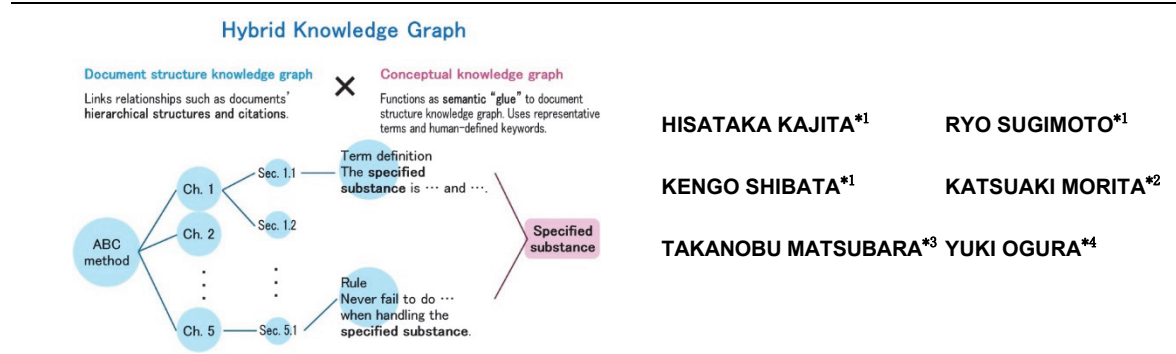


Industrial Application of Hybrid Knowledge Graph RAG Leveraging Structural and Conceptual Relationships in Industrial Documents



In corporate activities, many business operations rely on internal industrial documents, such as laws/regulations, standards, specifications, design standards, and manufacturing instructions, as their underlying evidence for decision-making. These documents require referencing across hierarchical structures, citation relationships, defined terms and relevant documents. As a result, conventional retrieval-augmented generation (RAG) may fail to retrieve necessary supporting evidence. This report presents a method that uses a hybrid knowledge graph, combining structural and conceptual relationships, to expand candidate references using initial search results obtained via traditional RAG as a starting point. Moreover, large language models are used to extract candidate node types and relationship types tailored to the target document set, thereby supporting the initial development of the ontology design. Verification conducted using test data designed for internal workflows confirmed improvements in both the rate of relevant document retrieval and answer accuracy across high-pressure gas-related laws/regulations, Foreign Exchange and Foreign Trade Act-related laws/regulations, and the EU Artificial Intelligence Act.

1. Introduction

In Mitsubishi Heavy Industries (hereinafter referred to as MHI) Group, initiatives are under way to enhance product and service value as well as internal workflows through cross-functional utilization of data across design, production, procurement and service. With the recent advancement of large language models (hereinafter referred to as LLMs), document-grounded knowledge utilization is of increasing importance. Specifically, it includes looking up legal documents, searching through standards, referencing design standards, and confirming production procedures.

A typical method used for this purpose is retrieval-augmented generation (hereinafter referred to as RAG). In this report, the traditional RAG refers to a framework in which text is segmented into chunks (as searchable units), before relevant chunks are obtained by vector or keyword search and are input into LLMs for answer generation. As shown in **Figure 1**, MHI Group has deployed a document-grounded knowledge utilization platform called the Artificial Intelligence BOosting Work (hereinafter referred to as AIBOW), which has been in use for multiple business operations such as design, quality, legal compliance and maintenance.

Industrial documents, however, require readers to go through texts, with the understanding of relational linkage of document structures and technical terms, because key information is dispersed across multiple locations. Therefore, as described in Chapter 2 below, search results from traditional RAG alone may not contain sufficient supporting evidence required for answer generation.

This report presents the hybrid knowledge graph RAG that leverages structural and conceptual relationships to expand the candidate pool using the initial search result with traditional RAG as the starting point. Moreover, LLMs are used to extract candidate node types and

*1 Digital Intelligence R & D Department, Digital Innovation Headquarters, Mitsubishi Heavy Industries, Ltd.

*2 Research Manager, Digital Intelligence R & D Department, Digital Innovation Headquarters, Mitsubishi Heavy Industries, Ltd.

*3 Manager, Digital Intelligence R & D Department, Digital Innovation Headquarters, Mitsubishi Heavy Industries, Ltd.

