

SK네트웍스 Family AI 과정 12기

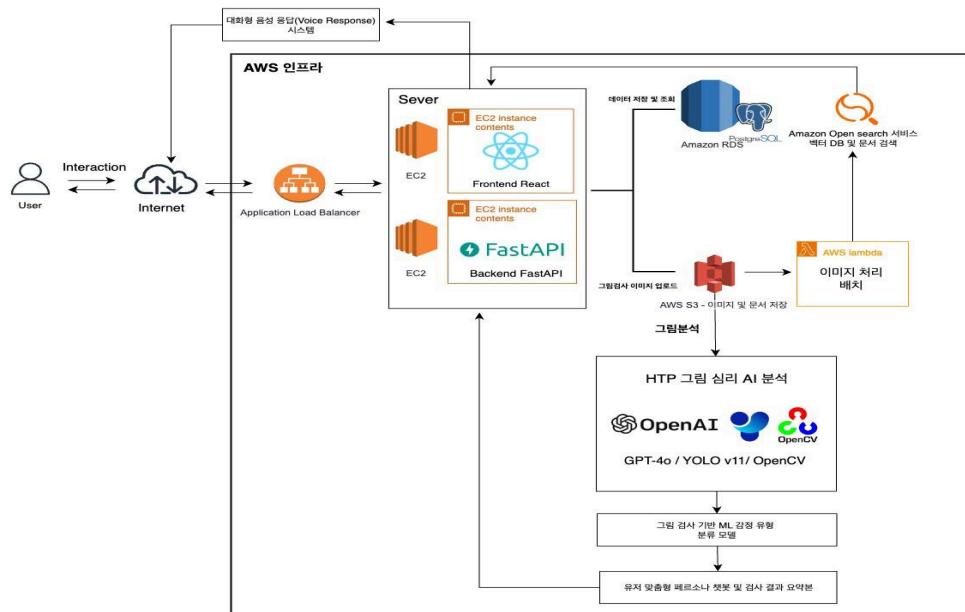
## 모델링 및 평가 시스템 아키텍처

산출물 단계	모델링 및 평가
평가 산출물	시스템 아키텍처
제출 일자	2025.07.20
깃허브 경로	<a href="https://github.com/SKNETWORKS-FAMILY-AICAMP/SKN12-FINAL-6TEAM">https://github.com/SKNETWORKS-FAMILY-AICAMP/SKN12-FINAL-6TEAM</a>
작성 팀원	김승학

### 1. 시스템 개요

본 시스템은 심리 유형 기반 챗봇 서비스를 제공하기 위해 AWS 클라우드 인프라와 Open AI, Yolo v11 등 AI 기술을 결합하여 설계하였다. 사용자는 React 기반 프론트엔드를 통해 HTP(House-Tree-Person) 그림 이미지 업로드 및 분석을 통해 자신의 심리 유형을 진단 받고, 그 결과에 따라 유형에 맞는 심리 챗봇과의 대화형 인터랙션을 수행하며, 백엔드에서는 FastAPI 중심으로 AI모델 연동, 문서 임베딩 및 검색, 음성 응답 기능을 처리한다. 본 시스템은 안정성과 확장성, 보안을 고려한 AWS 서비스 조합으로 구성되어 있다.

### 2. 시스템 구성도



본 시스템은 **AWS** 인프라를 기반으로 하여 **React** 프론트엔트와 **FastAPI** 백엔드, 그리고 다양한 **AI** 기술을 유기적으로 결합한 구조이다. 주요 구성 요소와 흐름은 다음과 같다.

## 1. 사용자 인터랙션 및 프론트엔드

- 사용자는 **React**로 구현된 프론트엔트를 통해 **HTP** 그림 이미지를 업로드 하고 분석 결과 기반 개개인 맞춤형 챗봇과 대화를 시작한다.
- 프론트엔드는 사용자와의 상호작용을 처리하며, **AWS** 인프라 내 여러 서비스와 통신한다.

## 2. AWS 인프라 내 주요 서비스

### 2.1 Application Load Balancer (ALB)

- 프론트엔드에서 발생한 **API** 호출을 받아 **FastAPI** 컨테이너 요청을 분산 처리한다.
- 트래픽 관리와 **SSL** 종료 역할을 수행하여 안정적인 서비스 운영을 지원한다.

