

# AI를 활용한 계약서 작성 자동화

중간발표: 모델링 및 설계

(주)빅뷰 장영민 대표  
김민석 · 김민수 · 박영찬 · 윤서영

**1**  
수행 배경 및 목표

**2**  
전체 시스템 구조

**3**  
시스템 모델 및 설계

**4**  
진행상황

**5**  
이슈사항 및 해결방안

**6**  
향후 일정

# 수행 배경 및 목표

## 문제 인식

- 구두 계약의 효력
- 계약서 작성의 필요성

## 프로젝트 목표

- 대화 음성 → 계약서 초안 생성
- 생성된 계약서 관리 가능 웹 서비스

1

수행 배경 및 목표

2

전체 시스템 구조

3

시스템 모델 및 설계

4

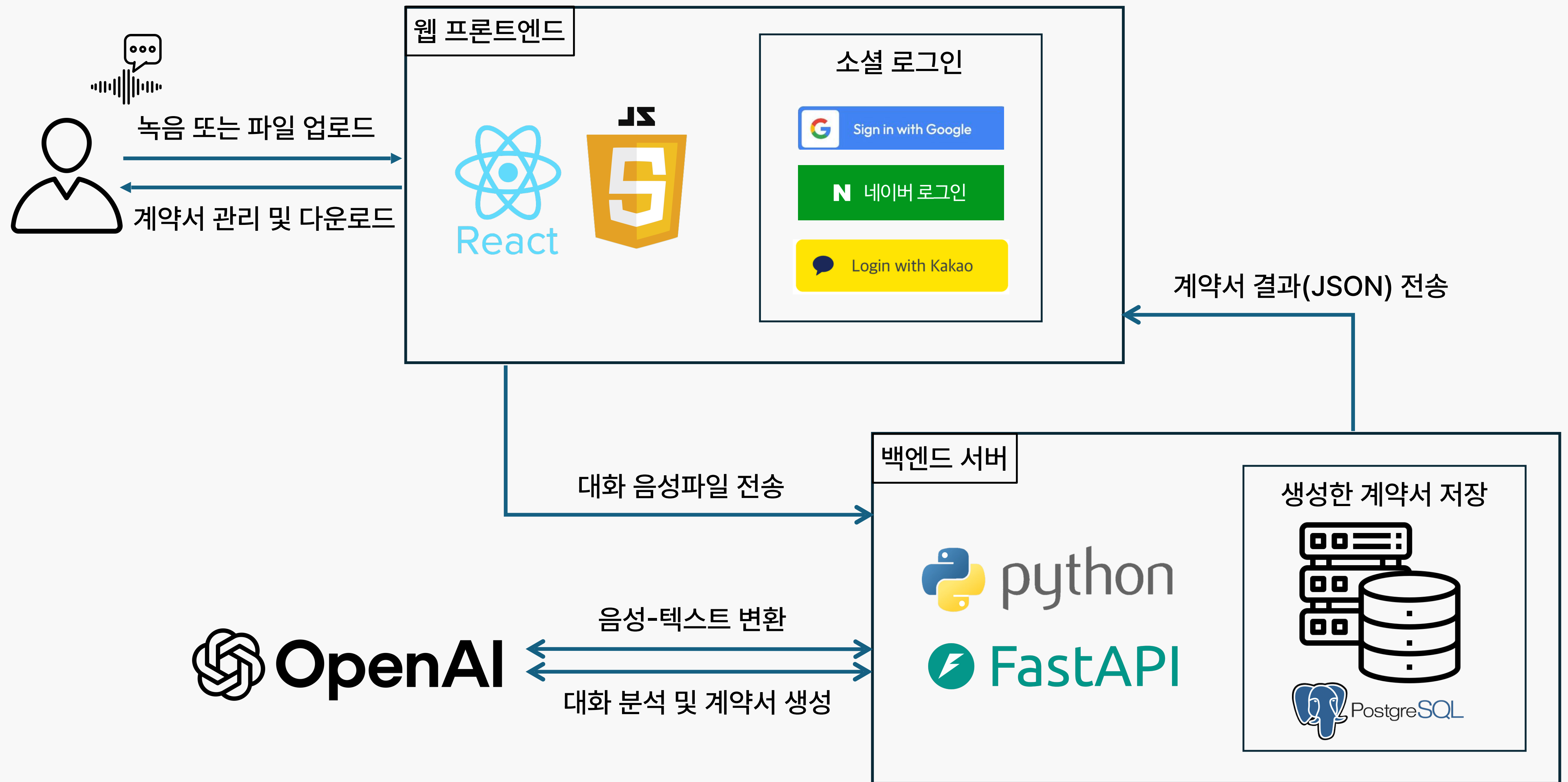
