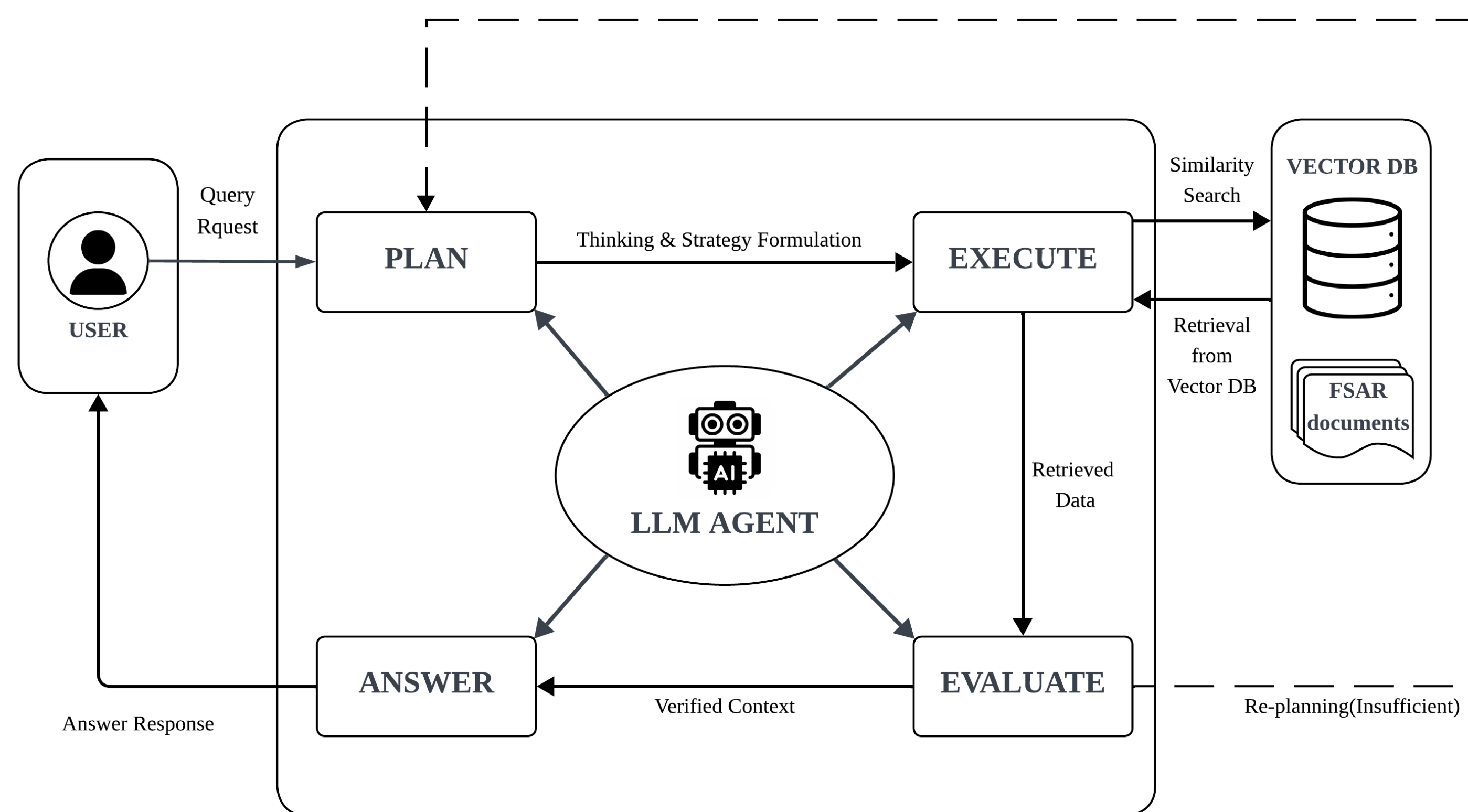


MOTIVATION

Recent advances in large-scale generative AI have revolutionized the paradigm of knowledge exploration and document analysis; however, the nuclear industry faces strict security constraints that make it difficult to directly utilize cloud-based AI systems. Technical documents such as the Final Safety Analysis Report (FSAR) contain tens of thousands of pages encompassing multi-domain information—including mechanical, core, and safety analysis—making it practically impossible for a single expert to respond to complex technical inquiries in real time. Conventional Vanilla RAG systems are limited by their single retrieval-generation structure, lack of transparent reasoning traceability, and dependency on external APIs that pose risks of sensitive information leakage. Therefore, this study aims to develop an Agentic Retrieval-Augmented Generation (Agentic RAG) framework operating in a secure on-premises environment, which emulates the human expert’s Plan–Execute–Evaluate problem-solving process to build a highly reliable and transparent question-answering system for FSAR documents.

