

음성 AI 기술, 실험 단계 지나 기업 운영의 핵심 인프라로 부상

사례포커스

▪ 투자·매출 지표로 확인되는 음성 AI 상용화

음성 AI 기업들의 연간 반복 매출 확대와 투자 유치 및 고객 확보가 동시에 나타나며, 기술 검증 단계를 넘어 수익 모델이 안정화되고 있음

| 주요 음성 AI 전문 기업 및 플랫폼 |

- Nuance: 헬스케어 및 엔터프라이즈 음성 인식 분야의 플랫폼으로, 임상 문서 작성, 고객 서비스 영역에 특화됨
- Deepgram: 컨택센터, 미디어, 대화형 AI를 위한 실시간 음성 인식 API를 제공함
- AssemblyAI: 개발자 및 기업을 대상으로 음성-텍스트 변환(STT), 자연어 처리(NLP), 감정 분석 기능을 제공함
- ElevenLabs: 엔터테인먼트, 게임, 오디오북 분야를 중심으로 시 음성 복제 및 음성 합성 기술을 선도함
- Cartesia: 실시간 상호작용에 적합한 초현실적 음성 생성과 빠른 응답에 특화됨

뉴스브리프

음성 AI 산업이 기술 실험 단계를 넘어 본격적인 산업 성장 국면에 진입했음. 2025년을 기점으로 음성 AI는 기존 IVR시스템의 “대화형 음성 응답” 단계를 넘어, 인간과 유사한 수준의 대화가 가능한 음성 AI 에이전트 중심 구조로 빠르게 전환 중임.

「State of Voice AI 2025」보고서에 따르면 기업들이 음성 기술을 핵심 인프라로 인식함에 따라, 음성 인터페이스가 디지털 상호작용의 주요 수단으로 자리 잡고 있음. 이런 변화는 주요 음성 AI 스타트업의 실적과 투자 지표에서 명확히 확인됨. 음성 AI 스타트업 일레븐랩스(ElevenLabs)는 2025년 기준 연간 반복 매출(ARR) 3억 3천만 달러(원화 약 4,741억 원)*를 기록했으며, 포춘 500대 기업을 포함한 고객층을 확보하고 있음. 이는 음성 생성-에이전트 기술이 실제 고객 응대와 콘텐츠 영역에서 빠르게 확산되고 있음을 보여주는 사례로 해석됨. 독일 스타트업 파를로아(Parloa)는 AI 기반 고객 서비스 플랫폼을 앞세워 3억 5천만 달러(원화 약 5,028억 원) 규모의 투자를 유치함. 파를로아는 1년만에 기업 가치가 세 배로 증가했으며, 음성

AI 기반 고객 서비스가 실질적인 비용 절감과 투자 대비 수익률(ROI)을 창출하는 영역으로 평가받고 있음을 의미함.

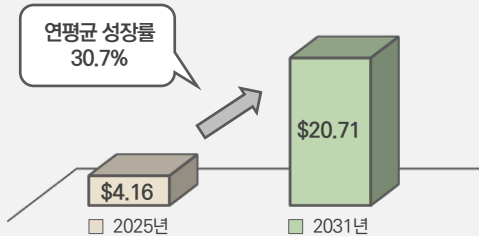
시장 전망 역시 고성장세를 보임. **생성형 AI 음성 시장은 2025년 41억 6천만 달러(원화 약 5조 9,767억 원)에서 2031년 207억 1천만 달러(원화 약 29조 7,541억 원)로 성장할 전망**이며, 연평균 성장률은 30.7%로 전망됨. 기업 수요 측면에서도 음성 AI는 ‘보편 기술’ 단계에 도달함. 2025년 북미 전역의 비즈니스 리더 400명을 대상으로 진행된 설문조사에 따르면 응답자의 97%가 이미 음성 기술을 활용 중이며, 84% 이상은 향후 1년 내 관련 예산을 확대할 계획이라고 응답함. 이는 음성 AI가 일시적인 트렌드를 넘어 기업 전략의 필수 요소로 자리잡고 있음을 보여주는 지표임.