진행상황

5

이슈사항 및 해결방안

6

향후 일정



1

수행 배경 및 목표

2

전체 시스템

3

시스템 모델 및 설계

4

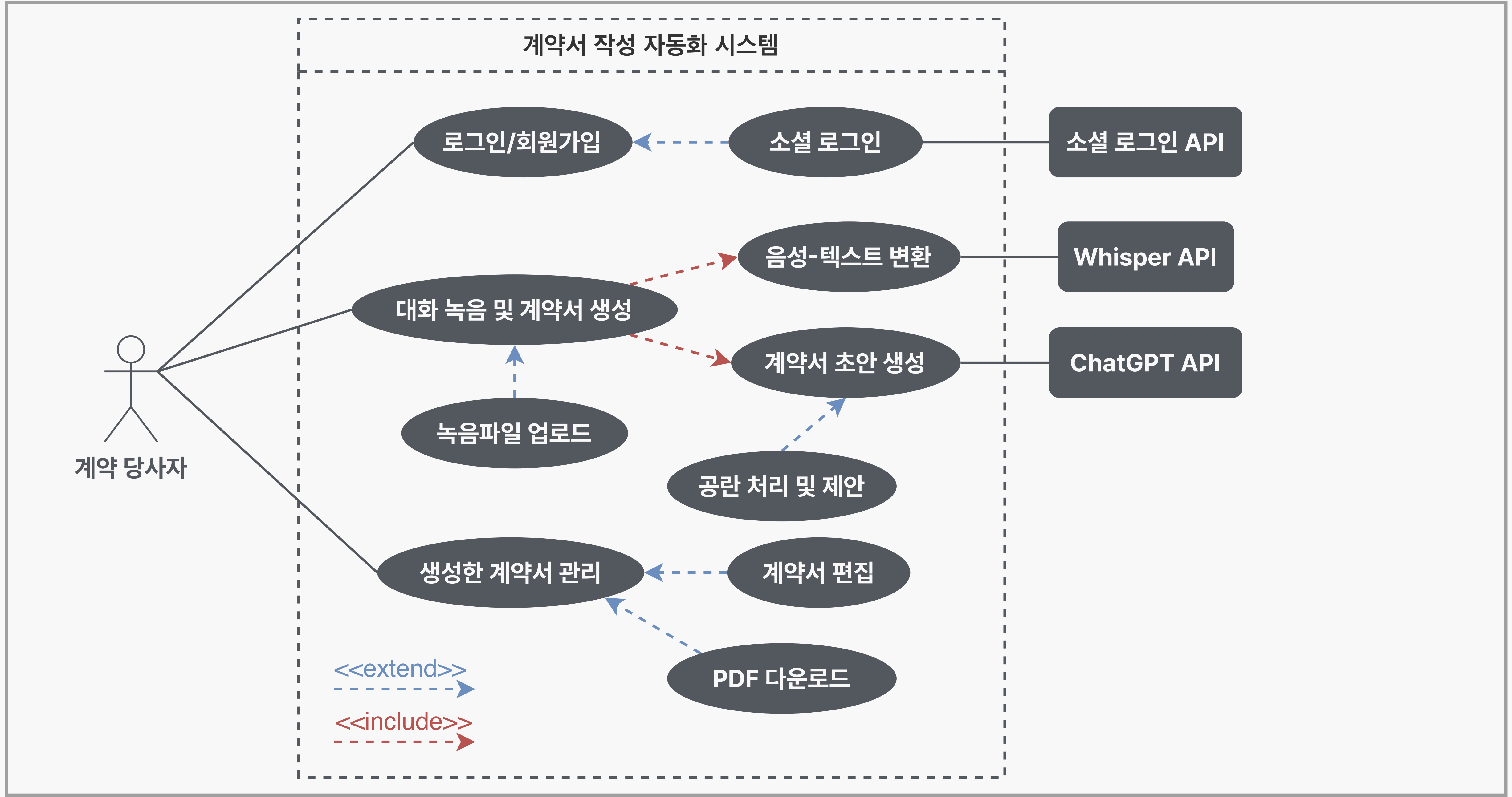
진행상황

5

이슈사항 및 해결방안

6

향후 일정



# 공란 처리 및 제안?

## 법률적 배경

- 계약 수단은 자유
- 유형 및 필수 요소는 정형화된 경우 많음
- 필수 요소 누락 시 무효

## 기능 상세

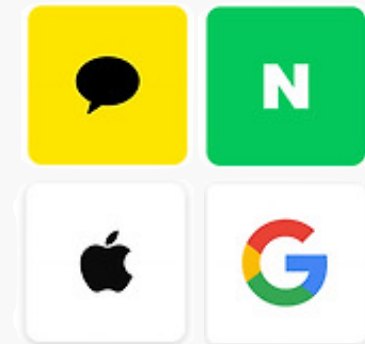
- 대화에서 필수 요소 누락 시 공란 처리 및 관련 정보 제안
- 예: "2025년 최저임금은 시간당 10,030원입니다" 표시