*4 Manager, DX Solution Division, Akari Inc.

relationship types tailored to a target document set, thereby supporting the initial development of the ontology design. This is also outlined in this report. The effectiveness of this methodology is evaluated through its application to legal documents, followed by discussion on its adaptability to design and production documents.

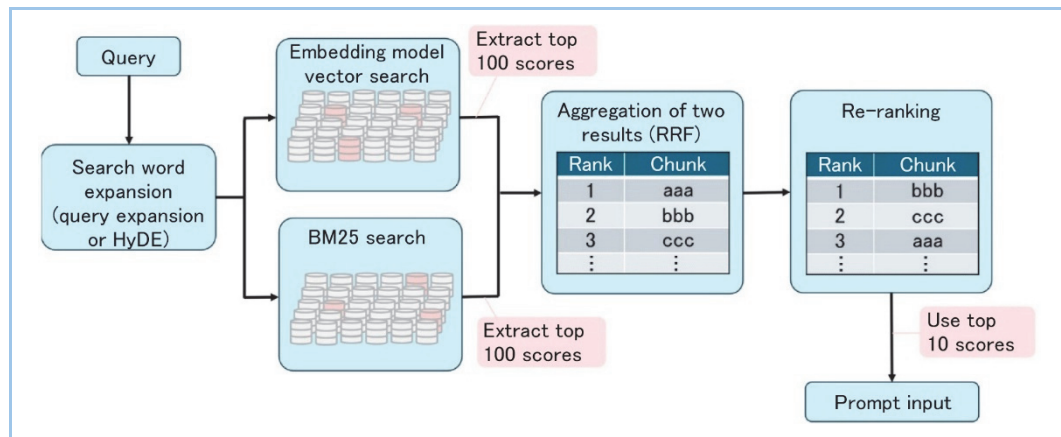


Figure 1 Architecture of RAG in AIBOW, MHI's internal knowledge utilization framework

2. Characteristics of industrial documents and knowledge utilization challenges

In industrial documents such as laws/regulations, specifications, customer's specifications, internal standards, design documents, manufacturing instructions and inspection records, a single chunk of text alone cannot determine the full meaning of description in many cases. The three primary reasons for this are as follows:

- Higher-level information within hierarchical structures, such as chapters, sections, articles, paragraphs, or process/task units, sets the scope and preconditions, which are not included within the chunk itself.
- Supporting evidence is dispersed across the main text, attached tables, appendices and related documents as per citation and reference relationships.
- As technical terms, defined terms and related terms are differently phrased in each document, search strings may be different for the same subject.

Moreover, establishing appropriate chunk boundaries and configuring search rules are difficult due to the following factors:

- As document formats and granularity levels of description vary across different operational tasks, horizontal deployment is difficult with fixed extraction rules.

Traditional RAG is effective in retrieving chunks close to a query. However, it has difficulty in simultaneously obtaining the information outside the chunks, such as preconditions, referenced texts, explanation of defined terms, and related descriptions with different wording. For example, in some cases of legal queries, the text of an article containing search terms can be retrieved, but another article or an attached table containing the definition referenced by the retrieved article may not be obtained. Similarly, in export control classification, when wording used in specifications is not identical to the legal terms, the articles necessary for the determination may not be included in the initial search result. Such search omissions decrease the capacity to provide accurate answers and the rationale.

Given these challenges, our new methodology is underpinned by the following ideas. The shortage of preconditions caused by hierarchical structures is dealt with by expanding the candidate pool based on parent-child relationships. Citation and reference relationships are used to address omissions of referenced texts and appendices. Wording variations in technical terms and defined terms are addressed using conceptual relationships. For the issues of diverse document formats and varying levels of granularity of description in each document, LLMs are used to extract candidate node types and relationship types, thereby supporting the initial development of the ontology design.

Aiding ontology design here does not mean fully automated creation of ontologies. The

process involves extracting and organizing candidate node types and relationship types based on the target document set and intended use, before completing with manual confirmation. This makes it easier to apply candidate expansion using structural and conceptual relationships, while suppressing the need for target-specific custom modeling.

3. Architecture of hybrid knowledge graph RAG

In this methodology, the initial search results obtained via traditional RAG serve as the starting point to expand candidate references on a hybrid knowledge graph for answer generation.

Figure 2 illustrates the architecture of this methodology.

