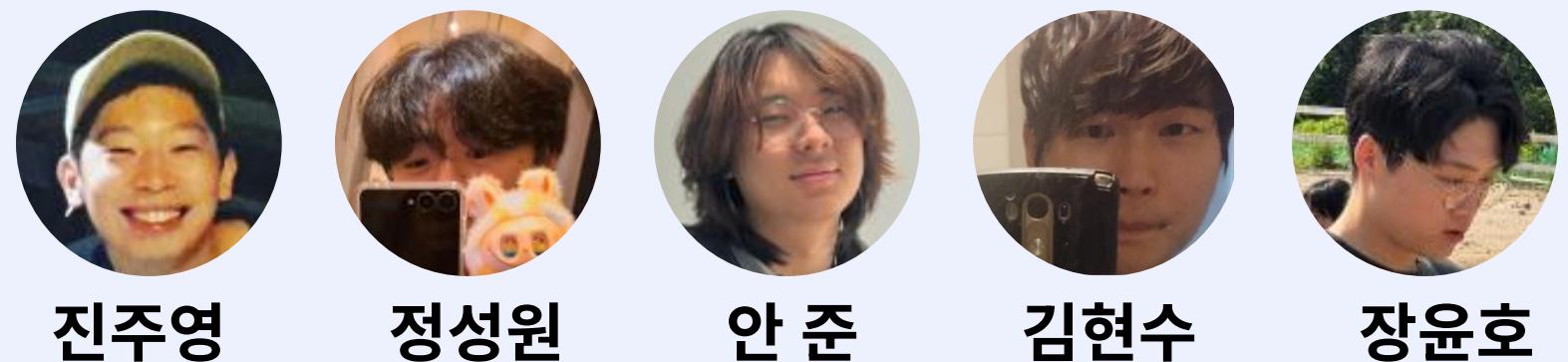




# Dupilot

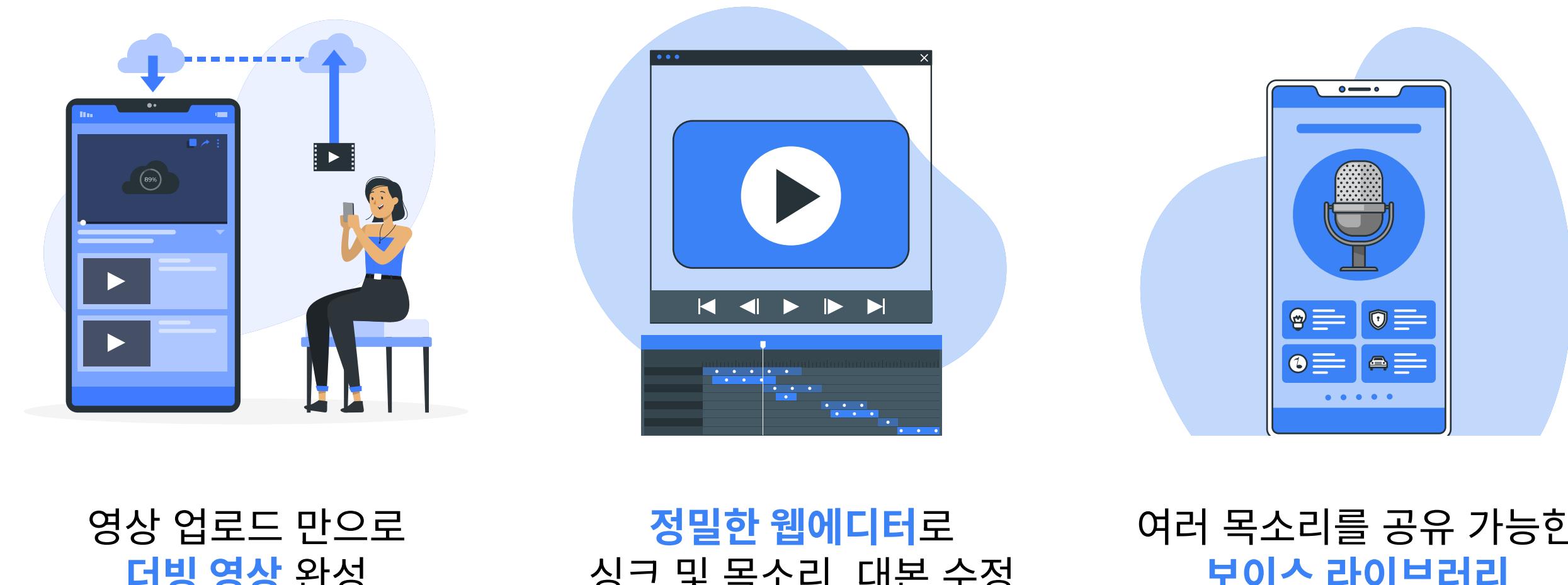


영상 번역부터 더빙 생성, 정밀 편집까지. Dupilot은 복잡한 더빙 워크플로우를 하나의 플랫폼으로 통합하여 작업 효율을 극대화합니다.