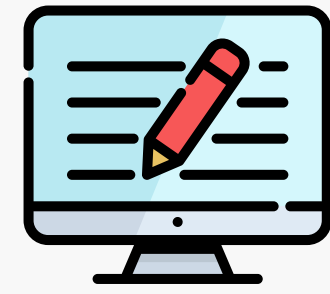
계약 당사자

2. 사용자 요청



소셜 로그인 API

1. 로그인

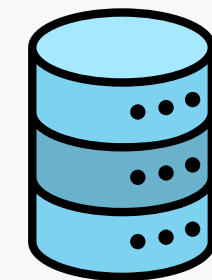


웹 프론트엔드

1. 로그인

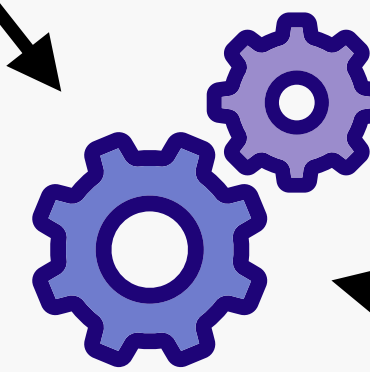


사용자 인증 DB



계약서 DB

6. 계약서 저장 및 관리



API 서버

4. 텍스트 전처리



텍스트 처리 모듈



프롬프트 생성 모듈

3. 음성-텍스트 변환



Whisper API

5. 계약서 생성



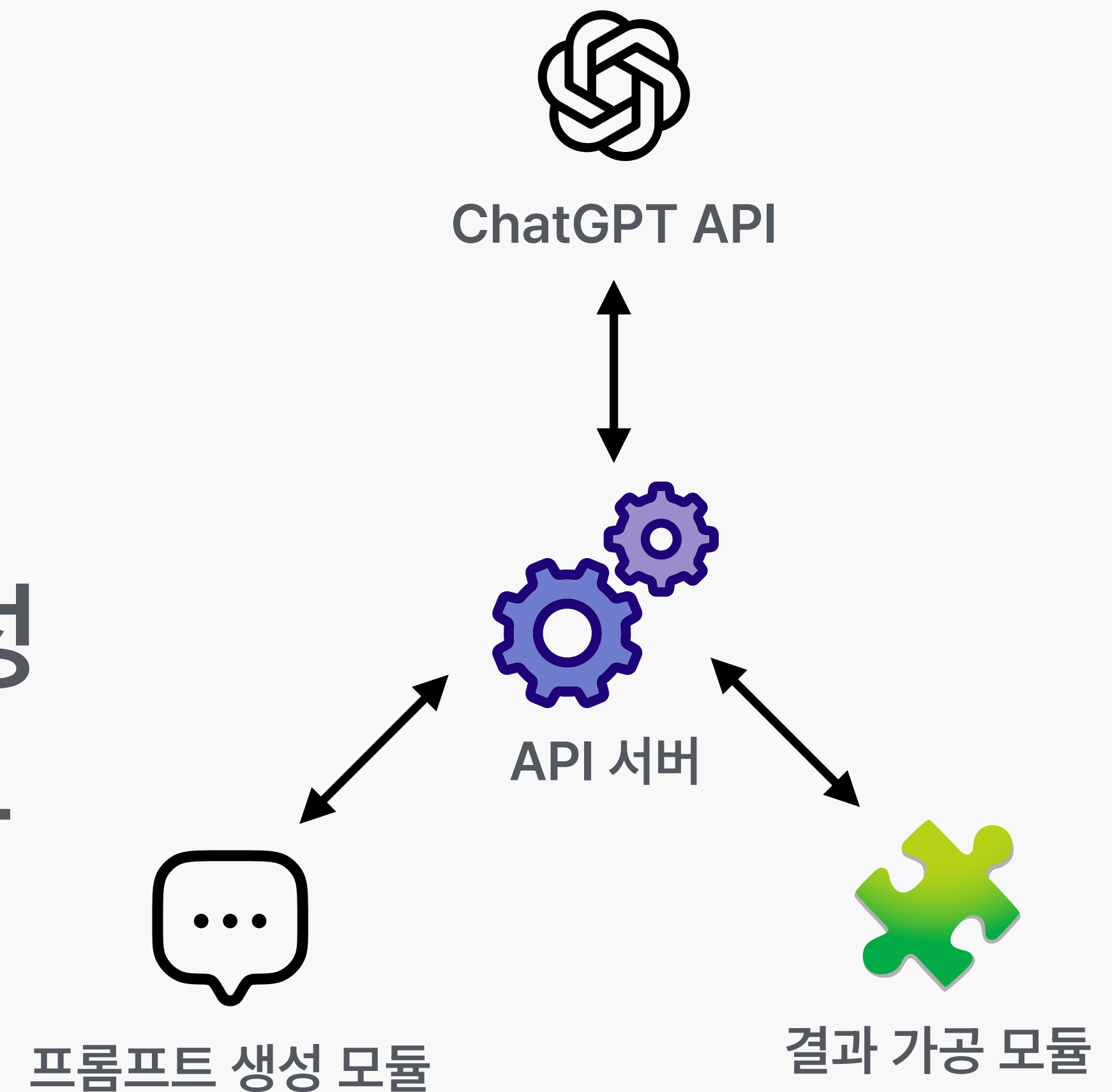
ChatGPT API



결과 가공 모듈

## 5. 계약서 생성

- 1) 텍스트 분석 및 계약 유형 판단
- 2) 키워드 추출 및 계약서 초안 생성
- 3) 누락 항목에 대한 법률정보 제안



1

수행 배경 및 목표

2

전체 시스템 구조

3

시스템 모델 및 설계

4

진행상황

5

이슈사항 및 해결방안

6

향후 일정

# AI (Prompt Engineering)

## 데이터 전처리 라이브러리

### 음성 녹음 파일 → 텍스트 변환 → GPT API에 입력

이 과정에서 녹음파일을 변환한 텍스트 파일을 그대로 GPT API에 입력으로 주면 불필요한 단어(추임새, 감탄사 등), 공백 을 제거하지 못한 채로 입력되기 때문에 토큰 최적화나 결과의 정확도 측면에서 문제가 있을수 있다.

따라서 이 텍스트 파일을 전처리 해주는 과정이 필요하다.

우리가 사용할 Whisper가 출력하는 결과값에는 다음과 같은 문제가 발생 할 수 있다.

