Development of an LLM-Based Framework for Research–Regulatory Alignment and Integrated Technology Mapping: A Case Study on TRU Fuel Deployment in Small Modular Reactors

SunWoo Park, Ho Jin Ryu *
Department of Nuclear and Quantum Engineering, Korea Advanced Institute of Science and Technology (KAIST), 291 Daehak-ro, Yuseong-gu, Daejeon 34141, Republic of Korea
**Corresponding author: hojinryu@kaist.ac.kr*

Keywords : LLM, AHP, framework development, research-regulatory alignment, technology mapping

1. Introduction

The nuclear industry is undergoing rapid change with the advancement of high-burnup technologies and the emergence of advanced reactor concepts. Research achievements accumulated across diverse technical areas have continuously enhanced the overall maturity of nuclear technology. To connect this technical maturity to licensing and industrial deployment, regulatory perspectives must be incorporated. The Sodium-cooled Fast Reactor (SFR) assessment conducted by Sandia National Laboratories (SNL) revealed gaps between the maturity of knowledge for specific technical phenomena and regulatory expectations. Even when advanced technologies are available, cases were identified in which additional regulatory requirements, such as validated experimental data, established quality assurance programs, and quantified uncertainty, were not fully satisfied [1]. This demonstrates that the technology evaluation frameworks in research and regulatory domains are fundamentally different.

This study establishes a text analysis and data processing framework based on a Large Language Model (LLM) and Python to analyze technologies in both research and regulatory domains. Through this framework, cross-referenced technology datasets applicable to both domains were generated, and a list of enabling technologies required to connect the two perspectives was derived. Additionally, a technology map explaining the entire field was created based on the derived technology list.

This framework was applied to the case of deploying TRU fuel in next-generation SMRs to identify priority differences and gap areas between research-oriented and regulation-oriented technologies. Based on this analysis, a list of gap-mitigating technologies and an overall technology map were generated. Through this case study, the applicability and performance of the framework were validated, demonstrating its ability to derive technology lists that connect advanced research outcomes with regulatory requirements based on large-scale data.

2. Methods and Results

2.1 Data Analysis Method

In this study, a Python-based LLM framework was established for large-scale literature analysis. OpenAI GPT-series models were employed for text processing, and embedding-based clustering was adopted as the primary analytical approach. Embedding is a technique that quantifies semantic relationships between texts within a high-dimensional vector space. Clustering is a method that groups semantically similar items based on the relative position and directional similarity of their vectors [2]. In this study, the text-embedding-3-large model was employed to embed the literature data into a 3,072-dimensional vector space, and topic-specific clusters were formed based on cosine similarity.

The gpt-4o model was used to extract information from the text and organize it into a standardized format, while gpt-5.2 was applied to analyze relationships among the extracted data and to generate the final results. To ensure the reliability of the LLM-based analysis, a Retrieval-Augmented Generation (RAG) approach was applied to all decision-making processes. RAG generates responses based on results retrieved from external databases and was adopted to enhance the traceability and reproducibility of analytical judgments.

Meanwhile, deriving technology lists independently each time can lead to redundancy issues. To address this issue, the Kneedle algorithm was applied to the cosine similarity distribution. The Kneedle algorithm detects a threshold point where the rate of change increases sharply [3]. In this study, both the inflection point and the rapid-rise region of the similarity curve were used to consolidate overlapping entries, thereby retaining only unique technologies.

2.2 Selection of SMR Types for TRU Fuel Deployment Based on AHP

Sodium-cooled Fast Reactors (SFRs), Molten Salt Reactors (MSRs), and Very High Temperature Reactors (VHTRs) were selected as candidate reactor types for TRU fuel deployment. The suitability of each TRU fuel–reactor combination was evaluated using the Analytic Hierarchy Process (AHP). The Analytic Hierarchy

Process (AHP) is a multi-criteria decision-making method that decomposes a problem into a hierarchical structure consisting of a goal, criteria, sub-criteria, and alternatives. Relative weights are then derived through pairwise comparisons among elements within the same hierarchical level. Pairwise comparisons were conducted based on Saaty's 9-point scale. The weights were derived using the principal eigenvalue method, and only results with a Consistency Ratio (CR) of less than 0.1 were accepted for analysis [4].