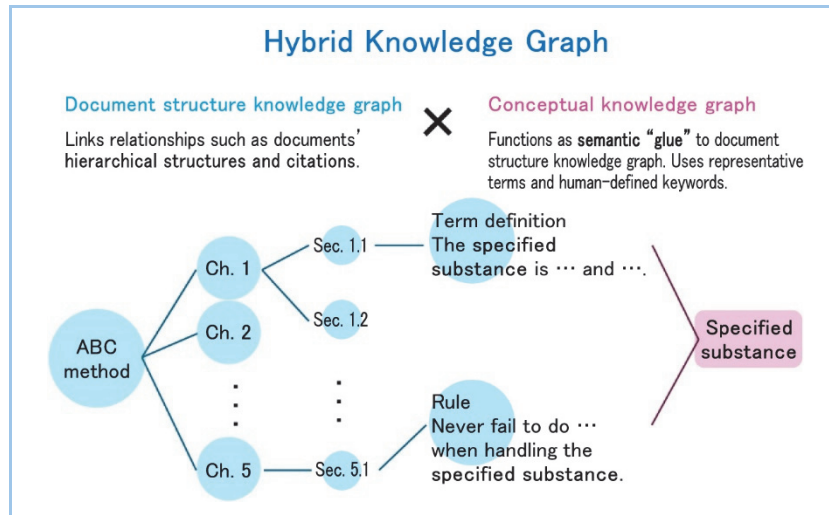


Figure 2 Architecture of proposed hybrid knowledge graph

3.1 Hybrid knowledge graph

In this report, the hybrid knowledge graph handles structural and conceptual relationships on the same graph. Structural relationships represent parent-child relationships (e.g., chapters, sections, articles and paragraphs), citation relationships and reference relationships. Conceptual relationships represent semantic linkages such as keywords, defined terms, synonyms and related terms.

The use of structural relationships makes it possible to include, based on the searched articles and sections, additional candidates from the higher/lower hierarchical levels, cited sources and referenced texts. Conceptual relationships can be employed to add as candidates the descriptions related to the same concept, even if they are located far apart within the document structure.

3.2 Ontology design support and graph creation

In this report an ontology refers to the definitions of node types and relationship types used in a given document group. In our methodology, LLMs are leveraged to extract and organize candidate node types and relationship types based on an overview of the target document set and its intended use. The obtained candidates are manually reviewed and established as the ontology, which is used to generate the knowledge graph.

Figure 3 gives an example of a hybrid knowledge graph. For data with a clear document structure like HTML, structural relationships are extracted using headings and hierarchical information. As the structure and wording in PDF and Word documents vary depending on the document, LLMs are used to extract candidate nodes and edges in alignment with the content of the document.

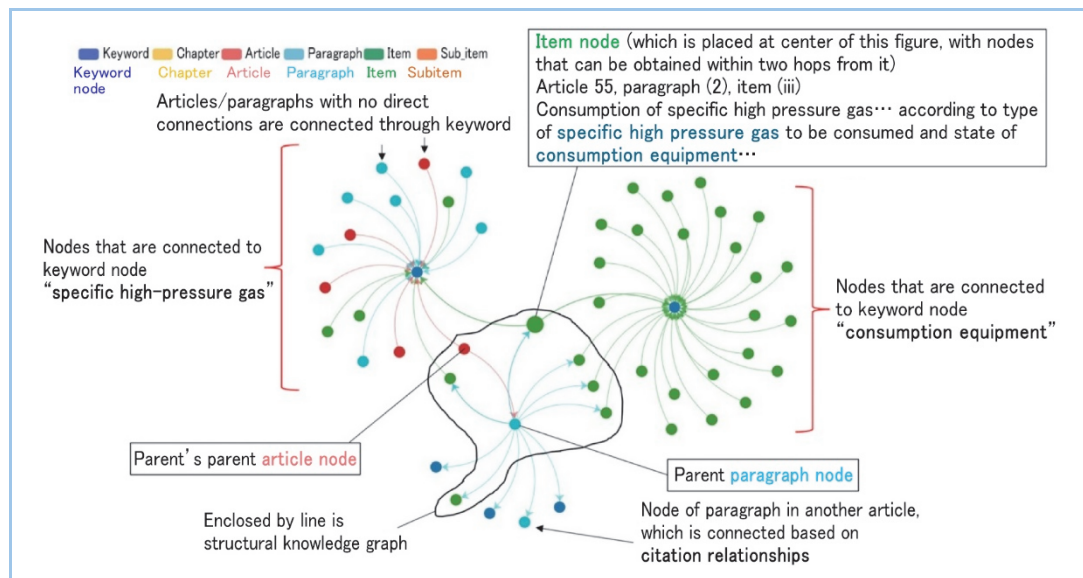


Figure 3 Hybrid knowledge graph based on High Pressure Gas Safety Act and its related laws/regulations