### 2.2 AWS S3

- 사용자가 업로드한 그림 검사 이미지와 문서 파일을 저장하는 객체 스토리지다.
- 프론트엔드와 백엔드 모두에서 접근 가능하며, 대용량 데이터 저장에 적합하다.
- 또한 챗봇 기능 중 하나인 **TTS(Text-to-Speech)**에서 생성된 음성 파일도 **S3**에 업로드하여 저장하고 관리한다.

### 2.3 EC2(FastAPI)

- **FastAPI** 기반 백엔드 서버가 컨테이너 형태로 실행된다.

- OpenAI GPT-4, YoLo v11 등 AI 모델과 연동하여 이미지 분석, 심리 유형 검사, 챗봇 대화 로직을 수행한다.
- 대화형 음성 응답 시스템(Voice Response)과 ML 기반 감정 유형 분류 모델도 이곳에서 작동 된다.

2.4 Amazon RDS(PostgreSQL)

- 사용자 정보, 세션, 메타데이터 등 관계형 데이터를 저장한다.
- FastAPI 백엔드와 연동되어 데이터의 안정적 관리와 조회를 지원한다.

2.5 Amazon OpenSearch Service

- 심리 유형 키워드를 기반으로 벡터 DB에 임베딩된 문서를 저장하고 검색하는 역할을 한다.
- FastAPI가 OpenAI 결과와 연계하여 관련 문서를 빠르게 검색할 수 있도록 지원한다.

3. 데이터 및 처리 흐름 요약

1. 사용자가 React 프론트엔드에서 HTTP 그림 이미지를 업로드한다.
2. 이미지는 AWS S3에 저장되고, 인증은 Amazon Congnito를 통해 처리된다.
3. FastAPI 백엔드는 ALB를 통해 요청을 받아 OpenAI GPT-4, Yolo v11, ML 등 AI 모델과 연동하여 이미지 분석과 심리 유형 분류를 수행한다.
4. 분석 결과를 바탕으로 심리 유형 키워드 기반 문서 검색을 Amazon OpenSearch Service에서 실행한다.
5. 챗봇 대화 및 대화형 음성 응답 시스템이 사용자에게 맞춤형 상담과 음성 피드백을 제공한다.
6. 모든 사용자 데이터와 메타데이터는 Amazon RDS에 안전하게 저장 및 관리된다.

3. 주요 컴포넌트 설명

컴포넌트명	역할 및 기능
Frontend(React)	사용자 인터페이스 제공, 이미지 업로드 및 챗봇 대화 인터랙션 처리
ALB	API 요청 분산처리, SSL 종료 및 트래픽 관리
AWS Fargate	백엔드 API 서버, AI 모델 연동, 심리 유형 검사 및 이미지 분석 수행

(FastAPI)	
AWS S3	이미지 및 문서 저장소, 프론트엔드와 백엔드 간 파일 공유
Amazon Cognito	사용자 인증 및 권한 관리, 토큰 발급 및 검증
Amazon RDS	사용자 정보, 세션, 메타데이터 등 관계형 데이터 저장
Amazon OpenSearch	벡터 DB 역할, 임베딩된 문서 저장 및 심리 유형 키워드 기반 문서 검색 지원
문서 사전처리 및 임베딩	감성 유형별 문서 <b>Chunking</b> 및 임베딩 생성, <b>OpenSearch</b> 에 저장
대화형 음성 응답 시스템	<b>GPT-4</b> 텍스트 응답을 음성으로 변환 하는 <b>TTS</b> 기능 제공
ML 감성 유형 분류 모델	이미지 기반 <b>ML</b> 모델로 감성 유형 분류 수행

