

지식 그래프를 활용한 도메인 특화 질의응답 시스템의 자동 생성 솔루션: QA Maker

옥지윤, 성주연, 김민서, 황기태

한성대학교

2271522@hansung.ac.kr, 2271269@hansung.ac.kr, sunmay@hansung.ac.kr, calafk@hansung.ac.kr

An Automated Solution for Domain-Specific Question Answering Systems Using Knowledge Graphs: QA Maker

Ji Yoon Ok, Ju Yeon Soung, Min Seo Kim, Ki Tae Hwang

Hansung University

요약

오늘날 LLM 기반 검색 시스템은 막대한 자본과 기술력을 바탕으로 전 세계 웹 정보를 검색의 대상으로 하며 뛰어난 성능과 추론 결과를 보이고 있다. 반면, 대학이나 기업, 공공기관 등 도메인을 운영하는 조직 내부의 검색 시스템은 키워드 기반 매칭에 의존하여 키워드가 일치하지 않는 경우 검색이 되지 않거나 질의 의도와 관계없이 엉뚱한 답을 내놓는 등 그 성능이 매우 낮다. 또한, 도메인 내 정보를 검색하는 우수한 검색 시스템 구축에는 막대한 비용이 요구된다. 본 논문은 도메인에 특화된 질의응답 웹 시스템을 자동으로 생성하는 QA Maker 솔루션을 제안한다. QA Maker는 도메인의 URL 하나만 입력받으면 도메인 내 모든 데이터를 자동 수집하여 지식 그래프와 지식 그래프 DB를 구축하고 GraphRAG 기법을 통해 검색 및 추론 성능을 극대화하였다. 3개의 도메인을 대상으로 한 실험에서 해당 도메인들의 현재 검색 시스템은 평균 19% 정확도에 머문 반면, QA Maker로 생성한 QA 시스템은 90% 이상의 정확도를 달성하였다. 본 논문에서 제안하고 구현한 QA Maker는 특정 도메인 내에 뛰어난 검색 및 추론 성능을 보이는 QA System을 자동으로 생성할 뿐만 아니라 낮은 비용과 짧은 구축 시간으로 생성할 수 있어 높은 실용적 가치를 가진다.

I. 서론

LLM을 기반으로 하는 ChatGPT, Google의 Gemini 등 검색 시스템은 빠르게 발전하여 인터넷에서 학습된 데이터를 기반으로 높은 수준의 답변을 제공한다. 그러나 대학교, 연구소, 기업 등 조직의 내부 자료는 ChatGPT 등이 접근할 수 없으며, 데이터가 제대로 정리되어 있지 않아 기존의 내부 검색 시스템조차 검색 성능이 떨어지는 문제점을 보인다. 이로 인해 대학을 비롯한 조직 내부 사용자들은 자신의 검색 시스템을 거의 사용하지 않거나 신뢰하지 않는다. [1]

본 연구는 이 점에 착안하여 특정 도메인 내의 웹페이지와 문서들을 지식그래프와 지식그래프 DB로 구성하고 이를 기반으로 뛰어난 검색 및 추론 성능을 가진 질의응답 시스템을 자동으로 생성해주는 QA Maker를 제안한다. 본 논문에서 제안한 QA Maker의 아이디어는 그림 1과 같다.

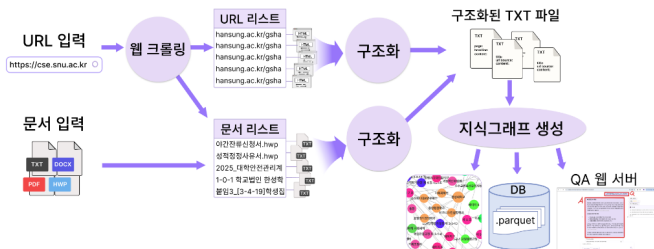


그림 1. QA Maker 시스템 동작 과정

QA Maker는 도메인의 대표 URL을 입력받으면 웹 크롤링을 통해 모든 하위 URL과 웹 페이지에 연결된 문서들의 리스트를 목록화한 후 이들을 지식 그래프 생성에 적합한 구조로 변환한다. 구조화된 데이터를 기반으로 GraphRAG 기술을 활용하여 지식그래프를 생성한다. [2] 미리 만들어진 템플릿화된 QA 웹 서버와 지식그래프, 지식그래프 DB를 결과물로 발생시킨다. 도메인 사용자는 결과로 제공된 웹을 통해 쉽게 질의하고 모든

질의에 대해 지식그래프 DB로부터 높은 수준의 응답과 추론 결과를 제공한다.

