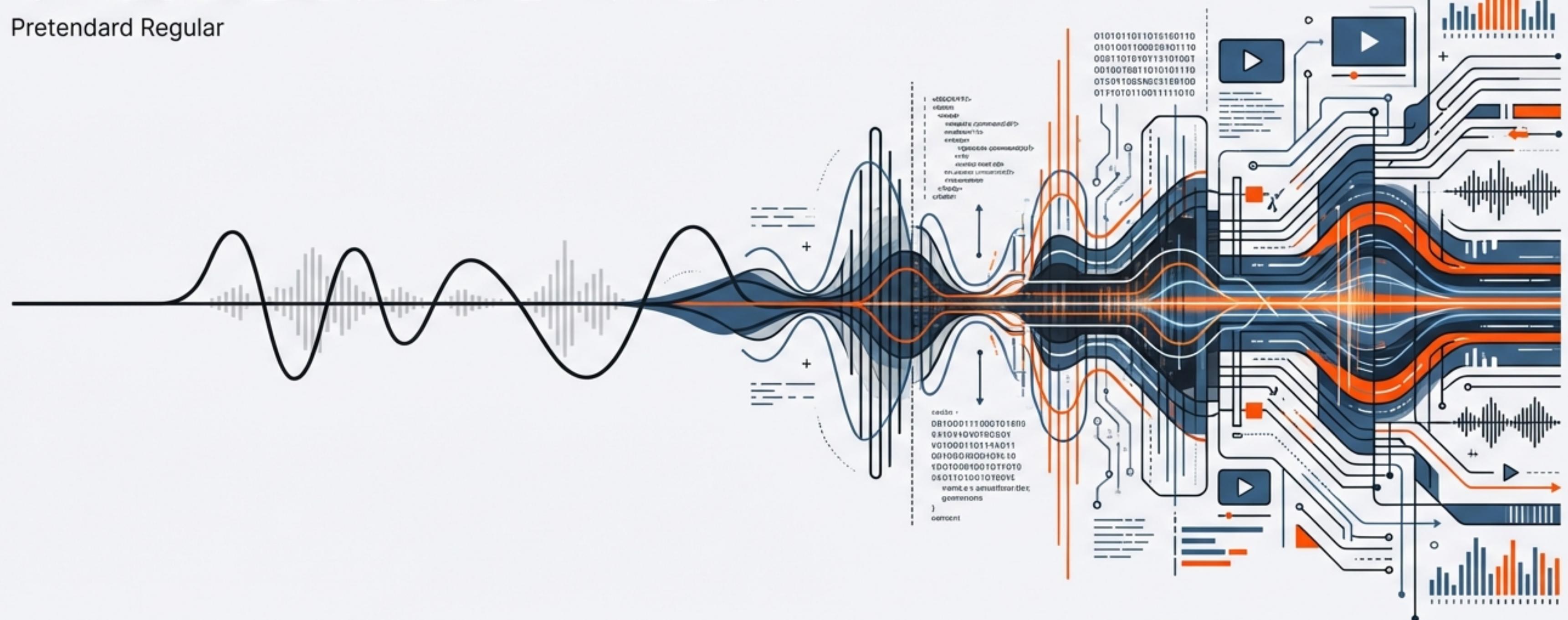


# ElevenLabs: AI 오디오의 미래

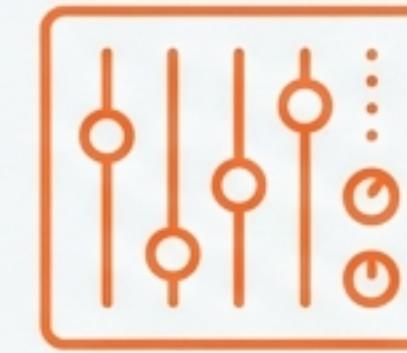
텍스트 음성 변환을 넘어선 '멀티모달 오디오 운영 체제'

Pretandard Regular



Voice (음성) → Production (제작) → Expansion (확장)

# ElevenLabs 플랫폼 개요



## The Foundation (기반)

Voice Generation & Cloning  
(음성 생성 및 복제)

## The Studio (제작)

Production & Multimedia  
(제작 및 멀티미디어 편집)

## The Expansion (확장)

Interaction & Deployment  
(상호작용 및 배포)

