

A Contract Analysis System based on Multi-hop Reasoning using a Knowledge Graph and LLM Fusion

지식그래프와 LLM을 융합한 멀티홉 추론 기반의 계약서 분석 시스템

박재연 최준호

INTRODUCTION

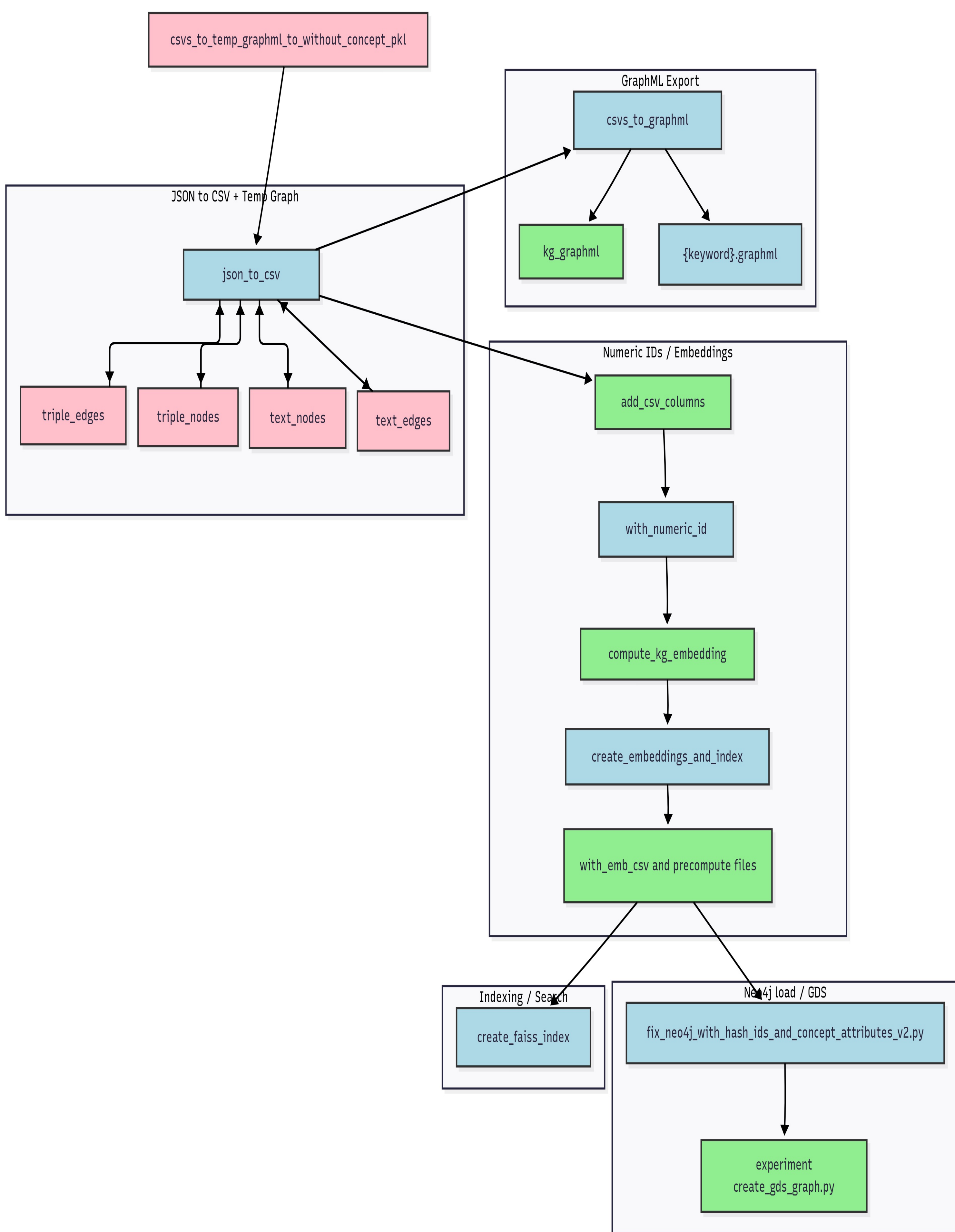
계약 관련 분쟁이 매년 급증하고(e.g., 한국공정거래조정원 접수 건수, '24년 역대 최대) 계약서의 복잡성이 심화됨에 따라, AI를 활용한 자동 검토의 필요성이 대두되고 있습니다. 하지만 키워드 및 유사도 기반의 기존 AI 모델은 개별 조항의 표면적 의미를 분석하는 데 그쳐, 조항 간의 복합적인 상호작용에서 발생하는 숨겨진 리스크를 탐지하는 데 명백한 한계를 보입니다.