## 서비스 소개

웹 기반 AI 더빙 솔루션

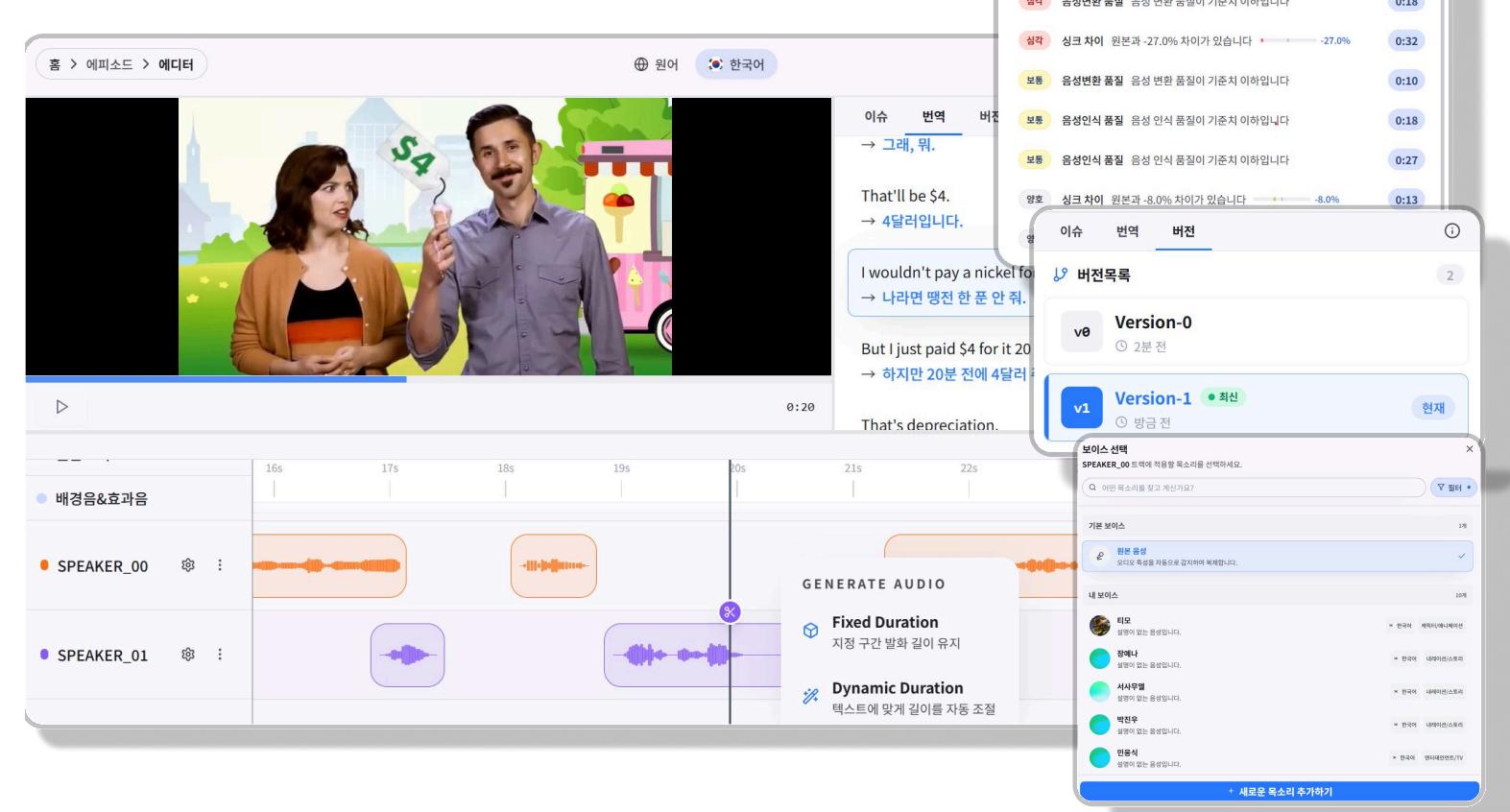


## 핵심기능

### 자동더빙

원본 영상 특성을 살린 자연스러운 더빙 생성

- 화자별 음성 구분 및 더빙
- 음색 임베딩 벡터를 추출 코사인 유사도 기반 보이스 매핑
- 더빙 파이프라인의 병렬처리 아키텍처 기반 다국어 지원