본 논문의 주요 기여는 다음과 같다. 첫째, 관리자가 URL만 제공하면 되는 생성 과정의 간결성이다. 둘째, 범용성 수준의 컴퓨터로도 짧게는 몇 시간에서 길게는 몇십 시간 내에 QA 시스템을 자동 생성할 수 있어 생성에 대한 비용이 거의 없다. 셋째, 생성된 QA 시스템은 기존 도메인 내 검색 시스템에 비해 월등한 성능을 보인다. 넷째, 전체 과정이 자동으로 동작하여 매우 편리하게 활용되어 실용적이다. 다섯째, 자연어로 질의하고 자연어로 응답한다.

II. 시스템 설계 및 구현

1. 시스템 구성 및 작동 원리

그림 2는 QA Maker의 구조와 생성된 QA 시스템의 구조를 보여준다.

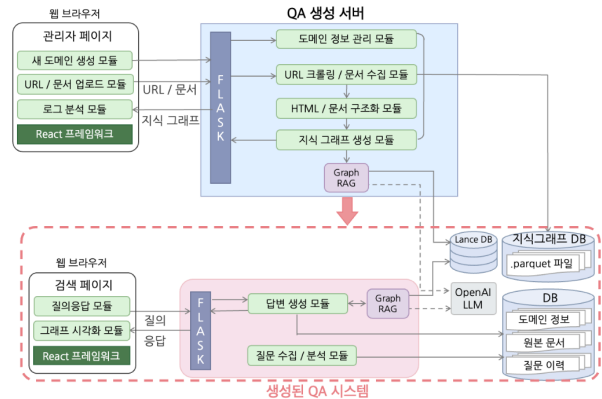


그림 2. QA Maker 시스템 구조도

QA 생성 서버는 Flask 웹 프레임워크 기반으로 구성되며 도메인을 등록

하는 도메인 정보 관리 모듈, 관리자가 입력한 URL을 기반으로 크롤링하고 업로드한 문서를 수집하는 URL 크롤링/문서 수집 모듈, 수집한 데이터를 구조화하는 HTML/문서 구조화 모듈, GraphRAG 기술을 활용해 구조화된 데이터를 지식 그래프로 생성하는 지식 그래프 생성 모듈이 연계되어 작동한다. QA 생성 서버의 최종 결과물은 지식그래프 DB와 Flask 기반 웹 서버이다. 사용자가 생성된 QA 시스템의 웹페이지에 질의를 하면 답변 생성 모듈은 지식그래프 DB에서 정보를 검색하여 사용자에게 정확한 답변을 제공한다.

2. 테스트 QA 시스템 생성

테스트를 위해 서울대학교 컴퓨터공학부 홈페이지를 대상으로 QA Maker를 활용하여 QA 시스템을 자동 생성하였다.

기존 서울대학교 컴퓨터공학부 검색 시스템은 “AI에 대해 상담하고 싶은데, 교수님 추천해줘”와 같은 자연어 질의에 대해 검색 결과를 전혀 제공하지 못했으며 다른 질의에서도 마찬가지였다. 반면 QA Maker로 생성된 QA 시스템은 그림 3과 같이 동일한 질의에 대해 정확한 답변을 제공할 뿐만 아니라 수많은 질의에 대해 충분하고 정확한 데이터를 잘 제공하고 있음을 확인하였다.

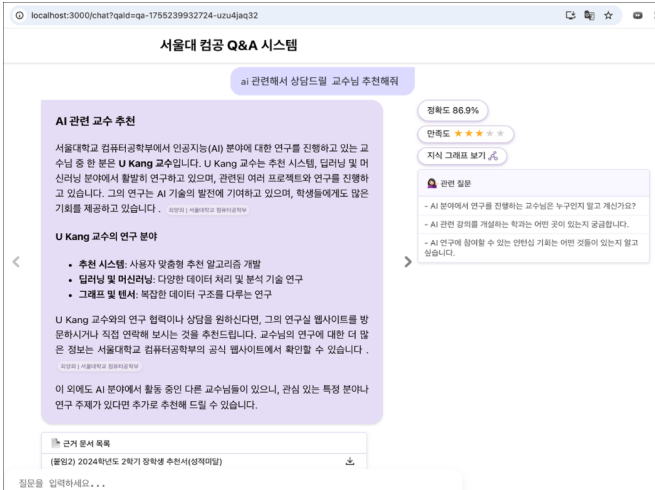


그림 3. QA Maker로 생성된 서울대학교 컴퓨터공학부 QA 시스템

III. 성능평가

이 절에서는 본 논문에서 제안한 QA Maker의 성능을 검증하기 위해 평가한 내용을 설명한다.