1. 한국어 띄어쓰기 미흡
2. 맞춤법 수정 필요
3. 화자 분리가 안됨
4. 전문용어(잘 사용하지 않는 단어)를 캐치 못할 수 있음
5. 불필요한 단어

1번과 2번, 3번은 데이터 전처리 라이브러리를 통해 해결

### 1번과 2번의 경우 사용가능한 라이브러리

- 한국어 경우
  - `py-hanspell` - 네이버 맞춤법 검사기를 활용

## 데이터셋

### ▶ STT 학습/검증

### ▼ 계약서 텍스트 및 키워드 추출

프롬프트 엔지니어링을 진행한 후에, 정확도 향상 및 개선을 위한 용도로 사용될 것 같음

#### COKE (Contract Knowledge Extraction)

- 내용: 계약서에서 추출한 의무, 권리, 조건 등의 키워드.
- 출처: [COKE Dataset Paper](#) (코드 공개).

### ▼ 템플릿 작성

#### CUAD (Contract Understanding Atticus Dataset)

- 내용: 500+ 개의 실제 계약서 (영어)에서 추출한 13,000+ 개의 질의응답(QA) 항목.
  - 계약서 조항 분석, 키워드 추출, 문장 생성에 적합.
- 출처: [Hugging Face CUAD](#) | 공식 사이트
- 용도:
  - 계약서 특정 조항(예: "해지 조건은 무엇인가?")을 찾는 QA 모델 학습.
  - 프롬프트 엔지니어링에 활용 가능.

# AI (Prompt Engineering)

## 계약서 규격화를 위한 조사

민법으로 정해진 15종의 전형계약 유형

### ▼ 매매

#### ▶ 여행계약

#### ▶ 현상광고

### ▼ 위임

# 키워드

# 당사자 정보

# 위임의 내용

- 행위의 범위와 한계를 명확히

# 위임의 기간

- 종료시점이 불분명하면 "본계

# 보수 유무

- 거의 대부분 유상으로 진행됨

- 유상의 경우 금액, 지급 시

# 위임장 제출여부

# 해지 및 손해배상 조항

# 특약사항

### 정형화 가능성(O)

매매는 당사자 일방이 재산을 상대방에게 이전할 것을 약정하고 상대방  
함으로써 그 효력이 생긴다(민법 제563조). 매매의 일방예약은 상대방이  
는 때에 매매의 효력이 생긴다(민법 제564조).

당사자들 간 합의가 '매매계약(특히 채무이행에 조건이나 기한을 붙인 매  
체결, 성립시키는 과정이 별도로 예정되어 있는 채권계약)' 중 어떤 것을 드

### 계약서 사항

계약당사자가 자유롭게 매매계약의 내용을 정할 수 있으므로 매매계약서에 정해진 양식은 없음→ [서울지방법원-양식-생활속의계약서](#) 를 참고할 수 있음.

부동산 중개업체를 통해 매매계약서를 작성할 시에는 추가 사항이 기재되어야 함.

항목	내용	상세
contract_type	계약 유형	매매 계약
property	부동산 표시	소재지, 지목, 면적, 건물내역
seller	매도인 정보	(or대리인) 이름, 주민번호, 주소, 연락처
buyer	매수자 정보	(or대리인) 이름, 주민번호, 주소, 연락처
sale_price	매매대금	매매대금(총액, 계약금, 중도금, 잔금), 지급날짜
ownership_transfer	소유권 이전, 인도	필요 서류 전달 날짜, 인도 날짜

### ▼ 프롬프트 설계 키워드

```
{
  "contract_type": "근로계약서",
  "employer": {
    "company_name": "",
    "representative_name": "",
    "address": "",
    "contact": ""
  },
  "employee": {
    "name": "",
    "resident_number": "",
    "address": ""
  }
}
```

### ▶ 프롬프트 설계 키워드

→ 키워드가 많고 다양하다 보니, 효율적인 데이터 처리 및 전송 방안을 고려해야 할 것 같음.

