



일레븐랩스, 99개 언어 지원으로 사람과 기술 상호작용 혁신

박종진 _ 전자신문 기자

[목차]

1. 들어가며

2. 모든 소리에 집중하는 글로벌 인공지능 데카콘 기업 '일레븐랩스'
3. 마티 스타니스제프스키 일레븐랩스 최고경영자(CEO)와 일문일답
4. 홍상원 일레븐랩스 한국지사장과 일문일답
5. 인터뷰아: 홍상원 지사장 소개



① 들어가며

인공지능 시대 음성 기술 활용도 확대·다변화되고 있다. 현재 음성 인공지능 기술은 단순 음성 인식 수준을 넘어 ‘사람의 말을 이해하고 의도를 파악, 업무를 처리하고 자연스럽게 응답’하는 실시간 인터페이스(interface)로 산업 전반에 활용되고 있다.

음성 인공지능의 산업적 가치는 ‘단순히 사람처럼 말한다’는 데 있는 것이 아니다. 핵심은 기업의 업무 처리 속도, 비용 구조, 접근성, 자동화 수준 자체를 바꾸는 인터페이스라는 점이다. 키보드·앱·웹 화면 중심이던 디지털 작업을 말로 대체하면 생산성과 고객 경험을 동시에 개선할 수 있다.

직접적인 효과는 비용 절감이다. 인공지능 컨택센터(AICC), 고객센터 영역에서 효과가 크다. 기존 콜센터는 상담사 채용, 교육, 이직률, 최대 24시간 운영 등 부담이 매우 컸다. 음성 인공지능은 반복 문의 대부분을 자동 처리할 수 있는 것이 강점이다. 배송 조회, 예약 변경, 단순 환불, 요금 안내, 계좌 조회 같은 업무는 인공지능이 처리하고 사람 상담사는 고난도 문제에 집중하는 구조다.

가트너는 에이전틱 인공지능(Agentic AI)이 2029년 일반 고객 서비스 문제의 80%를 자율 해결할 것으로 전망했다. 기업 운영비 절감 효과도 강조했다. 음성 인공지능을 단순 ‘자동응답기’가 아닌, 인건비 구조 자체를 바꿀 기술로 보는 것이다.

생산성 향상 측면에서도 효과가 크다. 사람은 손보다 입이 훨씬 빠르다. 평균적으로 노트북이나 스마트폰 자판을 치는 속도보다 말하는 속도가 3~4배 빠르다고 알려져 있다. 그래서 의료·법률·제조·물류 산업에서 효과가 클 것으로 기대된다. 의사가 환자와 대화하는 동안 인공지능이 진료 기록을 자동으로 작성하고, 현장 엔지니어는 장갑을 낀 상태에서 음성으로 설비 상태를 조회할 수 있게 되는 것이다. 편의성은 물론, 문서 작업 시간을 줄여 핵심 업무 집중도를 높이는 효과가 있다.

접근성도 확대된다. 음성 인터페이스는 디지털 기기 사용이 어려운 사람도 쉽게 접근할 수 있게 만드는 장점이 있다. 고령층, 장애인, 저숙련 사용자도 누구나 말만 하면 서비스를 이용할 수 있다. 금융·공공·보건 분야에서 음성 인공지능을 단순 사용자경험(UX)이 아닌 디지털 포용 기술로 평가하는 이유다.

데이터 축적과 분석에도 강점이 있다. 기존 콜센터 통화는 녹음만 하고 끝나는 경우가 많았지

만, 음성 인공지능은 고객 감정, 불만, 구매 의도, 이탈 위험을 실시간 분석할 수 있다. 예를 들어 보험사가 고객 통화를 분석해 해지 가능성 높은 고객을 조기 식별하거나, 유통사가 고객 문의 유형을 기반으로 상품 문제를 빠르게 발견하는 등의 방식이다. 음성 자체가 새로운 데이터 자산이 되는 것이다.

‘멀티모달 인공지능(Multimodal AI)’으로의 확장성도 있다. 최근 음성 인공지능은 단독 기능이 아닌 생성형 인공지능, 영상, 센서가 결합하는 추세다. 자동차에서는 운전자 음성과 차량 센서, 내비게이션이 합쳐지고 스마트팩토리(Smart Factory)에서는 음성과 디지털 트윈(Digital Twin), 사물인터넷(IoT)이 연결되는 등의 사례가 대표적이다. 음성 인공지능이 기업 인공지능 플랫폼의 입출력 창구 역할을 하게 되는 것이다. 또 브랜드와 서비스 경험을 차별화할 것으로 기대된다. 인공지능이 얼마나 자연스럽게 소통하고 일을 처리하는지가 경쟁력이 될 시대에 음성 인터페이스의 장점은 분명하다.

