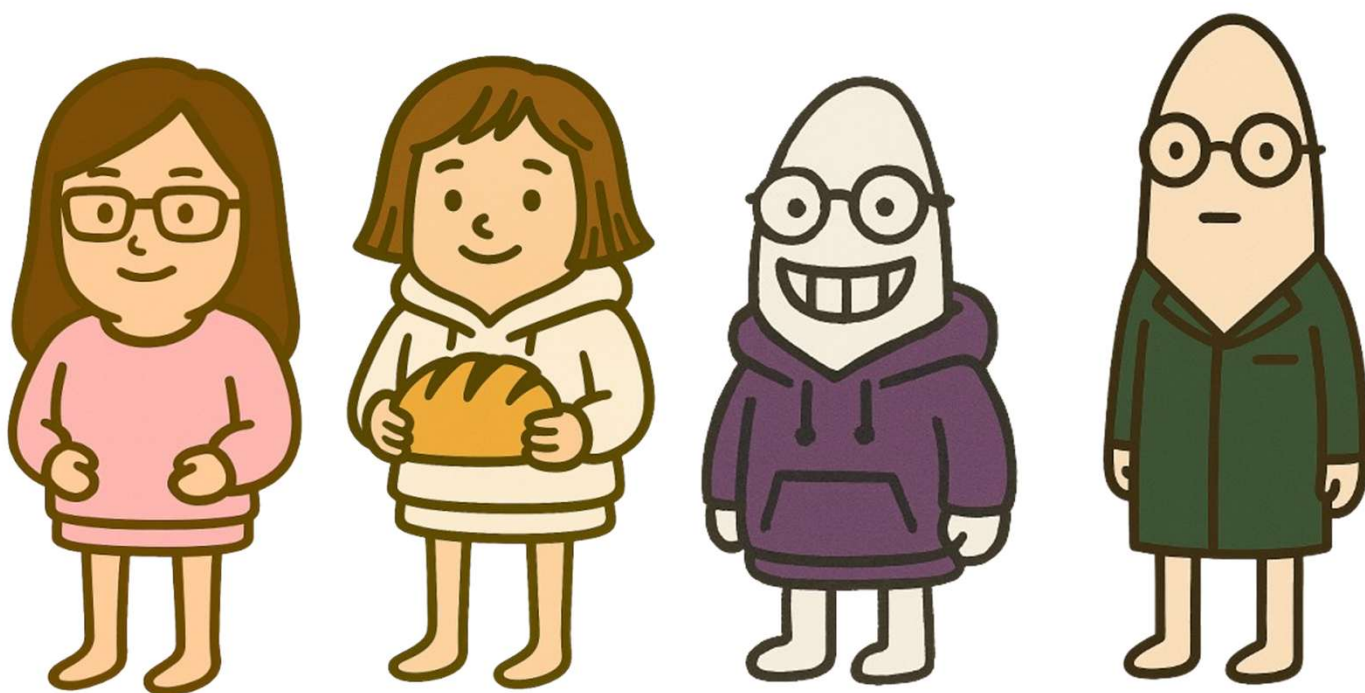


AI 기반의 개인 자동화 루틴 생성과 스마트 공간 케어

SW bootcamp 9기 A반 5팀

최하영, 정원, 박상현, 구승연



음성 기반 맞춤형 루틴 생성 시스템

개인정보 보호를 최우선으로 생각하여 모든 음성 데이터는 디바이스 내부에서 처리되고 암호화되어 외부로 유출되지 않으며, 네트워크 연결 없이도 완전한 음성 인식 서비스를 사용할 수 있습니다.

사용자의 음성을 인식하여 스마트홈을 즉시 활성화하여 집안의 조명, 온도, 음악 등 다양한 기기를 자동으로 제어합니다. 사용자의 음성 명령과 일상 생활 패턴을 인공지능이 분석해 최적화된 맞춤형 루틴을 생성함으로써 더욱 편리한 생활 환경을 제공합니다.

시스템 Flow

키워드 인식 (KWS)

BC-ResNet 기반 특정 키워드(예: "헤이 원") 검출 시스템으로 빠르고 가볍게 작동합니다.

프롬프트 엔지니어링

STT 결과를 LLM이 이해할 수 있도록 구조화하여 정확한 해석을 가능하게 합니다.



On-Device 음성인식 (STT)