본 연구는 이러한 한계를 극복하기 위해, 구조화된 **지식그래프(Knowledge Graph)**와 **대규모 언어 모델(LLM)**을 유기적으로 융합한 새로운 패러다임의 계약서 분석 시스템을 제안합니다. 제안 시스템은 **Graph RAG 기술**과 **다단계 추론(Multi-hop Reasoning)**을 통해, 텍스트의 의미를 넘어 계약의 법률적 '**관계**'와 '**맥락**'을 심층적으로 분석하는 것을 목표로 합니다.

OBJECTIVES

- 지식그래프 기반 계약서 모델링:** 비정형 계약서 텍스트로부터 법률 개체(Entities)와 관계(Relations)를 추출하고, 이를 구조화된 지식그래프로 자동 변환하는 파이프라인 구축
- 심층적 다단계 추론 구현:** 기존 수준을 넘어, 멀티 홉의 깊이로 조항 간의 숨겨진 연쇄 효과를 분석하여 복합적인 리스크를 탐지하는 추론 엔진 개발
- 실용적 리스크 평가 및 접근성 혁신:** 법률 전문가의 검토 효율성을 극대화하고, 나아가 중소기업 및 개인이 계약 리스크를 쉽게 식별하고 관리할 수 있는 실용적 시스템 개발

SYSTEM ARCHITECTURE



Input Layer: 계약서 원문 텍스트 입력

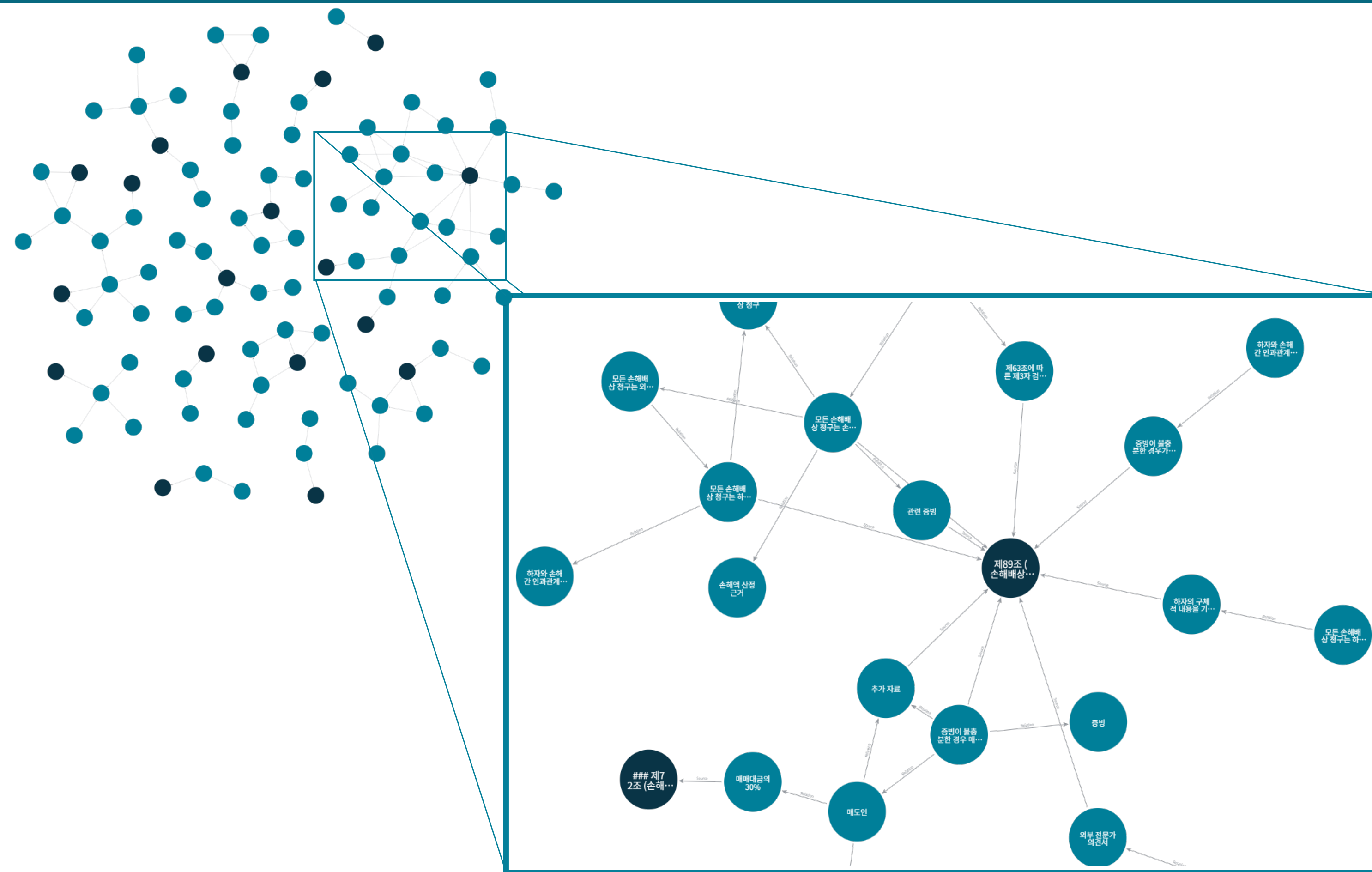
Knowledge Extraction Layer: NLP 기술을 통해 텍스트에서 트리플(Triples)을 추출하고, 개념을 추상화하여 노드(Node)와 엣지(Edge)로 구성

