**프로젝트 기획안**

**2025년 05 월 27 일**

* **과정명: 융합 메디컬 AI with 스마트 웰니스**
* **과정 기간: 2024-12-30 ~2025-07-21**

|  |  |
| --- | --- |
| **팀명** | 음성 인식 프로젝트 2팀 |
| **팀원** | \*김석호, 석지원, 이지훈 |
| **프로젝트 주제** | 음성 AI를 활용한 ADHD 조기 진단 지원 시스템 |
| **기술 키워드** | FastAPI, SvelteKit, TTS, STT, gTTS, Whisper, Zonos, AI Voice Cloning, 보이스 커스터마이즈 |
| **프로젝트 목적** | 현대 사회에서 설문조사는 중요한 데이터 수집 방법 중 하나이지만, 기존의 텍스트 기반 설문 시스템은 고령자, 어린이, 외국인 등 정보취약계층에게 높은 진입 장벽을 제공한다는 한계가 있습니다. 또한 비대면 환경의 확산, 디지털 전환 가속화로 인해 사용자 경험의 질과 몰입도가 더욱 중요해졌습니다. 이에 본 프로젝트는 AI 음성 합성 및 인식 기술을 활용한 **맞춤형 음성 기반 설문 시스템** 개발을 목표로 합니다.  본 시스템은 설문 문항을 사용자의 연령·특성에 맞춘 목소리로 친근하게 읽어주고, 사용자는 자신의 음성으로 손쉽게 설문에 응답할 수 있습니다. Whisper, gTTS, Zonos와 같은 최신 음성 AI 모델을 통합하여 다양한 화자의 보이스 클로닝 및 커스터마이즈 기능을 제공하고, 설문 몰입도를 크게 향상시킵니다. 이를 통해 사용자 편의성을 높일 뿐만 아니라, 실제 현장(공공기관, 병원, 학교, 키오스크 등)에서의 활용 가능성을 넓히고, 기존에 접근이 어려웠던 집단도 자연스럽게 데이터 수집에 참여하도록 만듭니다.  또한, 본 시스템은 관리자가 결과를 쉽고 직관적으로 분석할 수 있도록 설문 결과 요약·코멘트 자동화, CSV 다운로드, 데이터 시각화 등 다양한 부가 기능을 지원합니다. 맞춤형 음성 안내로 설문 참여의 문턱을 낮추고, AI 보이스 클로닝 기술로 차별화된 사용자 경험을 제공함으로써, 디지털 헬스케어, 공공 서비스, 교육 평가 등 여러 분야에서 서비스 혁신 및 새로운 비즈니스 기회 창출에 기여할 것으로 기대합니다. |
| **프로젝트 수행 방향**  **(개발일정 및 역할분담)** |  |
| **프로젝트 수행 도구**  **(개발 환경)** | * **운영체제:** Ubuntu Linux (18.04+) * **프로그래밍 언어:** Python 3.10, JavaScript, TypeScript * **백엔드 프레임워크:** FastAPI * **프론트엔드 프레임워크:** SvelteKit * **음성 합성/인식 라이브러리:** Whisper, gTTS, Zonos * **GPU 하드웨어:** NVIDIA RTX 4070Ti, Google Colab T4 GPU * **협업 및 문서화:** Git, Notion * **개발/실험 환경:** VSCode |
| **필수 기능**  **(주요 구현 기능)** | * **연령별/맞춤형 음성 안내:** 설문 문항을 사용자의 연령대, 선호에 맞춘 보이스로 TTS 생성 및 안내 * **음성 설문 응답 및 자동 인식:** Whisper, gTTS, Zonos 등을 활용한 음성 기반 설문 응답(STT) 및 실시간 변환/저장 * **보이스 클로닝 및 커스터마이즈:** 짧은 샘플 음성만으로 사용자의 목소리 스타일을 복제하는 제로샷 TTS 적용 * **설문 결과 분석 및 자동 코멘트:** 응답 결과 자동 점수 산출, 종합 평가 코멘트 생성 및 사용자 피드백 제공 * **결과/총평 페이지 구현:** 설문 종료 후 결과 및 총평을 직관적으로 확인할 수 있는 UI 제공, CSV 파일로 다운로드 지원 * **응답 결과 수정 및 UI/UX 개선:** 설문 응답의 클릭 기반 수정, 가독성·접근성 향상을 위한 UI/UX 개선 * **협업 개발 및 데이터 관리:** Git 기반 코드 협업, 프로젝트 산출물 자동화, 결과 데이터 관리 및 백업 |
| **수행 계획 및**  **예상 결과물** |  |