

고객 사례

LG CNS, Elasticsearch와 AI 검색 모델 통합으로 검색 정확도 고도화하다

지역

South Korea

산업

Software and Technology

솔루션

Elasticsearch



하이브리드 검색을 통해 검색 정확도 95% 수준 달성

- KeyLook AI모델을 KorQuAD 2 데이터 세트에서 테스트한 결과, 전문검색만으로 75% 수준이었던 정확도가 Elasticsearch의 전문검색, 벡터검색, 시맨틱 검색과 같은 세 가지 모듈에 KeyLook AI모델을 결합한 하이브리드 검색을 활용함으로써 20% 향상되어 95%를 달성했습니다.



방대한 데이터 검색 시간 절반으로 단축

- Elasticsearch의 역색인 (Inverted Index) 방식과 다양한 대용량 데이터 처리에 최적화된 기술들을 활용하여 11만 개, 질문 1만 개로 구성된 한국어 문서를 검색하는데 0.1 ~ 0.2초밖에 소요되지 않습니다. 이전에 비해 시간이 절반이나 단축되었으며 챗GPT와 같은 생성형 AI를 통합하여 답변 생성 속도는 더욱 빨라졌습니다.



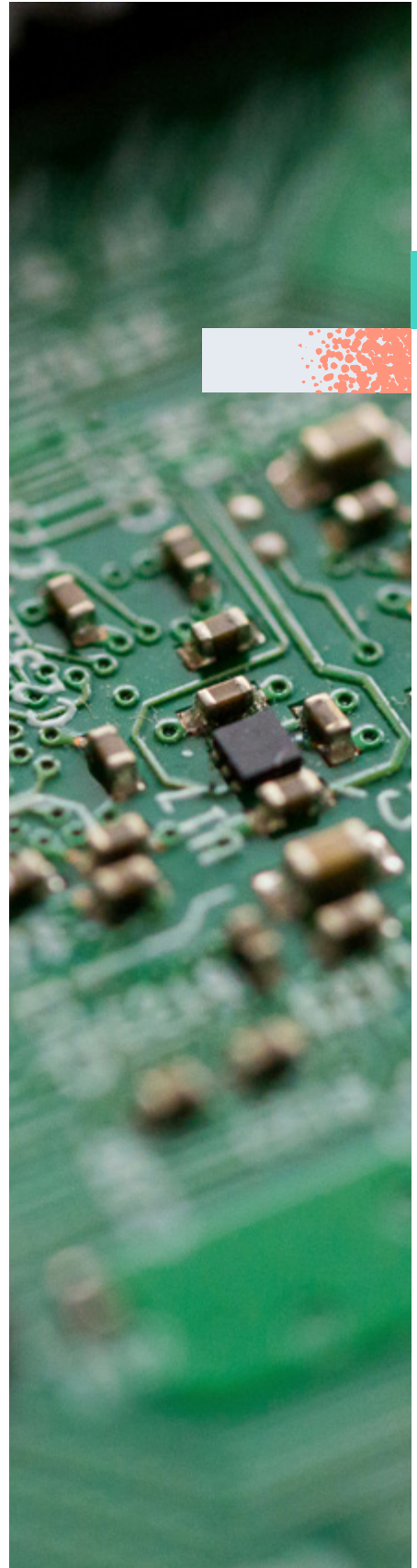
지식 관리 측면에서 임직원의 검색 경험 개선

- KeyLook의 주요 사용처는 흩어져 있는 기업 정보를 지식 관리(KM)로 전환하는 것입니다. LG CNS는 AI 모델을 통한 기업 문서 인코딩과 Elasticsearch의 색인화, 보안 기능 덕분에 지식 관리 모델을 발전시킬 수 있었습니다.

LG CNS는 자체 개발 중인 KeyLook AI 검색 알고리즘에 Elasticsearch를 결합하여 모델의 성능 및 검색 방식을 개선했습니다. 그 결과 원하는 검색 기능은 Elasticsearch를 활용하고 연구소는 모델 성능 개발에 더욱 집중할 수 있었습니다. 도입 이후 KeyLook AI 검색 알고리즘을 Elasticsearch와 통합하여 구현할 수 있었으며 대량 데이터 검색 시간 또한 0.1~0.2초 수준으로 50% 이상 단축되었습니다. 비즈니스 측면에서도 데이터를 지식 관리 방식으로 전환하여 임직원의 검색 경험이 개선되었습니다.

