

계층적 민법 구조를 반영한 모듈형 RAG 기반 한국어 법률 질의응답 시스템

신은성^{*1}, 소훈^{*2}, 윤영준^{*3}, 정찬우^{*4}, 김동연^{*5}, 정현준^{**}

Modular RAG-based Legal Question Answering System Reflecting the Hierarchical Structure of Civil Law

Eunsung Shin^{*1}, Hoon So^{*2}, Youngjun Yoon^{*3}, Chanwoo Jung^{*4}, Dongyeon Kim^{*5}, and
Hyunjun Jung^{**}

요 약

이 연구는 한국어 법률 문서, 특히 민법을 중심으로 신뢰도 높은 법률 답변 생성을 위한 모듈형 RAG(Retrieval-Augmented Generation) 기반 시스템을 제안한다. 장-절-조 계층 구조와 조문 간 참조 관계 등 법률 문서의 구조적 특성을 반영하고, 질의 분석, 검색, 생성 과정을 LangGraph를 통해 모듈화하여 각 단계의 정밀도와 유연성을 높였다. 다양한 질의 유형에 대응할 수 있도록 설계된 본 시스템은 직접 조문 참조, 개념 기반, 상황 기반, 복합 질의 등에서 모두 높은 성능을 보였으며, 전체 기준으로 검색 정확도(Retrieval accuracy) 88%, MRR 0.84, 응답 적절성(Response relevance) 92%를 달성하였다. 특히 직접 조문 참조 질의는 모든 항목에서 100%의 성능을 기록하였다. 그러나 유사 법률 용어 처리와 복잡한 참조 체인 추적에서 개선이 필요하며, 향후 다양한 법률 영역으로의 확장 연구가 요구된다.

Abstract

This study proposes a Modular Retrieval-Augmented Generation (RAG)-based system designed to generate reliable legal answers from Korean legal documents, with a focus on civil law. The system leverages the hierarchical structure of statutes chapters, sections, and articles and inter-article references, optimizing each stage of query analysis, retrieval, and generation through LangGraph-based modularization. It handles various query types, including direct article reference, legal concept, situation-based, and complex queries. Experiments show strong performance with 88% retrieval accuracy, 0.84 MRR, and 92% response relevance. Notably, the system achieved 100% accuracy across all metrics for direct article reference queries. However, improvements are needed in handling similar legal terms and complex reference chains, with future research required for expansion to diverse legal domains.

Keywords

RAG, legal search, hierarchical structure, LangGraph, modular architecture, Korean civil law

* 국립군산대학교 소프트웨어학과 학사과정

- ORCID¹: <https://orcid.org/0009-0008-2783-9662>

- ORCID²: <https://orcid.org/0009-0007-1198-5596>

- ORCID³: <https://orcid.org/0009-0005-1617-8393>

- ORCID⁴: <https://orcid.org/0009-0002-3580-3201>

- ORCID⁵: <https://orcid.org/0009-0006-8048-2696>

** 국립군산대학교 소프트웨어학과 교수(교신저자)

- ORCID: <https://orcid.org/0000-0002-6717-1395>

• Received: Aug. 11, 2025, Revised: Sep. 29, 2025, Accepted: Oct. 02, 2025

• Corresponding Author: Hyunjun Jung

Dept. of Software at Kunsan National University, 558, Daehak-ro,

Kunsan-si, Jeollabuk-do, Republic of Korea

Tel.: +82-63-469-8917, Email: junghj85@kunsan.ac.kr

I. 서 론

대규모 언어 모델(LLM, Large Language Model)의 발전은 자연어 처리 기술의 활용 가능성을 크게 확장시켰다[1]. 특히 법률 문서 분석과 질의응답 시스템 개발 분야에서 LLM 기술이 적극적으로 도입되고 있다[2][3]. 질의응답 시스템은 법률 정보 접근성을 높이고 비전문가의 이해를 지원할 수 있는 효과적인 도구로 주목받고 있다. 그러나 국내외 법률 문서는 장-절-조로 구성된 계층 구조와 조문 간 참조, 본문과 부칙의 구분 등 복잡한 구조적 특성을 가지고 있다. 법조문의 계층 구조와 상호 참조 관계로 인해 단순 텍스트 검색은 관련 조문을 완전히 파악하기 어렵다. 맥락 없는 부분적 검색은 법적 오해석을 유발하며, 일반 LLM 시스템으로는 정확하고 신뢰도 높은 법률 응답 생성이 어렵다.

한국어 법률 문서는 표현 양식과 구조 면에서 고유한 특성을 지닌다. 특히 민법의 경우 동일한 조문 번호가 본문과 부칙에서 중복되거나 조문 간 참조가 다층적으로 연결되어 있어 단순 검색이나 임베딩 기반 유사도 검색만으로는 응답 품질을 보장하기 어렵다[4]. 예를 들어, 민법 제1013조는 "상속인의 재산상 권리의무의 포괄적 승계"를 다루지만, 관련 조문들이 제1005조부터 제1019조까지 산재해 있으며 각 조문은 서로를 참조하는 복잡한 네트워크를 형성한다. 이러한 복잡성으로 인해 생성형 모델이 사실과 다른 응답을 출력하는 할루시네이션 문제가 자주 발생한다[5].

법률 분야에서의 할루시네이션은 특히 심각한 문제이다. 잘못된 법률 정보 제공은 실제 법적 분쟁이나 의사결정에 중대한 영향을 미칠 수 있기 때문이다. 상속 포기 기한을 잘못 안내하거나 존재하지 않는 조문을 인용하는 경우, 이용자에게 회복할 수 없는 손해를 입힐 수 있다. 따라서 법률 질의응답 시스템에서는 정확성과 신뢰성이 무엇보다 중요하다.

이러한 한계를 극복하기 위한 방안으로 검색 기반 생성(RAG, Retrieval-Augmented Generation) 아키텍처가 주목받고 있다[6]. RAG는 외부 문서 검색을 통해 모델의 지식 기반을 보완하고, 보다 정확하고 정당화 가능한 응답 생성을 가능하게 한다. 하지만

기존 RAG 기반 시스템은 대부분 단일 파이프라인 구조로 구성되어 있어 복잡한 법률 문서를 처리하는 데 한계가 있다[7]. 또한 질의 유형에 따른 차별화된 처리가 어렵다는 문제점이 있다[8]. 법률 도메인에서는 질의 유형에 따라 요구되는 정보의 형태와 검색 전략이 크게 달라지므로, 획일적인 처리 방식으로는 최적의 성능을 달성하기 어렵다.