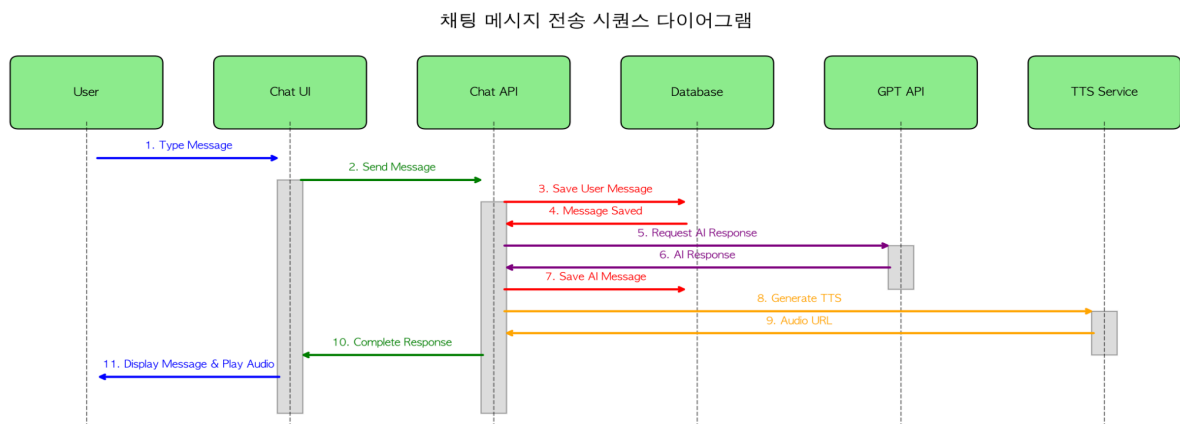
## 4. 시퀀스 다이어그램

### 1. 채팅 메시지 전송 시퀀스 다이어그램

개요

사용자가 채팅 메시지를 입력하고, AI 응답을 받아 음성으로 재생하는 전체

흐름



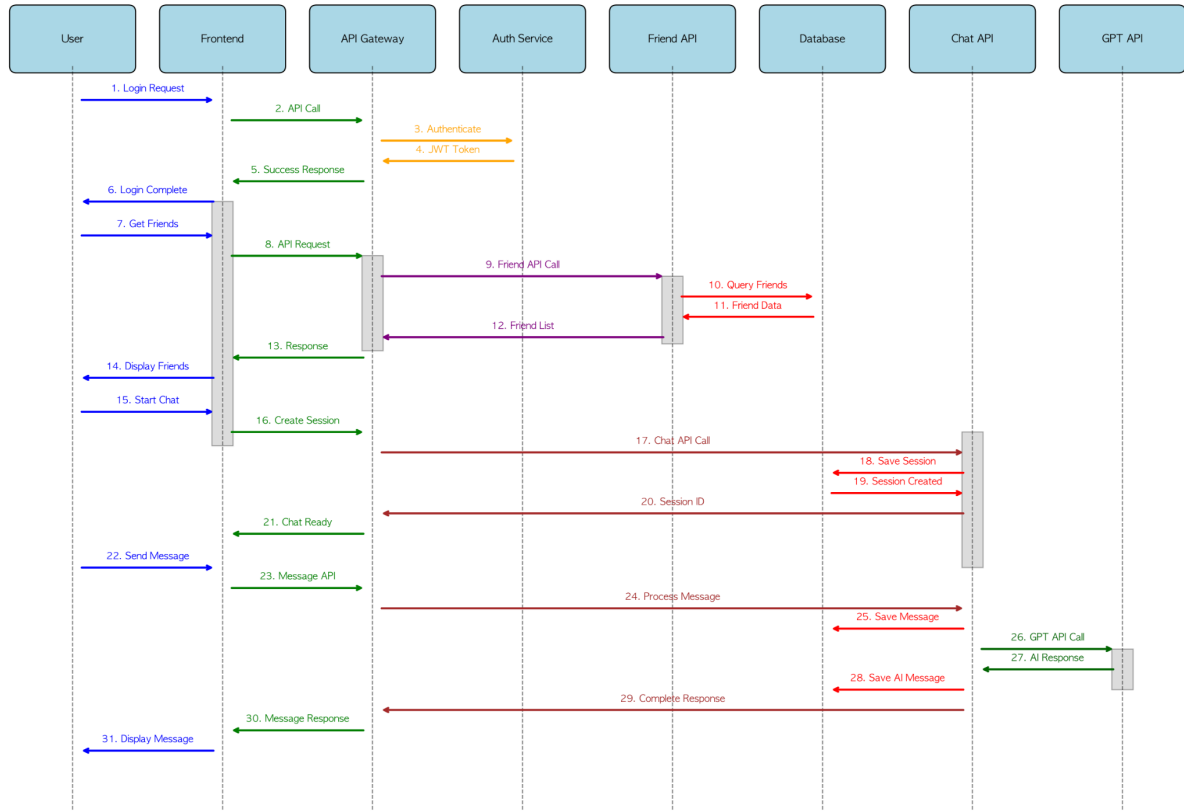
단계	주체	동작 내용	설명
1	User → Chat UI	메시지 입력	사용자가 채팅 메시지를 입력
2	Chat UI → Chat API	메시지 전송	UI에서 API로 메시지를 전송
3~4	Chat API ↔ Database	사용자 메시지 저장 및 확인	메시지를 DB에 저장하고 저장 완료를 확인
5~6	Chat API ↔ GPT API	AI 응답 요청 및 수신	GPT API에 메시지를 보내고 AI 응답 수신
7	Chat API → Database	AI 메시지 저장	AI가 생성한 메시지를 DB에 저장
8~9	Chat API ↔ TTS Service	TTS 생성 요청 및 오디오 URL 수신	텍스트를 음성으로 변환하고 URL을 수신
10	Chat API → Chat UI	응답 완료 알림	UI에 응답 완료를 알림
11	Chat UI → User	메시지 표시 및 오디오 재생	사용자에게 메시지와 음성을 출력

