Generative AI as a Good Questioner:A RAG-Based Question Transformation Approach for Eliciting Tacit Knowledge in Software Development Organizations

Soonnam Shin
Graduate School of System Design and Management
Keio University
Yokohama, Japan
e-mail: shinsoonnam@keio.jp

Takahiro Yakoh
Graduate School of System Design and Management
Keio University
Yokohama, Japan
e-mail: yakoh@keio.jp

Abstract—**Knowledge in software development organizations often depends on tacit knowledge possessed by experts, leading to challenges, such as insufficient documentation and knowledge silos. Furthermore, changes in work environments driven by remote work and the widespread adoption of generative Artificial Intelligence (AI) have made it increasingly difficult to externalize such knowledge through conventional question–answer processes. Rule-based Knowledge Management Systems (KMS) also struggle to prevent knowledge obsolescence and sustain continuous knowledge update. This study proposes a question transformation approach for eliciting tacit knowledge through an interactive knowledge sharing system that leverages generative AI and Retrieval-Augmented Generation (RAG). By positioning generative AI as an effective "good questioner," the proposed method detects knowledge gaps in the knowledge base and reformulates user queries into exploratory questions designed to elicit tacit knowledge from domain experts. The elicited knowledge is then structured and formalized, enabling iterative updates and reinforcement of the knowledge base. A QA chatbot mediating between knowledge seekers and knowledge providers was implemented and evaluated using a real-world dataset from an operational system within a DevOps-oriented software development organization. The results demonstrate that the proposed approach effectively supports sustainable and efficient knowledge sharing in software development environments.**

Keywords-Knowledge Sharing; Tacit Knowledge; RAG (Retrieval-Augmented Generation); Generative AI; Question Transformation.

I. INTRODUCTION

This section outlines the background of the study and clarifies the research problem and objectives addressed in this paper.

A. *Problem Definition*

The rise of generative Artificial Intelligence (AI) has prompted a reconsideration of how knowledge is utilized and shared within organizations. While individual use of generative AI effectively supports personal problem-solving and idea generation, the insights and learning outcomes derived from such use tend to accumulate internally within individuals. As a result, they are less likely to be systematically shared at the organizational level, increasing the risk that they remain as tacit knowledge siloed to individuals. In software development, tacit knowledge, such as design rationales, experience-based precautions, and troubleshooting know-how—tends to concentrate among experts.

This study focuses on knowledge gaps and tacit knowledge in software development that are not explicitly documented in formalized artifacts, such as specifications and manuals. Conventional Knowledge Management Systems (KMS) rely heavily on rule-based structuring and manual maintenance, which require substantial operational costs and hinder sustainable management.

At the same time, research focusing specifically on the design of questions to elicit tacit knowledge from humans remains limited. Generative AI has primarily been treated as an "answering system," while its potential as a "skillful questioner" has not been sufficiently explored.

B. *Research Objective and Contributions*

To address these challenges, this study aims to design a RAG-based knowledge sharing system with a conversational interface powered by generative AI for software development organizations, and to verify the effectiveness of a question transformation method that actively elicits tacit knowledge for the continuous updating of the knowledge base.

This study makes three primary contributions:

- Theoretical Contribution: It reconceptualizes generative AI not merely as an answer generator but as a questioner and mediator, proposing a new role for AI in the organizational knowledge creation process.
- Methodological Contribution: It proposes, a question transformation method specifically designed for tacit knowledge elicitation.
- Empirical Contribution: It implements and evaluates a RAG-based knowledge sharing system within a DevOps-oriented software development organization, demonstrating its effectiveness.

Through these contributions, this research provides a novel perspective on the formalization of tacit knowledge from the viewpoint of question transformation in the design of generative AI–enabled knowledge sharing systems.

The remainder of this paper is organized as follows. Section II reviews related work. Section III presents the proposed method and system implementation. Section IV describes the evaluation and results. Section V concludes the paper and discusses limitations and future work.

II. Related Work

This section reviews related research to position the present study within the existing body of knowledge. It first examines Knowledge Management approaches, followed by studies on tacit knowledge elicitation from experts and concludes with a review of research on RAG-based information augmentation and question transformation methods.