공통 유형: 계약 유형, 계약 당사자(갑, 을) 정보, 계약 설명, 계약 날짜, 서명

### ▼ 템플릿 구조

- 근로 계약서 예시

{{contract\_type}}

사업주(회사명) : {{contractee}}

대표자 성명 : {{representative\_name}}

사업장 주소 : {{company\_address}}

연락처 : {{company\_contact}}

근로자 성명 : {{contractor}}

주민등록번호 : {{employee\_resident\_number}}

주소 : {{employee\_address}}


연락처 : {{employee\_contact}}


1. 근로조건

- 직책 : {{position}}

# Frontend

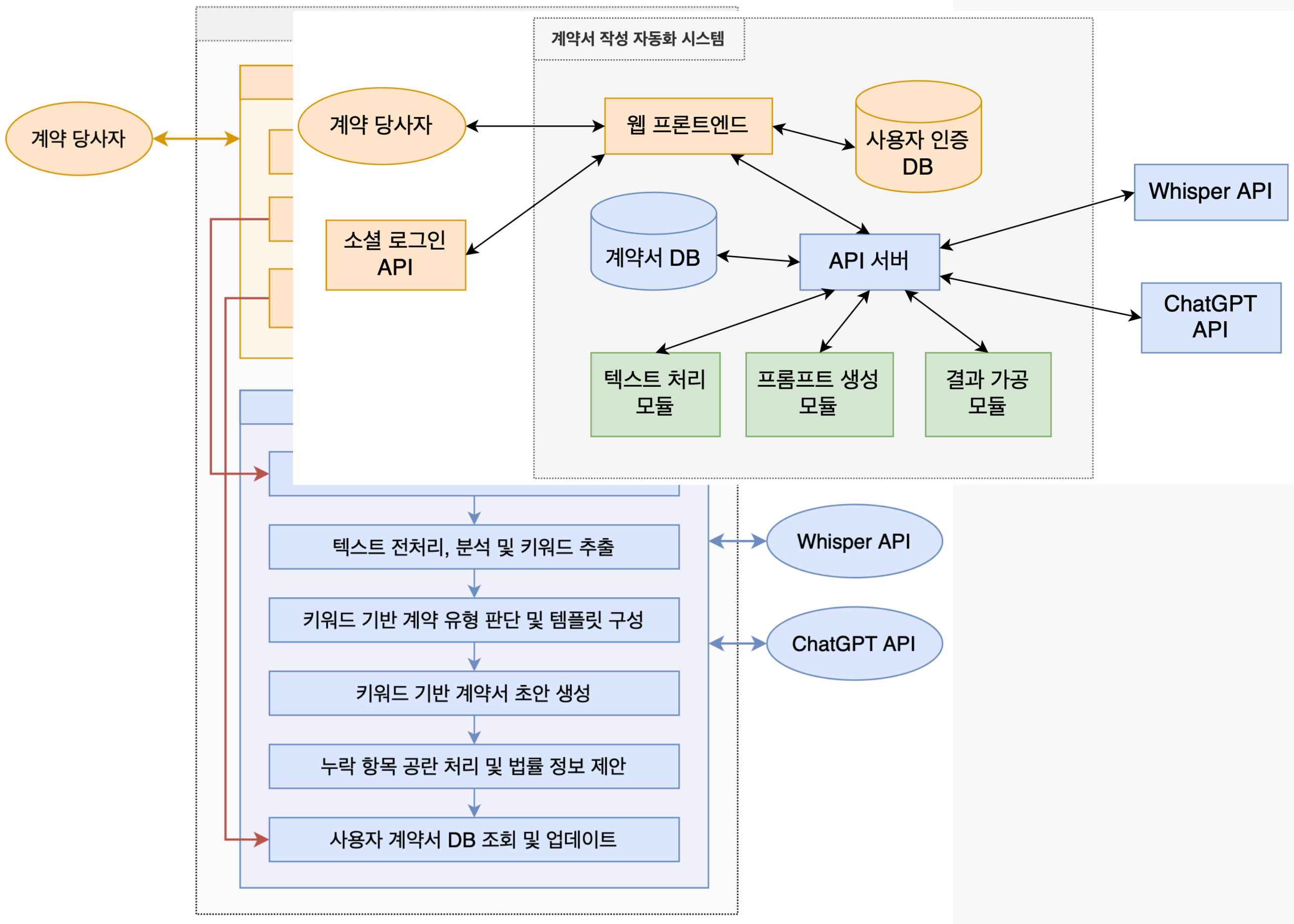
## 2. 화면(보고서) 설계

화면(보고서) ID	SC-TD3-001	화면(보고서) 명	로그인
화면(보고서) 경로	/login		
화면(보고서) 구성			
주요 항목 설명	- 주요 표시 항목 로그인 화면에서는 사용자가 ID와 비밀번호를 입력하여 시스템에 접속, 회원가입 버튼을 통해 신규 사용자가 계정을 생성, 비밀번호 찾기 기능을 제공하여 사용자가 비밀번호를 잊었을 경우 쉽게 복구 가능		
주요 체크 사항	- 로그인 기능 정상 작동 확인 올바른 ID와 비밀번호 입력 시 정상적으로 로그인되는지 점검 - 보안 및 접근성 점검		

화면(보고서) 경로	/Recording
화면(보고서) 구성	
주요 항목 설명	- 주요 표시 항목 실시간 음성 녹음을 진행 중일 때 사용자에게 상태를 안내하는 화면. 마이크 아이콘과 타이머, "듣는 중..." 문구를 통해 직관적으로 상태를 전달. 단순한 구성으로 집중도를 높임.
주요 체크 사항	- 녹음 중 사용자 조작(중단/취소) 가능 여부 및 버튼 상태 확인 뒤로가기, ESC, Stop 버튼 등 입력 시 적절히 반응하는지 - 음성 인식 로직과 실제 녹음 타이밍 간의 오차 또는 딜레이 여부 점검 사용자 입장에서 실제 녹음된 시간과 타이머 표시가 일치하는지 확인 - 네트워크 오류 발생 시 예외 처리 유무 확인 녹음 중 인터넷 연결이 끊겼을 경우 사용자에게 안내되는지