Although AHP is typically conducted through expert evaluation, this study utilized LLM to perform the entire process, from generating evaluation items to calculating weights. First, abstracts collected via the SCOPUS API using the keywords "TRU AND SFR", "TRU AND MSR", and "TRU AND VHTR" were embedded and clustered, and key variables were identified from the clustered results. The hierarchical relationships among the extracted variables were then analyzed to define sub-criteria, which were further grouped into broader criteria, thereby constructing the AHP evaluation structure.

The evaluation was conducted based on a total of 263 publications collected from SCOPUS and OSTI using the same keyword combinations. To distinguish the degree to which each alternative satisfies the evaluation criteria, a 0–100 scoring rubric was established as a reference for the LLM [5]. The relative ratios of the derived scores were then used to perform pairwise comparisons within the AHP framework. The rubric was structured into five levels: "Incompatible (0–20)," "Major Constraint (21–40)," "Conditional (41–60)," "Minor Modification (61–80)," and "Drop-in Ready (81–100)." These levels were defined to evaluate applicability based on design compatibility and fundamental suitability.

The derived evaluation structure consists of five main criteria: Neutronics, Thermal Behavior, Chemical Behavior, Fuel Cycle, and Source Term. These are further subdivided into a total of 19 sub-criteria. The highest weights were assigned to Neutronics (0.22) and Source Term (0.21), followed by Chemical Behavior (0.20), Fuel Cycle (0.19), and Thermal Behavior (0.18).

The final AHP scores were 0.37 for SFRs, 0.32 for MSRs, and 0.31 for VHTRs. Based on these results, SFRs and MSRs were selected as the target reactor types for further analysis of TRU fuel applicability. This result indicates that the evaluation was driven primarily by the compatibility between TRU fuel characteristics, including neutronic behavior, reactivity response, and radionuclide retention and release, and the reactor type, rather than by overall reactor design features.

2.3 Generation of Technology Lists and Maps

A technology list required for the deployment of TRU fuel in the SFRs and MSRs was derived. The scope of analysis was limited to metallic fuel for TRU-SFR and chloride-salt-based fuel for TRU-MSR. The technology list was derived by generating technical evidence required to satisfy the evaluation criteria. For this purpose, a RAG database was constructed based on a total of 597 publications. The database consists of 263 publications used in the AHP evaluation, 102 additional academic references, and 232 regulatory documents and related references, including NUREG-2246, NUREG/CR-7299, and NUREG/CR-7305.

Based on the RAG database, existing technologies and additional technologies required for TRU fuel deployment were identified. Scores were assigned according to the rubric to determine the development stage of each technology. The same procedure was applied consistently to both analytical perspectives. The first perspective reflects trends in the research domain. Accordingly, the technology list required to satisfy the AHP evaluation criteria was derived. From the second perspective, regulatory trends were reflected by deriving the technologies required to satisfy the evaluation criteria presented in Appendix A, Table A-1 of NUREG-2246 [6].

Cosine similarity and the Kneedle algorithm were applied to similar technologies within the generated lists to retain only unique technologies. As a result, 257 technologies were identified for TRU-SFR and 275 for TRU-MSR. Based on the selected technology lists, Bridge technologies were identified to address the gap between research and regulatory domains.

First, technologies from the research domain (FIELD) and the regulatory domain (NUREG) were reorganized by reactor category. For TRU-SFR, technologies were classified into Neutronics, Thermal Behavior, Fuel Integrity, and Fabrication/Cycle. For TRU-MSR, they were classified into Neutronics, Chemical Behavior, Materials, and Online Management and Safeguards. Each technology was further subdivided into Code, Experiment, and Methodology. Subsequently, clustering was performed within each sub-domain to group similar research topics and generate candidate sets for final classification. However, when technologies from only one domain were excessively concentrated within a specific cluster, this indicated an imbalance in knowledge accumulation rather than a true gap. In such cases, Bridge technologies derived from that cluster could lack meaningful cross-domain connectivity. To prevent this issue, a MIX index was defined to quantify the degree of balance between the two domains within each cluster.