METHODS AND EXPERIMENTS

1. System Architecture



- The proposed system is inspired by the cognitive process of human experts who solve complex problems by formulating plans and verifying information in stages. It employs a state-based management architecture built on LangGraph and operates through a four-step cyclic process: Plan → Execute → Evaluate → Replan.
- Plan Agent:** Receives a user query, establishes an overall strategy to reach the final answer, and defines the initial sub-goal for reasoning.
- Execute Agent:** Retrieves relevant information from the vector-indexed FSAR documents according to the predefined plan and generates an intermediate (sub-answer) based on the retrieved evidence.
- Evaluate Agent:** Assesses the sufficiency of the generated answer, and if it is deemed inadequate, formulates a revised plan incorporating previous results and re-executes the reasoning process.

2. Domain-specific Prompt Design for Nuclear Engineering

Agent	System Role	Prompt Template
PLAN	Multi-step Reasoning Planner	“You are a multi-step reasoning planner for nuclear FSAR documents. Design a reasoning pathway leveraging FSAR structure (sections, tables, figures)....” <i>Example Query: “Table 4.1-1” OR “design parameter” OR “control rod drop time” OR “§4.2.1.5”</i>
EXECUTE	FSAR Information Analyst	“You are an FSAR information analyst. Use only the retrieved text/tables/figures as evidence. For every statement, attach citations in the form [SOURCE N FSAR: Section/Table/Figure]. Quote technical values and units verbatim from the original text....” <i>Example Query: [SOURCE 2 Table 4.1-1] shows reactor pressure as 2250 psia(15.5 MPa)</i>
EVALUATE	Completeness Assessor	“You are a completeness assessor. Evaluate whether the gathered evidence sufficiently answers the question. Score completeness from 0.0–1.0....to final answer.”

2.1 Experimental Query Processing Example

NuScale SMR의 원자로냉각재(Reactor Coolant Water)의 화학적 조건(Chemistry Controls)에 대해 알려주세요.

답변

Reactor Coolant Water Chemistry Controls for NuScale SMR

- The chemical concentration limits for the reactor coolant water in the NuScale SMR are specified to ensure safe and efficient operation. These limits include:
 - Chloride concentration should be ≤ 0.05 ppm.
 - Fluoride concentration should be ≤ 0.05 ppm.
 - Dissolved oxygen concentration should be ≤ 0.005 ppm.
 - Sulfate concentration should be ≤ 0.05 ppm.
 - Hydrogen concentration should be maintained between 25 - 50 cc/kg.
 - Boron concentration should be within the range of 0 - 2000 ppm.
- These parameters are critical for maintaining the integrity of the reactor coolant system and preventing corrosion and other chemical-related issues.

출처문서

- The chemical concentration limits for chloride, fluoride, dissolved oxygen, sulfate, hydrogen, and boron are detailed in Table 5.2-5 of the FSAR [SOURCE 2 | FSAR: Table 5.2-5].
- Regulatory guidelines provide maximum allowable concentrations for chloride, fluoride, and dissolved oxygen, as outlined in RG 1.44 [SOURCE 3 | FSAR: Table 5.2-5].

Example Query:

- “What are the reactor coolant water chemistry control limits for NuScale SMR?”
- The system retrieved detailed chemical concentration limits from FSAR Table 5.2-5, including chloride (≤ 0.15 ppm), fluoride (≤ 0.15 ppm), dissolved oxygen (≤ 0.005 ppm), and sulfate (≤ 0.15 ppm). All numerical values and units were quoted verbatim from the original FSAR text, and the corresponding sources were explicitly cited in the format : [SOURCE 2 | FSAR: Table 5.2-5].
- This example demonstrates the system’s ability to provide accurate, traceable, and evidence-grounded answers by directly referencing FSAR anchors.

EVALUATION

3. Performance Evaluation Setup

- Experimental Setup:** We referenced about 500 pages from Chapter 1 and Chapter 5 of the APR1400 FSAR, and used 10 expert-validated nuclear safety queries provided by the Nuclear SMART Division as the evaluation dataset.

Metric	Vanilla RAG		Agentic RAG	
	GPT-4o / text-embedding-3-small (API)	Gemma3-27b-IT BGE-m3 (Onpremise)	GPT-4o text-embedding-3-small (API)	Gemma3-27b-IT BGE-m3 (Onpremise)
Hit@1	0.30	0.40	0.40	0.50
Hit@3	0.50	0.60	0.70	0.80
Precision	0.35	0.42	0.45	0.52
Recall	0.38	0.45	0.52	0.58
F1-Score	0.36	0.43	0.48	0.55

- Hit@k Metric:** Evaluates retrieval accuracy — the ratio of queries whose relevant documents appear within the top k search results.
- Results:** Agentic RAG achieved +10% improvement in Hit@1 and +20–30% in Hit@3, with precision, recall, and F1-score increasing by 10-12% overall. The Plan–Execute–Evaluate multi-step reasoning structure improved retrieval accuracy and reliability, demonstrating its practicality in secure environments.

CONCLUSIONS

In this study, we developed and validated an on-premises Agentic RAG system for nuclear safety document analysis. The proposed system can systematically analyze complex FSAR documents through a multi-step reasoning structure of Plan-Execute-Evaluate, demonstrating improved retrieval accuracy and response quality compared to conventional Vanilla RAG systems. Future research will focus on expanding system capacity to handle larger volumes of documents such as SMART nuclear design information and accommodating more diverse query types.