A. *Knowledge Management*

A prior study on ontology-based Knowledge Management tools for organizational knowledge sharing [1] provides a comprehensive review of case studies and implementation examples of ontology-based KMS. The study evaluates them using ten comparative criteria: motivation, domain, knowledge sources, types of knowledge, knowledge extraction processes, knowledge input processes, knowledge retrieval processes, knowledge sharing technologies, sources of ontology components, and ontology methodologies. However, challenges remain in addressing knowledge sharing problems that require person-to-person knowledge transfer, particularly in terms of technical approaches for extracting and retrieving knowledge from implicit sources across diverse knowledge domains. While ontology-based Knowledge Management tools are effective in organizing explicit knowledge, they face difficulties in extracting and updating tacit knowledge.

B. *Tacit Knowledge Elicitation from Experts*

To improve the collection of tacit knowledge in KMS, a prior study [3] proposes a storytelling-based approach for knowledge sharing. Compared to conventional interview-based or video-based methods, storytelling is suggested to reduce psychological resistance among knowledge holders and facilitate more natural knowledge sharing.

However, the study lacks large-scale experimental validation and quantitative evaluation, remaining issues regarding its practical effectiveness. Moreover, storytelling content tends to be subjective and unstructured, making knowledge standardization and systematic integration into knowledge systems difficult.

A case study on the use of Large Language Models (LLMs) in manufacturing environments [2] demonstrated their effectiveness for knowledge management and information retrieval support; however, mechanisms for eliciting tacit knowledge from domain experts were not sufficiently discussed.

Prior research has identified significant barriers to tacit knowledge sharing in software development teams, including team culture, trust, communication, and team dispersion [7]. These barriers limit access to tacit knowledge required for socio-technical tasks and contribute to project failures. However, existing studies primarily focus on identifying these barriers, with limited attention to mechanisms for actively eliciting tacit knowledge.

C. *Retrieval-Augmented Generation (RAG)*

A prior study analyzing the operational and validation challenges of systems based on RAG [4] systematically identifies seven Failure Points (FP) in RAG-based system design [5]: In particular, FP1 Missing Content highlights the problem of insufficient content caused by the absence of mechanisms for continuously maintaining and updating the knowledge base.

D. *Question Transformation*

A prior study on prompts for transforming ambiguous questions into more specific queries [6] proposes a novel prompting method called Ambiguity Type–Chain of Thought (AT-CoT). This approach enables LLMs to better understand user queries by identifying the type of ambiguity involved and generating clarification questions accordingly.

While existing research primarily focuses on query augmented techniques aimed at improving answer accuracy, relatively limited attention has been paid to question generation methods designed to elicit knowledge from humans (i.e., domain experts).

III. System Architecture and Implementation

The distinguishing feature of this study lies in proposing a knowledge sharing process centered on AI-mediated human interaction and question transformation for tacit knowledge elicitation. While conventional knowledge sharing and QA systems mainly focus on retrieving existing knowledge and generating answers, the proposed system extends this process by detecting knowledge gaps and transforming them into opportunities for tacit knowledge acquisition.

A. *Knowledge Sharing Model*

Figure 1 illustrates the Knowledge Sharing Model using the RAG-Based Question Transformation Method. The model consists of three interrelated flows: (A) Knowledge Sharing, (B) Knowledge Elicitation, and (C) Knowledge Acquisition.

Flow A, Knowledge Sharing, corresponds to a conventional QA process over a knowledge base. When a knowledge seeker submits a question, the system retrieves relevant information from the knowledge base through a RAG pipeline and generates a response in natural language.

Flow B, Knowledge Elicitation, is activated when the system detects that the knowledge base does not contain sufficient information to provide a reliable answer. In such cases, the AI agent GapNavigator generates exploratory questions through Exploratory Question Transformation (EQT). These questions are presented to a knowledge provider in order to elicit tacit knowledge that has not been explicitly documented.

Flow C, Knowledge Acquisition, structures and formalizes the elicited knowledge and incrementally integrates it into the knowledge base. Through this process, knowledge that was previously unavailable in routine QA interactions can be accumulated and reused in future knowledge sharing.

This model extends conventional QA-based knowledge sharing by introducing an explicit mechanism for identifying missing knowledge and converting it into opportunities for knowledge elicitation and knowledge base reinforcement.