이 연구에서는 한국어 민법을 중심으로 장-절-조 계층 구조와 조문 간 참조 관계를 반영한 모듈형 RAG 기반 법률 질의응답 시스템을 제안한다[9]. 제안 시스템은 질의 분석, 정보 검색, 응답 생성을 독립적인 모듈로 구성하고, LangGraph 기반 워크플로우를 통해 질의 유형에 따른 경로 최적화를 수행한다. 다양한 질의 유형에 대응할 수 있도록 설계되었으며, 성능 평가를 통해 높은 검색 정확도와 응답 적절성을 달성하였다. 또한 구조화된 법률 데이터를 기반으로 할루시네이션을 최소화할 수 있음을 실증하였다[10][11].

II. 관련 연구

2.1 모듈형 RAG 아키텍처

RAG 기반 시스템의 증가하고 있는 복잡성을 관리하기 위해 모듈형 RAG 아키텍처가 제안되었다. 이 아키텍처는 모듈, 서브모듈, 오퍼레이터로 구성된 3단계 설계를 통해 RAG 시스템을 통합되고 구조화된 방식으로 정의한다. 모듈형 RAG는 복잡한 RAG 시스템을 독립적 모듈과 전문화된 오퍼레이터로 분해하여 유연한 재구성이 가능한 프레임워크를 구축한다.

이 아키텍처는 전통적인 선형 검색-생성 프로세스를 넘어서 라우팅, 스케줄링, 퓨전 메커니즘이 통합된 형태로 발전시켰다. 모듈들과 오퍼레이터들의 오케스트레이션이 RAG 흐름을 형성하며, 이를 통해 현재의 RAG 방법들을 유연하게 표현할 수 있다. 또한 여섯 가지 대표적인 흐름 패턴을 요약하고 구체적인 방법들을 분석하여 모듈형 RAG의 실제 시나리오에서의 범용성을 검증하였다.

모듈형 RAG의 핵심 장점은 각 모듈의 독립성과

제사용성에 있다. 첫째, 질의 분석 모듈은 입력된 질의의 의도와 유형을 파악하는 역할을 수행한다. 둘째, 검색 모듈은 다양한 검색 전략을 구현할 수 있으며, 도메인 특성에 맞게 최적화될 수 있다. 셋째, 지식 통합 모듈은 검색된 정보를 효과적으로 조합하고 중복을 제거한다. 넷째, 응답 생성 모듈은 통합된 정보를 바탕으로 일관성 있고 정확한 응답을 생성한다.

이러한 모듈형 접근법은 시스템의 확장성과 유지보수성을 크게 향상시킨다. 새로운 기능이 필요한 경우 해당 모듈만 수정하거나 추가하면 되며, 각 모듈의 성능을 독립적으로 최적화할 수 있다. 또한 다른 도메인으로의 확장 시에도 핵심 구조를 유지하면서 도메인 특화 모듈만 교체하는 방식으로 효율적인 적용이 가능하다.

2.2 한국어 법률 도메인에서의 질의응답 및 RAG 응용

최근 대규모 언어 모델을 기반으로 한 법률 질의응답 시스템에 대한 연구가 활발히 진행되고 있다. 특히 한국어 법률 도메인에 특화된 질의응답 성능 평가 및 개선을 위한 시도들이 이루어지고 있다. 한국어 법률 문서에 대한 LLM 응답의 신뢰성과 정당성을 평가하기 위한 검색 기반 평가 프레임워크가 제안되었다[11].

이 연구는 외부 근거 문서를 기반으로 생성 응답을 정량적으로 평가할 수 있는 체계를 마련하였다. 실험을 통해 RAG 기반 응답이 기존 방식 대비 전문가 평가와 더 높은 일치도를 보이는 것을 실증하였다. 구체적으로, RAG 기반 시스템은 법률 전문가의 평가와 87%의 일치도를 보인 반면, 기존 LLM 단독 사용 시에는 62%의 일치도에 그쳤다.

이러한 연구는 한국어 법률 질의응답 시스템에서 RAG 아키텍처의 적용 가능성과 효과성을 실증적으로 뒷받침한다. 하지만 기존 연구들은 주로 단일 파이프라인 구조에 기반하여 법률 문서의 복잡한 구조적 특성을 충분히 활용하지 못했다는 한계가 있다.

법률 도메인의 특수성을 고려한 추가적인 연구들도 진행되었다. 법률 용어의 동의어 처리, 법률 개

념의 계층적 관계 모델링, 판례와 조문 간의 연결 관계 구축 등이 주요 연구 주제였다. 특히 한국어 법률 텍스트의 경우, 한자어 표현과 현대 한국어 표현이 혼재되어 있어 이를 효과적으로 처리하는 방법론도 연구되었다. 예를 들어, "포괄적 승계"와 "전면적 승계", "상속인"과 "승계인" 등의 유사 표현을 올바르게 매핑하는 것이 검색 성능에 중요한 영향을 미친다.

2.3 법률 문서의 구조적 특성 활용

법률 문서는 일반 텍스트와 달리 명확한 구조적 특성을 가지고 있다. 장-절-조의 계층 구조, 조문 간 참조 관계, 본문과 부칙의 구분 등이 대표적이다. 이러한 구조적 특성을 활용한 검색 시스템 연구들이 진행되어 왔다.

기존 연구들은 주로 법률 문서의 메타데이터를 활용하거나[12], 참조 관계를 그래프로 모델링하는 접근 방법을 취했다[13][14]. 그러나 대부분의 연구가 영어권 법률 문서를 대상으로 하였으며, 한국어 법률 문서의 특성을 충분히 반영하지 못했다는 한계가 있었다. 특히 한국 민법의 경우 조문 번호 체계가 복잡하고, 신설·개정·삭제가 빈번하여 이를 체계적으로 관리하는 방법론이 필요하다.

최근에는 법률 문서의 구조를 XML이나 JSON 형태로 구조화하여 저장하고, 이를 기반으로 검색 성능을 향상시키는 연구들이 진행되고 있다. 이러한 접근법은 단순 텍스트 검색보다 훨씬 정확한 결과를 제공할 수 있으며, 특히 조문 간 관계를 활용한 확장 검색이 가능하다는 장점이 있다.

하지만 기존 연구들은 대부분 단일 검색 전략에 의존하거나, 질의 유형에 따른 차별화된 처리가 부족하다는 한계를 보인다. 또한 법률 문서의 계층적 구조와 참조 관계를 종합적으로 활용하는 통합적 접근법은 아직 충분히 연구되지 않았다.

