

## 하루 읽기 : 음성 일기 및 요약 서비스

팀장 : 18011513 이우식 컴퓨터공학과  
[백엔드 리드 개발자]



팀원 : 20011607 이준수 컴퓨터공학과  
[프론트 리드 개발자 및 백엔드 개발자]



팀원 : 21011702 황영희 컴퓨터공학과  
[인공지능 리드 개발자]



팀원 : 21011700 전서현 컴퓨터공학과  
[UI/UX 디자이너 및 인공지능 개발자]



### 1. 개발 배경 및 중요성

#### 1.1. 종이 기록의 한계점과 디지털 기록의 보편화

문자 발명 이후, 문자를 활용한 기록은 전달되는 정보에 대해 추측과 해석이 필요한 그림과 달리 정확한 정보 전달이 가능하게 되었다. 나아가 문자 기록을

위한 보존 매체로 종이가 발명되었고, 이는 정보의 대량 생산 및 양산화로 이어져 책이나 신문과 같은 종이 인쇄 기록물이 대중화 되었다. 그러나, 종이 기록물은 기록을 저장하는 것에 한계가 있으며 기록물을 관리하는 것에 대한 체계적인 시스템을 갖추는 것에 어려움이 존재한다. 이에 디지털 기술이 발전하면서 앞선 전통적인 기록 방법의 단점을 보완함과 동시에 시공간적 제약에서 벗어날 수 있는 디지털 기록이 등장하게 되었다. 2020년 마크로밀엠브레인 트렌드 모니터의 일상기록 관련 조사에 의하면 44.9%가 일상 경험을 기록하고 있음을 확인할 수 있다. 이처럼 모바일 기기의 보급률 증가에 따라 기기 내 탑재된 소프트웨어 형태의 기록 서비스가 보편화되고 있다.

## 1.2. 일상적 기록의 중요성

기록 플랫폼을 향한 소비자들의 관심은 꾸준히 증가하고 있으며, 실제로 네이버 블로그 서비스는 ‘주간 일기 챌린지’를 시작한 이후, 2019년 대비 월간 활성 사용자 수가 30% 증가했다. 네이버 블로그 서비스를 총괄하는 김승언 대표의 인터뷰에 따르면, 사용자들은 기존의 관계 중심적인 소셜 서비스와 달리, 블로그를 통해 보다 온전한 개인 공간을 경험하며 심리적 안정감을 느낀다고 한다. 또한, 블로그와 같은 디지털 공간에서 일상을 기록하고 그 결과물을 축적함으로써 스토리텔링을 경험한다고 본다. 개인의 가치와 행복을 중심으로 하는 라이프 스타일을 추구하는 현대 사회에서, 기술의 발전은 일상 기록을 위한 다양한 매체의 선택을 가능하게 하며, 이를 통해 나타나는 스토리텔링은 전통적 의미를 넘어서 기록자의 ‘자아’를 찾아가는 과정으로 이해될 수 있다. 일상 기록을 통한 자기 성찰은 자신의 가치를 주관적으로 평가하고, 문제의 본질을 탐구함으로써 개인적 성장으로 이어질 수 있다. 아래 표1은 기록의 긍정적 경험에 대한 내용이다.

기록의 긍정적 경험	
감정 해소	<ul style="list-style-type: none"> <li>- 기록은 감정적으로 의지 할 수 있고, 자유롭게 표출 할 수 있는 유일한 수단임.</li> <li>- 기록은 감정에 대해 객관적으로 바라볼 수 있도록 하고, 이를 통해 안정감을 느낌.</li> <li>- 기록을 통해 감정에 몰입함으로써 편안함을 느낌.</li> </ul>
회상	<ul style="list-style-type: none"> <li>- 일기를 한 장씩 읽어보며 과거를 회상하는 추억여행을 경험함.</li> <li>- 기록 흔적을 통해 호뭇해진 기억을 입체적으로 상기할 수 있음.</li> <li>- 반복되는 일상을 다른 관점으로 기억할 수 있는 경험임.</li> </ul>
자아표현감	<ul style="list-style-type: none"> <li>- 하루에 대한 증명을 통해 만족감과 성취감을 느낌.</li> <li>- 꾸준한 기록 활동과 하루일과의 달성도를 통해 뿌듯함을 느낌.</li> <li>- 기록 내용을 회고하며 삶에 대한 성찰적 태도를 갖게 되었음.</li> <li>- 세월의 흐름에 따라 바뀌어가는 나의 가치관이나 생각을 모아볼 수 있음.</li> </ul>

표 1. 기록의 긍정적 경험

### 1.3. 음성 기록의 필요성과 피드백의 효과

#### 1.3.1. 음성 일상 기록 서비스 사례 분석을 통한 서비스 필요성 도출

음성 기록 서비스에 대한 사용자 후기를 분석해보면, 기존에 일기 작성이나 지속적인 기록에 대한 번거로움과 시간 투자에 대해 회의적이었던 이들이 이제 긍정적인 반응을 보이고 있다는 것을 확인할 수 있다. '쓰기'를 중심으로 한 기록에 어려움을 겪는 사용자들을 위해, 음성 인식 기능을 제공하는 서비스는 텍스트 기록에 비해 사용성이 높고 신속한 기록을 가능하게 한다는 장점을 통해 일상 기록에 대한 부담을 줄일 수 있다. 더욱이, '회고' 기능은 사용자 경험에 있어서 반드시 필요한 요소로 자리 잡고 있다. 우리는 이 '회고' 기능에 키워드 도출, 사용자의 하루를 요약하는 기능, 사용자가 기록 당시 느꼈던 감정을 반영하는 기능, 음성 피드백을 제공하는 기능을 추가함으로써, 텍스트 중심의 일상 기록 방식에 대한 개선안을 제시하고자 한다.

