata@bear.kobe-u.ac.jp, chensinan@gold.kobe-u.ac.jp
, saiki.sachio@kochi-tech.ac.jp, masa-n@cs.kobe-u.ac.jp

***Abstract*—Recent advancements in generative AI have attracted attention as a problem-solving approach in various domains, including natural language processing, debugging programming errors, and creative content generation. However, its effectiveness largely depends on the user's ability to design prompts, and its potential is often underutilized. Furthermore, dialogue logs generated during problem-solving interactions with generative AI, which could serve as valuable information resources, are frequently left unused and rarely revisited.**

This study proposes a service that evaluates dialogue logs with generative AI, accumulating and sharing them as problem-solving cases, including the contextual background of the questions. The aim is to effectively preserve and share dialogue logs, thereby enhancing knowledge sharing and problem-solving capabilities within organizations. A system was developed based on the proposed method, and its utility and proper operation were validated through practical use.

***Index Terms*—collective intelligence, generative AI, application development, large language model**

I. Introduction

Recent advancements in generative AI, represented by ChatGPT [1] and Claude [2], have drawn attention to dialogues with generative AI as a problem-solving approach in various fields, such as natural language processing, debugging in programming, and creative content creation. Early Large Language Models (LLM) struggled even to grasp the context of input sentences, but the emergence of Transformer models has rapidly accelerated the development of LLMs. Since the release of the GPT series in 2018, LLMs have demonstrated significant capabilities as problem-solving tools [3]. Furthermore, the OpenAI o1 model, announced in September 2024 [4], has enhanced reasoning capabilities, enabling it to address more complex problems, and further advancements in generative AI as a problem-solving tool are anticipated.

The problem-solving capabilities of generative AI are not limited to micro-level issues, such as debugging errors encountered during programming. They also extend to macro-level challenges, such as objectively analyzing the direction of organizations like companies or research labs through dialogues with generative AI. However, even if micro-level problems are resolved through dialogues with generative AI, the dialogue logs, which could be highly valuable to others within the same organization facing similar issues, are often left unused and rarely revisited. Similarly, for macro-level challenges, dialogues with generative AI could serve as highly valuable organizational assets for sharing directions, yet there are few examples of such dialogue logs being shared to align organizational goals.

Moreover, the effectiveness of generative AI heavily depends on the user's ability to construct prompts, and there are many instances where its potential is not fully realized. Effective utilization of generative AI requires appropriate prompt design. While services that provide templates for prompts are gradually increasing, there are still few platforms that evaluate and share which prompts were effective for specific problems.

To address these challenges, this study proposes a service that accumulates and shares problem-solving cases by saving dialogue logs with generative AI, including their context. By implementing this approach, dialogue logs generated daily through interactions with generative AI can be accumulated and shared as problem-solving cases, fostering **collective intelligence** within organizations. This aims to extract value from dialogue logs as organizational assets, enabling sustainable growth through the enhancement of the organization's knowledge base. This approach focuses on three key challenges:

(P1) **Disparity in problem-solving capabilities due to differences in prompt engineering skills for effectively utilizing generative AI**
(P2) **Lack of mechanisms for effectively sharing and utilizing dialogue logs within an organization**
(P3) **Challenges in ensuring the reliability of dialogue logs containing hallucinations or misinformation**

As a key idea, we developed a collective intelligence platform that effectively accumulates and shares problem-solving cases by saving dialogues with generative AI, including their context and usefulness, in the form of a web application. The proposed service is called **ChatHubAI** (CHAI).

II. Preliminaries

A. Collective Intelligence

Collective intelligence mechanisms refer to systems that integrate distributed information and cognitive abilities of multiple intelligent entities, such as humans or agents, through their interactions, thereby generating knowledge or intelligent behavior as a whole [5]. Sharing collective intelligence allows the knowledge base of an organization to accumulate, enabling sustainable growth. This facilitates the acquisition of broader insights that individuals alone cannot achieve, leading to more comprehensive problem-solving and decision-making.

