

LangChain 프레임워크 RAG 기반 개인정보 처리방침 분석 어시스턴트 개발 연구

강다빈¹, 강민진¹, 김한나¹, 김혜원¹, 이윤지¹
¹서울여자대학교 정보보호학과 학부생

kchabin@swu.ac.kr, rkdalswls789@swu.ac.kr, hanna2488@swu.ac.kr, kimhyewon415@swu.ac.kr,
 yunzxy@swu.ac.kr

A Study on the Development of a Privacy Policy Analysis Assistant Based on LangChain Framework and RAG

Da-Bin Kang¹, Min-Jin Kang¹, Han-Na Kim¹, Hye-Won Kim¹, Yun-Ji Lee¹, Sung-Wook Kim²
^{1,2} Dept. of Information Security, Seoul Women's University

요 약

개인정보 보호의 중요성이 증가함에 따라 기업의 개인정보 처리방침을 효과적으로 분석하고 개선하기 위한 방법을 제시하고자 한다. 개인정보 처리방침 평가 관련 데이터를 기반으로 RAG(검색 증강 생성, Retrieval-Augmented Generation) 기술과 LLM을 접목해 개인정보 처리방침 평가를 보조할 수 있는 AI 서비스를 개발한다. 이로서 평가위원들은 반복적인 노동에서 벗어나 전략적인 작업에 집중할 수 있으며, 각 기업은 개인정보 처리방침을 점검할 수 있으므로 전반적인 수준 향상이 기대된다.

1. 서론

지난 2023 년 개인정보보호위원회는 개인정보 보호법 개정을 통해 기업과 기관의 투명한 개인정보 처리와 책임성을 강화하기 위한 개인정보 처리방침 평가 제도를 도입했다. [1] 올해 6 월 평가 계획이 공개되었고, 첫 평가 시행을 앞둔 상태이다. [2]

개인정보 처리방침 평가를 위해 평가위원회를 구성하고 있으나 49 개 기업·기관을 대상으로 막대한 분석 작업이 필요하기에 현재 구성된 인적 자원만으로는 평가의 효율성이 떨어진다. 본 연구에서는 개인정보 처리방침의 정량적 평가를 돕는 RAG 모델 기반 개인정보 처리방침 분석 어시스턴트를 개발하고자 한다.

2. RAG 모델 기반 AI 서비스 제안

RAG(Retrieval-Augmented Generation)는 주어진 소스 데이터로부터 정보를 검색하고, 해당 정보를 활용하여 원하는 텍스트를 생성하는 과정을 수행한다. 이를 이용해 검색과 생성을 결합하여 보다 정확하고 의미 있는 텍스트를 생성할 수 있다. [3] 서비스 구성은 (그림 1)과 같다. 사용자가 평가 대상이 되는 개인정보 처리방침을 업로드하면 ChatGPT 는 해당 방침과 미리 불러온 작성지침을 참조하여 응답을 생성한다. 사용자가 개인정보 처리방침을 평가할 때, 일차적인 정량

평가와 요약 및 분석 등의 기능을 활용할 수 있는 AI 채팅 서비스를 제안하고자 한다.



(그림 1) 서비스 구성도

3. 개인정보 처리방침 분석 및 평가 연구 결과

분석 및 평가에 사용한 모델은 <표 1>과 같으며, 로컬 Visual Studio Code 환경에서 연구를 진행했다. [4]

LLM	GPT-4o ('gpt-4o-2024-05-13')
Embedding	text-embedding-3-small (CacheBackedEmbedding 적용)

<표 1> 사용한 모델

Knowledge Base 로 사용될 문서와 평가 대상이 되는 개인정보 처리방침을 모델에게 로드할 때에는 PDF 와 URL 방식을 사용할 수 있으며, 각각 PyPDFLoader, WebBaseLoader 를 사용했다. 연구에서는 빠른 응답 도출을 위해 개인정보 처리방침과 작성 지침의 내용을 텍스트 파일로 만들어 TextLoader 를 활용해 로드했다.