이러한 강점을 토대로 음성 인공지능은 산업 곳곳에서 활용되고 있다. 가장 빠르게 도입하는 분야는 금융이다. 은행과 보험사는 음성 인공지능을 단순 ARS(Automatic Response Service) 대체가 아닌 ‘인공지능 금융 창구’로 바꾸고 있다. 고객이 “이번 달 카드값 얼마나 나왔어?”라고 말하면 인공지능이 본인 인증 후 계좌, 카드, 대출 정보를 연결해 답변하고 추가로 ‘분할납부로 바꾸시겠어요?’까지 제안하는 구조다. 금융권에서는 목소리 분석 기술을 적극 도입하고 있다. 목소리 고유 특성을 인증, 이상거래 탐지에 활용하는 방식이다.

글로벌 음성인식 시장조사업체는 음성 인공지능이 금융, 건강 관리, 자동차 분야에서 폭발적으로 성장할 것으로 보고 있다. 실제 자동차 산업에서는 음성 인공지능이 차세대 사용자 경험(User Experience, UX)으로 자리매김하고 있다. 과거에는 “에어컨 켜줘” 정도의 명령을 실행하는 수준이었다면, 지금은 생성형 인공지능을 기반으로 “배터리가 부족한데 근처 충전소 중 가장 저렴한 곳을 찾아줘”, “탕수육을 파는 식당을 찾아줘” 같은 복합 명령을 처리하는 수준으로 진화했다. 전기차나 소프트웨어정의차량(SDV, Software Defined Vehicle)의 확산으로 차량 내부가 인공지능 인터페이스 중심으로 재편되며 음성 비중이 커지는 것도 특징이다.

미디어·콘텐츠 산업도 음성 인공지능을 적극 활용하는 영역이다. 오디오북, 광고, 더빙, 뉴스 진행자, 게임과 유튜브 해설 등 다양한 부분에서 활용된다. 더빙 등 글로벌 현지화에도 음성 인공지능이 투입되며, 제작 비용과 기간을 크게 줄이고 있다. 생성형 인공지능에 기반한 ‘텍스트 음성 변환(Text-To-Speech, 이하 TTS)’ 기술은 사람 수준의 억양과 감정 표현이 가능해지며 활

<그림 1> 일레븐랩스 로고



제공: 일레븐랩스

용도를 높이고 있다.

음성 인공지능은 산업 곳곳에서 기업 시스템과 연결된 음성 기반 업무 처리 인터페이스로 자리 잡았다. 단순 TTS 소프트웨어보다 실제 업무를 대신 수행하는 에이전틱 인공지능과 결합한 형태로 음성 인공지능이 빠르게 확산되고 있다. 음성 인공지능 분야 글로벌 선도기업 ‘일레븐랩스’ 사례로 기술과 산업 변화를 살펴보자.

② 모든 소리에 집중하는 글로벌 인공지능 데카콘 기업 ‘일레븐랩스’

일레븐랩스(ElevenLabs)는 전통적인 더빙 방식 한계를 극복하고자 설립된 음성 특화 인공지능 전문기업이다. 영국에 기반을 둔 글로벌 기업으로 전 세계 사용자 수는 월간 4,000만 명 이상, 창업 3년 만에 기업가치 10조 원 이상(110억 달러 규모)의 ‘데카콘(Decacorn) 기업’으로 자리매김했다.

음성이 기술 사용의 주요 소통 방식으로 자리잡으며 일레븐랩스는 음성 기반 상호작용을 보다 자연스럽게 인간적인 경험으로 만들어주는 음성 인공지능 기술을 개발하고 있다. 회사 설립 초기에는 명확하고 감정 표현이 풍부하며 다국어 지원을 지원하는 초현실적 음성 특화 인공지능 기술로 출발했다. 현재 음성, 음향 효과, 음악 전반에 걸쳐 인공지능 오디오 기술을 발전시키고 있다. 음성 인공지능을 일상적으로 사용하는 다양한 제품에 탑재하고 구현하는 글로벌 인공지능 오

〈그림 2〉 마티 스타니스제프스키(Mati Staniszewski) 일레븐랩스 최고경영자(CEO)



제공: 박종진 기자

디오 전문기업을 지향한다. 총 99개의 언어를 지원한다.