In particular, within organizations that share the same objectives, similar problems are likely to arise. Therefore, if problem-solving cases related to such issues can be accumulated within the organization, it becomes possible to foster higher-quality **organizational collective intelligence** tailored to the organization. Furthermore, beyond individual challenges, sharing dialogue logs with generative AI regarding macro-level issues within the organization enables the alignment of organizational directions. This can enhance the transparency of decision-making processes within the organization and improve the overall quality of organizational decision-making.

B. Generative AI

Generative AI refers to a general term for AI technologies that leverage large datasets and deep learning techniques to generate various types of content (e.g., text, images, audio). In this study, the default model used for interactions with generative AI is **ChatGPT-4o-latest** [6], which supports multimodal inputs, including text and images.

Regarding the performance of ChatGPT-4o, according to Chatbot Arena [7], ChatGPT-4o currently ranks third in the Overall category as of January 2025. It has received high evaluations for solving problems across various domains. However, in the Overall category, the first and second positions are occupied by Google's Gemini-2.5-Pro-Exp-03-25 model. This model also ranks highly in specialized categories such as mathematics and coding, making them the most advanced LLMs available. However, as indicated by the "Exp" in this name, the Gemini-Exp series is still in an experimental phase, and their specifications may change without notice.

On the other hand, ChatGPT-4o offers stable operation and relatively low API usage costs, making it highly cost-effective. This study prioritizes the Overall category evaluation, as it addresses problem-solving across various domains.

III. Focused Challenges

In this study, we focus on the following three problems when accumulating and sharing problem-solving cases based on generative AI:

P1: Disparity in problem-solving capabilities due to differences in prompt engineering skills for effectively utilizing generative AI

The response accuracy of generative AI heavily depends on the input prompt. To improve the quality of generative AI outputs, techniques such as **prompt engineering** are employed [8]. These include structuring the necessary information in the prompt or having the generative AI act as an expert agent to address the problem at hand. Such methods are primarily adopted by users who frequently utilize generative AI. However, users who are not adept at using generative AI may input prompts that are difficult for the AI to interpret, even if they are aware of these techniques, due to their limited ability to articulate their problems or lack of background knowledge. Consequently, they often fail to fully leverage the problem-solving capabilities of generative AI. This disparity has led to a division in problem-solving capabilities between frequent and infrequent users of generative AI. Such a situation poses a critical issue for a "technology" that aims to universally benefit society.

P2: Lack of mechanisms for effectively sharing and utilizing dialogue logs within an organization

While it is possible to share dialogue logs by sharing accounts of web-based tools like ChatGPT or Claude, simply sharing dialogue logs as a series of questions and answers often omits crucial contextual information, such as the problem being addressed or the usefulness of the responses. This makes it difficult for others to grasp the intent behind the dialogue logs. Moreover, the search functionality in ChatGPT only displays results if the search term is present in the dialogue, which becomes less effective as the number of dialogue logs scales. Additionally, when accounts are shared, it becomes unclear whose history is being accessed. These issues make simple account sharing insufficient for fostering collective intelligence. Furthermore, sharing accounts within an organization may violate OpenAI's terms of service, necessitating the use of appropriate methods for sharing dialogue logs within an organization.

P3: How to handle hallucinations and misinformation in dialogue logs as reliable information

While the usefulness of generative AI for problem-solving has been discussed, generative AI is known to exhibit a phenomenon called **hallucination**, where the AI fabricates information entirely and presents it as if it were factual [9]. Such outputs may appear plausible to non-experts in the field, necessitating fact-checking against other references. Therefore, dialogues with generative AI require careful scrutiny for hallucinations and misinformation, raising the issue of how to manage such mixed-quality information as reliable resources.

Considering these issues, the challenges to be addressed can be summarized as follows: By saving problem-solving cases based on dialogue logs with generative AI in a format that includes their intent and is comprehensible, it is possible to assist other users in interpreting the prompts that enabled the problem-solving. This approach aims to avoid the division in problem-solving capabilities caused by differences in prompt engineering skills. Furthermore, it is necessary to establish a mechanism for effectively and reliably utilizing problem-solving cases within an organization.