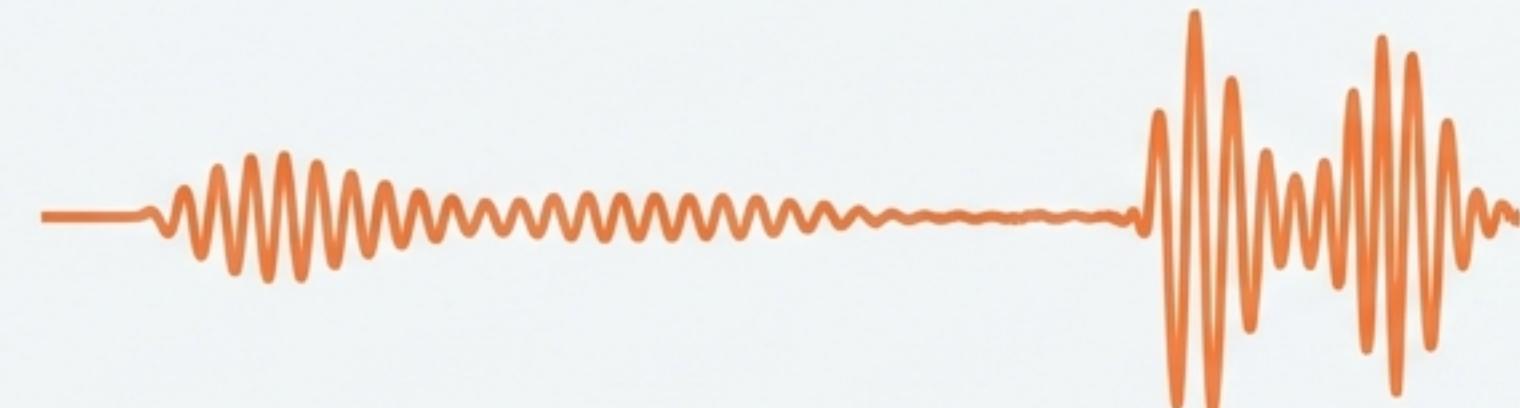
단순한 TTS 도구를 넘어선 포괄적인 오디오 창작 생태계

# Text-to-Speech (텍스트 음성 변환)

문맥과 감정을 이해하는 초현실적 음성 엔진 (Eleven v3)

- **기능:** 텍스트를 입력하여 즉시 고품질 오디오로 변환
- **핵심 기술:** Eleven v3 모델 (문맥 인식 및 감정 표현)
- **주요 특징:**
  - [whisper], [shout], [laugh] 등 태그를 통한 연기 지시
  - 32개 이상의 언어로 자연스러운 발화
  - 속도, 안정성, 스타일 과장 파라미터 조절 가능

그는 조용히 속삭였다. [whisper]  
'아무한테도 말하지 마세요.'



**활용:** 내레이션, 팟캐스트, 뉴스 리딩

# Voice Cloning (음성 복제)

짧은 샘플만으로 완벽한 디지털 트윈 생성



Audio Sample

## Instant Voice Cloning

- 1분 미만의 오디오 샘플 필요
- 즉각적인 생성 및 빠른 작업용
- 취미 및 간단한 프로젝트에 적합



## Professional Voice Cloning

- 30분 이상의 고품질 데이터 필요
- 미세한 뉘앙스와 말투까지 완벽 재현
- 디지털 트윈 및 고품질 콘텐츠용

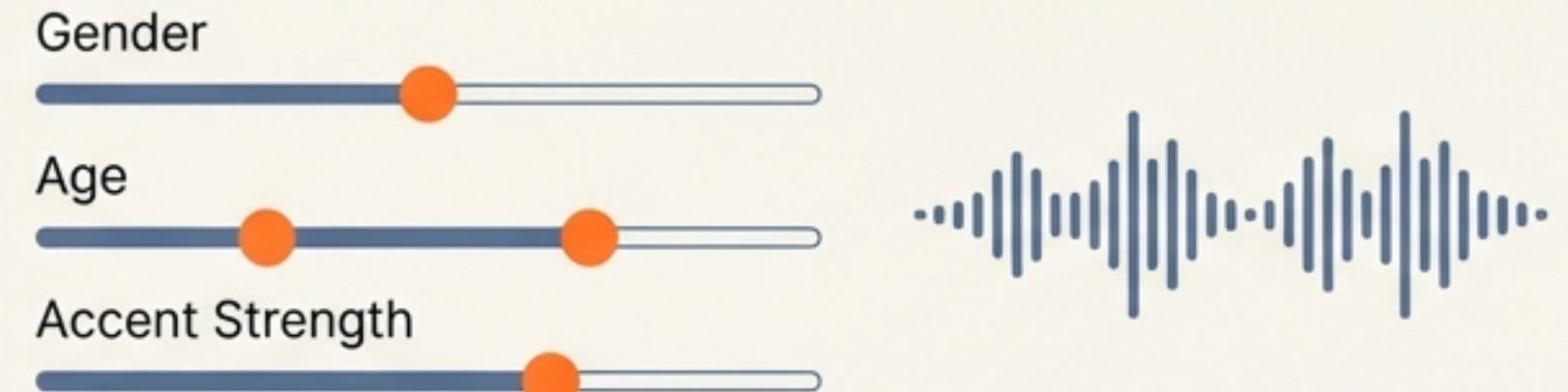
활용 사례: 저자 본인 목소리 오디오북, 환자의 목소리 디지털 보존 (Impact Program)