#### 1.3.2. 음성일기를 통한 말하기 능력 신장

〈음성 일기가 말하기 능력 신장에 주는 효과에 관한 사례 연구 : A Case Study on Voice Journals as a Way of Improving Speaking Skill〉 논문에서는 중학생들을 대상으로 음성 일기를 제출하고 교사가 피드백을 해주는 식의 실험을 진행하였다. 실험(컴퓨터에 장착된 녹음 장치를 이용하여 일기를 녹음하고 이메일로 수신하는 방법으로 약 6개월간 40회 진행)의 과정은 [표2](#)와 같다. 이는 우리가 설계하고자 하는 서비스의 로직과 매우 유사하다. 이 실험의 결론은 다음과 같다.

1. 음성 일기를 내용 면으로 보면 일화, 인상, 행위, 견해를 포함하는 다양한 범위의 실제 말하기 활동이라는 것을 알 수 있다.

2. 환경의 영향을 받지 않으므로 심리적으로 긴장감이 적다. 따라서 자유로운 발화가 가능해짐으로써 유의미하고 생산적인 말하기 학습 기회를 얻을 수 있다.

3. 녹음 일기의 효과 중, 녹음 시간이 줄고 단어 숫자가 늘어난다는 결과를 확인했다. 주저(hesitation)와 휴지(pause)가 현저하게 줄고 유창성, 정확성, 어휘 사용 등의 말하기 능력이 상승한 것이라고 분석한다.

위 실험은 음성 일기가 학습 방법으로서의 가능성을 가지고 있음을 보여준다. 하지만, 이를 학교 현장에서 적용하기는 실질적으로 어려움이 있다. 이유는 주로 많은 학생들을 지도하는 교사 한 사람이 모든 학생의 피드백을 처리하기

어렵기 때문이다. 이 문제를 인식한 해당 논문은 피드백을 제공하는 주체를 교사뿐만 아니라 학생 본인이나 다른 학생까지 확장함으로써 수업에서 활용할 수 있는 여러 대안을 제안한다. 우리는 이러한 접근을 통해, 음성 일기가 말하기 능력 향상에 긍정적 영향을 미칠 수 있음을 확인하였으며, 현장 적용의 어려움을 극복하고 학생들의 말하기 능력을 개선할 수 있는 방안으로서 ‘하루 읽기 서비스’가 유용할 것으로 기대한다.

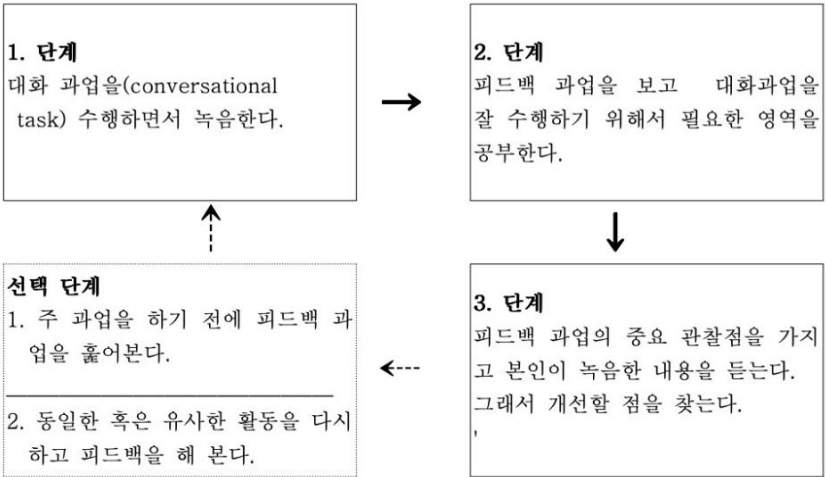


표 2. 음성 일기와 말하기 능력 간의 관계 도출을 위한 실험 과정

## 2. 개발 목표

본 서비스의 주요 개발 목표는 두 가지 핵심 기능에 초점을 맞추고 있다. 첫 번째는 사용자의 음성 데이터로부터 텍스트 데이터를 추출하고 이를 요약하는 기능이다. 이 과정을 통해 사용자는 자신의 일상이나 생각을 간결하게 정리된 형태로 볼 수 있으며, 이는 시간이 지나 회고할 때 중요한 사건이나 감정의 흐름을 쉽게 파악하는 데 도움이 된다.