IV. Proposed Method

A. Goal and Key Idea

The objective of this study is to effectively utilize past dialogue logs with generative AI, which have been scrutinized for hallucinations and misinformation, to avoid disparities in problem-solving capabilities and promote the creation and sharing of new knowledge.

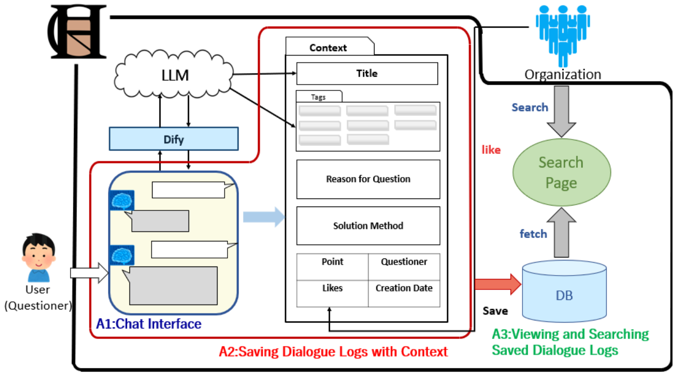

Fig. 1. Overall Architecture.

To achieve this objective, the key idea is to develop a web application, **ChatHubAI** (CHAI), which accumulates and shares dialogues with generative AI as problem-solving cases, including their context and usefulness, thereby fostering organizational collective intelligence. Furthermore, this study evaluates the utility of CHAI by deploying it within an organization.

Regarding organizational collective intelligence, this study limits the sharing of problem-solving cases to within the organization. This is because dialogues with generative AI may include non-public activities within the organization or personal information during problem-solving interactions, and thus, dialogue logs are kept private outside the organization.

The approach of this study categorizes the functionalities of CHAI as follows:

(A1) **Chat Interface**
(A2) **Saving Dialogue Logs with Context**
(A3) **Viewing and Searching Saved Dialogue Logs**

B. Overall Architecture

The overall architecture of CHAI, incorporating the approaches A1, A2, and A3, is shown in Fig 1.

The overall architecture is outlined below.

In CHAI, interactions between an organization's user and the LLM are conducted through the **A1: Chat Interface**. The dialogue logs are then saved in the database with the context defined in IV-D through **A2: Saving Dialogue Logs with Context**. Subsequently, in **A3: Viewing and Searching Saved Dialogue Logs**, users can view and search the saved problem-solving cases.

By referencing the viewed problem-solving cases, users can enhance their knowledge base and refine their prompts. This enables more effective problem-solving through the chat interface using appropriate prompts. By iterating this cycle within the organization, mutual interactions occur as collective intelligence, fostering the development of organizational collective intelligence.

C. A1: Chat Interface

To save dialogue logs with generative AI, a chat interface for interacting with the AI is indispensable. This chat interface utilizes the embedded chat interface of the open-source LLM application development platform **Dify** [10]. By doing so, it provides a user experience similar to using ChatGPT for problem-solving and facilitates connections with LLMs other than GPT-4o, as well as extensions to RAG functionality. Additionally, since Dify is open-source, it is well-suited for extensible development.

TABLE I
CONTEXT OF PROBLEM-SOLVING CASES

5W2H	Context
Who	**Questioner**
What	**Title**, **Tags**
Why	**Reason for Question**
When	**Creation Date**
Where	**Organization ID**
How	**Solution Method**
How many	**Points**, **Likes**

D. A2: Saving Dialogue Logs with Context

When accumulating dialogue logs with generative AI as problem-solving cases, it is essential that others can understand the intent behind these logs. Therefore, it is necessary to save not only the questions and answers with the generative AI but also their context. To achieve this, the required context for treating dialogue logs with generative AI as problem-solving cases is comprehensively defined based on the 5W1H (+How many) framework as follows TABLE I:

1) **Questioner** refers to the name of the person who generated the problem-solving case. By recording the information of the questioner, it becomes easier to understand their expertise and intent within the organization, facilitating the interpretation of dialogue logs. Additionally, not limited to CHAI, it becomes possible to provide direct feedback for problem-solving, creating opportunities for fostering collective intelligence.