B. *Knowledge Gap Detection in the RAG Pipeline*

A central component of the proposed system is the detection of knowledge gaps within the RAG pipeline. In this study, a knowledge gap is defined as a state in which the system cannot provide a sufficiently reliable or complete answer based on the current knowledge base.

Knowledge gaps are identified based on the following conditions:

1) Retrieved documents do not contain sufficient information to answer the query;

2) The generated response indicates uncertainty or the absence of relevant information;

When one or more of these conditions are satisfied, the query is classified as a knowledge-gap query, and GapNavigator is activated. In this way, the proposed system does not terminate at retrieval failure, but instead transforms failure into an opportunity for tacit knowledge elicitation.

C. *The Concept of "Generative AI as a Good Questioner" and EQT Method*

This study positions generative AI not merely as an answer generator but as a "good questioner" that facilitates the externalization of tacit knowledge. When a knowledge gap is detected, the proposed EQT method reformulates the original query into structured exploratory questions for knowledge providers. EQT is implemented through prompt engineering rather than model retraining or fine-tuning.

EQT is designed to elicit tacit knowledge in software development organizations from five predefined perspectives: (1) design philosophy and decision rationale, (2) knowledge provider heuristics and practical know-how, (3) dependencies and impact scope, (4) exceptional cases and failure knowledge, and (5) implicit rules and assumptions.

The prompt design strategy consists of four elements: role instruction, contextual augmentation, transformation constraints, and output constraints. Specifically, the LLM is instructed to act as an assistant for tacit knowledge elicitation, to use the original user query and retrieved context as input, to avoid simple paraphrasing, and to generate specific and answerable exploratory questions for knowledge provider.

The transformation procedure is defined in Algorithm 1. First, the system receives the original query and the retrieved context from the RAG pipeline. Second, when the available information is judged insufficient, the query is classified as a knowledge-gap query. Third, the relevance of the five tacit knowledge perspectives is evaluated based on the query and context. Fourth, the candidate perspectives are ranked, and up to the top three are selected. Finally, one exploratory question is generated for each selected perspective and presented to the knowledge provider.

The prioritization of perspectives is based on the semantic relevance to the original query, contextual relevance to the retrieved documents, the degree to which the missing information depends on knowledge provider judgment or experience, and the expected usefulness for knowledge base update. By explicitly constraining the transformation process in this way, EQT provides a transparent and reproducible mechanism for AI-mediated tacit knowledge elicitation while keeping the cognitive burden on knowledge providers manageable.

D. *Implementation Overview*

The system was implemented as a RAG-based chatbot using the generative AI development platform Dify. The knowledge base was constructed from documents related to an operational Intellectual Property Management System (IPMS) within a software development organization. The chatbot processes user questions through the RAG pipeline, while GapNavigator monitors the output and invokes EQT when a knowledge gap is detected.

Algorithm 1. Exploratory Question Transformation (EQT)

```
Input:   Original query Q, retrieved context C, tacit knowledge perspectives P
Output:  Exploratory questions E
         Receive Q and C from the RAG pipeline
         if C is insufficient then
             Classify Q as a knowledge-gap query
             for each p ∈ P do evaluate r(p) based on Q and C
             Rank perspectives by r(p) and select up to the top three P′
             for each p ∈ P′ do generate one exploratory question e based on Q, C, and p
             Add e to E
             Present E to the knowledge provider
         end if
```

