

# 국방 체계공학(System Engineerin) 고도화를 위한 검색증가생성(RAG, Retrieval-Augmentation-Generation)용 데이터 연구

지상용, 박순철, 김정훈  
국방기술품질원  
e-mail:updragon@dtaq.re.kr

## A Study on Retrieval-Augmented Generation (RAG) Data for the Advancement of Defense System Engineering

SANGYONG JI, SUNCHEOL PARK, JOUNGHUN KIM  
Defense Agency for Technology and Quality

### 요약

국방 무기체계 개발에는 체계공학(System Engineering)이 적용되어 사용자의 요구를 체계적으로 반영하고 있으나, 실제 양산 및 운용유지 단계에서는 여전히 다양한 문제점이 발생하고 있다. 이러한 문제는 품질 저하, 신뢰성 약화, 유지비용 증가 등으로 이어져 체계공학의 고도화 필요성을 더욱 강조한다. 최근 인공지능 연구가 활발히 진행되고 있지만, 국방 분야에 축적된 비정형 데이터는 보안성과 복잡성으로 인해 직접 활용하기 어렵다. 따라서 이를 효과적으로 적용하기 위해서는 데이터 정제와 구조화가 필수적이며, 특히 관련 전문가의 판단을 반영한 주석화 과정이 요구된다. 본 연구는 회전의 항공기 양산단계에서 식별된 문제점 분석을 토대로, 검색증가생성(Retrieval-Augmented Generation, RAG) 적용을 위한 데이터 구축 방법론을 제시한다. 이를 통해 국방 체계공학의 고도화를 지원하고, 양산 및 운용유지 단계에서의 문제 해결 능력을 향상시킬 수 있는 데이터 활용 가능성을 탐색한다.

## 1. 서론

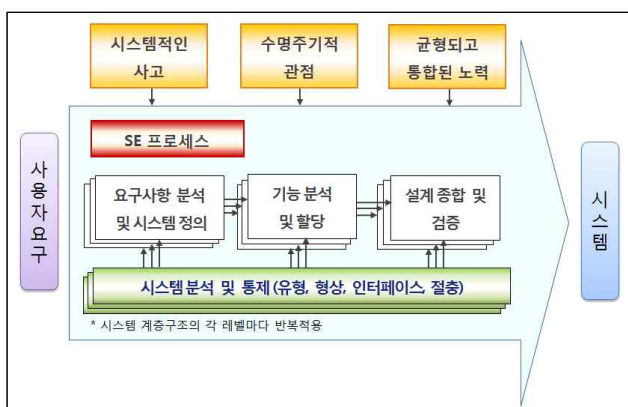
국방 분야의 무기체계 개발단계에서는 체계공학(System Engineering)이 적용되어 사용자의 요구도를 제품에 반영하고 있다.[1] 그러나 개발 이후 양산 및 운용유지 단계에서 여전히 다양한 문제점이 발생하며, 이를 개선하기 위한 활동이 지속적으로 요구되고 있다.

한편, 2022년 ChatGPT의 출시 이후 인공지능(Artificial Intelligence) 분야의 연구는 급격히 발전하고 있으며, 특히 검색증가생성(Retrieval-Augmented Generation, RAG) 기법은 방대한 데이터베이스에서 적합한 정보를 검색하여 생성 모델의 응답 품질을 향상시키는 방법론으로 주목받고 있다. 그러나 국방 분야에서 축적된 비정형 데이터는 곧바로 활용하기 어려운 특성이 있으며, 보안 요구와 전문가의 해석이 결합된 정제 과정이 필수적이다.

이에 본 연구는 회전의 항공기 양산단계에서 식별된 문제점에 대한 분석을 기반으로, 국방 체계공학 고도화를 위한 RAG 데이터 구축의 기초 연구 방법론을 제시한다. 이를 통해 국방 분야의 비정형 데이터를 체계적으로 관리·활용하여 양산 및 운용유지 단계의 문제 해결 능력을 제고하고자 한다.

## 2. 연구 방법

본 연구는 국방 분야에서 RAG용 데이터를 구축하기 위한 기초적 방법론을 다음과 같은 단계로 수행한다.