3.3 Flow of search and answer

During the search, traditional RAG is first used to obtain initial candidates. In this assessment, the improved RAG system for manufacturing documents⁽¹⁾, which was reported at the annual meeting of the Association for Natural Language Processing in Japan (NLP2025), was employed for the initial search. Specifically, vector search using a multilingual embedding model is combined with BM25 on normalized text chunks, followed by RRF and re-ranking to obtain the initial candidate set.

Next, candidate expansion is performed using the nodes corresponding to initial candidates as the starting point to follow structural and conceptual relationships. The obtained candidates are restructured into coherent units such as articles, paragraphs and items, before being provided as input to LLMs for answer generation.

4. Assessment of applicability to legal documents

4.1 Introductory information about assessment

This methodology was evaluated by extending the retrieval module of AIBOW, a document-grounded knowledge utilization platform used in MHI Group. The evaluation results were compared with those of the traditional RAG system featuring text normalization, hybrid search and re-ranking, as reported at NLP2025⁽¹⁾.

For the assessment, test data were prepared considering the actual internal workflows. Queries were generated assuming legal inquiries for high-pressure gas-related laws/regulations, and for export classification determination in the context of Foreign Exchange and Foreign Trade Act (hereinafter referred to as FEFTA)-related laws/regulations, respectively. The results were compared in terms of the rate at which the necessary relevant documents were retrieved within the top 10 search results. Furthermore, generated answers were evaluated for the validity of expected conclusions and the appropriateness of supporting evidence.

4.2 Application to Japanese legal documents

This methodology was applied to the high-pressure gas-related laws/regulations (comprising the High Pressure Gas Safety Act, and Regulation on Safety of General High Pressure Gas) and the FEFTA-related laws/regulations (including FEFTA, Foreign Exchange Order, and Export Trade Control Order). The primary results are presented in **Table 1**.

For the high-pressure gas-related laws/regulations, the retrieval rate of relevant documents improved from 59.1% to 88.6%, with an answer accuracy of 86.7%. In the FEFTA-related laws/regulations, the retrieval rate of relevant documents improved from 27.5% to 68.6%, achieving 100% accuracy for applicable items and 66.7% for non-applicable items.

For high-pressure gas-related laws/regulations, expanding the candidate pool by following hierarchical structures and citation relationships was effective. In contrast, for FEFTA-related

laws/regulations, conceptual relationships by which the phrases in specifications can be linked with legal concepts were effective. This has made it possible to provide rationales grounded in relevant articles and defined terms.

Table 1 Search and answer assessment results with Japanese legal documents

Laws/regulations	Purpose of use	Traditional RAG	Our methodology	Correct answer
High-pressure gas related	Legal referencing	59.10%	88.60%	86.70%
FEFTA related	Export classification determination	27.50%	68.60%	100% for applicable items 66.7% for non-applicable items

4.3 Application to EU AI Act

Our methodology was also applied to the EU Artificial Intelligence Act (hereinafter referred to as EU AI Act), with the aim of verifying the applicability to regulatory documents outside Japan. As the EU AI Act has many cross-references between the main articles and the annexes, requirements scattered across multiple locations need to be processed together. The results showed that this methodology made it easier to obtain candidates based on the citation relationships and referenced texts in annexes, resulting in improvements in retrieval and answer accuracy. It has thus been verified that our methodology is applicable to regulatory documents with many cross-references.

5. Adaptability to design/production documents and industrial application

5.1 Documents to be applied and how to approach

Having been verified with legal documents, this methodology is now aimed to be applied to another group of documents, that is, those related to design, manufacturing and cross-functional workflows. It is necessary to adjust the prioritized node types and relationship types depending on the target. These documents mentioned above are wide-ranging, including standards, specifications, design standards, work instructions, record sheets, reports, procedures, contracts and internal regulations.

