

Retrieval-Augmented Generation with Knowledge Graphs for Customer Service Question Answering

📅 Announcement Date	@2025년 2월 14일
≡ Conference Name	SIGIR 2024 Industry Track
⋮ Keywords	Knowledge Graph LLM RAG

1. 지식 그래프(Knowledge Graph)란?

지식 그래프(Knowledge Graph, KG)는 구조화된 방식으로 정보를 저장하고 검색할 수 있도록 설계된 데이터 구조입니다. KG는 **노드(Node)**와 **엣지(Edge)**로 이루어진 그래프로, 노드는 개체(엔터티)를, 엣지는 개체 간의 관계를 나타냅니다.

예시:

- 개체: "Apple Inc.", "Steve Jobs", "iPhone"
- 관계: "Apple Inc. - 창립자 - Steve Jobs", "Apple Inc. - 제품 - iPhone"

이러한 그래프 구조를 활용하면, 단순한 키워드 검색보다 더 정교한 의미 기반 검색이 가능하며, 정보 간의 연결을 이용해 보다 깊이 있는 분석이 가능합니다.

지식 그래프의 주요 특징

- 연결성:** 개체 간의 관계를 명확하게 정의하여 정보의 흐름을 이해할 수 있음.
- 의미론적 검색:** 단순한 키워드 매칭이 아닌, 의미 기반 검색을 지원함.
- 정보의 확장성:** 기존 지식을 바탕으로 새로운 정보나 관계를 추론할 수 있음.

2. 논문 요약 및 번역

논문 제목:

📌 Retrieval-Augmented Generation with Knowledge Graphs for Customer Service Question Answering

(고객 서비스 질문 응답을 위한 지식 그래프 기반 검색 증강 생성)

초록(Abstract)

고객 서비스 기술 지원에서는 과거 이슈(문제 해결 사례)를 신속하고 정확하게 검색하는 것이 중요한데, 기존의 검색 증강 생성(RAG) 방법에서는 과거 이슈 데이터를 단순한 텍스트로 처리하여 중요한 구조적 정보 및 이슈 간의 관계를 무시하는 한계가 있다. 본 연구에서는 **지식 그래프(KG)를 활용한 새로운 RAG 방법**을 제안하며, 과거 고객 서비스 이슈 데이터를 그래프 형태로 변환하여 검색 정확도를 향상시키고, 세분화된 문서로 인해 발생하는 정보 손실을 줄인다.

본 방법은 LinkedIn 고객 서비스팀에 6개월간 적용되었으며, 고객 문제 해결 시간의 중앙값(median)을 **28.6% 단축**하는 성과를 거두었다.

주요 성과

- 검색 성능 평가에서 MRR(Mean Reciprocal Rank)이 **77.6% 증가**
- BLEU 점수가 **0.32 향상**
- 고객 문제 해결 시간이 **28.6% 단축**됨

1. 서론(Introduction)

✓ 배경 및 문제점

- 고객 서비스에서 **과거 유사한 문제를 빠르게 검색하는 것**이 고객 만족도에 중요한 역할을 함.
- 기존 검색 증강 생성(RAG) 시스템은 **단순한 텍스트 기반 검색**을 사용하며, 이슈 간의 관계를 고려하지 않기 때문에 검색 정확도가 떨어짐.
- 이슈 설명이 긴 경우, **텍스트 세분화(segmentation)** 과정에서 중요한 정보가 손실될 수 있음.

✓ 기존 방법의 한계

1. **검색 정확도 저하:** 기존 RAG 방법은 이슈 간의 관계를 고려하지 않아 검색 성능이 낮음.
2. **답변 품질 저하:** 문제 해결 정보가 텍스트 분할 과정에서 분리되어 답변이 불완전할 수 있음.