1. QA 시스템 생성 시간 평가

표 1은 대학과 협회를 포함한 3개의 도메인에 대해 QA 시스템을 생성하여 테스트를 진행한 결과이다.

표 1. QA 시스템 생성 시간

도메인	URL 수	문서 수	웹 크롤링 시간	구조화 시간	인덱싱 시간	총합
서울대학교 컴퓨터공학부	802	220	1260초 (21분)	11,220초 (3시간)	7시간 1분	10시간
한성대학교 전체	11,937	4,417	3787초 (63분)	25,219초 (7시간)	31시간	39시간
KOSA	142	12	1200초 (20분)	820초 (13분)	4시간 35분	5시간

한성대학교 전체 도메인의 경우 11,937개의 URL과 44,167개의 문서를 수집하여 QA 시스템을 생성하는 데 총 39시간이 소요되었다. 이를 통해 소규모 도메인은 몇 시간 내, 대규모 도메인은 몇 십 시간 내에 질의응답 시스템을 생성할 수 있다.

2. 응답 정확도 평가

본 팀이 구현한 QA Maker를 통해 생성된 QA 시스템의 응답 성능을 검증하기 위해 서울대학교 컴퓨터공학부 도메인의 현재 검색 시스템과 QA 시스템에 50개의 질의를 선정하여 응답 정확도를 비교하였다. 현재 서울대학교 컴퓨터공학부 검색 시스템은 평가팀이 0-100점 척도로 주관적 평가를 수행하였고, QA Maker가 생성한 QA 시스템은 RAGAS 프레임워크를 활용하여 Faithfulness, Answer Relevance, Context Recall, Context Precision 네 가지 메트릭을 가중 평균하여 최종 Accuracy를 산출하였다.

$$Accuracy = \frac{35 \cdot F + 30 \cdot R + 20 \cdot P + 15 \cdot C}{100}$$

그림 4는 응답 정확도를 비교한 성능 평가 결과를 보여준다. 현재 서울대학교 컴퓨터공학부의 검색 시스템의 평균 정확도는 19.2점인 반면, QA Maker가 생성한 시스템은 90.6점을 달성하여 매우 높은 정확도를 보였다.

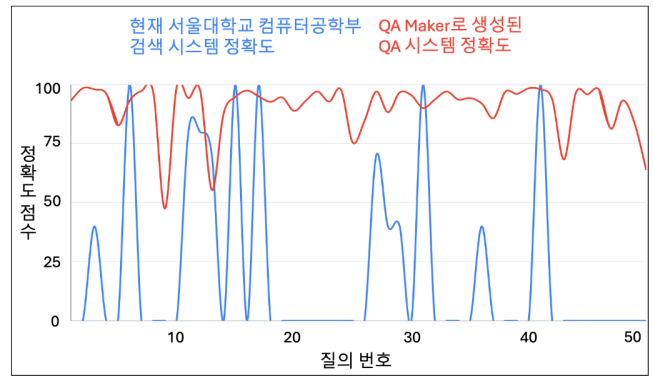


그림 4. 서울대학교 컴퓨터공학부의 응답 정확도 비교

IV. 결론

본 논문에서는 도메인 특화 질의응답 시스템을 자동 생성하는 QA Maker를 설계하고 구현하였다. QA Maker는 URL 입력만으로 웹 크롤링과 구조화, 지식 그래프 기반의 추론형 질의응답 시스템 생성을 자동화한다. 구현된 시스템은 여러 대학들을 대상으로 평가한 결과 응답의 정확도가 평균 90점에 달했고, 기존 도메인 검색 시스템과 비교한 결과 약 4.7배 높은 응답 정확도를 달성하였으며, 소규모 도메인은 수 시간 내, 대규모 도메인도 이틀 내에 시스템 생성이 가능함을 검증하였다. QA Maker는 낮은 구축 비용과 짧은 개발 시간, 높은 응답 정확도를 제공하며 내부 데이터를 보유한 모든 도메인 특화 환경에서 활용 가능한 솔루션임을 검증하였다.

ACKNOWLEDGMENT

본 연구는 과학기술정보통신부 및 정보통신기획평가원의 2024년도 SW중심대학사업의 결과로 수행되었음(2024-0-00049).

참 고 문 헌

- [1] W.Kim, et al., “Proposal of a Workflow Model for Generating a Construction Safety Guideline Knowledge Graph,” Journal of Construction Automation and Robotics, vol. 1, no. 4, pp. 46–54, 2022.
- [2] D. Edge, et al., “From local to global: A graph RAG approach to query-focused summarization,” arXiv preprint arXiv:2404.16130, 2024.