

② 공모 제안서 양식-실현 가능 서비스 부문

- ※ 글씨 크기 10pt, 서체 맑은 고딕으로 통일하여 작성해 주세요.
- ※ 실제 데이터가 개방 되어 있지 않은 경우는 반드시 데이터별 첨부된 '테크니컬 리포트'를 참고하여 아이디어를 제안해 주세요
- ※ 이미지, 동영상 등 자료 첨부 시, 본인이 저작권을 가지고 있는 자료를 사용하거나, 본인의 저작권이 없는 경우 반드시 저작권자 출처를 명시해 주세요.(URL포함)
- ※ 제안서 양식의 내용은 어디까지나 참고 자료입니다. 제안 시 자유롭게 아이디어를 제안해 주세요.

1. 인공지능 학습용 데이터 활용 아이디어 제목

법We법 (ai를 활용한 판례 검색 시스템, ai법률 상담소)

※ 인공지능 학습용 데이터를 활용 서비스 제목을 적어주세요. (50자 이내)

2. 인공지능 학습용 데이터 활용 아이디어 내용

법률은 일반인들에게 친숙하지 않은 분야이기 때문에 위법행위에 여부에 대해 알기 쉽지 않았고 판례검색도 어려웠다. 해당 어플은 변호사에게 직접 문의해야 할 무거운 사건들이 아닌, 주변에서 쉽게 일어날 만한 사건들을 ai를 활용하여 자신의 상황과 유사한 사건의 판례를 검색할 수 있고 요약해서 직관적으로 보여주는 서비스를 제공한다.

기존에는 NAV**사의 지식 in의 서비스를 활용하여 관련직에 있는 사람에게 직접 물어보거나 '종합법률정보' 사이트에서 키워드를 입력하여 수백개의 사건들을 일일이 찾아봐야 하는 불편함이 있었다. 해당 서비스는 기존의 문제점을 보완하기 위해 사용자가 직접 자신이 처한 상황을 어플에 등록하고 이와 유사한 판례만 보여줌으로써 보다 빠른 검색이 가능하게 함으로, 법에 대한 진입장벽이 낮아질 수 있고 판결을 보며 사용자가 스스로 법률적 문제여부를 판단할 수 있게 한다.