### 더빙 에디터

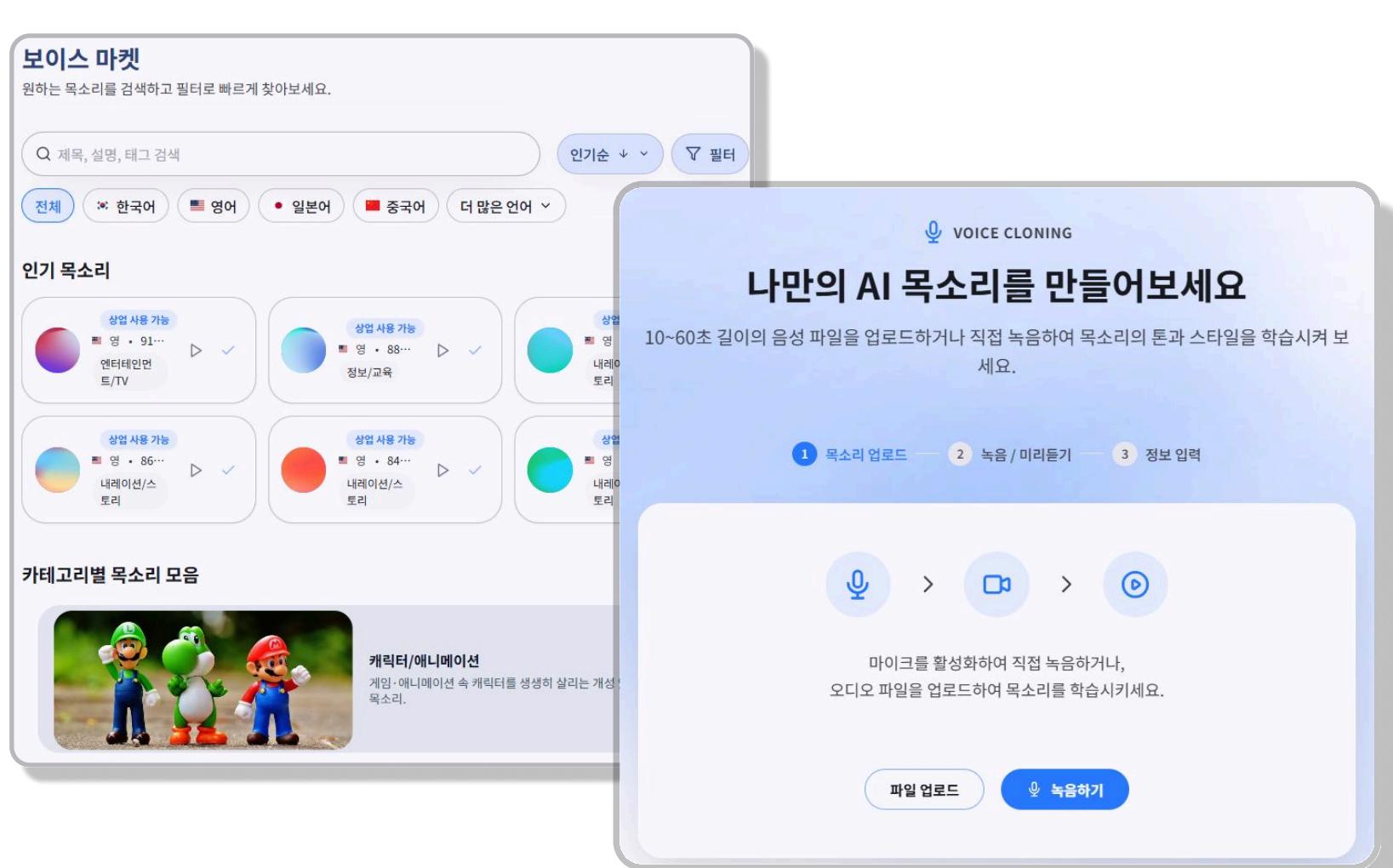
웹 기반 편집 도구로 직접 더빙 편집

- 오디오 편집
- STT 신뢰도 점수 및 싱크 차이를 이슈 리스트로 제공