2) **Title** is a summary of the dialogue log. Having a title allows users to infer the content of the dialogue log from the title, which can be helpful for organizing information even when a large number of logs are accumulated.

3) **Tags** are multiple keywords related to the dialogue log. By combining this with the search functionality, it becomes possible to efficiently aggregate logs related to specific themes or issues. Additionally, by saving the summary of the dialogue log as keywords, it becomes easier to grasp the content of the log.

4) **Points** represent a numerical evaluation of the usefulness of a log, with a maximum score of 10. Higher scores indicate greater usefulness, while lower scores suggest less usefulness. CHAI encourages saving dialogue logs regardless of their usefulness or lack thereof. While the value of useful logs has been discussed above, even logs deemed unnecessary can be significant. For instance, they may highlight problems that generative AI struggles with or prompts that are difficult for the AI to interpret. Accumulating such logs can help improve problem-solving methods. Therefore, by introducing a scoring system that evaluates both usefulness and lack thereof, the contextual richness of the logs can be enhanced.

5) **Reason for Question** is an item that describes the background or purpose of the question, such as the problem the user aimed to address through generative AI. Clearly stating why the question was asked makes the intent of the dialogue easier to understand.

6) **Solution Method** describes how the problem was resolved—or not resolved—through dialogue with generative AI. While some responses from generative AI can directly solve problems, in cases where the responses are adapted or applied to solve the issue, failing to document how they were applied would result in the loss of critical contextual information about the solution. Therefore, this item is one of the most important components of the context.

7) **Likes** indicates how valuable the problem-solving case is to other users. In CHAI, each user can give one like per log. Logs with many likes are likely to contain information that is useful to others and can be considered highly referential. Additionally, the like feature enables not only one-way but also two-way collective intelligence.

8) **Creation Date** indicates when the problem-solving case was created. Understanding the timeline helps users grasp the freshness and context of the information.

Among these eight items, the questioner, title, tags, and creation date can be set automatically without requiring the intent of the questioner, as they can be derived from the saved dialogue logs. Therefore, these items were implemented to be generated automatically.

While the above eight items are set for the entire log, each question and answer can also be assigned a necessity level with three categories: Critical, Necessary, Unnecessary. This implementation allows the user to specify particularly critical parts of the dialogue log among the responses, making the intent of the saver clearer. Additionally, the "Unnecessary" category serves as a feature to hide specific parts of the dialogue, such as questions containing information that should not be disclosed within the organization, by making only those parts private.

E. A3: Viewing and Searching Saved Problem-Solving Cases

By implementing the functionality to view and search saved dialogue logs, it becomes possible to promote knowledge sharing within the organization and foster collective intelligence.

In CHAI, saved problem-solving cases can only be viewed within the same organization, and searches can be conducted using three items: questioner, title, and tags. Additionally, this search functionality allows sorting by title, tags, questioner, points, creation date, and number of likes. As shown in Fig. 2, detailed filtering is also possible for each column using nine types of operators.

In the detailed view of problem-solving cases, questions and answers are displayed with highlights based on their necessity level, emphasizing critical parts. This detailed viewing and search functionality enables the promotion of knowledge sharing within the organization and the cultivation of collective intelligence.

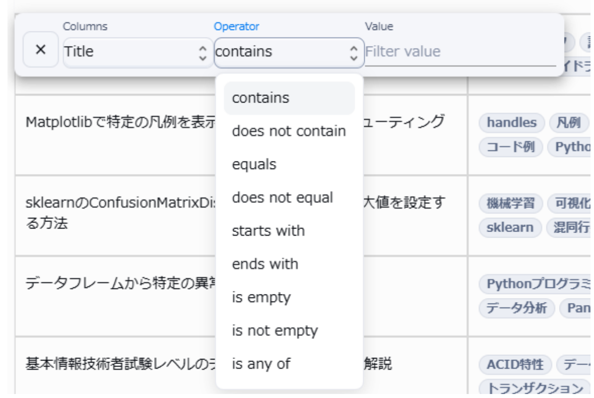

Fig. 2. Filter screen for problem-solving cases.