# Voice Library & Voice Design

새로운 목소리 창조 및 커뮤니티 자산 활용

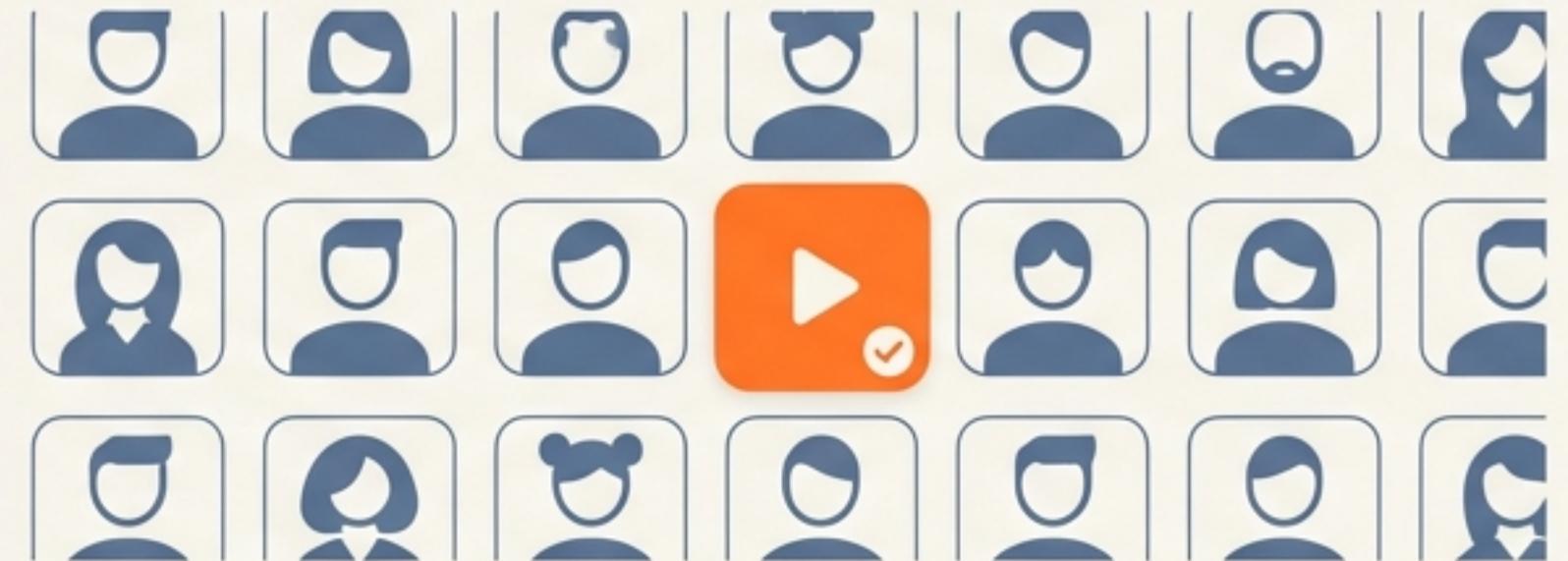
## Voice Design (음성 디자인)