이러한 기존 연구의 한계점들을 바탕으로, 이 연구에서는 한국어 민법의 구조적 특성을 종합적으로 반영한 모듈형 RAG 시스템을 제안하고자 한다. 특히 질의 유형별 최적화된 검색 전략과 계층적 구조 활용을 통해 기존 시스템의 한계를 극복하고자 한다.

III. 모듈형 RAG 기반 법률 문서 검색 시스템

3.1 시스템 아키텍처

이 연구에서 제안하는 모듈형 RAG 기반 법률 답변 생성 시스템은 한국어 민법의 장-절-조 계층 구조를 활용하여 법률 질의에 정확하고 신뢰성 있는 답변을 생성한다. 시스템은 질의 분석, 정보 검색, 지식 관리, 응답 생성의 네 가지 주요 과정을 독립된 모듈로 구성한 아키텍처이며, 각 모듈은 LangGraph를 기반으로 연결되어 최적화된 처리 경로를 구성한다.

그림 1에서 보이는 바와 같이, 전통적인 단일 파이프라인 구조의 RAG 시스템과 달리 모듈형 구조는 민법의 복잡한 계층 구조와 조문 간 참조 관계를 효과적으로 처리할 수 있다.

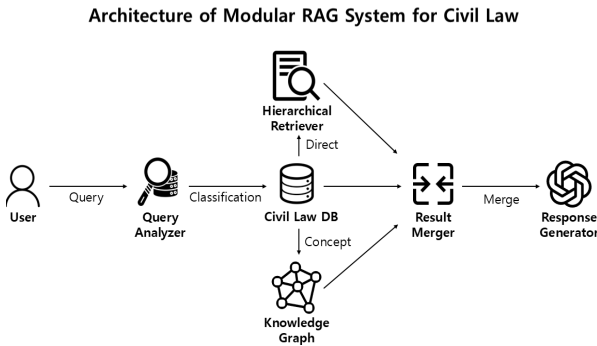


그림 1. 모듈형 RAG 기반 한국어 민법 질의 시스템 아키텍처

Fig. 1. Modular RAG-based Korean civil law query system architecture

시스템의 핵심 구성 요소는 다음과 같다. 첫째, 질의 분석 모듈은 사용자 질의를 분석하여 질의 유형을 판별하고, 조문 번호, 법률 개념, 핵심 키워드 등을 추출한다. 이 모듈은 정규표현식과 자연어 처리 기법을 결합하여 높은 정확도의 질의 분석을 수행한다. 둘째, 검색 모듈은 민법의 장-절-조 계층 구조를 활용한 다중 검색 전략을 구현하며, 질의 유형에 따라 직접 ID 검색 또는 임베딩 기반 의미 검색을 수행한다. 셋째, 지식 관리 모듈은 검색 결과를 통합하고 조문 간 참조 관계를 추적하여 관련 조문들을 종합적으로 제공한다. 넷째, 응답 생성 모듈은 검색된 조문과 관련 정보를 기반으로 질의 유형에

맞는 형식의 답변을 생성한다.

이러한 모듈 구조는 LangGraph 기반 워크플로우로 구현되며, 표 1과 같은 조건부 경로 설정을 통해 동작한다.

표 1. LangGraph를 활용한 워크플로우

Table 1. Workflow function using LangGraph

1. INPUT: User Query
2. NODE: Query Analyzer
 - Extract query type
 - (direct/concept/situation/complex)
 - Extract key entities and article numbers
3. CONDITIONAL ROUTING:
 - IF query_type = "direct" THEN
 - Hierarchical Retriever
 - ELIF query_type = "concept" THEN
 - Knowledge Graph Retriever
 - ELSE
 - Both Retrievers (parallel)
4. NODE: Knowledge Integration
 - Merge retrieval results
 - Remove duplicates
 - Rank by relevance
5. NODE: Response Generator
 - Generate structured legal response
 - Cite relevant articles
6. OUTPUT: Legal Answer with Citations

모듈형 구조와 LangGraph 기반 워크플로우의 조합은 복잡한 법률 질의에 대한 처리 과정을 명확히 하고, 각 모듈의 독립적인 개발과 최적화를 가능하게 한다. 또한 새로운 기능 추가나 기존 모듈의 개선이 전체 시스템에 미치는 영향을 최소화할 수 있다는 장점이 있다. 예를 들어, 향후 판례 검색 모듈을 추가하거나 다국어 지원을 위한 번역 모듈을 추가할 때도 기존 구조를 크게 변경하지 않고 확장이 가능하다.

3.2 법률 문서 계층 구조 처리

민법은 장-절-조로 이어지는 계층적 구조와 본문/부칙 구분, 조문 간 참조 관계 등 특수한 구조적 특성을 가지고 있다. 이 연구에서는 이러한 특성을 활용하기 위해 체계적인 데이터 구조화 방법을 개발하였다.

첫째, 고유 ID 체계를 도입하여 조문을 명확히 식별할 수 있도록 하였다. 본문 조항에는 "article_" 접두사를, 부칙에는 "s_article_" 접두사를 사용한 고유 ID 체계를 도입하였다. 각 조문은 소속된 장-절 정보를 함께 저장하여 계층적 탐색이 가능하도록 하였다. 예를 들어, 민법 제103조는 표 2와 같이 "article_103"으로, 부칙 제3조는 "s_article_3"으로 표현된다.

표 2. 민법 문서의 계층적 구조

Table 2. Hierarchical structure of civil law documents

<pre> { "id": "article_1013", "number": 1013, "title": "상속인의 재산상 권리의무의 포괄적 승계", "content": "상속인은 상속개시된 때로부터 피상속인의 재산에 관한 포괄적 권리의무를 승계한다.", "chapter": "제5편 상속", "section": "제1장 상속", "references": ["article_1005", "article_1019"], "referenced_by": ["article_1014", "article_1015"] } </pre>

둘째, 조문 간 참조 관계 추출을 위해 정규표현식 기반 패턴 매칭을 활용하였다. "제N조", "제M조부터 제N조까지" 등의 패턴을 인식하여 조문 간 참조 관계를 구성하였으며, 이를 통해 조문 참조 그래프를 구축하여 연관 조문을 효과적으로 추적할 수 있다. 참조 관계 추출의 핵심 패턴은 단일 조문 참조(제(n)조)와 범위 참조(제(n)조부터 제(m)조까지) 패턴을 포함한다.

