

# 고객의 요구(텍스트, 이미지, 동영상 등)에 최적의 결과물을 제공함에 있어 LLM, LMM기술 기반의 알고리즘의 정의와 데이터의 구조정의를 구현하는 기술에 대한 연구

김남현  
주식회사 엘젠 연구책임자  
e-mail:elgenict@elgen.co.kr

## A Study on Technology for Implementing Algorithm Definition and Data Structure Definition Based on LLM and LMM Technology in Providing Optimal Results to Customer Requirements

Nam-Hyeon Kim  
Research Director, ELGEN Co., Ltd.

### 요약

본 연구는 건설 산업의 생산성 저하, 안전 사고, 정보 관리 비효율성을 해결하기 위해 텍스트, 음성, 이미지, 동영상 등 다양한 형태의 데이터를 통합 처리하는 멀티모달 AI 전문가 시스템의 기술적 청사진을 제시한다. 제안 시스템은 온프레미스 환경의 대규모 멀티모달 모델(LMM)을 핵심 엔진으로 하며, 보안성과 확장성을 위해 엣지-온프레미스-클라우드 하이브리드 아키텍처를 채택한다. 데이터 처리 백엔드에는 그래프 RAG(Retrieval-Augmented Generation) 기반 하이브리드 데이터베이스를 구축하여 논리적 추론을 가능케 하고, 핸즈프리 멀티모달 UI/UX 및 자율적 멀티 에이전트 시스템을 통해 현장 업무를 능동적으로 해결하는 방안을 제시한다.

## 1. 서론

### 1.1 스마트 건설 분야 멀티모달 AI 전문가의 필요성

전 세계 스마트 건설 시장은 연평균 13.4% 성장 중이며, 국내 또한 16.5%의 높은 성장률이 전망된다. 현재 드론 및 IoT를 통한 현장 데이터 수집 인프라는 구축되었으나, 데이터가 BIM 도면, 현장 사진, CCTV 영상 등 비정형 멀티모달 형태로 존재하여 이를 통합 이해하고 질의에 응답할 수 있는 AI 전문가가 필요하다.

### 1.2 스마트 건설 분야에서 AI의 역할 확대

AI 기반 영상 분석은 실시간 위험 감지 및 사고 예방에 기여하며, 건설 특화 LLM은 현장 용어와 법규에 특화되어 업무 효율성을 제고한다. 본 연구는 다양한 데이터로부터 최적 해답을 도출하는 시스템 아키텍처 제시를 목표로 한다.

## 2. 이론적 배경 및 선행 연구 고찰

### 2.1 LLM 및 LMM 기술의 기본 원리와 동향

LMM은 텍스트, 이미지, 영상, 음성을 동시에 처리하는 능력

을 갖추며, 비전 인코더, 정렬 모듈, 고정된 LLM으로 구성된다. YOLOv8 등 객체 탐지 모델과 문서 지능화 기술을 통해 비정형 데이터를 정형화할 수 있다.

### 2.2 지식 기반 추론을 위한 기술

벡터 데이터베이스는 의미적 유사성 기반의 검색(RAG)을 지원하며, 도메인 온톨로지는 개념 간 논리적 관계를 정의하여 단순 검색 이상의 지식 추론을 가능하게 한다

### 2.3 선행 연구의 한계 및 차별성

기존 연구는 단일 과업에 집중된 반면, 본 연구는 하이브리드 배포, 그래프 RAG 기반 지식 백엔드, 자율 멀티 에이전트라는 세 가지 혁신적 통합 아키텍처를 제안한다.

## 3. 연구 방법론

### 3.1 전체 시스템 아키텍처 설계

본 시스템은 엣지(실시간 분석), 온프레미스(보안 및 심층 추론), 클라우드(모델 최적화)의 계층 구조로 설계되었다

### 3.2 단계별 구현 기술

1단계: 비정형 데이터를 MMKG(Multimodal Knowledge Graph) 형태로 변환 및 벡터화.

2단계: LMM 기반으로 시각적 데이터에서 심층 지식을 추출

하고 JSON 형태의 구조화된 데이터로 변환.

3단계: 그래프 RAG 알고리즘을 정의하여 온톨로지 추론 기반의 답변 생성 및 MCP(Model Context Protocol)를 통한 외부 도구 연동.

## 4. 실험 결과 및 분석

### 4.1 성능 평가 지표

객체 탐지 정확도 mAP 90% 이상, 문서 엔티티 추출 F1-Score 85% 이상, 시각적 질의응답 정확도 80% 이상을 목표로 한다.

### 4.2 그래프 RAG 및 에이전트 효율성

온톨로지 필터링을 통해 LLM 환각률을 50% 이상 감소시키며, 업무 시간은 기존 수동 작업 대비 70% 이상 단축될 것으로 예상된다.

## 5. 고찰 및 논의

하이브리드 LMM과 그래프 RAG는 보안과 추론 능력을 동시에 충족한다. 다만, 여전히 존재하는 환각 위험에 대비한 'Human-in-the-loop' 시스템과 모델 경량화 및 지식 그래프 업데이트(MLOps)가 향후 과제로 남는다 .

## 6. 결론

본 연구는 멀티모달 데이터를 처리하는 건설 AI 전문가 시스템의 알고리즘과 데이터 구조를 설계하였다. 향후 실제 엣지 디바이스에서의 실증 연구와 자율적 의사결정이 가능한 에이전트 고도화를 지속할 계획이다.

### 참고문헌

- [1] 백 터 데이터베이스와 그래프 데이터베이스: 차이점 이해 | Elastic Blog.
- [2] Ontology for a Knowledge-Based Deconstruction of Buildings Based on BIM Models and Linked Data Principles - MDPI.
- [3] LLM의 경량화와 성능 개선 방법에 대해 알아보자. - Marcus' Stroy - 티스토리.
- [4] 건설사도 스마트건설 기술 도입 총력 - 매일일보.
- [5] Vector Databases vs. Knowledge Graphs for RAG | Paragon Blog.

(이하 원문 참고문헌 목록 참조)