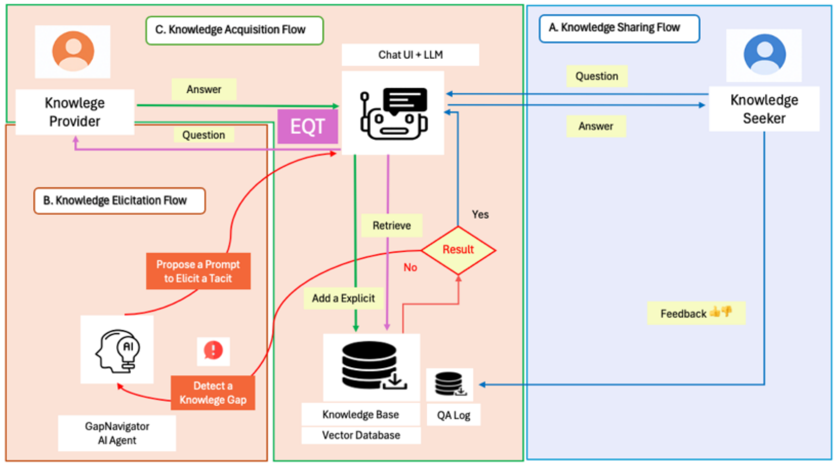

Figure 1. Knowledge Sharing Model with RAG-Based Question Transformation: from retrieval-based answering to knowledge gap detection, tacit knowledge elicitation, and iterative knowledge acquisition.

By integrating retrieval, gap detection, and tacit knowledge elicitation into a single framework, the proposed system supports not only access to existing knowledge but also the continuous acquisition and formalization of tacit knowledge for sustainable organizational knowledge sharing.

The current implementation is intended as a prototype to verify the feasibility of AI-mediated tacit knowledge elicitation in a software development context. Accordingly, the behavior of knowledge gap detection and EQT is influenced by the quality of the retrieved context and the prompt design. Therefore, the present implementation should be regarded as a design instantiation for feasibility validation rather than a universally optimized architecture.

IV. Evaluation and Experiment

This section evaluates the effectiveness of the proposed mechanism and knowledge sharing system in relation to the research objectives defined in Section I.

A. *Experimental Setup*

The evaluation was conducted using the proposed RAG-based chatbot implemented on the Dify platform. For the evaluation, the dataset was constructed from publicly available user documentation, including manuals, related to the IPMS, a real-world operational system in which the author is directly involved as a developer. These documents were converted into a structured knowledge base for the experiment

For the RAG configuration, the embedding model text-embedding-3-large was used, with the top-K retrieval parameter set to 2.

Using the IPMS dataset, two model conditions were compared: (1) a baseline model that generated general follow-up questions when a knowledge gap was detected, and (2) an EQT model that generated structured exploratory questions based on predefined tacit knowledge perspectives. Both models shared the same RAG pipeline and LLM configuration, while differing in the strategy used for question generation after knowledge gap detection. Both models employed the same LLM, gpt-5-chat-latest, with the temperature parameter set to 0.7 and the maximum token length set to 512. Table 1 summarizes the main differences in prompt design between the baseline and EQT models.

TABLE 1. COMPARISON OF PROMPT DESIGN BETWEEN THE BASELINE AND EQT MODELS.

Item	Baseline Model	EQT Model
Purpose	Generate general follow-up questions for knowledge-gap queries	Generate structured exploratory questions for eliciting tacit knowledge
Question generation style	General follow-up questioning	Exploratory questioning guided by tacit knowledge perspectives
Knowledge orientation	Focus on obtaining missing explicit information	Focus on eliciting tacit knowledge that is reusable and formalizable
Tacit knowledge perspectives	Not explicitly specified	Five predefined perspectives: design rationale, expert know-how, dependencies and impact, exceptions and failures, and implicit rules and assumptions
Multi-element query handling	Organize the content before generating questions	Break down the content step by step and generate exploratory questions

In the EQT model, the system prompt was configured to elicit tacit knowledge from five predefined perspectives commonly observed in software development organizations.

The evaluation was conducted within a DevOps-oriented software development organization. A total of 13 participants were recruited, representing diverse organizational roles, including system development, operations and support, sales, and management. The participants ranged from novice employees with less than two years of experience to senior professionals with more than 21 years of experience.

Participants entered the same original questions into both model conditions using both predefined and free-form queries and compared the generated outputs without being informed of the model names during the evaluation.

B. Metrics

The generated questions under the two model conditions were evaluated using a five-point Likert scale based on two criteria: (1) clarity in identifying and articulating knowledge gaps, and (2) effectiveness in eliciting tacit knowledge. In addition, participants conducted a comparative assessment to determine which model generated questions that were more effective in eliciting tacit knowledge.