셋째, 계층 구조 인덱싱 메커니즘을 구현하였다. 각 조문은 자신이 속한 장과 절의 정보를 메타데이터로 포함하며, 이를 통해 특정 장이나 절에 속한 모든 조문을 효율적으로 검색할 수 있다. 또한 역방향 인덱스를 구축하여 특정 조문이 참조하는 조문들과 특정 조문을 참조하는 조문들을 빠르게 찾을 수 있도록 하였다.

그림 2는 민법 조문들 간의 복잡한 참조 관계를 시각화한 것으로, 이러한 관계를 활용하면 단순 키워드 검색으로는 찾기 어려운 연관 조문들을 효과적으로 검색할 수 있다. 예를 들어, 상속 포기(제1041조)를 검색할 때 관련된 상속의 효력(제1005조), 상속인의 결격사유(제1004조) 등도 함께 제공할 수 있다.

Legal Document Hierarchical Structure and Reference Relations

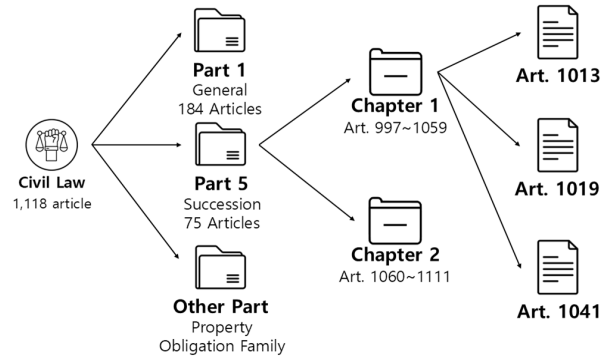


그림 2. 법률 문서 계층 구조 및 참조 관계 그래프

Fig. 2. Legal document hierarchy and reference relationship graph

넷째, 의미적 검색을 지원하기 위해 KLUEBERT 모델을 활용하여 조문 텍스트의 임베딩을 생성하고 FAISS 인덱스를 구축하였다. 이러한 계층 구조, 참조 관계, 의미 임베딩의 통합적 활용은 법률 문서의 맥락에 대한 이해를 가능하게 하며, 할루시네이션 문제를 최소화하는 데 기여한다.

3.3 질의 분석 및 유형 분류

효과적인 법률 답변 생성을 위해서는 사용자 질의의 정확한 분석과 유형 분류가 필수적이다. 이 시스템은 법률 질의를 다음의 5가지 유형으로 분류한다.

첫째, 직접 조문 참조 질의는 "민법 제1013조가 무엇인가요?"와 같이 특정 조문 번호를 명시적으로 포함하는 질의이다. 이러한 질의는 정규표현식을 통해 쉽게 식별할 수 있으며, 가장 단순한 처리 경로를 거친다. 조문 번호 추출 후 직접 ID 검색을 통해 해당 조문을 즉시 반환할 수 있다.

둘째, 조문 제목 기반 질의는 "상속인의 재산상 권리에 관한 조항은?"과 같이 특정 법률 개념이나 제목을 언급하는 질의이다. 이는 조문 제목과 내용에 대한 키워드 매칭과 의미적 유사도 검색을 병행하여 처리한다. 제목 인덱스를 구축하여 빠른 검색이 가능하도록 최적화하였다.

셋째, 법률 개념 질의는 "민법상 상속이란?"과 같이 추상적인 법률 개념에 대한 설명을 요구하는 질의이다. 이러한 질의는 여러 관련 조문을 종합하여 개념적 설명을 제공해야 하므로 복잡한 처리 과정

을 거친다. 개념 온톨로지를 구축하여 관련 조문들을 체계적으로 연결하였다.

넷째, 상황 기반 질의는 "이혼 시 재산분할은?"과 같이 특정 상황에서의 법적 처리 방법을 묻는 질의이다. 이는 상황과 관련된 여러 조문을 검색하고 통합하여 실용적인 답변을 생성해야 한다. 상황별 시나리오 데이터베이스를 구축하여 유사 상황에 대한 빠른 매칭을 지원한다.

다섯째, 복합 질의는 "상속 포기 후 채무 책임은?"과 같이 여러 법률 개념이 결합된 질의이다. 이러한 질의는 각 개념에 대한 조문을 검색한 후, 그들 간의 관계를 분석하여 종합적인 답변을 제공해야 한다.

질의 분석 모듈의 핵심 로직은 표 3과 같다.

표 3. 질의 분석 알고리즘

Table 3. Core algorithms for query analysis

INPUT: Query text Q
OUTPUT: Query type and extracted features
1. Preprocessing:
- Normalize Korean text (Hangeul normalization)
- Extract legal entities using NER
- Tokenize with legal domain dictionary
2. Pattern Matching (Priority-based):
- IF regex_match(Q, "제\d+조") THEN
type = "DIRECT", extract article_number
- ELIF semantic_similarity(Q, concept_patterns) > 0.8 THEN
type = "CONCEPT", extract key_concepts
- ELIF context_analyzer.detect_situation(Q) THEN
type = "SITUATION", extract context_features
3. Confidence Scoring:
- Calculate confidence based on pattern strength
- IF confidence < threshold THEN type = "COMPLEX"
4. Feature Extraction:
- Extract legal domain entities
- Identify temporal indicators
- Detect question intention

질의 분석의 정확성을 높이기 위해 법률 용어 사전과 동의어 매핑을 구축하였다. 예를 들어, "상속 포기"와 "상속 거부", "재산 분할"과 "재산 분배" 등의 유사 표현을 올바르게 인식할 수 있도록 하였다. 또한 각 질의 유형별로 신뢰도 점수를 부여하여, 불

확실한 경우 복수의 검색 전략을 병행할 수 있도록 하였다.

3.4 지식 그래프 기반 검색 전략

법률 문서의 복잡한 상호 참조 구조를 효과적으로 활용하기 위해 지식 그래프 기반 검색 전략을 구현하였다. 이 전략은 조문 간의 관계를 그래프 구조로 모델링하고, 그래프 탐색 알고리즘을 활용하여 관련 조문을 포괄적으로 검색한다.

그림 3과 표 4에서 보이는 것처럼, 지식 그래프는 다음과 같은 요소로 구성된다.

