

#로그연동 #Splunk #Elasticsearch #RESTAPI

Q. Splunk, Cloud 저장소 등과 연동 가능한가?

A. Elasticsearch 기반 외에도 다양한 외부 시스템과 RESTful API 연동이 가능하여 확장성 확보되어 있습니다.

5. 데이터 처리 및 파싱

5.1 문서 및 이미지 처리

#파싱 #OCR #이미지검색 #도면처리 #멀티모달

Q. 표와 이미지 파싱은 어떻게 하는지?

A. 전자문서(PPTX, Word, Excel, HWP 등)에 대한 파싱 기능을 지원하며, OCR 기술을 통한 이미지 내 텍스트 인식도 가능합니다.

Q. 이미지 검색 가능 여부?

A. 벡터 검색을 통한 TXT to IMG, IMG to IMG 검색 및 OCR을 통한 텍스트 추출 기반 검색은 지원됩니다.

Q. 이미지 기반 문서 청킹 가능 여부?

A. 청킹은 의미론적, 고정길이, 창 기반, 인접 문장 군집화 방식 등을 통해 진행되며, OCR로 추출한 텍스트에 대해 적용 가능합니다.

Q. OCR 지원 여부?

A. 전처리 자동화 과정에 OCR 기능이 포함되어 있어 이미지 기반 문서도 처리가 가능합니다. 오픈소스 기반의 OCR기능을 제공 하며 프로젝트 규모와 정확성을 요하는 작업의 경우에는 상용 OCR회사와 협업을 진행 하기도 합니다.