- YouTube Data API v3 유투브 배포

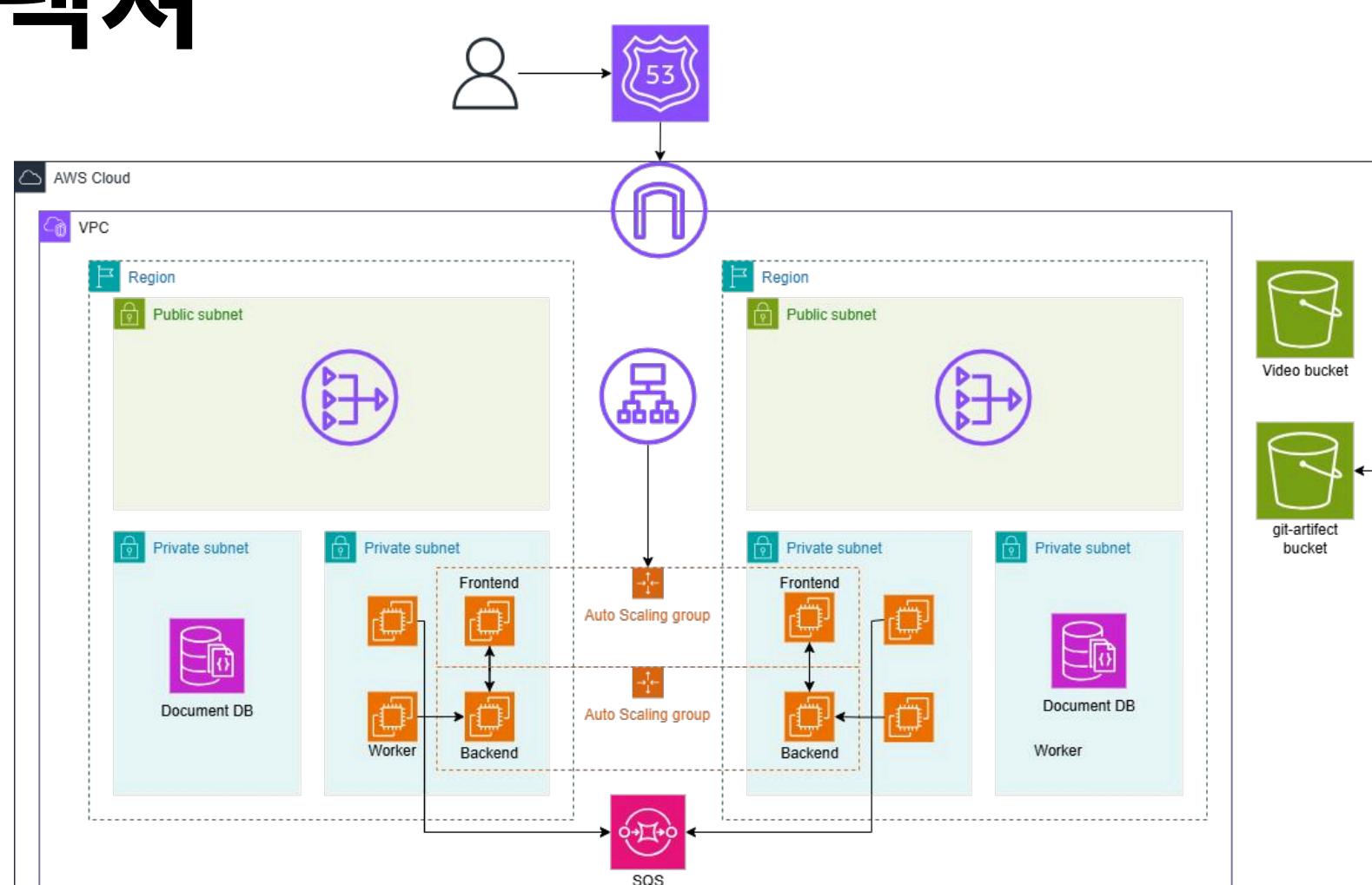
### 보이스 마켓

다양한 보이스 공유  
보이스 마켓 플레이스

- 다양한 음성 모델
- 보이스 클로닝