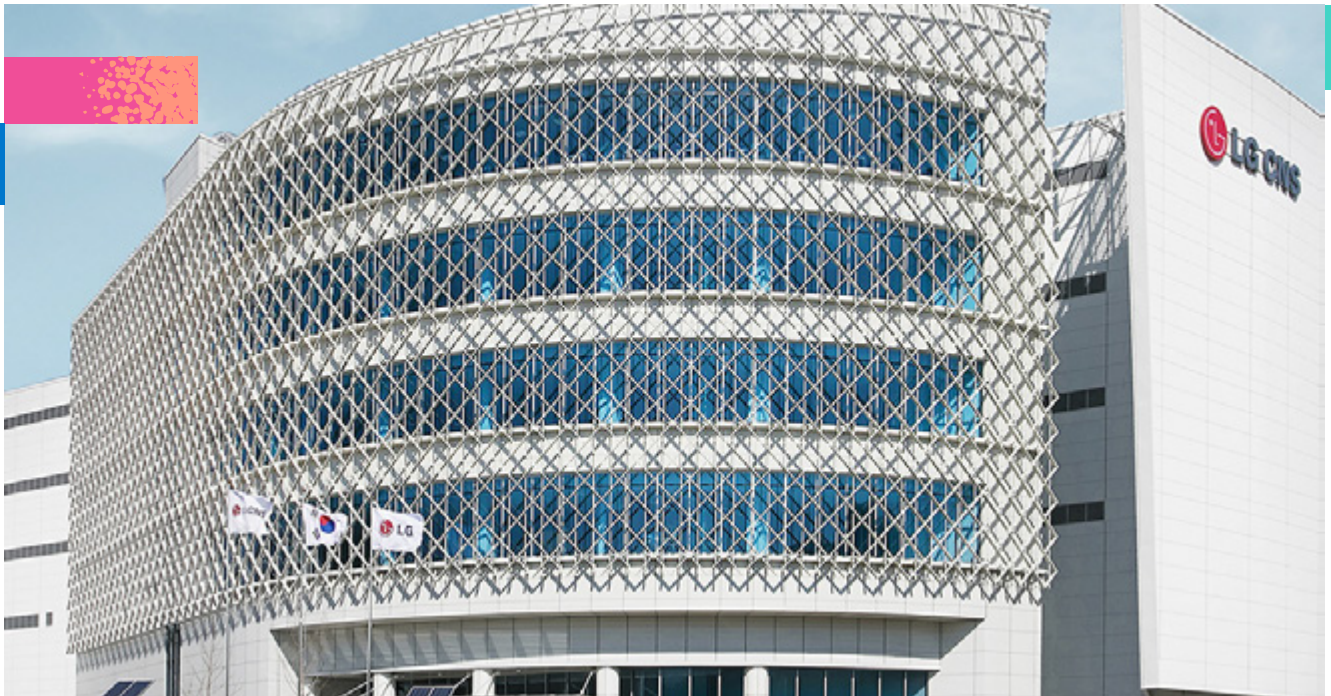
LG CNS는 대한민국을 선도하는 IT 서비스 기업입니다. 1987년 LG 그룹의 자회사로 설립되었으며 컨설팅, 시스템 통합, 네트워크 통합, 비즈니스 프로세스 아웃소싱 등의 서비스를 제공합니다. 대규모 공공 IT 인프라 프로젝트에 참여하는 등 업계를 선도하고 있으며 전 세계 개발 센터와 글로벌 법인을 기반으로 글로벌 시장에도 활발히 진출하고 있습니다.

[생성형 AI](#)가 등장한 이후 하루가 다르게 발전하고 있는 생성형 AI 기술을 따라잡기 위해 많은 기업이 다양한 관련 기술을 선보이고 있습니다. LG CNS 역시 생성형 AI로 프로젝트를 개발하여 기업에 서비스를 제공하고 있습니다. LG CNS D&A사업부 AI 연구소 언어 AI Lab에서 개발 중인 KeyLook AI 검색 알고리즘은 생성형 AI를 활용합니다. 향후 검색이 사용되는 모든 분야에 적용되어 전반적인 검색 기능을 향상하는 것이 목표이며 더 나아가 LG CNS는 이 검색 알고리즘을 통해 지식 관리(KM) 분야 혁신이라는 비전을 달성하고자 합니다. 하지만 이를 실현하기 전에 몇 가지 장애물을 극복해야 했습니다.