음성 기반 인공지능 에이전트의 기획·개발·배포·운영까지 한 번에 가능하도록 지원한다. 초저지연 실시간 재생, 양방향 대화, 다중 언어와 감정 처리, 안전성 관리, 품질 모니터링과 분석, 모바일·PC 웹 등 다양한 채널 연동을 포함한 전 과정 통합 툴체인(toolchain)을 제공, 기업 고객센터와 총괄 안내, 교육, 실시간 판매, 게임 등에서 음성 기반 인공지능 에이전트를 빠르고 안정적으로 출시·확장하는 기술을 제공하고 있다.

일레븐랩스는 인공지능 음성 및 오디오 분야에서 조사·연구에 기반해 세계 최고의 기술력을 뽐내는 신뢰받는 플랫폼 기업이다. 기술에 목소리를 부여하고, 세상의 지식과 이야기, 인공지능 에이전트에 생명을 불어넣는 일을 하고 있다. 특히 문자 정보와 상호 작용을 사람이 듣고 말할 수 있는 경험으로 바꿔놓고 있다.

사람처럼 자연스럽게 다국어 지원하며, 실시간으로 하나의 통합 오디오 플랫폼 위에서 작동하는 것이 특징이다. 크게 두 가지로 나뉘는데, 하나는 ‘일레븐크리에이티브(ElevenCreative)’

다. 음악, TTS, 음성인식(Speech-to-Text, 이하 STT)과 자막, 더빙, 음성 복제, 비디오·이미지 생성 등 창작에 필요한 자산을 한 플랫폼 안에서 모두 생성·관리하고 결과물을 확인할 수 있다. 또 음성 에이전트 ‘일레븐에이전트(ElevenAgents)’가 있다. 인공지능이 사람과 대화하며 맥락을 이해하고 과제를 주면 실제 수행하며 문제를 해결하는 에이전트다.

하나의 통합 플랫폼의 강점은 TTS로 연구한 감정 표현력이 그대로 에이전트 목소리가 되고, STT의 실시간 전사 정확도가 그대로 에이전트 경청 능력이 되기 때문이다. 일레븐랩스는 기반기술, 파운데이션 모델에 대한 연구를 지속하고 있다. 현재 주요 고객사는 월트디즈니(Walt Disney Company), 어도비(Adobe), 엔비디아(NVIDIA), 메타(Meta Platforms) 등 주요 산업별 글로벌 기업이다.

③ 마티 스타니스제프스키 일레븐랩스 최고경영자(CEO)와 일문일답

Q. 한국 AI 시장에 대해 어떻게 평가하고 있습니까?

한국은 인공지능 활용률이 높은 나라입니다. 영국보다 도입 속도가 빠릅니다. 일레븐랩스가 일상생활 인공지능 활용도를 분석한 결과 한국은 63%, 영국은 40~50%, 세계 평균은 32%였습니다. 일레븐랩스는 한국 시장에 인공지능 활용 수요가 넘치고 시장이 지속 성장하고 있다는 점을 고려, 진출을 결정했습니다. 한국에서 많은 기업이 빠르게 인공지능 기술을 도입하고 있고, 음성 인공지능 관련 기술자 등 훌륭한 인재가 많다는 점도 한국 진출을 결정하게 된 계기입니다. 언어적으로나 문맥적으로 까다로운 한국어 특화 음성 인공지능 모델도 최고 품질을 지원할 계획입니다.

Q. 영국과 한국 시장에서 일레븐랩스 인공지능 기술 활용 사례를 소개해주세요.

영국에서는 고객경험 지원에 음성 인공지능 에이전트를 활용하고 있습니다. 고객센터에 전화했을 때 음성 인공지능 에이전트가 자동 지원해 주는 경우가 대표적이죠. 영국 정부는 인공지능을 비자나 건강 관리 등 대시민 서비스 개선에 어떻게 활용할 수 있을지 지속 검토하며 개선하고 있습니다. 한국에서는 영화, 텔레비전, 게임에 음성 인공지능이 먼저 적용되고 있습니다. 게임과 사용자 간 상호작용을 지원하고 음성 인공지능 기반 번역·더빙 등으로 K-콘텐츠 글로벌

현지화를 지원하는 게 대표적이죠. 고객센터 등 기업의 대고객 서비스로 한국 내 타깃 시장을 확대하고 있습니다.