## 2. 심리상담 서비스 시퀀스 다이어그램

### 개요

사용자 로그인부터 유형별 챗봇(친구) 조회, 채팅 세션 생성, 메시지 전송 및 AI 응답까지 전체 상담 서비스 흐름

심리상담 서비스 시퀀스 다이어그램  
(사용자 로그인 → 친구 조회 → 채팅 세션 생성 → 메시지 전송)



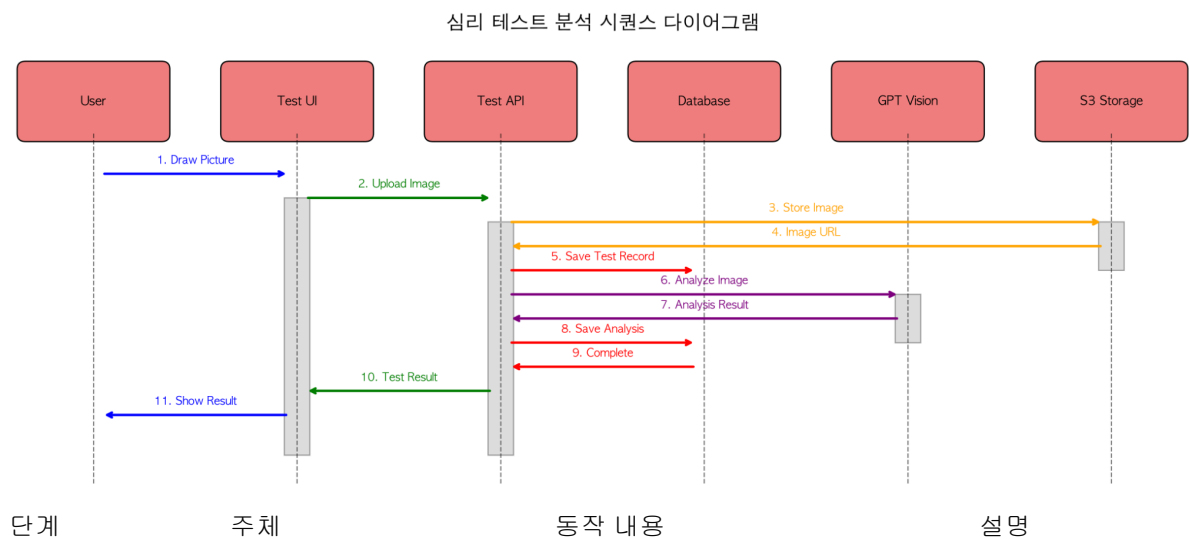
단계	주체	동작 내용	설명
1~6	User → Frontend	로그인 요청 및 완료	사용자가 로그인 요청 후 성공 응답 완료
7~14	User → Frontend → API Gateway → Friend API → Database	친구 목록 조회 및 표시	친구 API를 통해 친구 데이터를 조회하고 UI에 표시
15~20	User → Frontend → API Gateway → Chat API → Database	채팅 세션 생성 및 ID 수신	채팅 세션을 생성하고 세션 ID를 수신