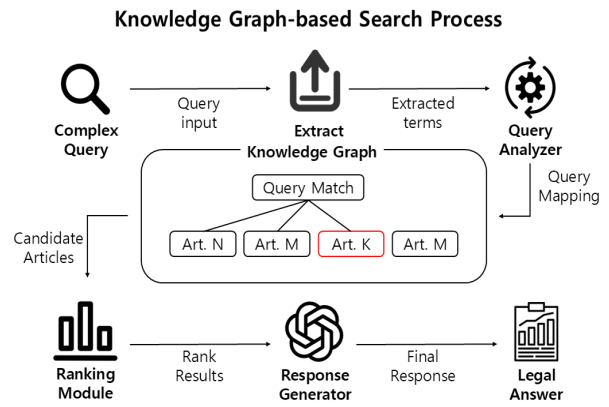


그림 3. 민법 지식 그래프 기반 검색 모듈 구조

Fig. 3. Civil law knowledge graph-based search module structure

표 4. 지식 그래프 구축 알고리즘

Table 4. Key methodologies for building knowledge graphs

Civil Law Knowledge Graph Algorithm
Input: Articles A = {a1, a2, ..., an}
Output: Graph G = (V, E)
Steps:
1. Node Creation: Create nodes with metadata (title, content, hierarchy)
2. Reference Extraction: Extract explicit references using regex, create directed edges
3. Semantic Edges: Calculate similarity between article pairs, add edges if above threshold
4. Structure Integration: Connect articles within chapters/sections, add inheritance relationships
Goal: Transform legal text into structured graph representing inter-article relationships.

그래프 탐색 알고리즘은 질의 유형에 따라 다르게 적용된다. 직접 참조 질의의 경우 단순히 해당 노드를 찾아 반환한다. 개념 기반 질의의 경우 관련 개념을 포함하는 모든 노드를 찾고, 그들 간의 관계를 분석하여 핵심 조문을 선별한다. 상황 기반 질의의 경우 너비 우선 탐색(BFS, Breadth First Search)을 통해 관련 조문을 확장적으로 검색한다.

이러한 그래프 기반 접근법은 단순 키워드 매칭이나 벡터 유사도 검색만으로는 찾기 어려운 간접적인 관련성을 가진 조문들도 효과적으로 검색할 수 있게 한다. 예를 들어, "미성년자의 상속 포기"라는 질의에 대해 미성년자 관련 조문과 상속 포기 관련 조문을 모두 찾은 후, 그들 간의 교집합과 주변 조문들을 함께 검색하여 종합적인 답변을 제공할 수 있다.

스코어 산출은 직접 참조 관계(0.4), 의미적 유사도(0.3), 계층적 거리(0.2), 키워드 매칭(0.1)의 가중합으로 계산된다. 최종 스코어가 임계값 0.5 이상인 조문만 검색 결과에 포함되며, 상위 10개를 스코어 순으로 반환한다.

3.5 응답 생성 및 품질 보증

응답 생성 모듈은 검색된 조문들을 기반으로 법률 문서의 계층 구조와 참조 관계를 반영하여 최종 응답을 구성한다. 이 과정에서 할루시네이션을 방지하고 법적 정확성을 보장하기 위한 여러 메커니즘을 적용한다.

첫째, 사실 기반 생성 원칙을 엄격히 적용한다. 모든 응답은 검색된 조문의 실제 내용에 기반해야 하며, 추론이나 해석은 최소화한다. 직접 조문 참조 질의에는 해당 조문의 전문을 정확히 인용하고, 개념 및 상황 기반 질의에는 관련 조문들을 통합하여 제공한다.

둘째, 응답 검증 메커니즘을 구현하였다. 생성된 응답이 원본 조문의 내용과 일치하는지 확인하고, 조문 번호나 내용의 정확성을 검증한다. 또한 참조된 조문들 간의 논리적 일관성을 확인하여 모순되는 정보가 포함되지 않도록 한다.

셋째, 구조화된 응답 형식을 채택하였다. 질의 유

형에 따라 적절한 응답 구조를 사용하여 정보의 명확성과 이해도를 높인다. 예를 들어, 개념 설명의 경우 정의-조문-예시의 구조를, 상황 기반 응답의 경우 관련 조문-적용 방법-주의사항의 구조를 사용한다.

이러한 체계적인 응답 생성 및 품질 보증 메커니즘을 통해 법률 질의에 대한 정확하고 신뢰성 있는 답변을 제공할 수 있으며, 사용자의 법률 정보 이해도 향상에 기여할 수 있다.

IV. 실험 및 비교 평가

4.1 실험 환경 및 데이터셋

이 연구에서 제안한 모듈형 RAG 기반 법률 문서 검색 시스템의 성능을 평가하기 위해 한국 민법 데이터셋을 사용하여 실험을 진행하였다[14].

실험 환경은 표 5와 같이 구성하였으며, 특히 대규모 임베딩 처리와 그래프 탐색을 위해 충분한 메모리와 GPU 자원을 확보하였다.

민법 데이터셋은 총 1,118개의 조문(본문 및 부칙 포함)과 그 계층 구조에 대한 정보를 포함하고 있다. 이 데이터셋은 시스템의 계층적 검색 로직과 참조 관계 추적 기능을 평가하는 데 적합하다. 평가 데이터셋 구축을 위해 다양한 유형의 법률 질의를 수집하였다. 질의 유형은 직접 조문 참조, 조문 제목 기반, 법률 개념 질의, 상황 기반 질의, 복합적 질의의 5가지 범주로 구분하여 각 범주별로 30개씩, 총 150개의 테스트 질의를 구성하였다. 각 질의에 대해 정답 조문과 기대 응답을 미리 정의하여 자동 평가가 가능하도록 하였다.

표 5. 실험 환경 구성

Table 5. Experimental environment configuration

Operating system	Ubuntu 22.04.1
Framework	Python 3.10.12
Library	LangGraph 0.0.15
	FAISS 1.7.2
	KLUEBERT
CPU	Xeon(R) Silver 4210R
GPU	NVIDIA RTX A6000x2
Memory	128 GB

4.2 성능 평가 결과 및 분석

시스템 성능 평가를 위해 검색 정확도(Retrieval accuracy), 응답 적절성(Response relevance), MRR (Mean Reciprocal Rank), 응답 시간(Response time)의 네 가지 지표를 사용하였다.

검색 정확도(Retrieval accuracy)는 질의에 대해 올바른 조문을 검색했는지를 평가하는 지표로, 정답 조문이 검색 결과에 포함된 질의의 비율을 다음과 같이 계산한다.

$$Accuracy = \frac{|Correct|}{|Q|} \quad (1)$$

응답 적절성(Response relevance)은 생성된 응답이 질의에 적절하게 답변하는지를 측정하는 지표로, 법률 전문가의 수동 평가와 BERT 기반 의미 유사도를 결합하여 0과 1 사이의 값으로 산출하였다.