성별, 연령, 억양, 악센트 강도를 조절하여  
세상에 없는 목소리 생성



## Voice Library (보이스 라이브러리)

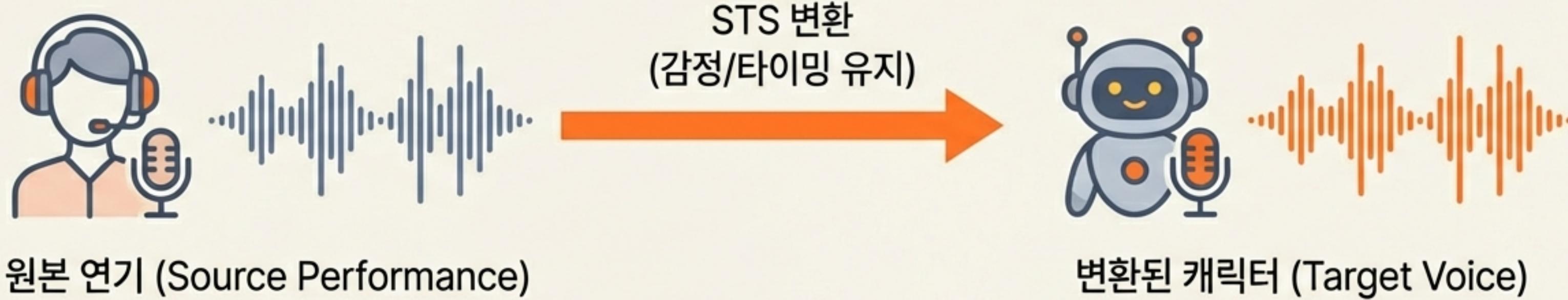
- 수천 개의 커뮤니티 보이스 탐색 및 활용
- Payouts: 자신의 목소리를 공유하여 수익 창출 가능



활용: 캐릭터 보이스 디자인, 마케팅 영상용 전문 성우 탐색

# Voice Changer (Speech-to-Speech)

감정과 연기를 유지한 채 목소리 톤만 변환



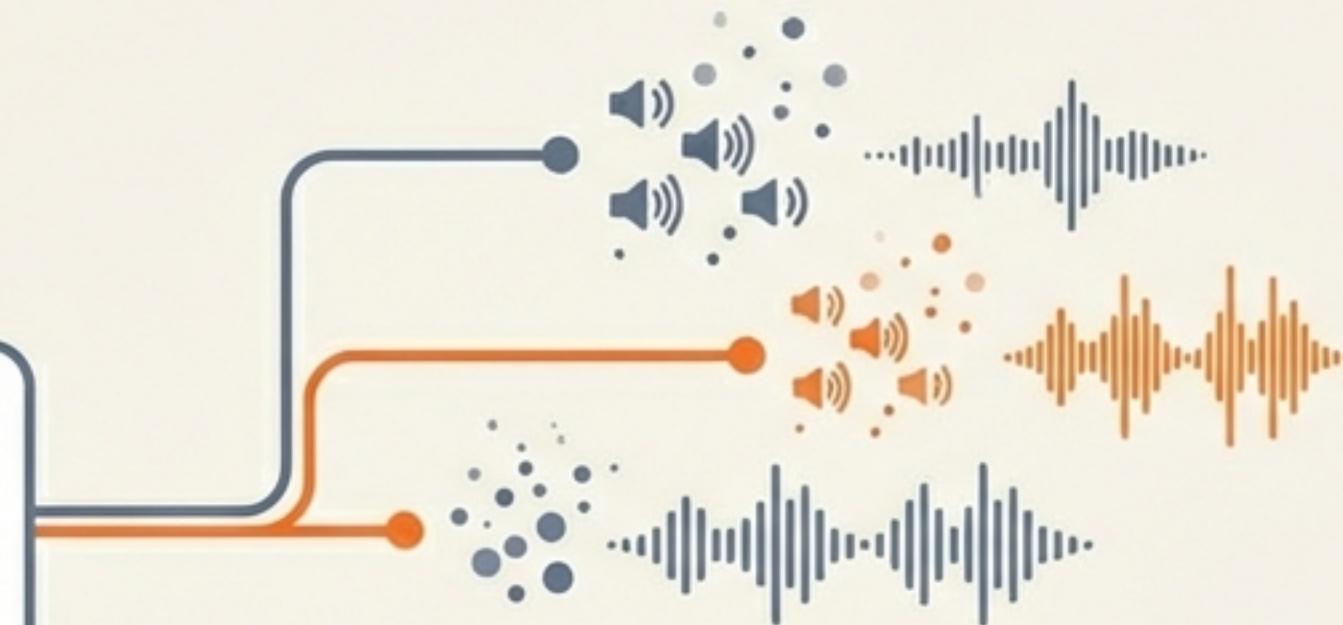
- 녹음 파일 또는 마이크 입력 지원
- 웃음, 한숨 등 비언어적 요소 완벽 보존
- 언어와 억양을 유지하며 화자(Speaker) 교체