(그림 2)와 같이 평가하려는 개인정보 처리방침의 특정 내용을 출력하라고 명령한 PDF 로 전달했을 때는 이어지는 내용의 다음 페이지에 있는 내용은 출력하지 못하였으나, 텍스트 파일로 전달했을 때는 제대로 출력하는 것을 확인할 수 있었다.

```
## 개인정보 보호책임자에 관한 사항 :
**개인정보 보호책임자** : 기획조정관 고은영, 02-2100-2440
**개인정보 보호 분야별 책임자** :
- 법무감사담당관 (포털 본인인증) : 김민기, 02-2100-2471
- 개인정보보호정책과장 (일일 권리) : 김직동, 02-2100-3051
**이메일** : ch50094@korea.kr
**구분 부서명 성명 연락처** :
- 부서명 : 개인정보포털 담당자
- 연락처 : 24. 7. 3. 오전 4:43 개인정보처리방침 | 개인정보 포털

#### Question: 개인정보처리방침에서 개인정보 보호책임자에 관한 사항을 출력해줘
#### Answer:
- **개인정보 보호 책임자** :
  - 기획조정관 고은영, 02-2100-2440
  - **개인정보 보호 분야별 책임자** :
    - 법무감사담당관 (포털 본인인증) : 김민기, 02-2100-2471
    - 개인정보보호정책과장 (일일 권리) : 김직동, 02-2100-3051
    - 데이터안전정책과장 (기밀정보) : 박민수, 02-2100-3071
    - 자료보호정책과장 (일일 요구, 정보주체권리행사, 교육, 전문강사) : 윤여진, 02-2100-3081
    - 분쟁조정과장 (분쟁조정) : 이경미, 02-2100-3141
- **개인정보 보호 담당자** :
  - 법무감사담당관 : 김민기, 02-2100-2471
  - **개인정보보호실무담당자** :
    - 법무감사담당관실 : 박다원, 02-2100-2487
```

(그림 2) PDF 로드 결과(상), TXT 로드 결과(하)

Retriever 초기 연구에서는 Vector Store 의 Retriever 를 활용하였으나, LLM 이 개인정보 처리방침과 관련된 질문에 정확한 답변을 내놓지 못했다. 성능 개선을 위해 Pre-Post Retrieval 방식을 도입하여 Pre-Retriever 로는 BM25-Retriever, FAISS Vector Store Retriever 로 구성된 Ensemble Retriever 를 사용하여 검색 효율성을 최적화하였으며, Post Retriever 로 순위 재정렬(Reranker)을 사용하여 생성한 후보 문서들을 유사도 기준으로 우선 순위를 재정렬하여 정확도를 높였다.

Prompt Learning 에 활용한 평가 지표 데이터셋은 ‘2024 년도 개인정보 처리방침 평가계획(안)’의 평가 항목 중 적정성 평가 항목의 정량적 평가 지표 24 개를 기반으로 LLM 에 모호한 질문이 없도록 수정하여 활용했다. 평가할 개인정보 처리방침은 {policy}, GPT 가 참고해야 할 작성 지침은 {guideline}으로 프롬프트에 명시하고 각각 Retriever 를 생성하여 Chain 에 연결했다. [5]

```
## 평가 목적 (하)
1. 평가 목적 1: 개인정보처리자가 제공하는 서비스에 대한 개인정보 처리 목적을 작성 지침에 기재대로 누락 기재하고 있는지, 개인정보처리방침 처리 목적과, 관련된 개인정보처리 목적 지침을 출력하세요.
- 평가 결과 : [비밀한 개인정보처리방침] 평가
- 근거 : [개인정보 처리 목적 적정성 평가 기준 ( guideline)]에서 찾아 출력하세요.
- 채택할 문장 : [개인정보 처리 목적 적정성 ( policy)]에서 찾아 출력하세요.
2. 평가 목적 2: 개인정보처리자가 제공하는 개인정보의 항목 및 보유기간이 지침에 맞게 출력하세요.
- 평가 결과 : [비밀한 개인정보처리방침] 평가
- 근거 : [ guideline]의 제4항은 개인정보 항목, 개인정보의 보유 기간, 개인정보의 처리 및 보유기간 적정성 평가 근거를 찾아 출력하세요.
- 채택할 문장 : [개인정보 처리 목적 적정성 ( policy)]에서 찾아 출력하세요.
3. 평가 목적 3: 개인정보 처리에 관한 사항을 출력 및 범용을 기재한 내용 출력하세요.
- 평가 결과 : [비밀한 개인정보처리방침] 평가
- 근거 : [ guideline]에서 개인정보의 처리 목적 및 범용에 관한 사항을 찾아 출력하세요.
- 채택할 문장 : [개인정보 처리 목적 적정성 ( policy)]에서 찾아 출력하세요.
```

(그림 3) 활용한 프롬프트

(그림 3)과 같이 GPT 가 형식에 따라 적절한 답변을 생성할 수 있도록 하고, 정확한 결과가 도출되는지 확인하기 위해 개인정보 포털의 개인정보 처리방침과 작성 지침 PDF 에서 ‘개인정보의 처리 목적’, ‘처리하는 개인정보의 항목 및 보유 기간’, ‘개인정보의 파기 절차 및 방법’ 세 가지 항목을 텍스트 파일 형태로 정리하여 사용했다.