MRR은 검색 결과에서 첫 번째 정답 조문의 순위를 고려한 지표로, 검색 결과의 순위 품질을 평가한다.

$$MRR = \left(\frac{1}{|Q|} \right) \sum_i \left(\frac{1}{rank_i} \right) \quad (2)$$

여기서 $|Q|$ 는 전체 질의 수, $rank_i$ 는 i 번째 질의에서 첫 번째 정답 조문의 순위이다.

표 6에 제시된 전체 시스템 성능 결과를 살펴보면, 검색 정확도(Retrieval accuracy) 88%, 응답 적절성(Response relevance) 92%를 달성하였다. MRR 값 0.84는 관련 문서가 검색 결과의 상위권에 포함되는 경향을 나타내며, 평균 응답 시간(Response time) 12.16초는 법률 전문가의 수동 조회 시간 대비 실용적인 수준이다.

표 6. 시스템 성능 평가 결과

Table 6. System performance evaluation results

Metric	Value
Retrieval accuracy	0.88
Response relevance	0.92
MRR	0.84
Average response time	12.16 sec

그림 4의 질의 유형별 상세 분석에서 가장 주목할 점은 직접 조문 참조 질의에서의 완벽한 성능이다. 모든 평가 지표에서 100%의 정확도를 달성한 것은 시스템의 기본적인 조문 검색 기능이 매우 안정적임을 보여준다. 이는 고유 ID 체계와 정규표현식 기반 조문 번호 추출의 효과성을 입증한다.

질의 복잡도가 증가할수록 성능이 점진적으로 감소하는 패턴을 보인다. 조문 제목 기반 질의(94% 검색 정확도)는 제목 인덱스와 키워드 매칭의 작동을 나타내며, 법률 개념 질의(92% 검색 정확도)는 여러 조문을 종합한 개념 설명 생성이 비교적 높은 성능을 보임을 나타낸다.

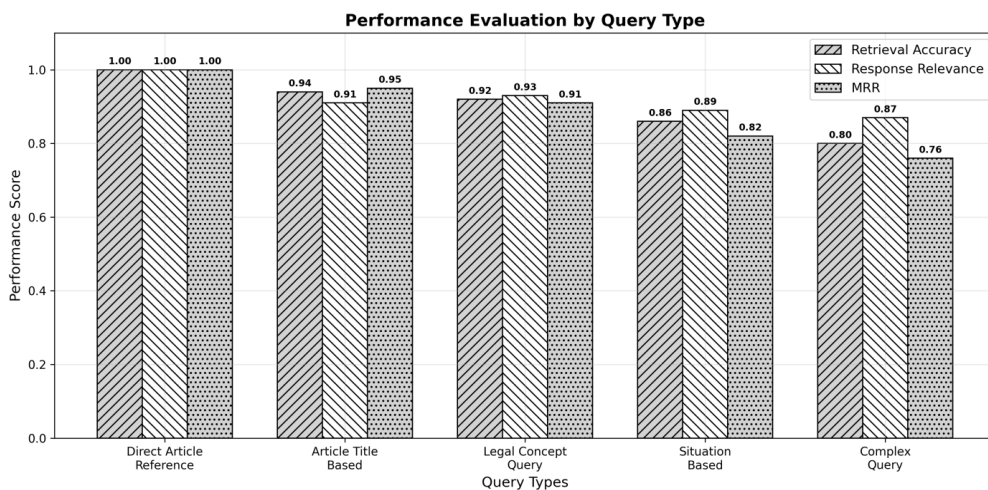


그림 4. 질의 유형별 성능 평가 결과

Fig. 4. Performance evaluation results by query type

상황 기반 질의(86% 검색 정확도)에서 성능 저하가 나타나는 것은 실제 상황과 법률 조문 간의 매핑이 가장 복잡한 추론 과정을 요구하기 때문이다. 복합 질의(80% 검색 정확도)의 경우 여러 법률 개념이 결합된 질의의 특성상 가장 낮은 성능을 보였지만, 여전히 실용적인 수준의 정확도를 유지하였다.

그림 4의 응답 시간 분석에서는 질의 복잡도와 처리 시간이 비례하는 합리적인 패턴을 보인다. 직접 조문 참조가 8.2초로 가장 빠르고, 복합 질의가 15.9초로 가장 느렸다. 각 모듈별 처리 시간 분석 결과, 지식 그래프 탐색이 전체 시간의 약 40%를 차지하여 주요 병목 지점으로 확인되었다.

4.3 베이스라인 시스템과의 비교

제안 시스템의 성능을 검증하기 위해 기존의 단순 키워드 기반 검색 시스템과 일반적인 RAG 시스템과의 비교 실험을 수행하였다. 키워드 기반 시스템은 BM25 알고리즘을 사용하였고, 일반 RAG 시스템은 단일 파이프라인 구조의 표준 RAG 구현을 사용하였다. 그림 5는 세 시스템 간의 성능 비교 결과를 보여준다.

비교 결과, 제안한 모듈형 RAG 시스템이 모든 품질 지표에서 높은 성능을 보였다. 특히 복합 질의에서의 성능 차이가 나타났는데, 키워드 기반 시스템은 68%의 검색 정확도를, 일반 RAG 시스템은 79%의 검색 정확도를 보인 반면, 제안 시스템은

88%의 정확도를 달성하였다. 이는 모듈형 구조와 지식 그래프 기반 검색의 효과를 보여준다.

응답 시간에서는 키워드 검색이 가장 빠르고 제안 시스템이 가장 느렸지만, 품질 대비 합리적인 처리 시간으로 평가된다. 법률 질의응답에서는 속도보다 정확성이 더 중요한 요소이므로, 약간의 시간 증가는 수용 가능한 범위이다.

개념 기반 질의에서도 제안 시스템은 92%의 정확도를 보인 반면, 일반 RAG는 83%, 키워드 검색은 71%에 그쳤다. 이는 법률 개념의 계층적 관계와 참조 구조가 복잡한 질의 처리에 활용되었음을 의미한다.

4.4 오류 분석 및 시스템 한계

전체 150개 질의 중 18개에서 오류가 발생하였다(오류율 12%). 오류 원인을 분석한 결과 다음과 같은 주요 패턴을 확인하였다.

첫째, 동음이의어 및 유사 표현 처리의 한계가 가장 큰 원인이었다(전체 오류의 44%). 예를 들어, "상속"과 "승계"와 같은 유사 개념이 질의와 조문에서 다르게 표현된 경우 검색 실패가 발생하였다. "포괄적 승계"와 "전면적 승계", "상속인"과 "승계인" 등의 유사 표현을 올바르게 매핑하지 못하는 경우가 대표적이다. 이는 법률 도메인 특화 동의어 사전의 구축 필요성을 시사한다.