벡터 검색 알고리즘 방식 도입의 필요성

LG CNS는 [검색증강생성\(RAG\)](#) 시스템 개발 중에 기존 검색 알고리즘이 사용자의 의도를 이해하지 못하는 문제를 해결해야 했습니다. 이를 위해 키워드 검색 기반의 알고리즘에서 벗어나 문맥 기반의 벡터 검색 방식을 도입해야 했습니다. 하지만 연구소는 검색 정확도와 관련된 모델 개발에 중점을 두었으며 그 외의 저장 및 검색 수행과 관련된 데이터베이스 검색 기능을 개선하기 위해서 외부 솔루션 도입을 검토했습니다. LG CNS D&A사업부 AI 연구소 언어 AI Lab (이하 'AI Lab') 김영민 총괄은 “기존의 검색 알고리즘에서는 ‘블록체인’을 검색했을 때 블록체인이 정확하게 들어있는 문서만 찾아볼 수가 있었다. 블록체인을 한글로 검색했을 때는 ‘Blockchain’이라는 영문 검색어 결과는 찾아오지 못하는 것이 문제였다”며 “동어의 사전 관리 등을 처리해서 해결할 수 있지만 이 작업이 생각보다 공수가 많이 드는 일이었다”라고 설명했습니다.



하이브리드 검색을 통해 검색 정확도 95% 수준을 달성

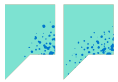
LG CNS는 AI 검색 모델과 함께 Elasticsearch를 도입해 검색 정확도와 속도를 개선할 수 있었습니다. 당시 [벡터검색](#)을 위한 검색엔진 솔루션을 선정 및 검토해본 결과 Elasticsearch가 컴퓨팅 자원도 효율적으로 사용하였고 벡터 검색 속도 향상을 위한 KNN 알고리즘의 성능도 가장 좋았습니다. 무엇보다 가장 결정적인 도입 이유는 Elasticsearch만이 LG CNS가 원하는 희소 벡터 모듈을 지원했기 때문입니다. AI Lab 김영민 총괄은 “검색엔진 검토 당시 Elasticsearch가 시맨틱 검색인 희소 벡터(Sparse vector) 모듈과 함께 한국어 전문검색을 지원하는 유일한 플랫폼이었다. 연구과정에서 개발한 희소 벡터 검색이 매우 좋은 성능을 보여주었기 때문에 우리는 이를 탑재할 수 있는 검색엔진을 원했다”고 설명했습니다. [Elasticsearch](#)의 희소 벡터 모듈을 활용하면 키워드가 정확히 일치하지 않거나 비슷한 의미의 단어더라도 심지어 오타가 있어도 검색이 가능합니다. Elastic 협력사의 지원을 받아 Elasticsearch를 도입한 결과는 놀라웠습니다.

LG CNS 관계자는 “모듈에 KeyLook AI모델을 KorQuAD 2 데이터 세트에서 테스트한 결과 Elasticsearch의 전문검색, 벡터검색, [시맨틱 검색](#)과 같은 세 가지 모듈에 KeyLook AI 모델을 결합함으로써 검색 정확도가 75%에서 95%로 20%나 향상되었다”며 “희소 벡터 검색은 검색어의 동의어들이 있는 문서도 함께 검색될 수 있도록 해주며 밀도 벡터 검색은 구어체 문장을 인식하고 의도를 이해하는 데 초점을 맞춘다. 따라서 이러한 모듈 결합 및 하이브리드 검색을 통해 성능을 개선할 수 있었다”고 답했습니다.

대용량 데이터 검색 시간 절반 이상 단축

검색 기능 개선은 정확도 향상에만 그치지 않았습니다. 대용량 데이터를 처리하기 위해서는 검색 원하는 정보를 빠르게 가져올 수 있어야 합니다. 초기에 LG CNS는 데이터가 늘어남에 따라 느려지는 검색 속도 최적화 및 효율적인 컴퓨터 리소스 활용의 중요성을 인지하였습니다. 따라서 Elasticsearch 도입 시 플랫폼의 효과적인 대량 데이터 처리와 규모 확장 가능성에 주목했습니다. LG CNS는 주로 고객사의 데이터를 활용하기 때문에 온프레미스 설치형으로 Elasticsearch를 도입했으며 최적화 및 하이브리드 모듈을 활용하여 검색 속도를 개선했습니다

AI Lab 김영민 총괄은 “메모리 관리 측면에서 Elasticsearch 플랫폼의 혜택을 봤다. 데이터가 많아지면 많아질수록 Elasticsearch의 처리 속도가 훨씬 빨랐다.” 고 설명했습니다. 또한 “Elasticsearch의 최적화 및 [하이브리드 검색\(Hybrid search\)](#) 지원으로 한국어 웹문서 11만 개의 질문 1만 개의 문서를 검색하는 데 걸리는 시간은 0.1초에서 0.2초로 이전보다 절반 이상 단축되었다”고 덧붙였습니다.