V. Implementation

This section describes the implementation of the CHAI.

For the development of CHAI, Java/SpringBoot was adopted for the backend, and TypeScript/React/Next.js was used for the frontend.

Furthermore, with the potential for open-source development in the future, Docker [11] was employed as the infrastructure foundation. This enabled the construction of the development environment using Docker Compose.

The login functionality implemented with Spring Security allows administrators to set an organization ID and password. Users can create unique user IDs using the organization ID and log in with their user ID and the organization password.

The generative AI model used for the chat interface is ChatGPT-4o. Since the problem-solving cases collected in this study aim to address issues across various domains, ChatGPT-4o was adopted for the reasons described in Section II-B. Additionally, LangChain [12] was used for connecting with generative AI outside the chat interface.

Based on these technologies, the functionalities described in Sections IV-C, IV-D, and IV-E were implemented.

A. Usage

First, when accessing CHAI, the login screen shown in Fig. 3 is displayed. This screen conveys the purpose and features of CHAI, fostering awareness of collective intelligence cultivation. If the user has not registered, they can register their user information along with the organization ID on the Sign-Up page. Afterward, they can log in by entering the registered user ID and the organization password on the Sign-In page. The organization ID and password are set by the organization's administrator.

After logging in, the screen displaying the list of accumulated problem-solving cases within the organization, as shown in Fig. 4, is displayed. On this screen, clicking on the title of a saved log displays the details of the corresponding problem-solving case. Additionally, a search function is implemented

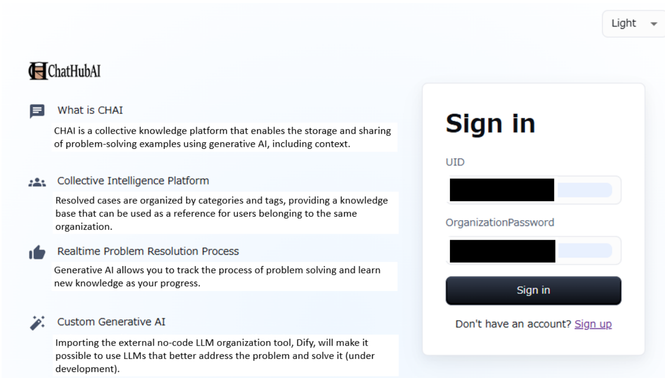

Fig. 3. CHAI Login Screen.

Fig. 5. CHAI Chat Screen.

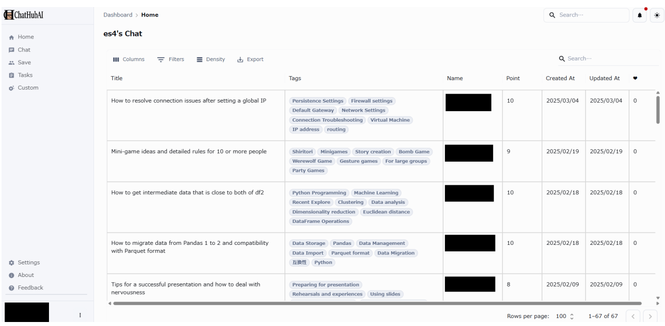

Fig. 4. CHAI Log List Screen.