(1) $$MIX = \min(n_F, n_N)/\max(n_F, n_N)$$

n_F and n_N denote the numbers of FIELD and NUREG technologies within a given cluster, respectively. In this study, only clusters with MIX $\geq$ 0.4 were selected as candidates for gap technology identification.

For each technology pair within the selected clusters, the extent to which regulatory requirements were not satisfied was defined as a "dissatisfaction level." This was evaluated using six integrated indicators: validated code coupling, end-to-end validation, uncertainty propagation, traceable data lineage, acceptance criteria definition, and margin-to-limit demonstration. Each technology combination was scored on a 0–1 scale for these indicators. If the average score was 0.5 or higher, the pair was considered not to sufficiently satisfy regulatory requirements and was therefore selected as a candidate for Bridge technology derivation.

In the TRU-SFR case analysis, 140 FIELD technologies, 117 NUREG technologies, and 96 Bridge technologies were identified. The FIELD and NUREG technologies showed average scores of 62.23 and 63.14, respectively, falling within the 41–80 range. This indicates that core research activities are primarily focused on strengthening design bases and advancing model verification and validation. Bridge technologies exhibited an average score of 75.16 and were predominantly distributed within the 81–100 range. This suggests that the fundamental gap does not lie in the understanding of individual phenomena, but rather in the lack of integrated validation and qualification frameworks.

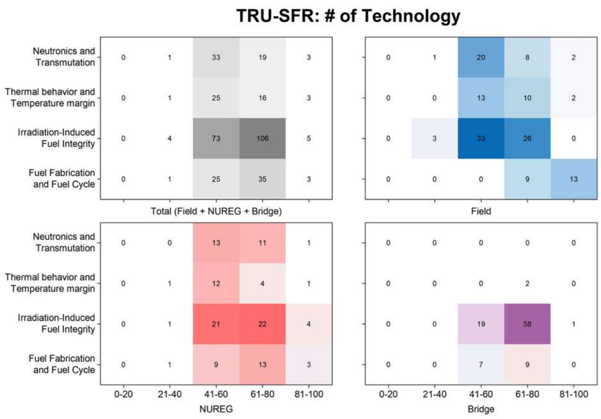

Fig. 1. Number of technologies identified in the FIELD, NUREG, and Bridge categories for the TRU-SFR case. The x-axis represents the score ranges, and the y-axis corresponds to the technical domains.

In the TRU-MSR analysis, 134 FIELD technologies, 141 NUREG technologies, and 112 Bridge technologies were identified. The FIELD and NUREG technologies showed average scores of 56.54 and 57.94, respectively, concentrated within the 41–60 range. This indicates that current research is primarily focused on developing data and models capable of consistently describing the coupled behavior among salt composition, thermophysical properties, corrosion, and source term behavior. Bridge technologies exhibited an average score of 74.18 and were predominantly distributed in the higher score range. This indicates that establishing integrated validation and qualification packages is a necessary condition for meeting regulatory expectations and advancing toward design-level applicability.

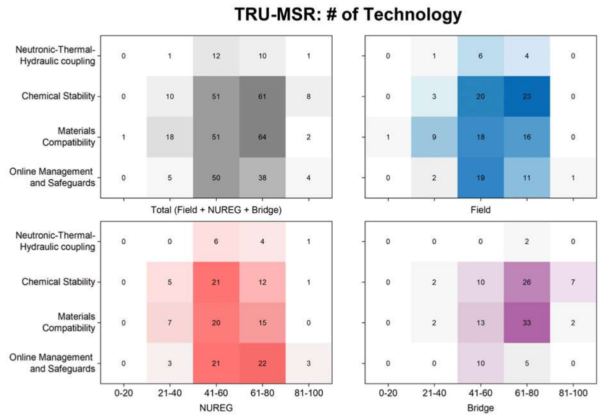

Fig. 2. Number of technologies identified in the FIELD, NUREG, and Bridge categories for the TRU-MSR case. The x-axis represents the score ranges, and the y-axis corresponds to the technical domains.

The technology lists derived through the above procedure were arranged according to score and technology types to construct a technology map. This map simultaneously visualizes the stages of research required to advance each domain and the additional elements necessary to transition to integrated validation and qualification levels. In particular, the score-based positioning of each technology enables intuitive identification of whether a given item remains at the research stage, has reached partial design integration, or is approaching integrated validation and qualification. The proposed technology map is expected to serve as a decision-support framework for designing development strategies in nuclear technology programs.