1) Tacit Knowledge (TK) Elicitation Rate

The TK Elicitation Rate was used as an auxiliary metric to assess whether the exploratory questions generated by EQT were answerable by knowledge providers in practice. In this study, a generated question was regarded as valid if an answer could be provided by a participant acting as a knowledge provider. The metric was therefore based on actual response behavior rather than on a post hoc judgment of question quality.

The TK Elicitation Rate is defined as the proportion of answerable EQT-generated questions among all exploratory questions generated from knowledge-gap queries, as shown in (1). Because the EQT model generated up to three exploratory questions for each original knowledge-gap query, the denominator is defined as the number of original knowledge-gap queries multiplied by three.

The validity of generated questions was assessed through actual responses provided by participants according to their areas of expertise. When a participant was not able to answer a question because it fell outside their domain, the question was referred to another participant with relevant domain knowledge. This metric should therefore be interpreted as an operational indicator of the practical answerability of EQT-generated elicitation questions, rather than as a direct measurement of tacit knowledge itself. Therefore, no separate rubric table was used for this metric. In addition, no inter-rater reliability analysis was conducted, because validity was not determined by post hoc labeling across multiple evaluators, but by actual answerability in practice.

$$TK\ Elicitation\ Rate = \frac{Number\ of\ valid\ Questions}{3 \times (Number\ of\ Generated\ Questions)} \quad (1)$$

C. Results

Figure 2 presents the results of the model comparison using the Wilcoxon signed-rank test. For Knowledge Gap Clarifying Level, the EQT model showed slightly higher scores than the baseline model, although the difference was not statistically significant. For Tacit Knowledge Elicitation Level, the EQT model showed clearly higher scores, with a statistically significant difference ($p = 0.002$). In the direct comparative assessment, 93.3% of participants selected the EQT model as more effective for tacit knowledge elicitation ($p = 0.0034$).

These results suggest that the proposed EQT method more effectively generated questions that were perceived as useful for eliciting tacit knowledge from knowledge providers. The findings support the validity of transforming knowledge-gap queries into exploratory questions rather than treating retrieval failure as the end point of the interaction.

Using the EQT model, 102 exploratory questions were generated from 34 original knowledge-gap queries. Among these, 99 questions received actual responses from participants and were therefore regarded as valid, resulting in a TK Elicitation Rate of 97.1%. This result indicates that most EQT-generated questions were practically answerable by knowledge providers in the present experimental setting.

This metric should be interpreted with caution. It reflects the practical answerability of EQT-generated elicitation questions, rather than the completeness or quality of tacit knowledge eventually obtained. Taken together, the statistical comparison and the TK Elicitation Rate suggest that EQT was effective as a structured question transformation mechanism for tacit knowledge elicitation in the present study.

The results should also be interpreted in light of the limited experimental scope, as the evaluation was conducted in a single organization with a domain-specific dataset and a relatively small number of participants.

V. CONCLUSION AND FUTURE WORK

This section summarizes the main findings of this study, discusses its limitations, and outlines directions for future research.

A. Conclusion

While existing studies primarily focus on knowledge retrieval and QA systems, this study introduces a question transformation approach that positions generative AI as a skillful questioner for eliciting and formalizing tacit knowledge.

Using a real-world dataset from an operational system, the experiments showed that knowledge providers could externalize tacit knowledge through EQT-generated exploratory questions, thereby strengthening the knowledge base and supporting organizational knowledge sharing.