Q. 일레븐랩스 인공지능 기술의 강점은 무엇입니까?

일레븐랩스는 인공지능 신뢰성과 투명성을 중시하는 영국 기업답게 안전한 기술 개발과 고품질 기술 확보를 목표로 하고 있습니다. 안전한 음성 인공지능 서비스를 위해 생성물 출처를 추적하도록 하고, 사기나 부정행위를 확인하면 중재나 조율에 나서는 역할까지 맡고 있어요. 일레븐랩스가 제공하는 모든 인공지능 모델은 자체적으로 아무 것도 없는 기초 상태(From Scratch)에서 개발했습니다. 음성 인공지능 분야에서도 거대언어모델(Large Language Models, LLM) 수준의 혁신 서비스가 가능하도록 기술을 고도화하고 있는데요. 지난해 한 글로벌 벤치마크 테스트 결과, 일레븐랩스의 음성 인공지능 모델이 구글(Google) ‘제미나이(Gemini)’ 등 모든 음성 인공지능을 뛰어넘는 기술력을 입증했습니다.

Q. 한국 시장 진출 전에 특별하게 준비한 바가 있다면요?

일레븐랩스가 지원하는 99개 언어 중에서도 정확한 한국어 지원을 위해 한국 기업과 데이터 제휴를 진행했구요. 정보 분류 작업을 통해 무엇이 발화됐는지 전사하고 억양·어투까지 잘 들리게 말하는 작업을 선행했습니다. 문맥에 따라, 같은 문장도 다른 뜻으로 이해되는 한국어만의 어려움이 있는데요. 그 어려움 해소를 위해 의미 주석을 달고 반복 학습 등 미세조정까지 거듭했습니다.

Q. 인공지능이 ‘버블[Bubble]’이라는 일각의 지적에 대해서는 어떻게 생각하세요?

인공지능은 최소한 인터넷과 맞먹거나 인터넷을 능가할 정도의 변화를 불러올 기술이라는 점에서 ‘버블’이라고 보기는 어렵습니다. 일레븐랩스는 정확한 음성 인공지능 기술과 서비스를 제공하는 것은 물론, 고객 경험을 고도화해 개인 맞춤형 고객 응대까지 가능하도록 기술 진화를 이뤄낼 것입니다. 언어에 구애받지 않고 기술과 사람이 상호작용하는 방식을 혁신하겠습니다.



제공: 박종진 기자

④ 홍상원 일레븐랩스 한국지사장과 일문일답

Q. 일레븐랩스의 핵심 음성 인공지능의 경쟁력은 무엇입니까?

일레븐랩스 경쟁력은 크게 세 가지입니다. 첫째, 사람의 호흡과 감정 곡선을 그대로 담아내는 표현력입니다. ‘일레븐 v3’는 단순히 텍스트를 읽는 것이 아니라 문맥을 이해하고 감정 연기를 수행합니다. 둘째, 99개 이상 언어를 동일한 수준의 자연스러움으로 처리하는 다국어 엔진입니다. 셋째, 음성인식(STT), 음성합성(TTS), 음성복제(Voice Cloning), 일레븐에이전트까지 음성 전 영역을 단일 플랫폼에서 제공한다는 점입니다. 그 결과 ‘포춘 500(Fortune 500)’에 포함된 기업 다수가 일레븐랩스를 선택하고 있습니다.

Q. 일레븐랩스가 감정 표현의 자연스러움, 억양의 정교함, 그리고 끊김이 없는 호흡 처리 등 기술적 강점을 확보한 핵심 요인은 무엇입니까?

이것도 세 가지 요인이 있는데요. 첫째, 모델 설계 방식 자체가 텍스트의 의미와 문맥을 함께 학

습하도록 설계돼 있습니다. 같은 문장도 상황에 따라 다르게 발화합니다. 둘째, 호흡·한숨·망설임 같은 비언어적 요소를 별도 처리가 아닌 자연 발화의 일부로 학습시켰습니다. 셋째, 전 세계 전문 성우 데이터를 정식 사용 허가를 받아 학습에 활용하고 자체 연구팀이 모델을 지속 고도화합니다. 연구 조직과 제품 조직이 한 팀으로 움직인다는 점이 다른 회사와의 출발점부터의 차이입니다.