활용: 1인 다역 연기, 게임 캐릭터 대사 제작

# Sound Effects (효과음 생성)

텍스트 프롬프트로 세상의 모든 소리 구현

“타오르는 모닥불 소리,  
장작이 튀는 소리”



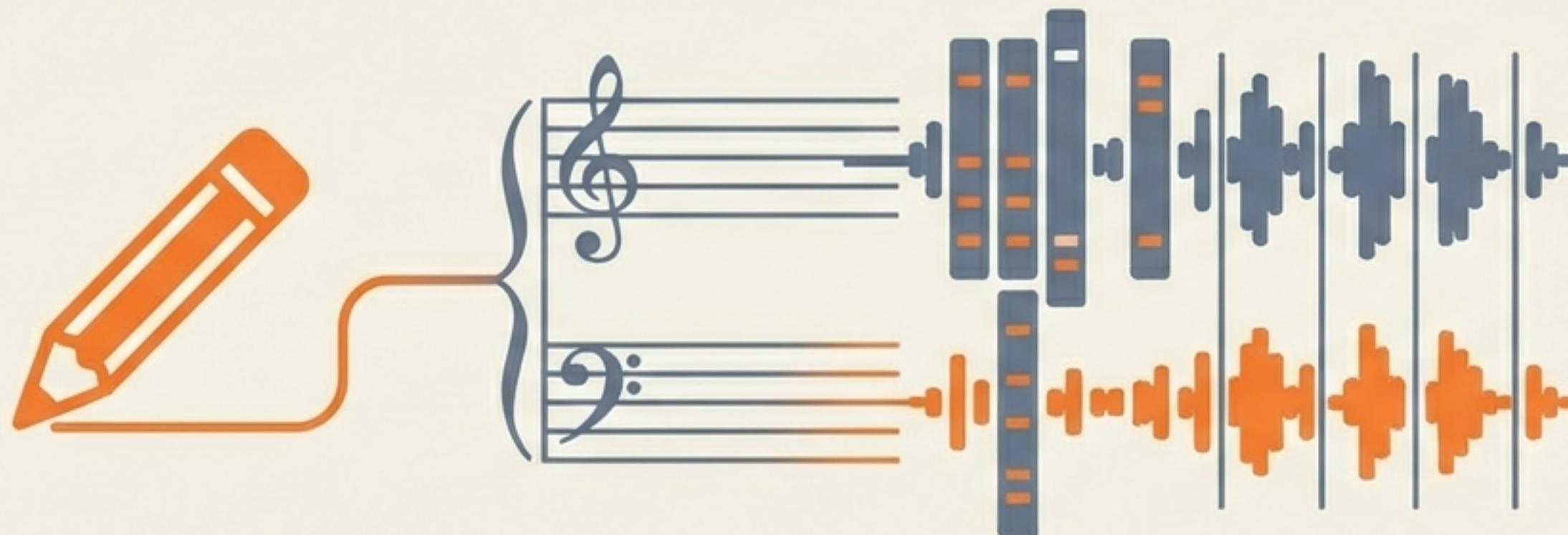
- 텍스트 묘사만으로 효과음(SFX) 및 배경음(Ambience) 생성
- 영상 길이에 맞춘 길이 조절 가능
- 여러 생성 옵션 중 최적의 사운드 선택

생성 산출물:

- 로열티 프리 효과음 (WAV/MP3)
- 영상용 폴리(Foley) 사운드

# Eleven Music (AI 음악 생성)

프롬프트를 통한 배경 음악 및 가사 곡 작곡



- 장르, 분위기, 악기, 템포를 텍스트로 지정
- 가사 입력 시 보컬이 포함된 노래 생성 가능
- 하나의 프롬프트로 여러 곡 리믹스 및 생성

활용: 팟캐스트 인트로, 영상 BGM, 데모 곡 제작

# Image & Video Generation (Beta)

오디오에 완벽하게 어울리는 시각적 스토리텔링

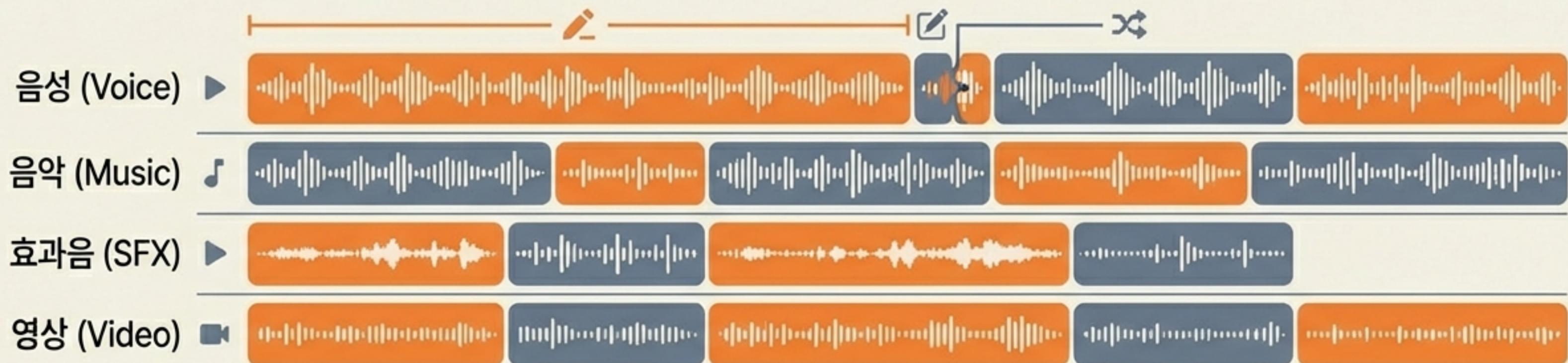


- 텍스트 프롬프트로 고품질 이미지 및 비디오 클립 생성
- 생성된 캐릭터에 오디오를 입혀 립싱크(Lip-sync) 적용
- 스토리보드, 썸네일, 쇼츠 영상 소스 제작