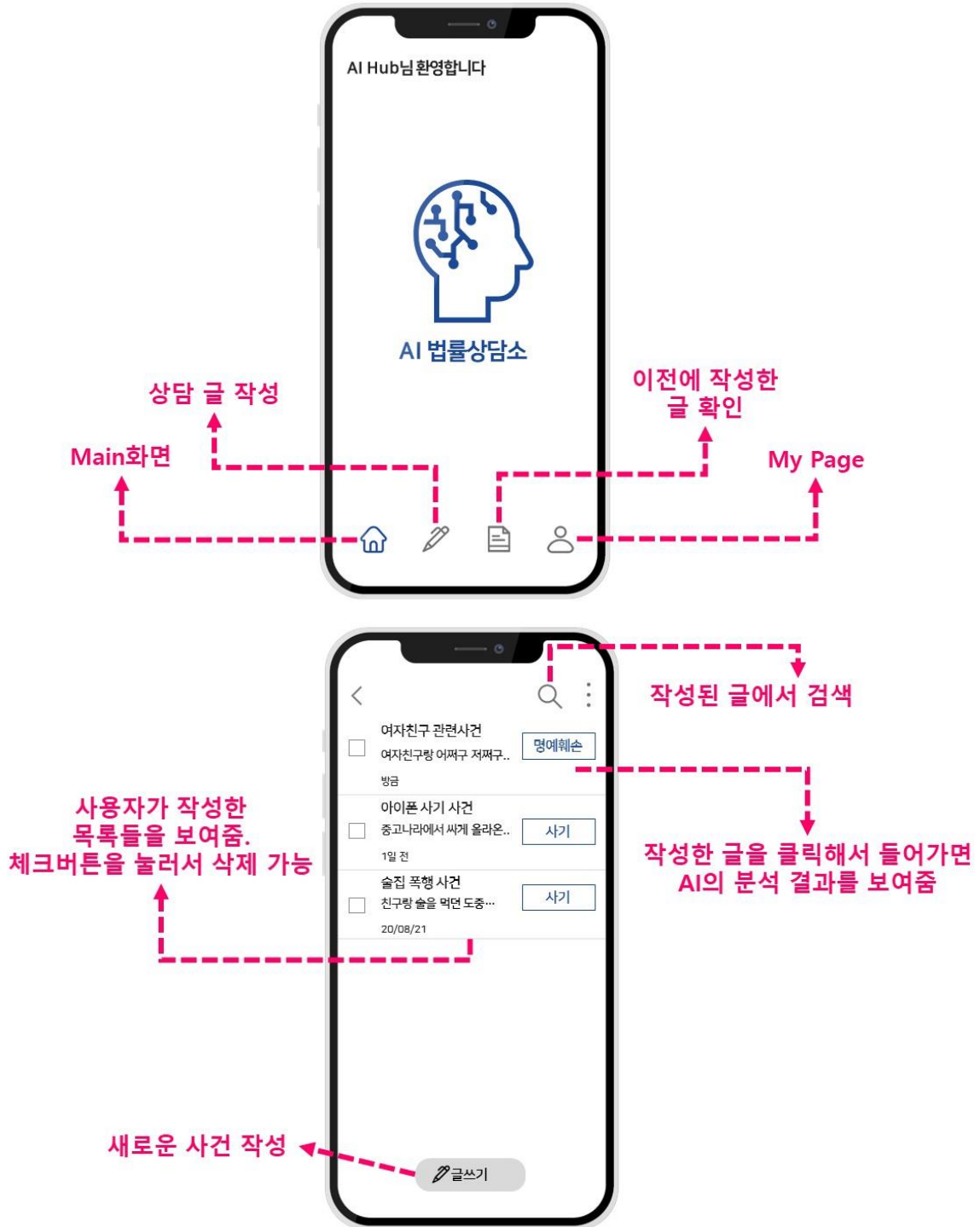
==서비스 제공 절차==

해당 어플의 서비스 범위는 '교통사고, 층간소음, 창업 인허가, 퇴직금, 이혼, 사이버공간에서의 명예훼손/모욕'으로 한정함.

1. 사용자가 자신의 상황과 가까운 불법행위 카테고리(위에서 언급한 교통사고 외 5가지)를 선택함.
2. 사용자가 해당 상황에 관련된 자료를 등록함.
 - 2-1. 해당 상황을 글로 모두 적어서 등록함.
 - 2-2. '사이버 공간에서의 명예훼손/모욕' 카테고리를 선택한 경우, 텍스트가 적힌 이미지 파일(해당 상황과 관련된 카카오톡이나 문자 메시지, SNS 상의 캡처본), 녹음 파일(해당 상황에서 녹음된 대화 파일)을 추가로 등록할 수 있음.
3. 해당 자료가 위법한 내용이 포함이 되어있지 않다면 '관련 없음'으로 알려줌
4. 위법한 내용이 포함되어 있다면, 부합하는 판례를 요약하여 아래와 같은 형식으로 보여줌.
 - 4-1. 해당 사건과 관련성이 높은 관련 사건들을 연관성이 높은 5개의 사례의 사건번호를 보여줌.

4-2. 해당 사례들에 대한 평균 결과를 페이지 하단에 “징역 _년, 집행유예 _년, 벌금 _원”의 형식으로 보여줌.

4-3. 출력된 5 개의 판결문 중 자세하게 보고 싶은 판례가 있다면, 해당 판례를 클릭하면 법률 판결문을 요약해서 제공함.



<명예훼손/모욕>

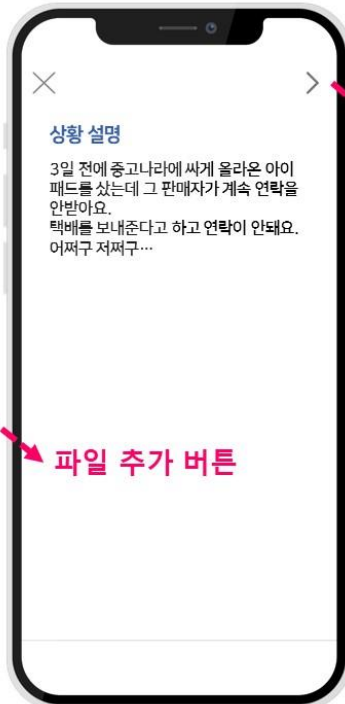
<나머지>

이전 화면

카카오톡이나 메시지, SNS상의 글 같은 텍스트 이미지를 첨부. 사진을 꼭 누르면 삭제 가능

대화내용이 녹음된 녹음 파일 첨부

텍스트로 말하고자 하는 상황을 설명함



텍스트로 변환된 내용이 정확한지 사용자가 다시 한 번 확인하고 수정할 수 있게 함



- ※ 인공지능 학습용 데이터를 활용 서비스 개요를 요약하여 적어주세요 (자유양식, 2,000자 이내)
- ※ 아이디어의 이해를 돕기 위한 다양한 이미지를 활용하여 설명하셔도 됩니다.