현재 음성 AI 산업은 성숙 단계에 진입한 상황이며 ‘도입 여부를 고민하는 기술’에서 ‘어떤 솔루션을 선택할 것인가’를 결정하는 단계로 평가됨. **향후 경쟁의 핵심은 유연성과 적응성**으로 이동할 가능성이 높으며, 음성 AI는 기업 운영 전략의 핵심 인프라로 작동할 전망이다.

* 1달러=1,436.70원(2025.02.02. KEB 하나은행 최초 매매기준율 적용)

글로벌 생성형 음성 AI 분야의 시장 규모

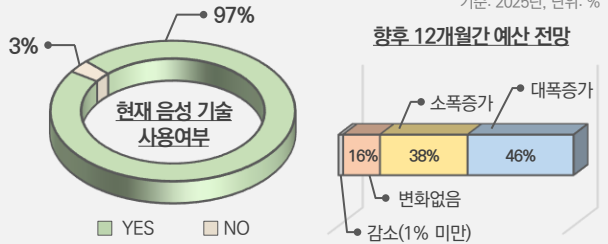
기준: 2025년&2031년, 단위: 십 억



출처: Markets and Markets, "AI Voice Generator Market", 2025.12., <https://www.marketsandmarkets.com/PressReleases/ai-voice-generator.asp>

북미 음성 AI 활용 현황 설문 조사

기준: 2025년, 단위: %



출처: Deepgram, "Introducing 'State of Voice AI 2025': The Year of Human-like Voice AI Agents", 2025.04.08., <https://deepgram.com/learn/state-of-voice-ai-2025>

본 보고서의 주요내용

음성 AI 산업은 2025년을 기점으로 실험 단계를 넘어, IVR 중심 구조에서 인간과 유사한 대화가 가능한 음성 AI 에이전트 중심의 산업 성장 국면에 진입함

본 보고서의 시사점

음성 AI는 이제 ‘도입 여부를 고민하는 기술’에서 ‘어떤 솔루션을 선택하고 어떻게 전략적으로 활용할 것인가’를 결정하는 단계로 전환됨

* 용어 설명: 1) IVR(Interactive Voice Response): 대화형 음성 응답으로 고객이 전화로 기업과 상호작용할 수 있도록 하는 자동 전화 시스템
2) ARR(Annual Recurring Revenue): 연간 반복 매출의 약자로 비즈니스에서 매년 반복적으로 발생하는 수익을 나타내는 핵심 지표
3) ROI(Return on Investment): 투자 대비 수익률의 약자로 투자 대비 얼마나 이익을 얻었는지 측정하는 재무 지표

* 내용 출처: 4PSA, "Voice AI's Defining Moment: Insights from the State of Voice AI 2025 Report", 2025.11.15., <https://blog.4psa.com/voice-ai-defining-moment-insights-state-voice-ai-2025-report/>
TechCrunch, "ElevenLabs CEO says the voice AI startup crossed \$330M ARR last year", 2026.01.13., <https://techcrunch.com/2026/01/13/elevenlabs-ceo-says-the-voice-ai-startup-crossed-330-million-arr-last-year/>
Reuters, "German AI startup Parloa triples valuation to \$3 billion in latest fundraising", 2026.01.15., <https://www.reuters.com/business/german-ai-startup-parloa-triples-valuation-3-billion-latest-fundraise-2026-01-15/>
Deepgram, "Introducing 'State of Voice AI 2025': The Year of Human-like Voice AI Agents", 2025.04.08., <https://deepgram.com/learn/state-of-voice-ai-2025>

* 통계 출처: Markets and Markets, "AI Voice Generator Market", 2025.12., <https://www.marketsandmarkets.com/PressReleases/ai-voice-generator.asp>
Deepgram, "Introducing 'State of Voice AI 2025': The Year of Human-like Voice AI Agents", 2025.04.08., <https://deepgram.com/learn/state-of-voice-ai-2025>