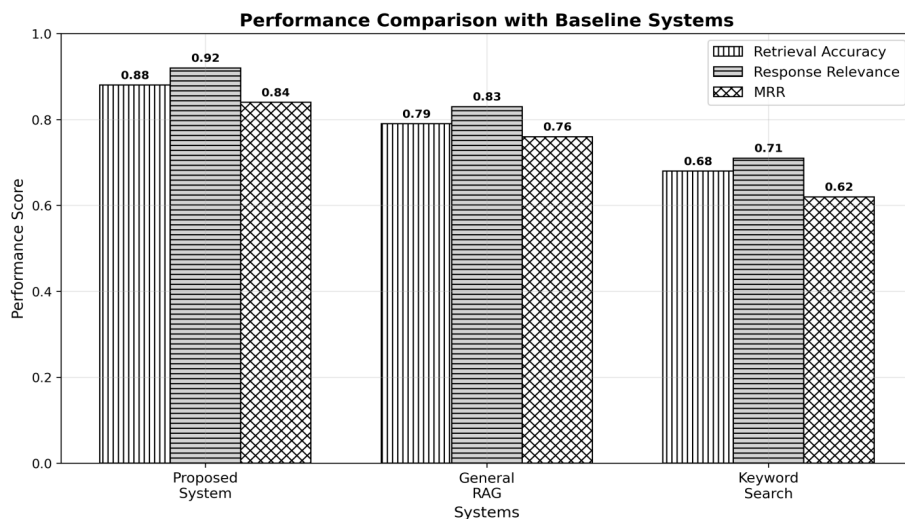


그림 5. 베이스라인 시스템과의 성능비교

Fig. 5. Performance comparison with the baseline system

둘째, 복잡한 참조 관계의 불완전한 추적이 주요 문제였다(전체 오류의 33%). 현재 시스템은 2단계까지의 참조만을 추적하도록 설계되어 있어, 더 깊은 다단계 간접 참조를 놓치는 경우가 있었다. 이는 그래프 탐색의 계산 복잡도 증가와 응답 시간 제약을 고려한 설계 결정이었으나, 법률 조문 간의 복잡한 연관 관계를 완전히 포착하는 데 한계로 작용하였다. 예를 들어, A조문이 B조문을 참조하고, B조문이 C조문을 참조하는 3단계 관계에서 A-C 간의 연관성을 찾지 못하는 경우가 발생하였다. 특히 상속 포기과 관련된 복잡한 참조 체인에서 이러한 문제가 두드러졌다.

셋째, 시간적 변화에 대한 고려 부족도 한계로 지적되었다(전체 오류의 23%). 민법은 지속적으로 개정되는데, 현재 시스템은 최신 버전만을 다루고 있어 과거 버전이나 개정 이력에 대한 질의에는 대응하지 못한다는 한계가 있다. 또한 부칙과 본문 간의 충돌 처리에서도 일부 미흡한 부분이 발견되었다.

V. 결론 및 향후 과제

이 연구에서는 한국어 민법을 중심으로 모듈형 RAG 아키텍처를 활용한 법률 답변 생성 시스템을 구현하였다. 법률 문서의 장-절-조 계층 구조를 활용하여 정확한 법률 답변을 생성하는 시스템을 개발하였으며, LangGraph를 통해 질의 유형별 최적화된 응답 생성 워크플로우를 구현하였다.

실험 결과, 제안된 시스템은 직접 조문 참조 질의에서 모든 평가 지표에서 100%의 정확도를 달성하였으며, 전체적으로 검색 정확도(Retrieval accuracy) 88%, 응답 적절성(Response relevance), 92%를 보였다. 계층 구조와 참조 관계를 활용한 접근 방식은 법률 답변 생성 시 할루시네이션을 방지하는 결과를 보였다. 특히 상황 기반 질의와 복합 질의에서 베이스라인 시스템 대비 높은 성능을 기록하였다.

이 연구의 주요 기여는 다음과 같다. 첫째, 한국어 법률 문서의 구조적 특성을 체계적으로 모델링하고 활용하는 방법론을 제시하였다. 장-절-조 계층 구조와 조문 간 참조 관계를 지식 그래프로 구축하고, 이를 기반으로 한 검색 전략을 개발하였다. 둘

째, 모듈형 RAG 아키텍처를 법률 도메인에 적용하여 높은 정확도와 신뢰성을 달성하였다. 각 모듈의 독립성과 재사용성을 확보하여 시스템의 확장성과 유지보수성을 향상시켰다. 셋째, 질의 유형별 최적화된 처리 경로를 구현하여 다양한 법률 질의에 대응할 수 있는 시스템을 개발하였다. 질의 분석부터 응답 생성까지 각 단계를 최적화하여 높은 성능을 달성하였다.

다만 이 연구에는 몇 가지 한계점이 있다. 첫째, 동음이의어 및 유사 표현 처리에서 한계가 확인되었다. "상속"과 "승계"와 같은 유사 개념이 질의와 조문에서 다르게 표현된 경우 검색 실패가 발생하는 경우가 있었다. 둘째, 복잡한 참조 관계의 불완전한 추적 문제가 나타났다. 현재 시스템은 2단계까지의 참조만을 추적하도록 설계되어 있어, 더 깊은 다단계 간접 참조를 놓치는 경우가 발생하였다. 셋째, 시간적 변화에 대한 고려가 부족하다. 민법의 지속적인 개정에 대응하기 위한 버전 관리 메커니즘이 필요하다.

향후 연구 방향으로서는 다음과 같은 과제들이 있다. 첫째, 법률 용어 동의어 사전 구축을 통한 표현 다양성 처리 능력 향상이 필요하다. 한국어 법률 텍스트의 다양한 표현 방식을 체계적으로 매핑하여 검색 성능을 개선해야 한다. 둘째, 다단계 참조 관계 추적 알고리즘 개선이 요구된다. 현재 2단계 제한을 확장하여 더 복잡한 참조 체인을 효과적으로 처리할 수 있는 방법론 개발이 필요하다. 셋째, 다양한 법률 영역으로의 확장 연구가 필요하다. 민법에서 형법, 상법 등 다른 법률 영역으로 답변 생성 범위를 확장하여 종합적인 법률 질의응답 시스템을 구축하는 연구가 요구된다.

또한 실제 서비스 적용을 위해서는 추가적인 개선이 필요하다. 응답 시간 단축을 위한 최적화 기법 적용, 사용자 피드백을 반영한 지속적인 모델 개선 체계 구축, 그리고 법률 전문가와의 협업을 통한 시스템 신뢰성 검증이 중요한 과제이다.