KG Construction Layer: 추출된 정보를 바탕으로 계약서에 대한 지식그래프 구축

AI Reasoning Layer: 사용자 질의를 바탕으로 Graph RAG가 지식그래프에서 관련 '지식 뭉치(Chunk)'를 검색하고, 이를 LLM에 전달하여 다단계 추론 수행

Output Layer: 분석된 리스크, 관련 조항 등을 종합하여 최종 결과 생성

KNOWLEDGE GRAPH



EXPERIMENTS

실험 데이터 및 평가 문항

[데이터] 실제 고난도 계약서

- 선정:** 4종의 복잡한 투자매매계약서를 분석 대상으로
- 사전 분석:** 각 계약서에 포함된 주요 독소 조항, 논리적 함정, 잠재적 분쟁 요소를 사전에 모두 분석하여 '기대 정답' 데이터베이스를 구축

[문항] 심층 추론 능력 측정

- 구성:** 계약서 별 100개, 총 400개의 심층 질문 세트를 개발
- 핵심:** 단순 사실 확인을 지양하고, '조항 조합'에 기반한 추론 능력을 평가하는 데 집중했다. 질문은 난이도에 따라 다음과 같이 분류
 - Level 1 (관계 추론):** 2~3개 조항의 상호작용 및 충돌 여부 판단
 - Level 2 (복합 추론):** 4개 이상 조항의 의존 관계를 분석하여 숨겨진 리스크 식별
 - Level 3 (최고난도 추론):** 조항 간 시간 제약 계산, 순환 정의 등 구조적 모순을 통해 '권리의 실질적 무력화' 여부 판단

실험 절차 및 평가 방법

(1) 모델 비교

- 비교 분석을 위해 **일반 상용 LLM**과 **지식그래프 기반의 AI**를 실험 모델로 선정

(2) 질의응답 자동화

- 400개의 질문을 각 AI 모델에 입력하고, 답변을 자동으로 추출 및 저장

(3) 성능 평가

- AI가 생성한 답변과 사전에 구축된 '기대 정답'의 논리적 일치 여부를 비교하여 채점을 통한 평가

| | 계약서1 | 계약서2 | 계약서3 | 계약서4 | 총점 |
|--------------|------|------|------|------|---------|
| 범용 LLM | 70 | 73 | 79 | 84 | 306/400 |
| 지식그래프 기반 LLM | 88 | 89 | 87 | 93 | 357/400 |

CONCLUSIONS & FUTURE WORK

결론

본 프로젝트는 지식그래프와 LLM의 성공적인 융합을 통해, 기존 AI의 한계를 뛰어넘는 **심층 법률 추론**이 가능함을 실증했습니다. 제안된 시스템은 실제 문제 해결 중심의 접근을 통해 **정량적인 리스크 관리 능력**을 입증했으며, 나아가 법률 서비스의 **접근성을 혁신**할 수 있는 새로운 가능성을 제시했다는 점에서 중요한 의의를 가집니다.

향후 연구

동적 지식그래프 구축: 법제처 API 등과 실시간 연동하여 법령/판례 개정 사항을 그래프에 자동 업데이트하는 연구

예측적 리스크 분석: 계약 외부 변수(시장 상황, 당사자 정보 등)까지 통합하여 미래의 분쟁 발생 확률을 예측하는 모델 개발