Design documents mainly focus on the connections between requirement specifications, design criteria, acceptance criteria, and related standards, while manufacturing documents mainly focus on processes, target parts, equipment used, verification items, criteria, and precautions. Cross-functional workflows involve checking across multiple groups of documents and confirming the rationale with these different sources. It is therefore important to change the designs of structural and conceptual relationships depending on the given target documents.

5.2 Trial application to welding-related documents

As an example of manufacturing documentation, welding-related documents were selected for trial application. In these documents, information such as work procedures, target areas, equipment used, verification items and criteria are scattered across multiple documents. In addition to varying document formats and description granularity, processes or tasks with identical names may be used in different contexts.

This trial confirmed the insufficiency of being based solely on fixed hierarchical structures as in the case of legal documents and the importance of relating entities (e.g., processes, workpieces, conditions, precautions and criteria) based on task context in manufacturing documents. Furthermore, it confirmed the necessity of restructuring search results into coherent units based on document granularity, rather than using them directly in chunk units.

5.3 Key points for industrial application

This methodology is considered to have high applicability in supporting legal and internal regulation inquiries, regulatory compliance verification, export control, quality assurance, searching for supporting evidence across standards, specifications and contracts, and referencing support for design documents and manufacturing procedures.

In practical operations, it is effective to gradually introduce structural/conceptual relationships in the operation as needed, while making use of the existing traditional RAG system. It is also necessary to consider access rights management, version control, and how to keep up with

law amendments and document revisions. This methodology can be deployed as a common framework for evidence retrieval across the domains of design, quality, production and compliance, while enabling LLM-based ontology design to minimize the load of initial conceptualization when source documents are added.

6. Conclusion

This report presents a methodology to expand initial search results obtained via traditional RAG into supporting evidence, using a hybrid knowledge graph combining structural and conceptual relationships. Furthermore, LLMs are leveraged to extract and organize candidate node types and relationship types thereby supporting the initial development of the ontology design tailored to the target document set.

In the application to legal documents using test data prepared in consideration of actual internal workflows, improvement in both retrieval and answer quality was confirmed for high-pressure gas-related laws/regulations, FEFTA-related laws/regulations, and the EU AI Act. It has thus been demonstrated that this methodology is effective for supporting decision-making that involves citation relationships, defined term alignment, and referencing across multiple documents.

This methodology also has potential for application to industrial documents beyond legal documents. Looking forward, MHI Group will conduct further verifications on documents such as standards, specifications, work instructions, procedures, contracts and internal regulations to confirm its applicability for supporting practical operations.

References

- (1) K. Shibata et al., Improving Retrieval-Augmented Generation for Real-World Data in Manufacturing, Proceedings of the 31st Annual Meeting of the Association for Natural Language Processing (NLP2025) C5-3 (2025) (in Japanese)
- (2) Wang, H. et al., Retrieval-Augmented Generation for Large Language Models: A Survey, arXiv:2312.10997 (2023)
- (3) Robertson, S. and Zaragoza, H., The Probabilistic Relevance Framework: BM25 and Beyond, Foundations and Trends in Information Retrieval, Vol.3 No.4 (2009)
- (4) Glass, M. et al., Re2G: Retrieve, Rerank, Generate, Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (2022)
- (5) Jagerman, R. et al., Query Expansion by Prompting Large Language Models, arXiv:2305.03653 (2023)
- (6) Gao, L. et al., Precise Zero-Shot Dense Retrieval without Relevance Labels, Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (2023)
- (7) Chen, J. et al., M3-Embedding: Multi-Linguality, Multi-Functionality, Multi-Granularity Text Embeddings Through Self-Knowledge Distillation, Findings of the Association for Computational Linguistics: ACL 2024 (2024)
- (8) Cormack, G. V. et al., Reciprocal Rank Fusion Outperforms Condorcet and Individual Rank Learning Methods, Proceedings of the 32nd International ACM SIGIR Conference on Research and Development in Information Retrieval (2009)
- (9) Nogueira, R. and Cho, K., Passage Re-ranking with BERT, arXiv:1901.04085 (2019)
- (10) Edge, D. et al., From Local to Global: A Graph RAG Approach to Query-Focused Summarization, arXiv:2404.16130 (2024)
- (11) Jimenez Gutierrez, B. et al., HippoRAG: Neurobiologically Inspired Long-Term Memory for Large Language Models, Advances in Neural Information Processing Systems 37 (2024)