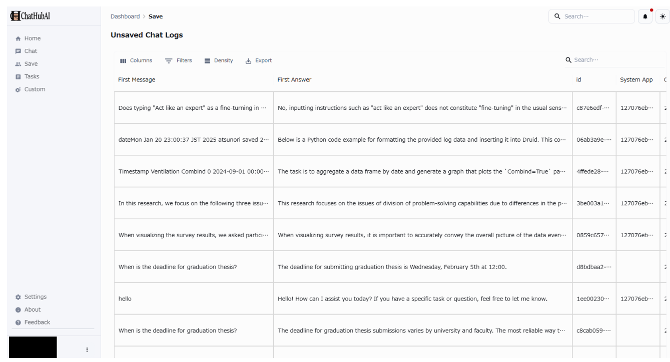

Fig. 6. List of Unsaved Logs in CHAI.

on this screen. By default, the cases are sorted by creation date, but sorting by the number of likes or points is also possible. Furthermore, detailed filtering for each column is available. By clicking the Chat button from the side menu in Fig. 4, the chat screen shown in Fig. 5 is displayed.

As the user continues to engage in problem-solving dialogues on the Chat page, each dialogue is displayed as an "Unsaved Chat" on the Save page, as shown in Fig. 6. These unsaved logs are not publicly accessible. By clicking on a log from this list and entering the context defined in Section IV-D on the screen shown in Fig. 7, the problem-solving case can be saved. Note that while the title and tags are automatically generated from the dialogue log, the user must manually input the other fields.

The saved problem-solving cases are published within the organization on the problem-solving list page, enabling users to view the dialogue logs that led to the solution along with their context.

VI. Experiment

In this study, an experiment was conducted to verify whether CHAI can effectively facilitate knowledge sharing and whether problem-solving cases can be saved without issues when deployed within an organization. The experiment was conducted in the Nakamura Laboratory at Kobe University, where the author belongs (number of members: 32). Each member was issued an organizational login ID and password. CHAI was deployed on January 17, 2025, and its operational performance was evaluated up to April 4, 2025 (the time of writing this paper).

As a result of operating CHAI from January 17, 2025, to April 4, 2025, a total of 32 problem-solving cases were accumulated. Focusing on the distribution of "points" assigned to the context of these problem-solving cases, the results are shown in Fig. 8. Additionally, the average number of points and the median is shown in TABLE II, as derived from the collected data.

TABLE II
Summary Statistics of Saved Problem-Solving Cases

Statistic	Value
Mean Points	6.625
Median Points	8.000

VII. Discussion

CHAI accumulated 32 problem-solving cases over 2.5 months without major technical or operational issues, demonstrating its practicality within an actual organization. The contextual annotations—such as reasons for questions, solution methods, and usefulness scores—helped organize individual insights and made accumulated dialogue logs more reusable than conventional unstructured chat histories.

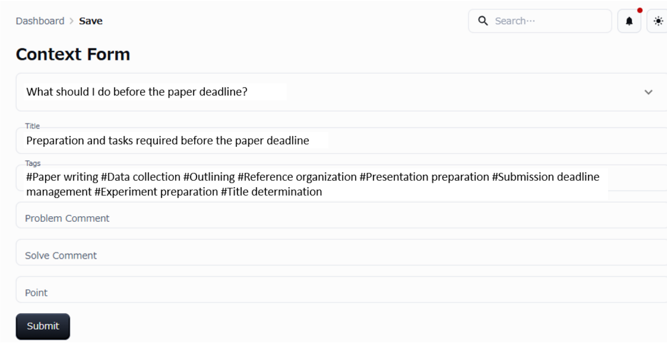

Fig. 7. Context Input Screen in CHAI.

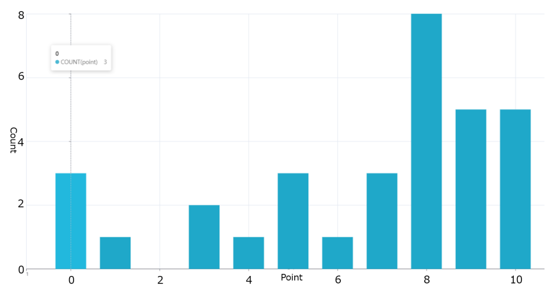

Fig. 8. Point Distribution of Saved Logs.