✓ 본 연구의 기여

- **지식 그래프(KG)를 활용한 RAG 시스템 개발**
- **이슈 간의 관계를 보존**하여 검색 성능 향상
- **고객 서비스팀에 실제 적용**하여 문제 해결 시간 단축

2. 관련 연구(Related Work)

✓ 기존의 QA 시스템 분류

1. **검색 기반 QA:** 관계 추출 및 분산 표현을 활용하여 질문에 대한 답변을 검색.
2. **템플릿 기반 QA:** 미리 정의된 템플릿을 사용하여 질문을 매핑.
3. **의미 분석 기반 QA:** 질문을 논리적 표현으로 변환하여 지식 그래프에서 답을 찾음.

✓ 지식 그래프와 LLM(대형 언어 모델)의 통합 연구

- LLM을 예측자(Predictor), 인코더(Encoder), 정렬자(Aligner)로 활용하는 연구 진행됨.
- 의료 및 식품 분야에서도 LLM과 KG를 활용한 QA 시스템 개발 사례 존재.

3. 방법론(Methods)

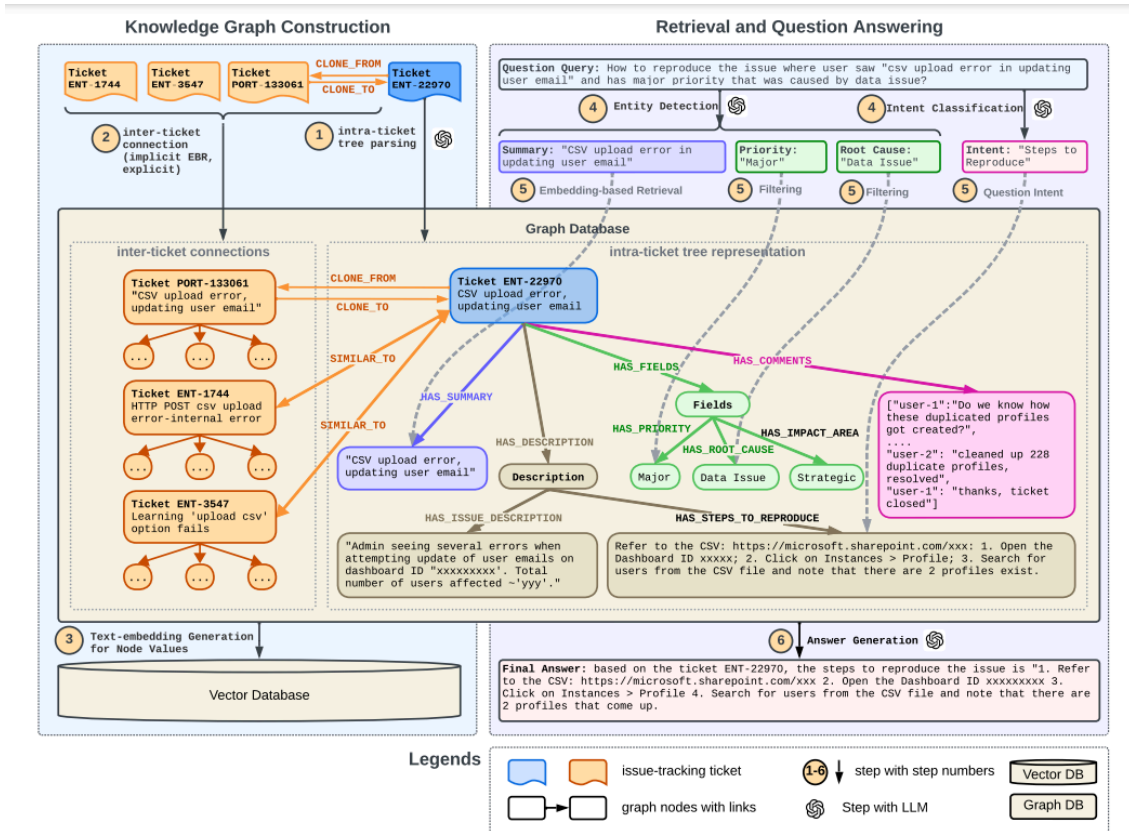


Figure 1: An overview of our proposed retrieval-augmented generation with knowledge graph framework. The left side of this diagram illustrates the knowledge graph construction; the right side shows the retrieval and question answering process.