21~23	User → Frontend → API Gateway → Chat API	메시지 전송 및 처리	사용자가 메시지를 보내고 API가 처리
24~29	Chat API ↔ Database, GPT API	메시지 저장, AI 호출, 응답 저장 및 완료	메시지 저장, GPT 호출, AI 응답 저장 후 완료 처리
30~31	API Gateway → Frontend → User	메시지 응답 및 표시	최종 메시지를 사용자에게 표시

### 3. 심리 테스트 분석 시퀀스 다이어그램

개요

사용자가 그림을 그리고, 이를 업로드하여 이미지 분석 후 결과를 받는 심리 테스트 분석 흐름



1~2	User → Test UI → Test API	그림 그리기 및 이미지 업로드	사용자가 그림을 그리고 API에 이미지를 전송
3~4	Test API → Database, S3 Storage	이미지 저장 및 URL 수신	이미지를 DB와 S3에 저장하고 URL을 수신

5~9	Test API ↔ Database, GPT Vision	테스트 기록 저장, 이미지 분석, 결과 저장 및 완료	이미지 분석 요청, 결과 수신, 저장 및 완료 처리
10~11	Test API → Test UI → User	테스트 결과 전송 및 표시	분석 결과를 UI에 전달하고 사용자에게 표시

## 요약 및 활용 방안

시퀀스 다이어그램들은 심리상담 서비스의 핵심 기능들이 어떻게 구현되고 상호작용하는지를 체계적으로 보여준다.

1. 명확한 책임 분리  
각 컴포넌트(사용자 인터페이스, **API** 게이트웨이, 인증 서비스, 데이터베이스, **AI** 서비스 등)가 고유의 역할을 담당하여 시스템 유지보수와 확장에 유리한 구조를 갖추고 있습니다. 이는 개발 및 운영 효율성을 높이고, 장애 발생 시 원인 파악과 대응을 용이하다.
2. 일관된 에러 처리 및 상태 관리  
모든 단계에서 메시지 저장, 세션 생성, **AI** 응답 처리 등 주요 작업에 대해 성공/실패 여부를 명확히 관리할 수 있도록 설계되어, 안정적인 서비스 운영이 가능합니다. 예외 처리 로직이 포함되면 사용자 경험 저하를 방지하고 신뢰도를 높일 수 있다.
3. 확장 가능한 구조  
**API** 중심의 모듈화된 설계 덕분에 새로운 기능 추가나 외부 서비스 연동이 용이합니다. 예를 들어, **TTS** 서비스나 **GPT API** 같은 외부 **AI** 서비스가 독립적으로 관리되며, 향후 다른 **AI** 모델이나 분석 도구로 교체 및 확장이 가능하다.
4. 사용자 중심 설계  
사용자 입력부터 결과 출력까지의 흐름이 직관적이고 빠르게 처리되도록 구성되어 있어, 실시간 응답성과 사용자 만족도를 높인다. 특히 음성 재생, 유형별 챗봇(친구) 조회, 심리 테스트 결과 표시 등 다양한 사용자 경험 요소가 반영되어 있습니다.

## 활용 방안

- 시스템 설계 및 개발 가이드  
각 시퀀스 다이어그램은 개발팀이 기능별 책임과 데이터 흐름을 명확히 이해하는 데 도움을 줍니다. 이를 기반으로 **API** 명세서 작성, 모듈별 개발 우선순위 설정, 테스트 케이스 설계가 가능하다.



- 운영 및 장애 대응  
단계별 상태와 데이터 저장 흐름이 명확하므로, 장애 발생 시 어느 단계에서 문제가 발생했는지 빠르게 진단할 수 있습니다. 또한, 로그 수집 및 모니터링 지점 설정에도 활용할 수 있다.
- 확장 및 신규 기능 추가  
기존 구조를 유지하면서 새로운 AI 모델, 분석 서비스, 사용자 인터페이스 개선 등을 유연하게 통합할 수 있습니다. 예를 들어, 심리 테스트 분석에 추가적인 이미지 처리 기능이나 상담 세션에 다중 사용자 지원 기능을 쉽게 추가할 수 있다.
- 사용자 경험 개선  
실시간 메시지 처리, 음성 변환, 결과 표시 등 사용자 인터랙션이 원활하도록 설계되어 있어, UI/UX 개선 시 참고 자료로 활용할 수 있다.