However, the relatively low number of saved cases suggests that many users continued to engage with generative AI through individual, private interactions, without using CHAI. This indicates that the knowledge-sharing mechanism of CHAI had not yet been fully internalized during the experimental period.

Importantly, not all dialogues are suitable for organizational sharing. Some may include sensitive topics such as confidential strategies, personal concerns, or queries containing private information. To address this, CHAI implements a labeling system that allows users to mark each question-answer pair as "Critical," "Necessary," or "Unnecessary." The "Unnecessary" label hides specific parts of the dialogue from others, enabling selective sharing. This feature supports flexible knowledge management by allowing users to share only relevant and non-sensitive content, while protecting privacy. Nonetheless, it may introduce a bias toward sharing more technical or impersonal topics, which could affect the diversity of shared organizational knowledge.

Regarding evaluation, this study primarily relied on log counts and user-assigned usefulness scores. While these provide a basic indicator of engagement, more comprehensive evaluations—such as measuring actual reuse frequency or surveying perceived impact on task performance—will be needed to assess the true effectiveness of CHAI in promoting organizational learning.

VIII. Conclusion

This study proposed a method to foster organizational collective intelligence by accumulating and sharing dialogue logs with generative AI, including contextual information such as reasons, solution methods, and usefulness scores. To implement this, we developed CHAI, a web application that facilitates the saving, viewing, and searching of structured problem-solving cases. The system was deployed within a real organization and operated stably, with 32 cases accumulated during a 2.5-month period.

CHAI enabled users to preserve valuable dialogue logs and encouraged them to reflect on the problem-solving process. The selective sharing feature using necessity labels also helped balance privacy protection with knowledge sharing.

To further enhance user participation, future work will consider incorporating gamification elements such as badges, rankings, and contribution-based feedback. Visualizing individual and collective contributions could strengthen users' motivation and help recognize active contributors. We also plan to implement features such as reference tracking, comment threads, and notification systems to promote ongoing interaction with shared logs.

In addition, a multifaceted evaluation will be conducted to assess the actual behavioral impact of CHAI on organizational knowledge creation and collaboration. Through these enhancements, we aim to establish CHAI as a sustainable foundation for fostering collective intelligence.

IX. Acknowledgment

This research was partially supported by JSPS KAKENHI Grant Numbers JP25H01167, JP25K02946, JP24K02765, JP24K02774, JP23K17006, JP23K28091, JP23K28383.

References

[1] "Chatgpt," https://chatgpt.com/, visited on 2025-4-30.
[2] "Claude," https://claude.ai/new, visited on 2025-4-30.
[3] G. Orrù, A. Piarulli, C. Conversano, and A. Gemignani, "Human-like problem-solving abilities in large language models using chatgpt," *Frontiers in Artificial Intelligence*, vol. 6, 2023.
[4] T. Z. et al., "Evaluation of openai o1: Opportunities and challenges of agi," 2024. [Online]. Available: \url{https://arxiv.org/abs/2409.18486}
[5] "Collective knowledge mechanisms for society," https://www.jst.go.jp/sis/scienceagora/2021/session/07-c15.html, visited on 2025-4-30.
[6] O. et al., "Gpt-4o system card," 2024. [Online]. Available: \url{https://arxiv.org/abs/2410.21276}
[7] "Chatbot arena," https://lmarena.ai/?leaderboard, visited on 2025-4-30.
[8] N. Knoth, A. Tolzin, A. Janson, and J. M. Leimeister, "Ai literacy and its implications for prompt engineering strategies," *Computers and Education: Artificial Intelligence*, vol. 6, p. 100225, 2024. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2666920X24000262
[9] "Openai is pursuing a new way to fight a.i. 'hallucinations'," https://www.cnbc.com/2023/05/31/openai-is-pursuing-a-new-way-to-fight-ai-hallucinations.html, visited on 2025-4-30.
[10] "Dify," https://dify.ai/, visited on 2025-4-30.
[11] "Docker," https://www.docker.com/, visited on 2025-4-30.
[12] "Langchain," https://www.langchain.com/, visited on 2025-4-30.