두 번째 목표는 사용자가 선택한 감정과 발화 스타일을 고려하여 새로운 음성을 합성하는 기능이다. 이를 통해 제공되는 사용자 맞춤형 피드백은, 일기를 작성한 본인의 목소리가 아니라 다양한 감정과 스타일을 반영한 합성 음성으로 들을 수 있게 함으로써, 사용자가 자신의 기록을 보다 객관적이고 다양한 관점에서 평가할 수 있게 한다.

이 두 가지 기능은 사용자가 자신의 생각과 경험을 보다 효과적으로 정리하고, 자기 성찰과 개인적 성장을 지원하는 데 중점을 두고 있다. 요약 기능은 정보의 핵심을 파악하게 도와주고, 감정과 발화 스타일을 반영한 피드백은 일기 내용에 대한 새로운 시각과 깊이 있는 이해를 제공한다. 자세한 개발 목표는 다음과 같다.

1. 음성-텍스트 변환 기능 및 요약: 사용자가 녹음한 일기는 우선 STT 모델을 통해 텍스트로 변환된다. 이 과정에서 사용자는 텍스트로 변환될 것을 염두에 두지 않고, 마치 친구나 가족과 대화하듯 자유롭게 하루 동안 있었던 일들을 말한다. 이렇게 변환된 텍스트는 이후 GPT와 같은 모델을 통해 적절한 형태로 요약된다. 일기에서 선택된 감정은 그날의 전반적인 정서를 나타내며, 음성 합성에 사용되는 것 외에도 통계 분석을 통해 사용자가 한 달 동안 느낀 감정의 분포를 파악할 수 있게 해, 사용자가 자신의 대략적인 정서 상태를 이해하는데 도움을 준다.

2. 감정 및 발화 스타일을 동시에 고려하여 개인화된 음성 피드백 제공: 이 기능은 사용자가 경험한 감정 상태와 선호하는 발화 스타일을 기반으로 사용자에게 맞춤형 피드백을 생성한다. 사용자는 자신의 감정과 원하는 발화 스타일을 선택함으로써, 개인화된 음성 피드백을 받을 수 있다. 이 과정은 사용자의 선호와 감정 상태를 정밀하게 반영하여, 보다 깊이 있는 이해와 공감을 가능하게 하는 맞춤형 피드백을 제공한다.

3. 아카이브된 일기 접근: 지정된 날짜에 작성된 일기는, 원본 음성 데이터에서 추출되고 요약된 텍스트, 선택된 감정 상태, 그리고 개인화된 음성 피드백을 포함하여 보존된다. 이러한 과정을 통해, 사용자는 해당 일기를 검토함으로써, 당시를 회고할 수 있는 기회를 갖게 된다.

본 서비스는 세 가지 주요 기능을 통해 사용자의 일기 작성 및 회고 경험을 효과적으로 개선한다. 첫째, 사용자의 일기는 음성에서 텍스트로 변환되며, 이후 사용자의 감정 선택에 따라 요약되어 개인의 정서 상태 이해를 돕는다. 둘째, 사용자의 감정 상태와 발화 스타일을 기반으로 한 맞춤형 음성 피드백을 제공함으로써, 보다 깊이 있는 자기 성찰과 공감을 촉진한다. 마지막으로, 아카이브된 일기는 사용자가 언제든지 접근하여 과거의 자신을 회고할 수 있는 개인화된 기록을 보존한다. 이 세 가지 기능은 함께 작동하여 사용자에게 보다 풍부하고 다층적인 일기 경험을 제공한다. [그림 1](#)과 [그림2](#)는 본 서비스의 개발 목표를 설명한다.

