

신한캐피탈 투자신청의견서 정형데이터 추출 시스템 도입

제안요청서

2024년 6월

(주)신한캐피탈

1. 사업개요

1.1. 사업명 : 투자신청의견서 정형데이터 추출 시스템 구축

1.2. 사업 배경

- 1) 투자신청의견서는 여신심사 과정에서 매우 중요한 역할을 함. 이는 대출이나 투자를 신청한 기업 또는 개인에 대한 심사와 평가를 통해 투자 결정을 내리기 위해 작성되는 문서이며 투자신청 의견서 내용 검토 및 확인 후, 여신심사보고서 작성에 이용 됨
- 2) 투자신청의견서 문서 인식을 통하여 작성된 투자신청의견서의 내용을 자동 추출하여 투자신청의견서의 내용의 라벨링을 통하여 정형데이터 추출하여 이후 업무 처리에 이용할 수 있는 정형데이터 추출 자동화 시스템을 구축하고자 함

1.3. 사업 추진 방향

- 1) 투자신청의견서 문서에 작성된 내용을 솔루션을 통하여 추출하여 자동으로 시스템에 저장함으로써 업무 시간 단축 및 투자신청의견서 정보 정확성 증대
- 2) 의견서 수취 후 솔루션을 통한 추출 후 여신심사에 필요한 추출 정보와 투자 Term 뿐만 아니라 각종 투자 정보에 대한 문구 차이를 인식하여 정합성 검증 진행 및 관리
- 3) 문서 인식 및 정합성 검증, 비표준화 정보에 대한 표준화가 가능하도록 자연어 처리 등 AI 기술을 이용한 문자인식, GUI 형태 뿐 아니라 API 제공으로 사용자 및 내부 시스템 연동 가능한 솔루션 도입

1.4. 사업 구축 범위

1) 투자신청의견서 정형데이터 추출 솔루션 도입

1) AI 기반 문서인식 솔루션

- 자연어 처리 등 AI 기술을 이용하여 비표준화된 투자신청의견서에서 비정형화된 투자정보 및 재무, 사업계획, 신용, 담보, 보증 등의 정보 추출
- 신청자 마다 다른 형식의 투자신청의견서를 표준양식으로 변환하여 자동 처리 및 필수 기재 정보 관리
- GUI 형태 뿐 아니라 API 제공으로 사용자 및 내부 시스템(서버) 동시 사용이 가능하도록 서버 용 솔루션 제공
- -

2) 여신심사보고서 관리

- 투자신청의견서 표준화
- 신청서 종류 : 000, 000, 000
- 적용 범위
 - 투자신청의견서 구조 및 내용 학습
 - 투자신청의견서 용어집 등록
 - Structured Product 관련 향후 적용이 가능하도록 기본 업무 기능 제공
- 투자신청의견서 범위에 대한 요건 정의 및 검토
- 계약서에 표기되는 상품 정보 인식 및 상품 정보 표준화 가능
- 구축 시스템에 대한 데이터 관리/문자인식/상품 정보 표준화 등 업무 프로세스 설계

2) 내부시스템 연동

1) 0000 와 연동

- 데이터 / 파일 인터페이스 개발
- 여신심사보고서 초안 자동 작성 프로세스 구축
- 정합성 검증 프로세스 구축

1.5. 제안 시 주안점

- 1) 신한캐피탈의 투자신청의견서 정형화를 위한 최적 방안 제시
- 2) 시스템 도입에 대한 업무 프로세스가 적절하게 설계되었는지 여부를 결과로서 입증할 수 있는 검증 방안을 명확하게 제시하고, 향후 시스템 운영 과정에서 정기적인 검증 방안을 구체화하여 제시
- 5) 본 사업과 관련하여 필요한 모든 항목을 포함하여 제시하고, 제안 요청서에 제시되지 않는 사항에 대한 추가 제안 사항도 제안 가능

1.6. 사업 수행기간

: 프로젝트 시작일로부터 00 개월 이내에 오픈하여 안정화

※ 사업기간은 사업진행과정 및 계약일에 따라 변동 될 수 있음

요구사항 상세

1) 기능요구 사항

요구사항 분류		기능 요구사항
상세 설명	정의	다양한 문서 형식 지원
	세부 내용	<ul style="list-style-type: none"> ○ HWP, HWPX, DOC, DOCX 문서를 취급할 수 있어야 한다 ○ 원본의 어떠한 변경 없이 처리가 가능해야 한다 ○ 소요기관에 맞춤형 형태로 진행 할 수 있어야 한다
요구사항 분류		기능 요구사항
상세 설명	정의	문서 라벨링 방안
	세부 내용	<ul style="list-style-type: none"> ○ 문서 라벨링 처리 프로세스는 아래와 같이 진행함 학습 -> 문서처리 -> 문서보정처리 ○ 정형데이터 추출 :문서 내 문단, 표의 각 셀을 개별적으로 라벨링이 가능해야 함 ○ 문서구조 : Tree 형태 문서 구조를 벡터에 반영할 수 있어야 한다 ○ 유사문장 : 문서 구조의 각 노드별 유사 문장을 벡터에 반영 할 수 있어야 한다 <ul style="list-style-type: none"> - 유사도 판단 기준 값 설정을 통해 문서 특성에 맞게 최적화가 가능해야함

요구사항 분류		기능 요구사항
상세 설명	정의	문서처리 요청 및 데이터 전송
	세부 내용	<ul style="list-style-type: none"> ○ 문서 처리 요청 접수가 시스템을 통해 가능해야 함 ○ 솔루션은 요청된 문서를 Fetch 하여 정형 데이터 추출 및 전송 ○ 추출된 데이터는 XML 형태로 변화하여 제공 해야 함
요구사항 분류		기능 요구사항
상세 설명	정의	템플릿 관리 기능
	세부 내용	<ul style="list-style-type: none"> ○ 템플릿 목록을 보여야 하며 필요에 따라 사용자가 선택하여 사용 할수 있도록 구성해야 함 ○ 템플릿을 추가 또는 삭제가 가능 해야함. ○ 템플릿에 대한 형상관리가 가능하도록 해야함

요구사항 분류		기능 요구사항
상세 설명	정의	투자신청의견서에서 주요 계약조건 추출 기능
	세부 내용	<ul style="list-style-type: none"> ○ 다양한 형식의 투자신청의견서에서 기 정의한 주요 계약 조건을 key value 형식으로 추출하는 기능. ○ key값을 표준화된 key로 변경하고 value도 기 정의한 형식, 즉 날짜, 단위 등으로 표준화할 수 있음.
요구사항 분류		기능 요구사항
상세 설명	정의	추출된 key value 검증 및 수정 기능
	세부 내용	<ul style="list-style-type: none"> ○ 투자신청 의견서에서 추출 및 정형화된 결과를 사용자 친화적인 UI에 표시하는 기능. ○ 투자신청 의견서와 결과를 함께 보면서 검증 및 수정할 수 있는 기능. ○ 최종 검증한 결과를 저장하는 기능.
요구사항 분류		기능 요구사항
상세 설명	정의	검증된 추출 결과를 기반으로 투자신청 의견서 자동 생성
	세부 내용	<ul style="list-style-type: none"> ○ 검증된 결과를 기반으로 필요한 문서 자동생성 ○ 생성되는 문서는 표준화 및 데이터베이스화하여 관리 가능. ○ 생성된 문서를 거래상대방에게 이메일로 발송