“KeyLook 알고리즘과 Elastic 검색 플랫폼의 통합은 Gen.AI 기반 대화 형식을 통해 업무에 필요한 지식을 검색하고 활용하는 데 중요한 역할을 했습니다. 그 과정에서 서비스 개발을 위해 필요한 다양한 기능을 Elasticsearch를 활용해 제공할 수 있었습니다. 이는 사용자 경험을 향상하고, 신속하고 정확한 응답으로 사용자 만족도를 높였습니다. Elastic은 검색 시장에서 높은 품질의 서비스를 제공하는 플랫폼으로, 다양한 고민을 통해 완성한 탄탄한 솔루션임을 실감할 수 있었습니다.”

김영민

LG CNS AI연구소 언어 AI Lab 총괄 consultant

지식 관리 측면에서 임직원의 검색 경험 개선

Elasticsearch의 도입은 비즈니스 측면에서도 실질적인 이익으로 나타났습니다. KeyLook AI의 알고리즘은 기업 데이터를 인코딩하고 이를 Elasticsearch에서 색인화합니다. 사용자가 검색 사이트에서 질문을 입력하면 KeyLook AI는 관련 문서를 식별하여 챗GPT에 답변을 전달합니다. 그런 다음 챗GPT는 사용자 친화적이고 이해하기 쉬운 답변을 종합하여 검색 결과를 사용자에게 제시합니다. LG CNS는 Elasticsearch의 강력한 보안과 색인화 기능 덕분에 원활하게 검색 기능을 구현할 수 있었으며 현재 지식 관리(KM) 부문에 이를 시범 적용했습니다. 관리자는 보안 기능을 활용해 등급별로 액세스 권한에 따라 정보를 필터링하여 검색 결과로 제공함으로써 전반적인 기업 지식 검색 성능을 개선했습니다. 또한 일반적인 전문(폴텍스트) 검색 방식을 사용하면 관련 없는 정보도 모두 포함되어 속도가 느려지는 반면 Elasticsearch는 역색인 방식을 활용하여 원하는 정보가 포함된 문서를 바로 찾을 수 있어 빠른 검색 속도를 자랑합니다. LG CNS 관계자는 “우리는 AI 모델을 통한 기업 문서 인코딩과 Elasticsearch의 특별한 역색인 방식 덕분에 사용자가 원하는 정보를 빠르게 검색할 수 있어서 지식 관리(KM) 서비스 비즈니스 모델을 발전시킬 수 있었다”고 설명했습니다.



LG CNS는 대한민국 선도하는 IT 서비스 기업으로, 컨설팅, 시스템 통합, 네트워크 통합, 비즈니스 프로세스 아웃소싱 등의 서비스를 제공합니다.

Elasticsearch를 활용하여 차세대 지식 관리 서비스 산업으로 사업 확대 예정

Elasticsearch를 통한 개발 혁신 과정에서는 편의성과 사용자의 만족도가 가장 중요합니다. AI Lab김영민 총괄은 KeyLook AI 검색 알고리즘의 성능과 검색 방식 개선 여정을 회상하며 Elasticsearch의 다양한 기능과 지원에 깊은 신뢰를 보였습니다. 그는 “Elasticsearch 사용이 굉장히 만족스러웠다. Elasticsearch가 쿼리를 통해 제공하는 다양한 기능들과 LG CNS가 집중했던 AI 검색 모델의 결합으로 좋은 시너지를 확인했다.”며 “Elastic의 Aggregation 기능 등을 활용해 실시간 데이터 처리와 분석이 가능하기 때문에, 사용자가 검색한 상위 Top 10 검색어, 사용자의 질문과 유사한 질의 등 최신 정보에 대한 빠른 접근이 가능하였다. 이는 기업의 의사결정 과정을 지원하며, 신속한 대응을 가능하게 했다.”라고 설명했습니다.

또한 LG CNS의 향후 계획에 대해 김영민 총괄은 “Elasticsearch에서 지원하는 더 많은 기능을 활용해 KeyLook AI의 속도와 정확도를 높이는 방법을 고민하고 있다”며 “나아가 LG CNS는 흩어져 있는 정보를 가치 있는 자산으로 ‘자산화’하는 차세대 지식 관리 서비스를 새로운 먹거리로 보고 관련 사업을 확대하고 있다”고 덧붙였습니다. 한편, LG CNS는 기술 확보 로드맵에도 명시하였듯 KeyLook AI 기능으로 한국어 외에 다국어 기술 지원을 제공할 예정이며 이러한 AI 서비스의 글로벌 시장 확대에 이를 실현 할 수 있을 것으로 기대하고 있습니다.

업계에서 가장 많이 사용되는 벡터 데이터베이스를 사용하여 혁신적인 AI 검색 경험을 구축하세요.

지금 시작하기