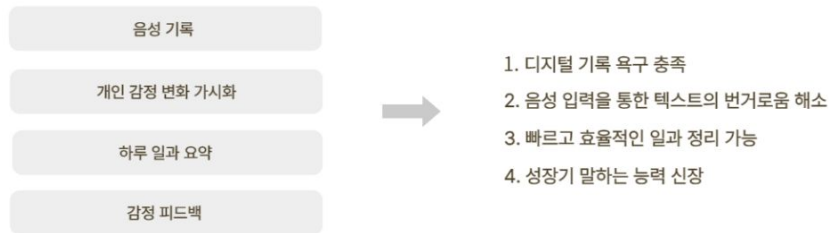


그림 1. 개발 목표

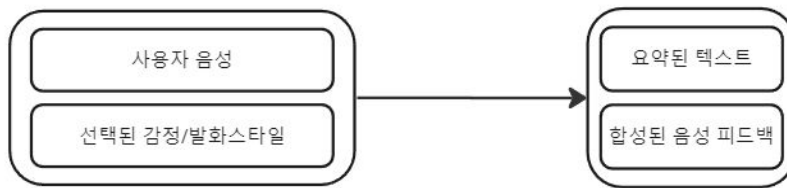


그림 2. 전체 시스템의 Input/Output

### 3. 차별성

#### 3.1. VUI 기술 추세

현재 시장에서 제공되는 다양한 음성 사용자 인터페이스(VUI) 기술들은 사용자의 음성 데이터를 분석하여 주파수와 크기를 파악하거나, 자연어 처리(NLP)를 이용해 음성 명령어를 해석하고 그에 맞는 작업을 수행하는 데 중점을 두고 있다. 대부분의 VUI 기술이 명령 실행이나 단순한 정보 검색에 초점을 맞추고 있는 반면, 본 서비스는 사용자의 일상적인 순간들이 담긴 음성 데이터가 복잡하고 다양한 감정과 사건으로 가득 차 있음에도 불구하고, 고급 자연어 처리(NLP) 기술을 활용하여 이야기의 핵심을 파악하고 요약하는 고도화된 프로세스를 제공한다. 특히, 사용자의 말이 문법적으로 완벽하지 않아도, 서비스는 의미 있는 정보를 효과적으로 추출하고 요약한다. 이는 단순한 음성 인식을 넘어서는 차별화된 접근 방식이며, 사용자가 자신의 경험과 감정을 음성으로 자유롭게 표현할 수 있게 하여, 이를 자동으로 텍스트화하여 요약 및 저장함으로써 디지털 시대의 일상 기록 및 보관 방식을 한 단계 더 발전시킨다.

또한, 본 서비스는 최근 MZ 세대를 포함한 다양한 사용자들 사이에서 자신의 일상을 기록하고 보관하는 행위의 중요성이 점점 더 강조되는 경향과 맞물려 있다. MZ 세대는 인스타그램 보관함, 페이스북 타임라인, 클라우드 기반 드라이브 등과 같은 다양한 디지털 플랫폼을 활용하여 자신의 생활을 기록하고, 나중에 이를 회상하는 것을 선호한다. 이러한 경향은 여러 논문과 연구에서도 잘 문서화되어 있으며, 개인의 경험과 추억을 디지털 아카이브 형태로 보존하는

것이 현대 사회에서 어떻게 중요한 문화적 현상으로 자리 잡았는지를 보여준다. 사용자는 자신의 경험과 감정을 음성으로 자유롭게 표현하고, 요약된 텍스트 일기 및 음성 피드백을 통해 그날의 주요 사건과 감정을 효과적으로 회상할 수 있으며, 원본 음성 데이터 보관 기능을 통해 과거의 자신과 재연결할 수 있는 기회를 갖게 된다.

요약하자면, 본 서비스는 사용자의 음성 기반 일상 기록을 자연스러운 텍스트 일기로 요약하고 피드백 음성을 제공하는 독창적인 방식을 통해, 개인화된 기록과 아카이브의 새로운 형태를 제공한다. 이는 현대 디지털 시대에 걸맞은 개인화된 기록 및 아카이브 솔루션으로, 사회의 다양한 구성원들이 자신들의 이야기를 남기고 공유할 수 있는 소중한 수단이 되며, 기존 VUI 시스템과의 명확한 차별점을 제공한다.

## 3.2. 경쟁 서비스 개요

### 3.2.1. '소곤소곤' 어플

음성 다이어리 "소곤소곤"은 사용자에게 그날의 감정을 선택하고 음성으로 일기를 녹음해 해당 녹음 파일을 날짜별로 정리하여 보관하는 기본적인 기능을 제공한다. [그림3](#)과 [그림4](#)은 '소곤소곤'의 사용자 화면이다. 이러한 방식은 사용자가 자신의 감정과 생각을 음성 데이터로 기록하고 시간 순으로 접근할 수 있게 하지만, 내용 분석이나 요약, 음성 피드백 제공 등의 추가적인 가치를 제공하지는 않는다.

반면, 본 서비스는 단순한 보관을 넘어서는 고급화된 접근 방식을 채택한다. 우선, 고급 자연어 처리(NLP) 기술을 활용하여 사용자의 음성 데이터를 분석한다. 이 과정에서 음성 데이터 내의 핵심 내용을 정확하게 식별하고, 그 내용을 기반으로 요약된 텍스트 형태의 일기를 생성한다. 이 요약 기능은 사용자가 자신이 녹음한 내용의 핵심을 빠르게 파악하고, 특정 날짜나 주제에 대한 기록을 쉽게 검색하고 재방문할 수 있게 만들어 준다.

본 서비스는 또한 사용자의 음성 데이터에서 추출된 정보를 텍스트화하고 이를 통해 생성된 일기에 감정 태그나 키워드 태그를 자동으로 부여하는 기능을 포함한다. 이는 사용자가 나중에 특정 감정이나 사건에 대한 기록을 검색할 때 매우 유용하다. 더 나아가, 이 텍스트화 과정은 사용자가 자신의 일기를 공유하거나 다른 디지털 형태로 전환하는 데도 용이하게 해준다.

결국, "소곤소곤"과 같은 기존 음성 다이어리 서비스가 제공하는 단순한 음성

기록 보관 기능과 달리, 본 서비스는 사용자의 음성 데이터를 보다 심도 있게 분석하고 그 안에서 가치 있는 정보를 추출해 내어, 사용자에게 보다 풍부하고 접근성 높은 일기 경험을 제공한다. 이러한 차별화된 기능은 사용자가 자신의 생각과 감정을 보다 깊이 있게 기록하고, 시간이 지나도 그 의미를 쉽게 회상할 수 있게 해준다.