- 실시간 TTS



### 아키텍쳐



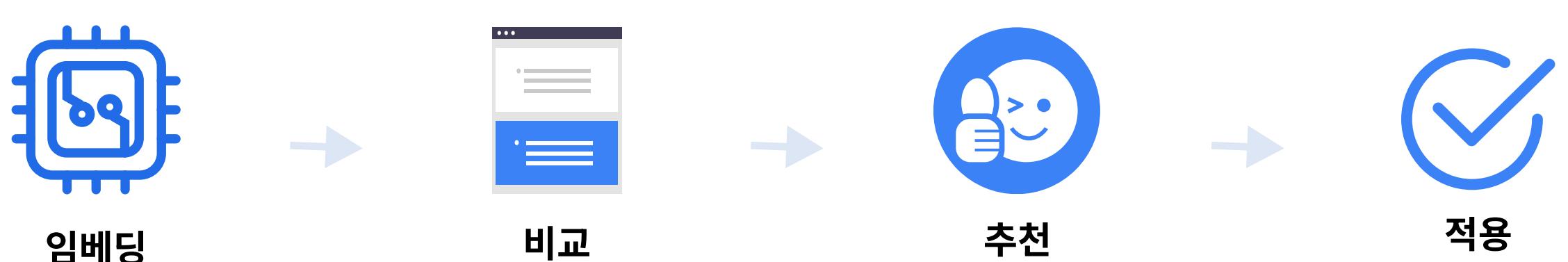
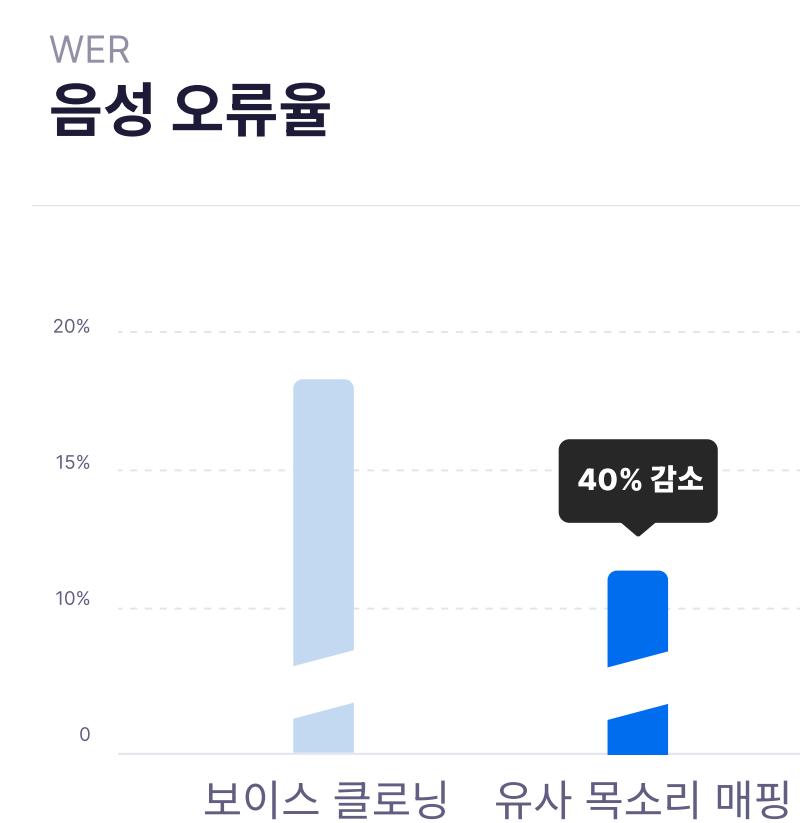
## 기술적 챌린지