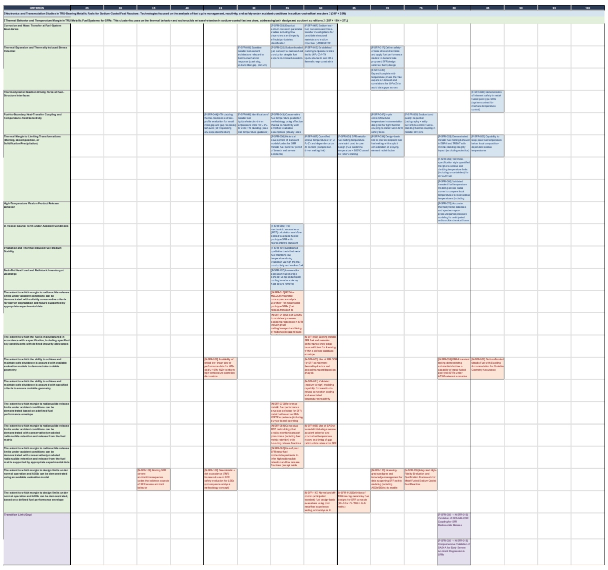

Fig. 3. Example of the technology map for the Thermal Behavior domain in the TRU-SFR case. Blue, red, and purple items represent FIELD, NUREG, and Bridge technologies, respectively.

3. Conclusions

In this study, an LLM-based methodology was proposed to generate a technology list that connects research and regulatory domains, along with a comprehensive technology map that integrates these elements. Based on 597 publications, 257 and 275 technologies were identified for the TRU-SFR and TRU-MSR cases. In addition, 96 and 112 Bridge technologies were derived for each case. The concentration of Bridge technologies within the 81–100 score range indicates that future research should prioritize the advancement of integrated validation and qualification frameworks. This implies a shift from improving individual performance metrics toward package-based technology development that links testing, modeling, analysis, and quality assurance within a coherent framework.

However, the present analysis is based solely on publicly available literature and reports, which limits its ability to reflect real-time research and industrial developments. In addition, LLM-based assessments are subject to structural constraints, as the outcomes depend on the scope and quality of the input data, the selection criteria applied, and the design of the scoring rubric. Future work should incorporate industrial demonstration data and structured feedback from both research and regulatory domains to refine the analysis model.

By incorporating demonstration data and improving performance through cross-validation, the proposed Bridge technology identification and technology map generation framework is expected to be broadly applicable to the formulation of development strategies for advanced reactors and fuel cycle technologies.

Acknowledgement

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Ministry of Science and ICT (MSIT) of the Republic of Korea (RS-2024-00420956, RS-2025-02353161, RS-2025-25463436), Korea Institute of Energy Technology Evaluation and Planning (KETEP), and the Ministry of Trade, Industry & Energy (MOTIE) of the Republic of Korea (RS-2025-02633904).

REFERENCES

[1] Sandia National Laboratories (SNL), Sodium Fast Reactor Safety and Licensing Research Plan-Volume I, SAND2012-4260, Albuquerque, NM, U.S., 2012.

[2] T. Bleckmann, P. Tschisgale, Evaluating NLP Embedding Models for Handling Science-Specific Symbolic Expressions in Student Texts, arXiv preprint arXiv:2025.17950v2, 2025.

[3] V. Satopää, J. Albrecht, D. Irwin, B. Raghavan, Finding a “Kneedle” in a haystack: Detecting knee points in system behavior, 31st International Conference on Distributed Computing Systems, IEEE, pp. 166–171, 2011.

[4] T.L. Saaty, J.M. Katz, How to make a decision: The Analytic Hierarchy Process, European Journal of Operational Research. 48, 9–26, 1990.

[5] P. Dawson, Assessment rubrics: towards clearer and more replicable design, research and practice, Assessment and Evaluation in Higher Education, 42, pp. 347–360, 2017.

[6] U.S. Nuclear Regulatory Commission (NRC), Fuel Qualification for Advanced Reactors, NUREG-2246, Washington, DC, 2022.