[그림 1] 체계공학 개념[2]

### 3. 후속 계획

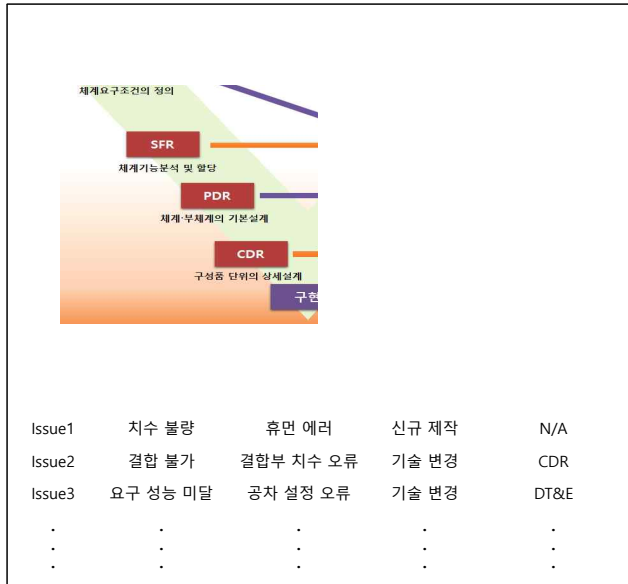
본 연구는 국방 체계공학 고도화를 위한 RAG 데이터 구축의 기초적 방법론을 제시하는 단계에 머물러 있다.

향후 연구에서는 우선 회전익 항공기를 넘어 고정익 항공기, 무인기 등 다양한 무기체계로 연구 대상을 확장하여 보다 보편적이고 활용 범위가 넓은 데이터셋을 마련하고자 한다.

또한 구축된 데이터셋을 실제 질의응답 시스템에 적용하여 RAG 기반 응답의 정확성과 신뢰성을 정량적으로 검증하고, 이를 통해 문제 해결 지원 효과를 평가할 계획이다.

아울러 요구사항 관리, 위험 관리, 품질 보증 등 체계공학 활동과의 연계를 강화하여 RAG 데이터 활용이 실무적 차원에서 체계공학 활동을 지원할 수 있는 방안을 모색할 것이다.

마지막으로, 국방 분야 특성상 반드시 고려되어야 하는 보안과 윤리적 측면에 대한 가이드라인을 수립함으로써 안전하고 지속 가능한 데이터 활용 기반을 마련하고자 한다.



[그림 2] 결합 데이터의 개발단계 연계[4]

#### 2.1 데이터 수집

회전익 항공기 양산단계에서 발생한 문제점 및 개선 사례를 중심으로 데이터를 수집한다. 주요 자료는 품질 보증 활동 기록, 기술 문서 등이다.

#### 2.2 데이터 정제 및 구조화

수집된 비정형 데이터에서 중복, 불필요한 표현, 민감 정보를 제거하고 핵심 요소(문제 발생 원인, 조치 과정, 개선 효과 등)를 추출한다. 이를 메타데이터화하여 체계적인 검색이 가능하도록 구조화한다.

#### 2.3 전문가 주석화(Annotation)

관련 분야 전문가의 판단을 반영하여 데이터에 주석을 부여한다. 예를 들어 결합 유형 분류, 영향도 평가, 품질 특성 매핑 등을 통해 데이터의 의미 기반 활용도를 강화한다. 그 중에서도 문제 발생 원인에 대한 개발단계 연계성을 데이터화 하는 것을 주 목표로 한다.

#### 2.4 RAG 파이프라인 설계

정제·구조화된 데이터를 기반으로 검색 모듈을 구축하고, 이를 생성 모델과 연계하여 응답 품질 향상을 검증한다. 이 과정을 통해 국방 분야 문제 해결 지원 도구로서 RAG의 가능성을 확인한다.

#### 참고문헌

- [1][2] 국방기술품질원, SE기반 기술검토회의 가이드북, 방위사업청, pp. 8, 2024.
- [3] M. Klesel and H. F. Wittmann, "Retrieval-Augmented Generation (RAG)," Business & Information Systems Engineering, Vol.67, pp. 551-561, 2025.
- [4] 국방기술품질원, SE기반 기술검토회의 가이드북, 방위사업청, pp. 16, 2024.