그림 3. ‘소곤소곤’의 감정 선택 화면

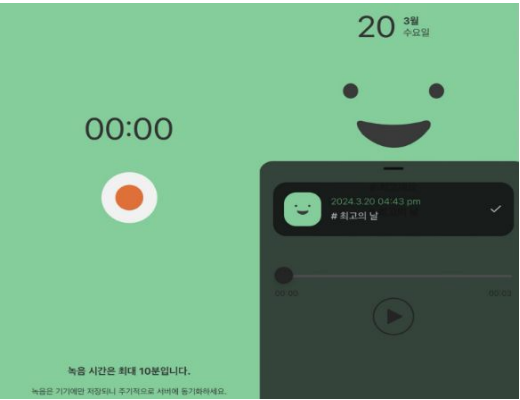


그림 4. ‘소곤소곤’의 음성 녹음 화면

## 4. 개발 방법 및 체계

팀은 체계적인 협업과 프로젝트 관리를 위해 Git과 GitHub를 활용하여, 코드 버전 관리와 협업을 수행할 예정이다. GitHub를 통해 개발 진행 상황을 공유하고, 이슈 트래킹을 통해 발생하는 문제들을 효과적으로 관리하며, 코드 리뷰를 통해 코드 품질을 지속적으로 개선한다. 팀원들은 정기적인 미팅을 통해 개발 진행 상황을 공유하고, 서로의 작업에 대한 피드백을 제공한다. 이를 통해, 개발 과정에서 발생할 수 있는 문제들을 신속하게 해결하고, 프로젝트 목표 달성을 위한 전략을 지속적으로 조정한다. 본 팀은 크게 프론트엔드, 인공지능, 백엔드의 세 주요 분야로 나누어져 각각의 서브 모듈과 팀원들의 역할이 구체적으로 정의되어 있으며, 내용은 다음과 같다.

### 4.1. 프론트엔드



프론트엔드 개발은 사용자가 직관적으로 시스템을 사용할 수 있도록 사용자 경험(UX)과 사용자 인터페이스(UI) 디자인에 중점을 둔다. 이를 위해 React나 Vue.js와 같은 현대적인 웹 프레임워크를 사용하여, Android나 iOS 등 플랫폼에 상관 없이 사용할 수 있는 웹 애플리케이션을 구축할 예정이다. 사용자 인터페이스는 일기 열람 및 음성 재생, 작성을 위한 녹음 등의 기능을 사용할 수 있도록 설계된다.

디자인적 요소에는 일기 열람 및 작성을 위한 캘린더, 음성 녹음 및 재생, 요약된 텍스트 표시 등을 위한 필드가 포함된다. 이러한 요소들은 사용자가 서비스가 제공하는 기능을 직관적으로 이해하고 사용할 수 있도록 고안된다.

프론트엔드는 UI 컴포넌트 개발, 사용자 입력 처리, 백엔드 서버와의 API 통신을 통한 데이터 교환 등을 담당한다. 이 과정에서 사용자 경험을 최적화하기 위한 다양한 기술적 접근이 적용된다. 또한, 사용자가 서비스를 이용하는 과정에서 발생할 수 있는 다양한 시나리오를 고려하여, 오류 처리 및 사용자 피드백 메커니즘도 중요한 디자인 요소로 포함된다. 이는 사용자가 서비스 사용 중 발생할 수 있는 문제를 쉽게 인지하고, 필요한 조치를 취할 수 있도록 돕는다.

최종적으로, 프론트엔드는 사용자에게 매력적이고 직관적인 사용자 경험을 제공함으로써, 사용자가 음성 일기를 녹음하고, 열람하는 일기의 요약된 텍스트와 합성된 음성을 쉽게 받아 볼 수 있도록 만든다.

## 4.2. 인공지능

인공지능 파트의 개발은 주로 사용자로부터 받은 요약된 텍스트 데이터를 바탕으로 감성과 발화 스타일을 고려한 맞춤형 음성 피드백을 생성하는 작업에 집중된다. 이 과정에서 수행해야 하는 주요 작업은 다음과 같다.

### 1. 모델 학습을 위한 데이터 준비 및 전처리

음성 데이터를 텍스트로 변환하는 작업은 백엔드에서 수행된 후, 이 텍스트 데이터는 요약되어 인공지능 시스템에 전달될 예정이다. 인공지능 시스템을 학습시키기 위해서는, 먼저 대량의 음성 데이터를 텍스트로 변환하여 이를 바탕으로 감성과 발화 스타일을 분석할 수 있는 데이터셋을 구축해야 한다. 이 데이터셋은 음성 합성 모델의 학습에 사용되며, 모델이 다양한 감성과 발화 스타일을 인식하고 반영할 수 있도록 한다.