Q. 일레븐랩스는 또 특정 화자의 말투·음색·경향을 학습해 그 사람처럼 말하는 음성을 생성하는 '음성 복제' 기술의 높은 정확도로 주목받고 있잖아요. 향후 해당 분야 사업 방향은 어떻게 설정하고 있습니까?

음성 복제는 단순 복제가 아닌 정체성과 자산의 영역으로 진화하고 있습니다. 방향은 세 갈래입니다. 첫째, 엔터프라이즈(Enterprise)에서 브랜드 목소리, 성우, 아나운서의 디지털 자산화를 지원합니다. 경인방송과 함께한 고(故) 김광한 DJ의 목소리 복원 사례가 대표적입니다. 둘째, 건강 관리 영역에서 발성 장애인을 위한 목소리 복원에 10억 달러 규모 무상 지원을 약속했습니다. 셋째, 목소리 저장소를 통한 성우 보상 생태계를 확장해 권리자가 정당하게 보상받는 구조를 키워가고 있습니다.

Q. 한국과 일본 시장을 모두 담당하고 계신 만큼, 두 시장의 주요 특징과 차이점을 어떻게 평가합니까?

양국 모두 콘텐츠와 엔터프라이즈에서 음성 인공지능 수요가 폭발적입니다만, 서로 결이 다릅니다. 일본은 상장 대기업과 통신사 중심으로 기술검증(Proof of Concept, PoC)을 거쳐 단계적으로 확장하는 보수적으로 도입하는 경향이 강하고, 콜센터 BPO(Business Process Outsourcing) 시장이 거대합니다. NTT(Nippon Telegraph and Telephone), 도쿄해상(東京海上ホールディングス), 트랜스코스모스(Transcosmos)가 좋은 예입니다. 한국은 의사결정 속도가 빠르고 게임, 미디어, 핀테크 영역에서 새로운 기능을 가장 먼저 시도하는 얼리어답터 성향이 강합니다. 다만 두 시장 모두 자국어 음성 품질에 대한 눈높이가 세계 최고 수준이라는 공통점이 있습니다.

〈그림 4〉 올해 5월 서울 삼성동 코엑스에서 열린 'AI 엑스포 2026' 일레븐랩스 부스



제공: 박종진 기자

Q. 한국과 일본 시장에서의 주요 고객군, 파트너십 전략, 그리고 사업 추진 방향은 각각 무엇인가요?

한국에서는 미디어, 게임, 콜센터, 통신사를 핵심 분야로 보고 있습니다. 네이버, 크래프톤 등과 협업 경험을 바탕으로 엔터프라이즈 직판과 채널 파트너십을 병행 검토하고 있습니다. 일본에서는 NTT 도코모(NTT docomo), 도쿄해상, 트랜스코스모스 같은 산업별 리더와의 심층 협업 모델을 추진하고 있고, 통신사와는 네트워크 통합형 음성 어시스턴트도 검토 중입니다. 두 시장 공통으로 직판, 파트너 경유, ISV 임베드(ISV Embed) 세 갈래로 시장 진입 전략을 운영·검토하고 있습니다.

Q. 특히 미디어·콘텐츠 산업 분야에서 일레븐랩스 기술 활용 사례가 많을 것 같은데요?

일레븐랩스는 다양한 언어를 지원합니다. 특정 하나의 국가나 지역에서 생성된 콘텐츠를 글로벌

별 시장으로 가져갈 수 있다는 것이 사업적으로 효과가 큼니다. 음성을 생성하는 분야에서 다 언어를 지원하지만, 더빙 프로젝트 사업도 영위하고 있습니다. 입력된 영상과 음성을 토대로 원하는 언어로 전환해주고 있는데요. 일레븐랩스 서비스 중 하나로 ‘더빙’을 제공하고 있습니다. 더빙은 내부적으로는 초기에 출시된 서비스로, 조만간 차세대 서비스로 새 단장을 계획하고 있습니다. 기존에는 더빙을 해도 완벽한 결과물을 얻을 수는 없었습니다. 사람이 투입돼 번역이 제대로 됐는지, 말하는 동안 번역된 언어가 영상 속 화자의 입술 모양과 맞는지 점검해야 했어요. 부자연스러운 부분을 해소하는 것 역시 사람 몫이었습니다. 수정 버전은 사람의 개입을 최소화하는 방향으로 보완될 것입니다. 더빙으로 지원하는 언어는 현재 총 32개인데요. 더 확장할 것입니다.