### 교차 언어 발음 왜곡 해결

**Why** 원본 화자 목소리 적용시 원어만 가능

**Process & Solution** 목소리 임베딩 후 유사도 판정 후 원본 유사 목소리의 목표 언어 음성을 적용

**Result** 목소리 유사성 유지하며, WER 지표 감소

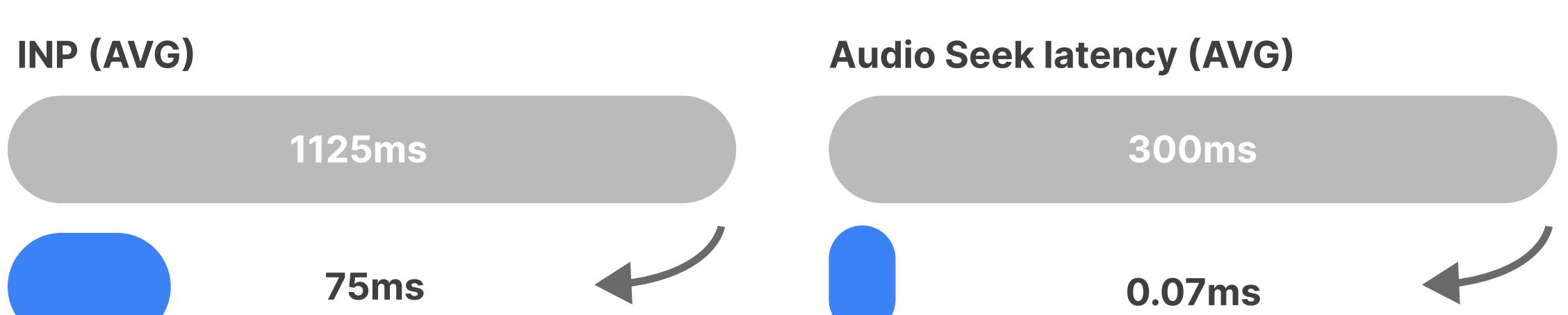


### 웨이터 렌더링 최적화

**Why** 영상/오디오 실시간 동기화 시, 브라우저 부하로 프레임 드롭 및 UI 렉 발생

**Process & Solution** 비가시 영역 메모리 가상화: 화면 밖의 데이터는 메모리에서 해제  
오디오 청크 이진 탐색: 전체 탐색 대신 필요한 부분만 빠르게 검색

**Result** 오디오 탐색 지역: 99.9% 감소  
INP(반응 속도): 1125ms → 75ms



### 자동 더빙 처리 속도 최적화

#### Why

초기 순차 처리로 180초 영상 기준 평균 500초가 소요되어  
영상 길이 대비 약 2.8배 시간 소요

#### Process & Solution

##### Before

TTS 단계에서 수백 개의 세그먼트를 순차 합성  
GPU 리소스가 유휴 상태로 남아 병목이 발생

##### Solution

세그먼트를 정해진 청크 단위로 나누어 분산 전송  
각 청크를 독립적으로 병렬처리를 진행  
모든 청크를 완료하면 MUX 진행

##### Result

180초 영상 기준 처리 시간을  
500초에서 250초로 단축 (약 0.5배)