### 2. 감성 및 발화 스타일 기반 음성 합성

인공지능 시스템은 백엔드로부터 전달받은 요약된 텍스트와 사용자가 선택한 감성 및 발화 스타일 정보를 기반으로, 사용자에게 제공할 맞춤형 음성 피드백을 생성한다. 이 과정에서는 멀티 모달 음성 합성 기술을 활용하여, 텍스트 내용과 감성 정보를 동시에 고려하여 음성을 생성할 예정이다. 멀티 모달 음성 합성 모듈은 사전에 정의된 감성과 발화 스타일 레이블을 갖는 풍부한 음성 데이터셋을 기반으로 딥러닝 모델을 훈련시켜, 백엔드에서 처리된 요약 텍스트와 사용자 지정 감성 데이터를 결합하여 음성 피드백을 생성한다. 이 과정은 텍스트-투-스피치(TTS) 기술과 감정 인식을 위한 조건부 생성 신경망을 활용할 예정이다. 음성 합성 모델의 예성 구조는 [그림 5](#)와 같다.

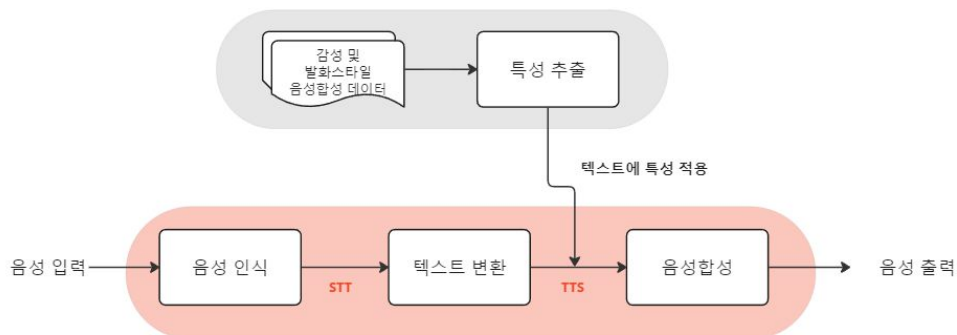


그림 5. 음성 합성 모델의 예성 구조

#### 4.3. 백엔드

백엔드 서버 개발은 사용자 음성 데이터와 관련 정보의 관리, 처리, 그리고 인공지능 서버와의 효과적인 통신을 중심으로 진행된다. 사용자로부터 수집된 음성 데이터와 상황 정보를 안전하게 저장하고, 이를 인공지능 서버로 전송해 음성 분석 및 합성의 기반 자료로 삼는다. 또한, ChatGPT API를 활용해 사용자 일기를 분석하는 기능도 백엔드에서 처리한다.

백엔드 개발은 현대적 웹 프레임워크의 대표 주자인 Python 언어의 Django 웹 프레임워크를 기반으로 구현할 예정이다. Django는 웹 개발에 있어서 다양하고 강력한 기본 라이브러리를 보유하고 있어 빠른 속도로 웹 서비스를 개발할 수 있으며, 확장성이 매우 뛰어나 인공지능 서비스와 연동하여 다양한 작업을 수행하기에 유리하다.

데이터베이스의 경우 MySQL을 사용할 예정이다. Django에도 기본적으로 제공하는 Sqlite라는 데이터베이스가 존재하고 활용 가치가 있기는 하지만, 데이터에 관계도가 중요해지거나 분량이 많아지면 운용하는데에 어려움이 있기에

MYSQL을 외부 연결할 예정이다. 많은 대형 어플리케이션들이 활용하는 MYSQL의 경우 많은 데이터가 복잡한 관계도로 들어와도 충분히 버틸 수 있는 데이터베이스이며, 웹 프레임워크나 외부 스토리지등 다양한 서비스와 연동하는 데에도 어려움이 없다.

서비스의 특성상 우리는 음성데이터 또한 저장해야 하는데, 이러한 파일 형식의 데이터를 데이터베이스에 저장하는 경우는 매우 비효율적이다. 따라서 백엔드는 음성 파일을 외부 저장소인 AWS의 S3에 저장할 계획이다. 해당 서비스는 유료 제공이긴 하지만, 일정 용량까지는 무료로 이용이 가능하다. 음성 파일이 매우 큰 용량을 차지하지는 않을 것이라 예상하여 해당 외부 스토리지를 이용하고자 한다. 프로젝트의 진행 상황에 따라 언제든 이용 가능한 용량을 추가로 이용할 수 있다는 점이 가장 큰 장점이다.

백엔드 시스템의 핵심은 사용자의 요청 처리, 데이터 관리 및 보안, 인공지능 서버와의 효율적인 데이터 교환을 위한 고도화된 시스템 구축에 있다. 이를 위해 최신의 백엔드 기술과 프레임워크를 활용하여, 사용자와 인공지능 서버 간의 매끄러운 데이터 흐름과 실시간 처리를 지원하는 안정적이고 확장 가능한 백엔드 아키텍처를 개발할 예정이다. 결과적으로, 이 백엔드 시스템은 인공지능 모델이 사용자에게 최적화된 음성 합성 결과를 제공하는데 필수적인 역할을 수행할 것이다.

