프로젝트 기획안

	사용자 맞춤형 RAG 기반 챗봇 서비스 기획(안)
서비스명 및 개요	서비스명 : 삼성 갤럭시 핸드폰 설정 정보 열람 서비스
	서비스 개요:
	- 삼성 갤럭시 핸드폰 설정에서 필요한 기능을 바로 찾기가 어렵고, 자체 검색 기능을 사용하더라도 느립니다. Samsung Members 에서 제공하는 AI 챗봇 서비스 또한 휴대폰 뿐만 아니라 다른 모든 삼성 제품에 대해 정보가 들어가 있어 Query 에 대한 pair 이 정확한 데이터를 제공해 주지 못해 곧바로 원하는 결과를 찾을 수 없고, 여러 단계를 거쳐야 해서 시간이 오래 걸렸으며, 부적합한 결과를 내는 경우가 잦았습니다.
	- 목표를 삼성 갤럭시 핸드폰의 설정에만 집중하고, 사용자 질문에 대해 바로 추론하여 곧바로 매뉴얼의 비슷한 항목들에 대해 호출하는 형식으로 사용자 접근성 및 편의성을 높일 수 있을 것입니다.
	-타겟 사용자 : 이는 핸드폰 설정에 관련한 내용이므로 삼성 갤럭시 핸드폰을 구매한 사용자가 타
목표 및 기대효과	서비스 목표:
	- 각 삼성 갤럭시 핸드폰의 기종에 대한 설정 질문에 대해 빠르고 정확한 답변을 제공하는 것 - 매뉴얼을 찾아보는 과정을 쉽게 하도록 챗봇에 Q/A 기능을 도
	기대효과: - 사용자는 긴 매뉴얼을 보지 않고 간단한 키워드만으로 필요한 정보를 얻을 수 있고, 비슷한 내용으로 추론하여 가장 근접한 답변을 빠르게 찾아 낼 수 있습니다.

타겟 사용자	타켓 사용자: - 개인 삼성 갤럭시 핸드폰 사용자이며, 핸드폰의 설정에 어려움을 겪는 사람(어느 메뉴에 어떤 옵션이 있는지 잘 모르거나, 원하는 옵션의 정확한 용어를 몰라 검색을 하지 못하는 사람)을 대상으로 서비스를 제공할 예정입니다. - 매뉴얼을 찾아 방법을 해결하는 것이 아닌 챗봇에게 질문을 하면 매뉴얼의 부분을 가지고 와서 답변을 드려 해결책을 제시함
데이터 구성 및 처리 방법	원천데이터 소스: 삼성 갤럭시 제품 매뉴얼 문서 원천 데이터 형식: PDF 문서 파일 데이터 처리 방법: ● 데이터 수집: ○ https://www.samsungsvc.co,kr 페이지에서 각 제품별 사용 설명서 PDF 다운로드 ○ https://usermanual.wiki/search.php 페이지에서 웹크롤링을 통한 데이터 수집 ● 데이터 전처리: ○ Document Parse 를 활용하여 PDF 파일을 HTML, Markdown 형태의 텍스트 데이터로 변환. ○ 변환한 데이터에서 HTML 및 Markdown 태그와 문법을 정제하여 순수 텍스트 데이터로 변환. ○ 텍스트 데이터로 변환. ○ 텍스트 데이터와 제품 이름 Pair 를 JSONL 형식의 데이터로 변환하여 QA 데이터셋 생성 준비.