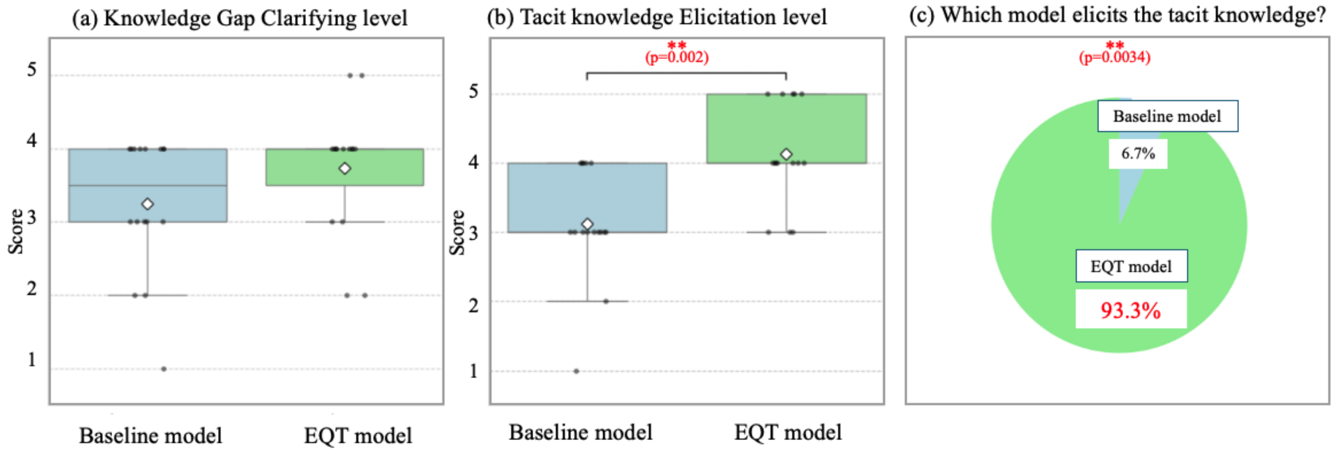

Figure 2. Model comparison: Wilcoxon signed-rank test.

The significance of this study lies in showing that tacit knowledge in software development can be externalized through ordinary Q&A interactions without relying solely on costly, non-routine methods such as interviews or workshops. Thus, the study contributes to knowledge sharing system design by advancing tacit knowledge formalization from the perspective of question transformation.

B. *Limitations*

This study does not aim to comprehensively examine all possibilities of generative AI–enabled knowledge sharing, nor does it evaluate the intrinsic performance of generative AI models or RAG architectures themselves. Accordingly, quantitative comparisons of retrieval accuracy, answer generation performance, and differences across LLM models were outside the scope of this research.

The proposed system focuses on tacit knowledge that is linguistically expressible but not spontaneously articulated; embodied skills and highly intuitive expertise were outside the scope of this study.

C. *Future Work*

Generative AI has the potential to provide a dynamic and interactive platform for managing knowledge sharing within organizations. However, careful consideration must be given to operational design and the potential burden placed on knowledge providers. Determining the appropriate degree of AI-agent intervention represents an important area for future investigation. In addition, organizational challenges remain, including how to evaluate and recognize the contributions of knowledge providers within AI-mediated knowledge sharing processes.

As a direction for future work, expanding the applicability of the proposed method is an important priority. By integrating knowledge sources of varying contents and formats, such as design review records, incident response logs, and chat histories—it may become possible to complementarily elicit a broader range of tacit knowledge.

REFERENCES

[1] M. A. Osman, S. A. M. Noah, and S. Saad, “Ontology-based knowledge management tools for knowledge sharing in organization—a review”, IEEE access, vol. 10, pp. 43267-43283,2022.

[2] S. K. Freire et al., “Knowledge sharing in manufacturing using LLM-powered tools: user study and model benchmarking.”, Frontiers in Artificial intelligence, vol. 7, 1293084, 2024.

[3] N. Shaw and P. Liu, “A knowledge management system (KMS) using a storytelling-based approach to collect tacit knowledge.”, IEEE, pp. 1-6, 2016.

[4] P. Lewis et al., “Retrieval- augmented generation for knowledge-intensive nlp tasks.”, Advances in neural information processing systems, vol. 33, pp. 9459-9474, 2020.

[5] S. Barnett, S. Kurniawan, S. Thudumu, Z. Brannelly, and M. Abdelrazek. “Seven failure points when engineering a retrieval augmented generation system.”, In Proceedings of the IEEE/ACM 3rd International Conference on AI Engineering-Software Engineering for AI, pp. 194- 199, 2024.

[6] A. Tang, I. Soulier, and V. Guigue, “Clarifying ambiguities: on the role of ambiguity types in prompting methods for clarification generation.”, In Proceedings of the 48th International ACM SIGIR Conference on Research and Development in Information Retrieval, pp. 20-30, 2025.

[7] E. Mtsweni and K. Gorejena, “Team barriers to tacit knowledge sharing in software development project teams.”, Electronic Journal of Knowledge Management, vol. 21(1), pp. 59-72, 2023.