#### 4.4. 전체적인 개발 구조

아래 그림 6은 사용자, 서버, AI 모델, 및 데이터베이스 간의 상호 작용을 명료하게 도식화한 시스템 아키텍처 다이어그램을 나타낸다. 이 구조에서 사용자는 일기의 음성 데이터와 함께 선택한 감정 및 발화 스타일 정보를 서버로 전송한다. 서버는 첫 단계로 STT(Speech-to-Text) 모델을 활용해 음성 데이터를 텍스트로 변환한 뒤, ChatGPT와 같은 자연어 처리 모델을 이용하여 해당 텍스트를 요약한다. 변환 및 요약된 텍스트는 이후에, 사용자가 지정한 감정과 발화 스타일을 반영하는 맞춤형 음성 피드백 합성을 위한 AI 모델에 의해 처리된다. 합성된 음성은 서버의 저장 공간에 보관되며, 데이터베이스는 합성된 음성이 어떤 사용자의 몇 월 며칠에 해당하는 일기의 음성인지, 어떤 감정 및 발화 스타일을 기반으로 합성되었는지, 그리고 원본 텍스트가 무엇인지 등의 정보를 관리한다. 사용자는 특정 날짜의 일기를 열람할 때, 이러한 상세 정보와 함께 합성된 음성을 통해 맞춤형 피드백을 받는다.

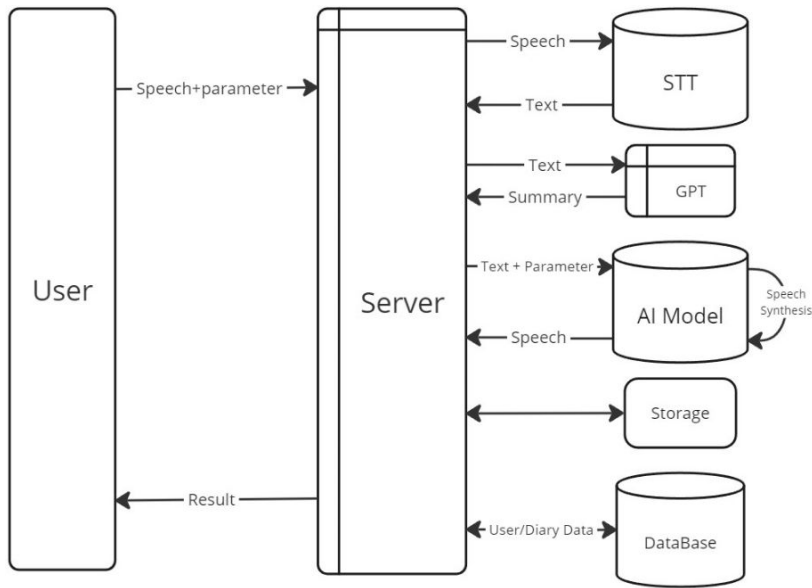


그림 6. 전체 시스템 아키텍처

## 5. 개발 추진 계획

초기에는 팀원들 간 회의를 통해 프로젝트가 나아가야 할 방향과 그를 이루기 위하여 각 파트에서 이루어야 하는 내용들을 이야기하고, 이후에는 각자 파트 별 개발을 하며 정기적으로 진행 상황을 공유하는 방향성을 가지고자 한다. 아래 표 3은 전체적인 개발 추진 예상 계획이다.

ID	Task	담당	Duration	주차													
				1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	AI 모델	전서현, 황영희															
1.1	논문 리뷰		2주														
1.2	요구 사항 분석 및 명세서 작성		1주														
1.3	프로토타입 설계		1주														
1.4	데이터 수집 및 전처리		2주														
1.5	모델 학습 환경 설정		1주														
1.6	모델 학습 및 평가		1주														
1.7	피드백 반영 및 모델 조정		1주														
1.8	최적화 및 성능 개선		1주														
1.9	백엔드 및 프론트엔드의 통합 테스트		1주														
1.10	최종 결과 및 발표		1주														
2	서버	이우식															
2.1	개발 환경 구성		1주														
2.2	모델 파일 생성		1주														
2.3	views 파일 및 url 설정		1주														
2.4	프론트 파일의 웹 연결		1주														
2.5	S3 스토리지 연동		1주														
2.6	AI 파트 서버 및 DB 연동		3주														
2.7	서버스 이용 내역 DB 기록		3주														
2.8	최종 구현 및 유지 보수		1주														
3	Database	이우식															
3.1	DB - 서버 연동		1주														
3.2	DB 구축 및 구조 확정		1주														
3.3	용량/연결 모니터링 및 유지보수		14주														
4	프론트엔드	이준수															
4.1	기초 디자인 및 계획		1주														
4.2	컴포넌트 개발		2주														
4.4	프론트엔드 기능 구현		2주														
4.5	백엔드 연동 및 테스트		1주														
4.6	사용성 테스트		1주														
4.7	최종 리뷰 및 비그 수정		4주														
5	기획																
5.1	요구사항 분석																
5.2	피드백 및 제안서 수정																