활용: 오디오북 삽화, 소설 미디어용 말하는 캐릭터 영상

# ElevenLabs Studio (스튜디오 편집기)

음성, 영상, 오디오를 통합하는 온라인 타임라인

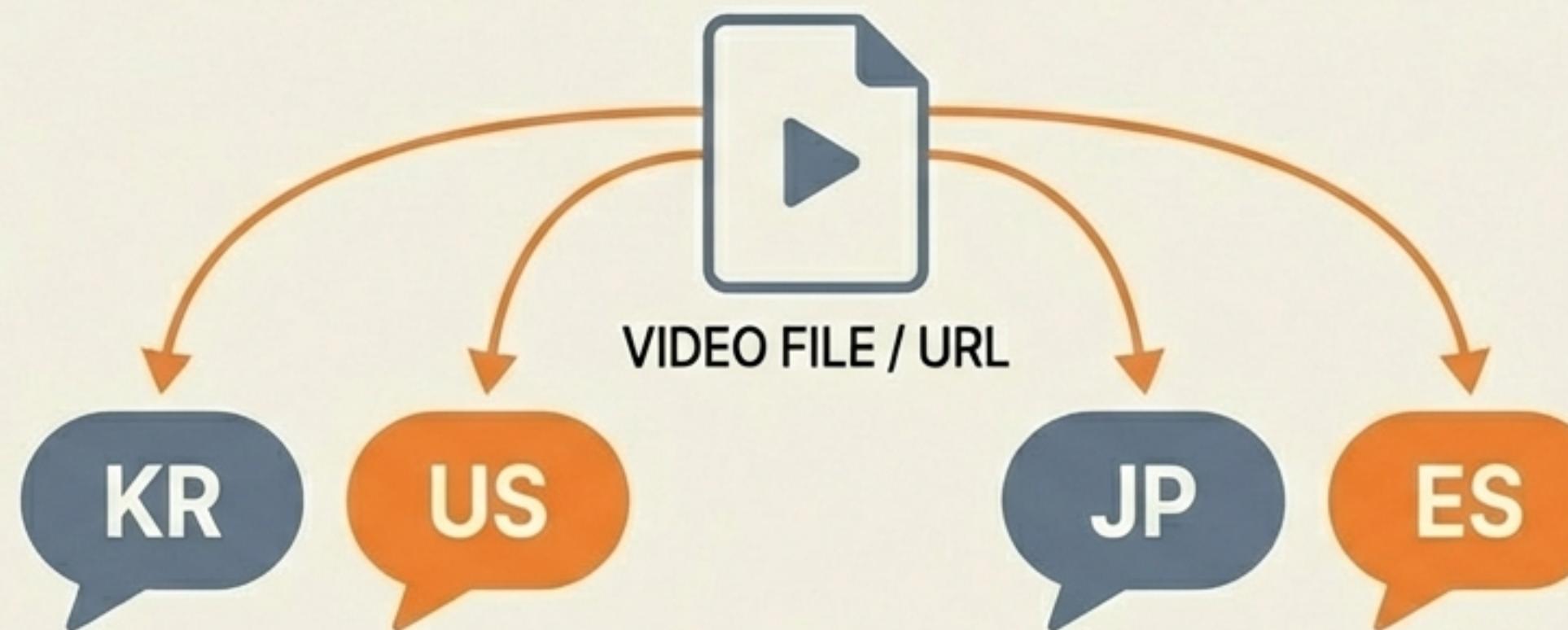


- 타임라인 기반의 정밀한 멀티미디어 동기화
- 대본 수정 시 해당 구간 음성만 즉시 재생성 (Regenerate)
- 자동 자막 생성 및 번역 기능
- 다수 화자(Multi-speaker) 할당 및 대화형 콘텐츠 제작

산출물: 완성된 오디오북, 팟캐스트, 영상 콘텐츠

# AI Dubbing Studio (AI 더빙 스튜디오)

원래 목소리를 유지하며 29개국 언어로 확장



- 영상 파일 또는 URL(유튜브/틱톡) 입력으로 자동 더빙
- 다화자 감지 및 화자별 목소리 분리 변환
- 대본 편집기를 통한 오역 수정 및 타이밍 미세 조정
- 원작자의 감정과 목소리 톤 보존 (Voice Preservation)

활용: 글로벌 교육 콘텐츠, 다국어 유튜브 채널 운영

# Conversational AI Agents (대화형 AI 에이전트)

코딩 없이 구축하는 실시간 음성 대화 봇

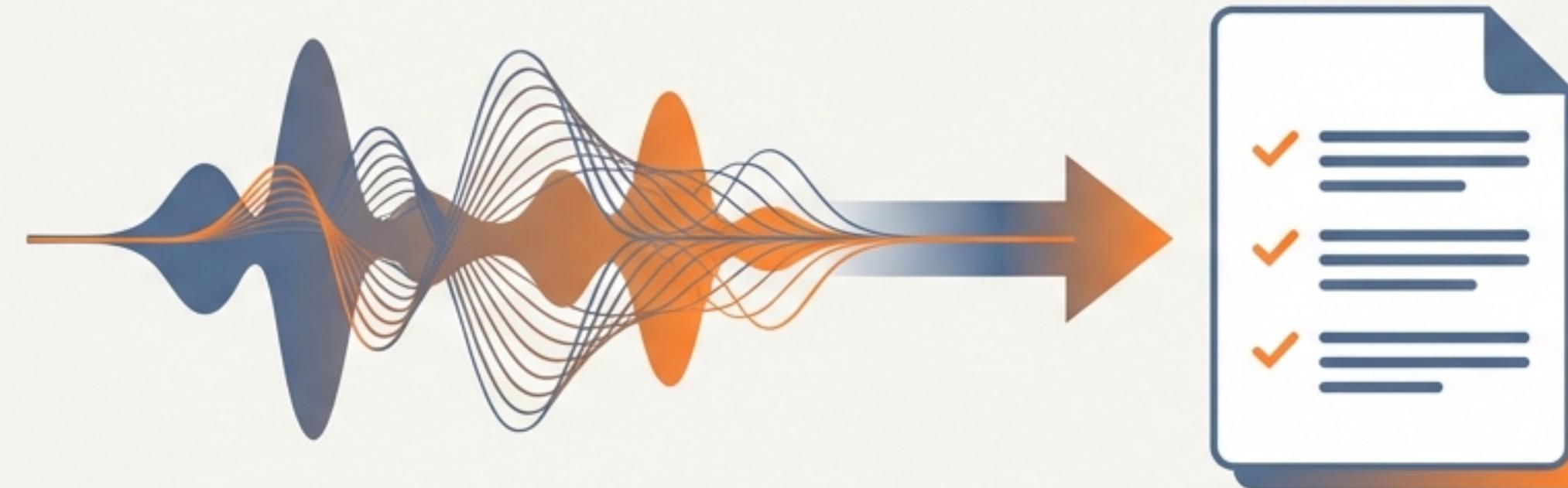


- 웹사이트 위젯 또는 전화 시스템 통합
- 문서/링크 업로드를 통한 맞춤형 지식 학습
- 사용자 의도 파악 및 자연스러운 턴테이킹(Turn-taking)
- 대화 분석 및 데이터 수집