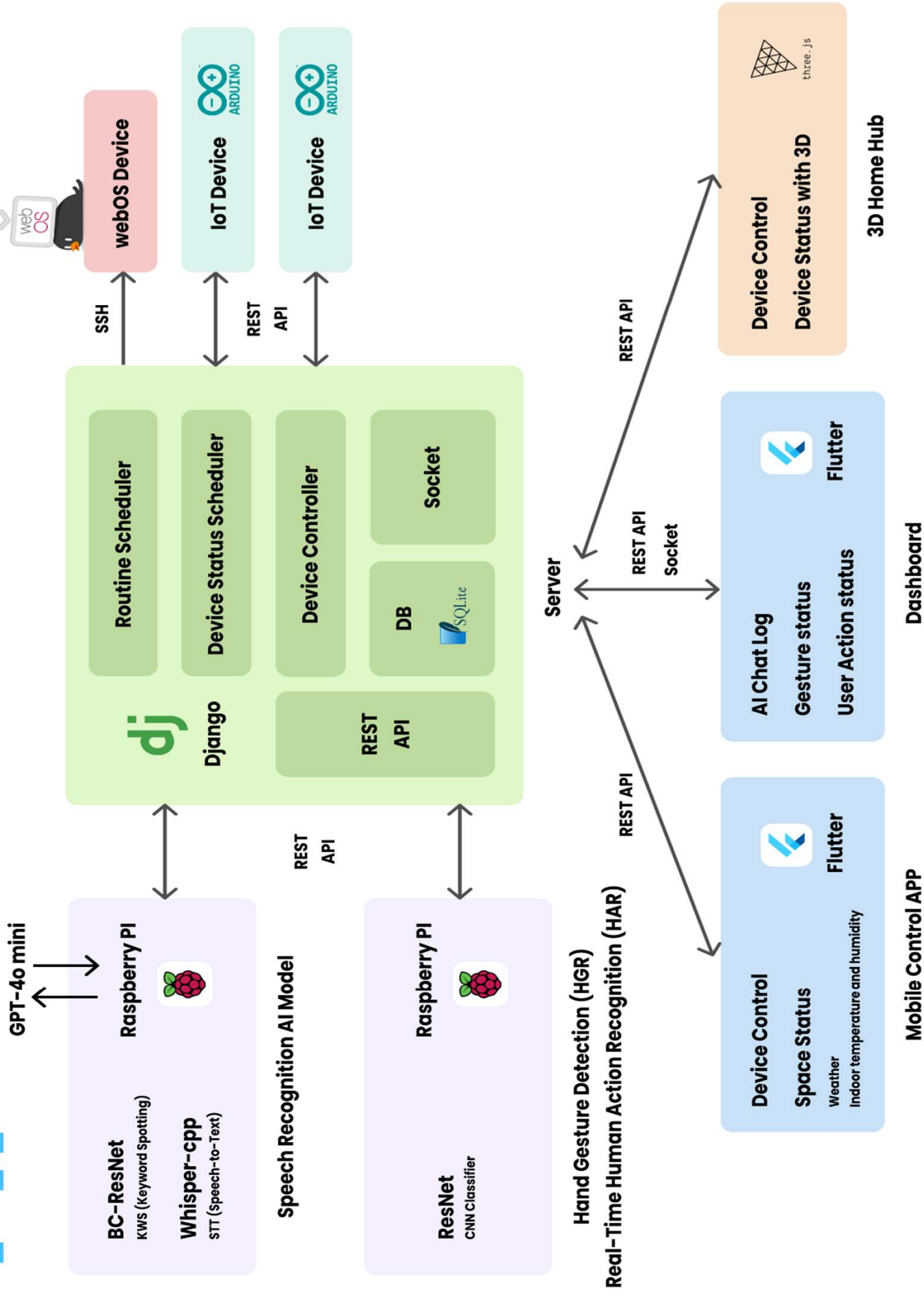
Whisper-cpp 기반 디바이스 내 음성-텍스트 변환으로 개인정보를 보호하며 신속하게 처리합니다.

LLM 기반 루틴 생성

사용자 정보와 발화를 분석하여 조건과 제어로 구성된 맞춤형 스마트홈 루틴을 자동 생성합니다.

음성인식 성능 비교 표

모델	환경	평균 추론 시간 (5초 오디오)	비고
whisper-tiny (PyTorch)	CPU	2~5초	느림, 정확도 동일
whisper-tiny (PyTorch)	GPU	0.5~1초	GPU에서 매우 빠르고 정확함
faster-whisper-tiny	CPU/R.Pi	1~2초	ONNX Runtime 하드웨어 최적화
whisper.cpp-tiny	CPU/R.Pi	1~2초	C++로 구현, SIMD(AVX/NEON) 등 CPU 명령어 최적화



영상 분석 기반 webOS 제어 시스템

영상 분석 기반 webOS 제어 시스템은 카메라를 통해 사용자의 제스처를 인식하고 이를 통해 TV나 다른 webOS 기반 기기를 직관적으로 제어할 수 있게 해주는 기술입니다. Raspberry Pi 5와 연결된 카메라가 사용자의 제스처를 실시간으로 분석하여 다양한 명령을 수행합니다.

사용자 액션	인식 유형	기능
👉 (OK 사인)	ok sign	알림 확인
👉 (손가락 하나)	one	재생/멈춤
👐 (그립)	grip (left, right)	컨텐츠 전환 (이전, 이후)
😴 (졸음)	졸음 여부	영상 재생/멈춤

이 시스템은 사용자의 제스처뿐만 아니라 졸음 상태까지 감지하여 사용자 경험을 최적화합니다. 예를 들어, 사용자가 TV 시청 중 졸음이 감지되면 자동으로 영상을 일시 정지시켜 중요한 장면을 놓치지 않도록 합니다



시스템 기능

- 전원 켜기/끄기: 특정 제스처를 통해 기기의 전원을 제어
- 채널 변경: 손 제스처를 통한 직관적인 채널 전환
- 재생/일시정지: OK 사인을 통한 미디어 컨트롤
- 알림 상호작용: Toast 알림에 대한 OK/Cancel 응답

대시보드를 통해 실시간으로 제스처 인식 결과와 졸음 감지 결과를 확인할 수 있어, 시스템의 작동 상태를 직관적으로 모니터링할 수 있습니다. 이는 사용자 경험 개선뿐만 아니라 시스템 디버깅 및 개발에도 유용하게 활용됩니다.