표 3. 전체적인 개발 추진 계획

## 6. 기대 효과

본 서비스는 사용자의 음성을 정확히 텍스트로 변환하고, 사용자가 선호하는 감성과 발화 스타일에 기반하여 음성 데이터를 처리하는 기술을 채택한다. 이러한 과정은 음성 데이터의 효율적인 텍스트화와 더불어 개인적 기록에 대한 세심한 피드백(예: 위로, 칭찬 등)을 제공한다. 이를 통해 사용자는 자신의 일상을 기록하고, 해당 기록에 대한 공감을 바탕으로 하루를 의미 있게 마무리할 수 있다. 또한, 사용자는 자신의 과거 기록을 언제든지 회고하고, 당시의 상황과 감정에 더욱 몰입하여 상기할 수 있는 기회를 갖는다. 장기적으로 서비스를 활용함에 따라, 사용자는 자신이 겪는 감정 변화의 전반적인 경향을 명확히 인식할 수 있게 되며, 이는 자아 인식의 향상과 개인 성장을 촉진한다. 본 서비스의 기대 효과는 구체적으로 다음과 같다.

1. 개인 일상 기록: 최근 20대를 중심으로 다이어리 작성이나 '다꾸'(다이어리 꾸미기)가 유행하고 있다. SNS 등을 통해 자신의 일상을 공유하며 받는 좋아요나 관심에 지친 많은 MZ세대들이 스스로의 일상을 통한 꾸준한 실천과 생산적인 삶을 추구하는 경향이 증가하고 있다. 그럼에도 불구하고, 바쁜 일상 속에서 글을 직접 적는 과정이 번거롭거나 귀찮아 생략되는 경우가 종종 있다. 이러한 상황에서 본 서비스는 사용자가 단순히 말로 자신의 이야기를 하고 감정을 선택하는 것만으로도 개인 기록의 욕구를 편리하게 충족시킬 수 있는 해결책을 제공한다.

2. 개인 감정 변화 가시화 : 본 서비스는 사용자가 단순히 일기를 작성하거나 음성 메시지를 남기는 것을 넘어, 그 날의 하루가 어땠는지에 대한 감정 상태를 직접 기록할 수 있는 기능을 제공한다. 이를 통해 사용자는 일주일 단위부터 한 달 이상의 기간에 걸쳐 자신이 경험한 시간과 그 시간 동안의 감정 변화를 한눈에 파악할 수 있다. 특히, 일상에서 벗어난 감정 상태가 기록되어 있다면, 해당 날짜에 어떤 특별한 사건이 발생했는지 쉽게 확인할 수 있다. 이러한 과정은 사용자가 자신의 감정 변화를 명확하게 인지하고, 필요한 경우 적절한 대응을 할 수 있게 도와준다.

3. 하루 요약 기능 : 개인 기록을 선호하는 이들 사이에서도, 자신이 친구에게 이야기하듯 말로 전한 긴 내용을 모두 다시 듣는 것은 번거로울 수 있다. 이러한 문제를 해결하기 위해, ChatGPT API를 활용하여 입력받은 내용을 간결하게 요약해 들을 수 있는 기능을 도입할 예정이다. 이 기능 덕분에 사용자는 단순한 녹음을 넘어서, 해당 날짜에 어떤 일이 있었는지를 신속하게 파악할 수 있다.

4. 개인 맞춤형 음성 피드백 : 하루 동안 있었던 일들을 이야기한 후에는, 그 순간의 감정이 긍정적이든 부정적이든 자연스레 떠오르게 된다. 이에 본 서비스는 사용자가 녹음을 마친 후 자신의 감정 상태를 선택하면, 선택된 감정에 부합하는 메시지를 제공한다. 예를 들어, 사용자가 '슬픔'을 선택하면 위로의 메시지를, '기쁨'을 선택하면 축하의 메시지를 출력하는 식이다. 이는 사용자가 하루를 보다 긍정적으로 마무리할 수 있도록 지원한다. 개발 배경에서 언급했듯, 일기 등의 개인적인 이야기에 대해 감정적인 피드백을 받는 것은 긍정적인 효과를 유발한다. 그러나 이러한 개인적인 이야기를 타인에게 공유하는 것은 쉽지 않으며, 들어줄 사람을 찾는 것도 어려운 일이다. 따라서 본 서비스는 이러한 긍정적인 효과를 가장 효율적으로 제공할 수 있는 방법이 될 수 있다.

5. 성장기 말하기 능력 향상 : 실험 결과(1.3.2.)에 따르면 자신의 일상을 정리하여 말하는 행위를 통해 말하기 능력을 학습 시킬 수 있다. 본 서비스는 심리적으로 긴장감이 적은 상태에서 자신의 이야기를 자유롭게 할 수 있게 하는 환경을 제공한다. 이러한 환경은 주저와 휴지가 줄고 유창성, 정확성, 어휘 사용 등의 말하기 능력 향상의 효과를 얻을 수 있다.

본 서비스는 이러한 긍정적 영향을 최대한 많은 사람들이 높은 접근성을 가진 간편화된 음성 녹음을 통해 경험할 수 있도록 한다. 이는 사용자가 자신의 일상을 더욱 풍부하게 기록하고, 개인적인 성장과 성찰을 위한 중요한 도구로 작용할 것이다.