# Backend



## 보안 적용 범위 구분

연결 대상	예시	보안 필요 여부	적용 이유
외부 API	Whisper, ChatGPT	필요	민감 데이터가 외부로 전송되므로 TLS 전송, 토큰 보호, 요청 로그 차단 등 필요
DB	사용자 인증, 계약서 DB	필요	저장되는 데이터가 민감하므로 접근 제어, 암호화 저장, 인증 필요
내부 모듈	프롬프트 생성, 결과 가공 등	조건부 필요	민감 데이터를 처리하거나 로그로 남길 경우, 내부 모듈도 보안 고려 대상

## API 서버 내 하위 모듈(FastAPI + Python)

- 단순 Python 함수 or 클래스 기반 호출의 경우
  - 예: API 서버에서 `generate_prompt(text)` 호출
  - 보안 처리는 별도 필요 없음. 단일 프로세스 내 함수 호출임
  - 로그 기록, 예외 처리, 캐시 등에는 민감한 정보가 포함될 가능성
- 모듈을 FastAPI의 라우터나 백그라운드 태스크로 분리 구성
  - 예: `@router.post("/generate-prompt")` 형태로 나뉘거나, `async background_task()` 처리
  - 이 경우 API 서버와 동일 서버 내더라도 '모듈 간 통신' 형태가 되므로,
    - 요청 파라미터 검사
    - 민감 데이터 로깅 제한
    - 인증 헤더 검사 등 기본 보안 절차가 필요할 수 있음

1

수행 배경 및 목표

2

전체 시스템 구조

3

시스템 모델 및 설계

4

진행상황

5

이슈사항 및 해결방안

6

향후 일정



# 요구사항 변동

## 구현 범위 확장

- 로그인 및 사용자 기능 추가

## 대응 상황

- DB/보안 추가 설계 및 시스템 재구성 중
- 계획 검토, 설계 기간 조정

1

수행 배경 및 목표

2

전체 시스템 구조

3

시스템 모델 및 설계

4

진행상황

5

이슈사항 및 해결방안

6

향후 일정

# 향후 일정

**~4.13 (~4.23)**

- 설계 마무리 (API 명세, DB 설계, 필요한 모듈 함수 정의)

**~5.11**

- 본격적인 기능 구현, 시스템 통합

**~5.25**

- 테스트 및 성능 개선, 배포 작업