3.1 지식 그래프 구축(Knowledge Graph Construction)

✓ 구성 요소

- **이슈 내부 구조(Intra-issue Tree):** 각 고객 서비스 티켓을 트리(tree)로 변환하여 세부 구조를 유지.
- **이슈 간 관계(Inter-issue Graph):** 티켓 간의 관계(예: 복사됨, 관련됨)를 반영하여 연결된 그래프를 생성.

✓ KG 구축 과정

1. 티켓을 트리 형태로 변환 → 기본 필드를 정의하고, LLM을 활용해 추가 정보를 자동 추출.
2. 이슈 간 관계를 그래프에 반영 → 유사한 티켓을 임베딩 기반으로 연결.

3.2 검색 및 질문 응답(Retrieval and Question Answering)

✓ 질문 처리 단계

1. 질문 속 개체(Entity) 및 의도(Intent) 추출

- 예: "사용자가 LinkedIn에 로그인할 수 없는 문제를 해결하는 방법은?"
→ 개체:
로그인 문제, 사용자가 로그인할 수 없음
→ 의도:
해결 방법 찾기

2. 유사한 과거 이슈 검색

- 개체를 기반으로 코사인 유사도를 계산하여 가장 관련성 높은 티켓을 검색.
- 검색된 티켓에서 관련 정보를 추출하여 LLM이 답변 생성.

4. 실험(Experiment)

4.1 실험 설계(Experiment Design)

- **데이터셋**: 고객 서비스 문의 및 해결 티켓을 포함한 "Golden Dataset"을 구축.
- **비교군**: 기존 텍스트 기반 검색(RAG) vs. 지식 그래프 기반 검색(RAG-KG).
- **모델**: LLM(GPT-4), 임베딩 모델(E5).

✓ 평가 지표

- **검색 성능**: MRR(Mean Reciprocal Rank), Recall@K, NDCG@K
- **QA 성능**: BLEU, ROUGE, METEOR 점수 비교

4.2 결과 및 분석(Results and Analysis)

✓ 검색 성능 비교 (Table 1)

모델	MRR	Recall@1	Recall@3	NDCG@1	NDCG@3
기존 방식	0.522	0.400	0.640	0.400	0.520
제안 방법	0.927	0.860	1.000	0.860	0.946

✓ QA 성능 비교 (Table 2)

모델	BLEU	METEOR	ROUGE
기존 방식	0.057	0.279	0.183
제안 방법	0.377	0.613	0.546

5. 실제 적용(Production Use Case)

✓ LinkedIn 고객 서비스팀에 실제 배포됨

- 기존 수작업 방식 대비 **문제 해결 시간 28.6% 단축**
- 평균 해결 시간 **40시간 → 15시간으로 감소**

6. 결론 및 미래 연구(Conclusion & Future Work)


✓ 연구 기여

- KG 기반 RAG 시스템이 검색 및 QA 성능을 향상.
- 실제 업무 환경에서 문제 해결 시간을 단축하는 효과 검증.

✓ 향후 연구 방향

1. 자동화된 그래프 템플릿 생성 연구
2. 실시간 지식 그래프 업데이트 방법 개발

3. 고객 서비스 외 다양한 도메인으로 확장

 **요약:** 본 연구는 **지식 그래프를 활용한 검색 증강 생성(RAG-KG) 시스템**을 개발하여 고객 서비스 QA 성능을 대폭 향상시켰으며, 실제 업무 환경에서 **문제 해결 시간을 28.6% 단축**하는 성과를 거두었다. 