이 연구는 법률 답변 생성에 모듈형 RAG 아키텍처를 적용한 시도로서, 일반 사용자의 법률 정보 이해도 향상에 기여할 수 있는 기반을 제공하였다. 향후 이 연구가 더욱 발전하여 실제 법률 서비스에

활용되고, 법률 정보의 접근성을 크게 향상시킬 수 있기를 기대한다. 특히 법률 지식의 민주화와 사법 서비스의 효율성 향상에 기여할 수 있을 것으로 예상된다.

References

- [1] T. Brown, et al., "Language Models are Few-Shot Learners", *Advances in Neural Information Processing Systems*, Online, Vol. 33, pp. 1877-1901, Dec. 2020. <https://doi.org/10.48550/arXiv.2005.14165>.
- [2] N. Guha, et al., "LegalBench: A Collaboratively Built Benchmark for Measuring Legal Reasoning in Large Language Models", *Advances in Neural Information Processing Systems*, New Orleans, LA, USA, Vol. 36, pp. 1-14, Dec. 2023. <https://doi.org/10.48550/arXiv.2308.11462>.
- [3] D. M. Katz, M. J. Bommarito, S. Gao, and P. Arredondo, "GPT-4 Passes the Bar Exam", *Philosophy of Science*, Vol. 382, No. 2270, Article 20230254, Apr. 2024. <https://doi.org/10.1098/rsta.2023.0254>.
- [4] H. Kim, S. Lee, and M. Choi, "Korean Legal Document Processing with Hierarchical Structure Analysis", *Proc. of the Korean Institute of Information Scientists and Engineers Conference*, Jeju, Korea, pp. 245-247, Jun. 2022.
- [5] S. Ji, T. Zhang, and R. Jia, "Survey of Hallucination in Natural Language Generation", *ACM Computing Surveys*, Vol. 55, No. 12, pp. 1-38, Mar. 2023. <https://doi.org/10.1145/3571730>.
- [6] P. Lewis, E. Perez, A. Piktus, F. Petroni, V. Karpukhin, N. Goyal, H. Küttler, M. Lewis, W. Yih, T. Rocktäschel, S. Riedel, and D. Kiela, "Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks", *Advances in Neural Information Processing Systems*, Online, Vol. 33, pp. 9459-9474, Dec. 2020. <https://doi.org/10.48550/arXiv.2005.11401>.
- [7] S.-Q. Yan, J.-C. Gu, Y. Zhu, and Z.-H. Ling, "Corrective Retrieval Augmented Generation", *arXiv preprint arXiv:2401.15884*, Jan. 2024. <https://doi.org/10.48550/arXiv.2401.15884>.
- [8] A. Asai, Z. Wu, Y. Wang, A. Sil, and H. Hajjishi rzi, "Self-RAG: Learning to Retrieve, Generate, and Critique through Self-Reflection", *Proc. of the Twelfth International Conference on Learning Representations (ICLR)*, Vienna, Austria, May 2024.
- [9] Y. Gao, Y. Xiong, M. Wang, and H. Wang, "Modular RAG: Transforming RAG Systems into LEO-like Reconfigurable Frameworks", *arXiv preprint arXiv:2407.21059*, Jul. 2024. <https://doi.org/10.48550/arXiv.2407.21059>.
- [10] Y. Gao, Y. Xiong, X. Gao, K. Jia, J. Pan, Y. Bi, Y. Dai, J. Sun, and H. Wang, "Retrieval-Augmented Generation for Large Language Models: A Survey", *arXiv preprint arXiv:2312.10997*, Dec. 2023. <https://doi.org/10.48550/arXiv.2312.10997>.
- [11] C. Ryu, S. Lee, and J.-Y. Sohn, "Retrieval-based Evaluation for LLMs: A Case Study in Korean Legal QA", *Proc. of the Natural Language Processing Workshop 2023*, Singapore, pp. 132-137, Dec. 2023. <https://doi.org/10.18653/v1/2023.nllp-1.13>.
- [12] OASIS Legal XML Member Section, "XML Court Document 1.1 Proposed Specification", *OASIS Standards*, 2021. <https://www.oasis-open.org/committees/legalxml-courtfilling/>.
- [13] D. Francesconi, U. Passerini, and D. Tiscornia, "Building Semantics into Legal Information", *Legal Knowledge and Information Systems*, pp. 171-180, IOS Press, Nov. 2005.
- [14] A. Boer, R. Hoekstra, and R. Winkels, "METALex: Legislation in XML", *Legal Knowledge and Information Systems*, pp. 1-10, IOS Press, Dec. 2002.
- [15] Ministry of Government Legislation, Republic of Korea, "National Legal Information Sharing Service", *National Law Information Center*, 2024. <https://www.data.go.kr/data/15000115/openapi.do>. [accessed: Dec. 15, 2024]

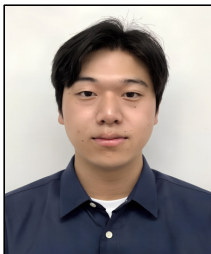
저자소개

신 은 성 (Eunsung Shin)



2021년 3월 ~ 현재 :
국립군산대학교 소프트웨어학과
학사과정
관심분야 : 보안, LLM, RAG

소 훈 (Hoon So)



2020년 3월 ~ 현재 :
국립군산대학교 소프트웨어학과
학사과정
관심분야 : LLM, 게임

윤 영 준 (Youngjun Yoon)



2022년 3월 ~ 현재 :
국립군산대학교 소프트웨어학과
학사과정
관심분야 : LLM, 백엔드, 블록체인

정 찬 우 (Chanwoo Jung)



2022년 3월 ~ 현재 :
국립군산대학교 소프트웨어학과
학사과정
관심분야 : LLM, 블록체인, 웹

김 동 연 (Dongyeon Kim)



2021년 3월 ~ 현재 :
국립군산대학교 소프트웨어학과
학사과정
관심분야 : LLM, 강화학습

정 현 준 (Hyunjun Jung)



2008년 3월 : 삼육대학교
컴퓨터과학과(공학사)
2010년 3월 : 숭실대학교
컴퓨터학과(공학석사)
2017년 9월 : 고려대학교
컴퓨터·전파통신공학과(공학박사)
2017년 8월 ~ 2020년 8월 :
광주과학기술원 블록체인인터넷경제연구센터 연구원
2021년 3월 ~ 현재 : 국립군산대학교 소프트웨어학과
교수
관심분야 : 블록체인, 데이터 사이언스, 센서 네트워크,
사물인터넷, 머신러닝