활용: 24시간 고객 응대, 인터랙티브 튜터, 예약 시스템

# Scribe (음성 텍스트 변환)

최고의 정확도를 자랑하는 전사(Transcription) 모델



- 99개 이상의 언어 고정밀 텍스트 변환
- 화자 분리(Speaker Diarization) 및 식별
- 타임스탬프가 포함된 정확한 자막 파일 생성

활용: 회의록 작성, 영상 자막(SRT) 생성, 인터뷰 정리

# ElevenReader App (일레븐 리더 앱)

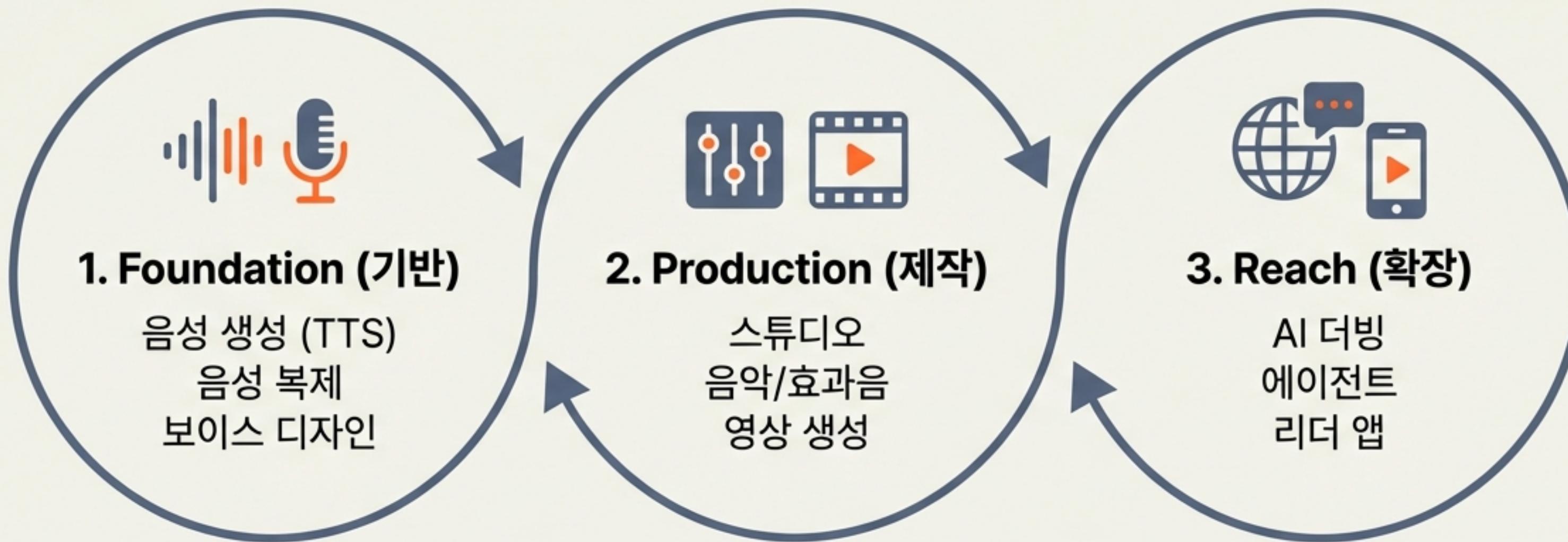
모든 텍스트를 고품질 AI 음성으로 소비



- 기사, PDF, ePub 등 텍스트 문서를 오디오로 청취
- 이동 중 뉴스, 이메일, 소설 등 콘텐츠 소비
- 유명인(Judy Garland 등)의 상징적 목소리 라이브러리 제공
- 스캔한 텍스트 이미지의 음성 변환 지원

활용: 출퇴근길 리스닝, 시각 장애인 정보 접근성 향상

# ElevenLabs 에코시스템 요약



단순한 TTS 도구를 넘어, 기획부터 제작, 배포까지  
가능한 통합 오디오 창작 플랫폼입니다.