3. 아이디어를 실현하기 위해 필요한 인공지능 학습용 데이터

<활용된 인공지능 학습용 데이터>

1. 자연어 데이터 - 한국인 대화 음성
2. 자연어 데이터 - 한국어 음성
3. 자연어 데이터 - 문서요약 텍스트
4. 지식베이스 - 법률

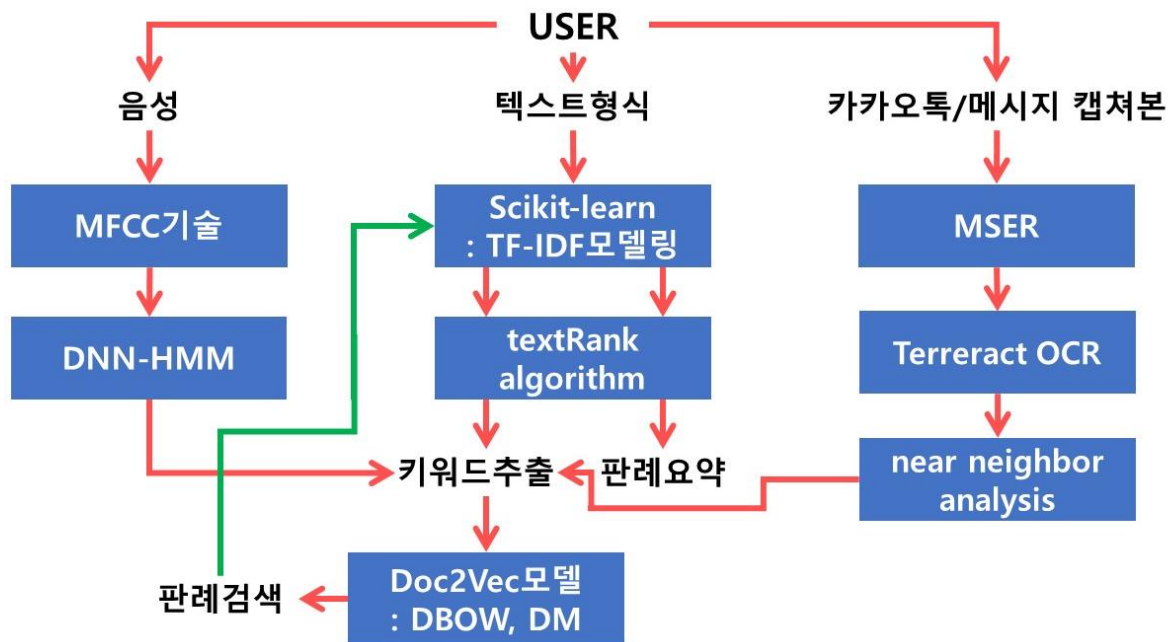
<본 서비스를 위해 가공이 필요한 인공지능 학습용 데이터>

1. 명예훼손 및 모욕 법률데이터(위의 4번에 해당하는 법률 데이터의 확장을 위한 데이터)
(출처: 대한민국 법원 종합법률정보 <https://glaw.scourt.go.kr/>)

- ※ 아이디어를 실현하기 위해 필요한 인공지능 학습용 데이터를 입력해 주세요.
- ※ AI 허브에 제공되고 있지 않은 학습용 데이터도 필요한 경우 적어주셔도 됩니다.

4. 인공지능 학습용 데이터 학습방법

1. 음성-텍스트 변환
 - 한국인의 대화음성 ai학습 데이터 사용
 - MFCC기술을 사용하여 음성특징을 추출
 - DNN-HMM기반으로 음성인식 진행
 - wFST를 사용하여 디코딩
2. 카카오톡/문자메세지 캡처본에서 텍스트추출
 - 한국어 글자체 이미지 ai데이터사용
 - MSER으로 등록된 텍스트 이미지를 전처리(preprocessing)함
 - Tesseract OCR을 이용하여 image 인식하여 명예훼손/모욕에 관한 image 파일에 적힌 텍스트 검출
 - near neighbor analysis이용하여 해당 이미지의 텍스트를 post processing함 (이후 이미지와 위에서 분석한 텍스트를 동시 출력하여 사용자가 오타자를 수정할 수 있도록 함)
3. 텍스트형식으로 표현된 문장에서 키워드 추출
 - 문서요약 텍스트 데이터 사용
 - Scikit-learn을 이용하여 TF-IDF 모델링 수행
 - textRank알고리즘 사용해서 word/sentance graph 기반 모형 제작
4. 법률 데이터를 활용하여 판례 불러오기
 - Doc2Vec모델 이용
 - dbow과 dm 방법으로 문맥 학습



※ 아이디어를 실현하기 위해 인공지능 학습용 데이터의 라벨링 및 학습방법을 기술해주세요. 본 아이디어의 실현가능성을 확인하기 위해 활용됩니다. (2,000자 이내)

※ 필요한 경우 분석 알고리즘에 대해 기술해주세요.