그 결과 (그림 4)에서 볼 수 있듯이 PDF 내용을 복사한 텍스트 파일을 로드했을 때보다 높은 정확도의 결과를 도출하는 것을 확인할 수 있었다.

```
### 평가 결과
1. 평가지표 1: 개인정보처리자가 제공하는 서비스에 대한 개인정보 처리 목적을 작성 지침에 기재된대로 누락 없이 기재하고 있는지, 개인정보 처리방침의 처리 목적과, 관련된 개인정보 처리 방침 작성 지침을 출력하세요.
- 평가 결과: 개인정보 처리방침은 개인정보의 처리 목적을 구체적으로 기재하고 있으며, 정보주체가 이해하기 쉽게 작성되었습니다. 다만, 일부 항목에서 목적이 다소 모호하게 기재된 부분이 있습니다.
- 근거:
  - 작성지침:
    ...
  # 2. 개인정보의 처리 목적
  * 개인정보처리자가 개인정보를 처리하기 위한 목적을 기재함
  * "개인정보의 처리"란 개인정보의 수집, 생성, 연계, 연동, 기록, 저장, 보유, 가공, 편집, 검색, 출력, 정정(訂正), 복구, 이용, 제공, 공개, 파기(破棄), 그밖에 이와 유사한 행위를 말함(개인정보 보호법 제2조 제2호)
  <중략>
  ...
- 처리방침 문장:
  ...
  # 개인정보의 처리 목적
  ① 개인정보위는 다음의 목적을 위하여 최소한의 개인정보를 수집하여 처리합니다. 처리하고 있는 개인정보는 다음목적 이외의 용도로는 이용되지 않으며, 이용 목적이 변경되는 경우에는 「개인정보 보호법」 제 18 조에 따라 별도의 동의를 받는 등 필요한 조치를 이행할 예정입니다.
  ② 개인정보위가 개인정보 보호법 제 32 조에 따라 등록·공개하는 개인정보파일의 처리목적은 다음과 같습니다.
  - 등록·공개하는 개인정보 : 교육서비스 제공 사용자 정보
  - 처리목적 : 개인정보보호 온라인 교육 서비스 제공, 이력관리, 수요증 발급
  - 개인정보 처리 사이트 : www.privacy.go.kr
  <중략>
```

(그림 4) 연구 결과

4. 결론 및 향후 연구

본 논문에서는 GPT-4o 에 검색 증강 생성(RAG) 기법을 적용해 개인정보 처리방침의 정량적 평가에 효과적인 방법을 제시한다. 시스템이 작성 지침을 참고해 개인정보 처리방침을 평가하고 내용을 출력하는 모습을 확인하였다. 이를 통해, 개인정보 처리방침의 평가를 보조하는 RAG 기반 AI 서비스를 구축함으로써 평가 위인 인적 자원의 한계를 극복하는 데 기여할 수 있을 것이다.

향후 연구에서는 공공기관의 정형화된 처리방침 뿐만 아니라 다양한 형식의 방침에 대해서도 평가 모델의 적용 가능성을 탐색하고, 답변의 정확도를 높이기 위해 보다 많은 데이터셋을 학습시킨 후 발전된 데이터 지표를 제공하고자 한다. 단순히 개인정보 처리방침의 내용을 출력하는 것을 넘어 점수, 등급 등 정량적 평가 결과를 도출해 낼 수 있는 RAG 모델로 발전시켜 나갈 것이다.

※ 본 논문은 과학기술정보통신부 대학디지털교육역량강화 사업의 지원을 통해 수행한 ICT 멘토링 프로젝트 결과물입니다.

참고문헌

[1] <https://www.pipc.go.kr/np/cop/bbs/selectBoardArticle.do?bbsId=BS074&mCode=C020010000&nttId=10244>.
 [2] 개인정보위원회, “2024 년도 개인정보 처리방침 평가계획(안)”, 2024
 [3] 정천수, “LLM 애플리케이션 아키텍처를 활용한 생성형 AI 서비스 구현: RAG 모델과 LangChain 프레임워크 기반.” 지능정보연구 29(4), 129-164, 2023.
 [4] <https://github.com/kchabin/pbl>.
 [5] <https://wikidocs.net/book/14314>