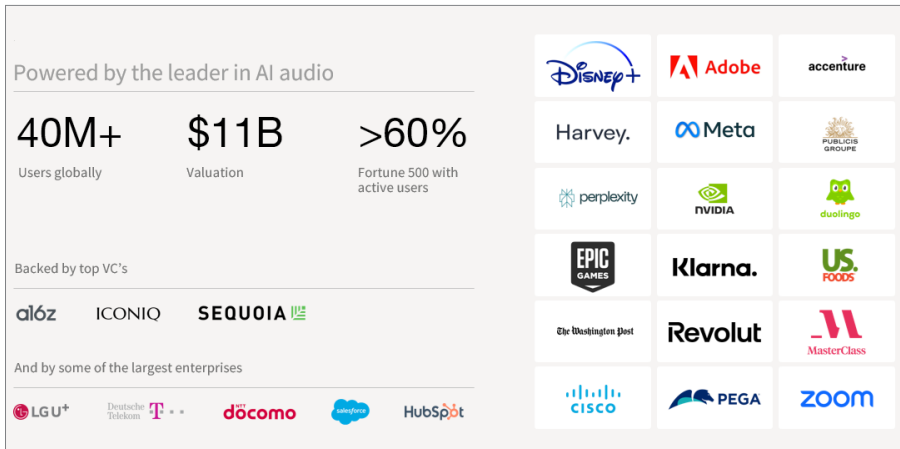
Q. 콘텐츠 현지화는 어느 단계에서 이뤄지나요? 작품 제작 완성 후인가요?

보통 새롭게 만드는 콘텐츠는 글로벌 시장을 염두에 두고 제작을 시작합니다. 스푼랩스(Spoon-Labs) 같은 팟캐스트·오디오북이라든지 인공지능으로 영상 자체를 만드는 것도 포함입니다. 미리 음성을 넣는 게 기술적으로 용이해서는 아니고요. 제작을 시작하기 전에 현지화를 기획하는 게 절차적으로 덜 복잡하기 때문입니다. 기존에 완성된 콘텐츠를 해외로 수출하려면 해당 콘텐츠에 대한 권리자와 관계자가 많아서 수익 등의 분배가 힘들다는 단점이 있습니다. 그래서 현지화할 콘텐츠라면 기획 초기부터 시작하도록 권장합니다. 현지화가 이뤄지면 더 넓은 시장을 목표할 수 있다는 것 역시 강점입니다.

Q. 국내에서 음성 인공지능 수요가 많은 분야가 있을까요?

한국 기업은 교육자료나 영어 등 언어학습에 TTS를 사용합니다. 영상을 만들 때 음성을 TTS로 생성해서 넣는 경우도 있습니다. 이스트소프트(ESTsoft)와의 협업이 대표적입니다. 눈에 보이는 결과물은 아니지만, 시스템에 통합해 활용하는 수요가 가장 많습니다. 인공지능 에이전트 같은 경우는 서비스 제공 시에 시스템에 내재화돼 음성 인프라를 제공하기도 합니다. 또 한국과 일본에서 가장 많이 나타나는 활용 사례 중 하나인데요. 바로 회의 자료 정리입니다. 회의에서 논의된 내용을 속기록처럼 정리하는 게 아니라 텍스트를 자세히 구조화해서 뽑아내는지가 중요한데 그런 측면에서 기술을 지원하고 있습니다.

<그림 5> 일레븐랩스 주요 지표와 고객사



제공: 일레븐랩스

Q. 인간 수준에 가까운 자연스러운 감정 표현이 가능한 음성 인공지능은 언제쯤 실현될 것으로 보십니까?

이미 일정 수준에서는 인간과 구분이 어려운 단계에 도달했다고 봅니다. ‘일레븐 v3’의 경우 짧은 발화에서 블라인드 평가로 구분이 거의 불가능하며, 올해 2월 공개한 익스프레스티브(Expressive) 모드는 분노, 망설임, 즐거움 같은 감정 전이를 실시간 대화에서 재현합니다. 다만 장시간 대화에서 일관된 인격, 상대 발화에 대한 즉각적 공감 반응, 비언어적 신호의 정밀 제어까지 모두 갖추는 것은 여전히 과제입니다. 향후 12개월 안에 일상 대화 영역에서 인간 수준에 매우 가깝게 도달할 것으로 보고 있습니다.

Q. 음성 복제 기술이 딥페이크나 보이스피싱 등으로 악용될 수 있다는 우려가 있습니다. 이러한 문제를 방지하기 위한 대응 방안과 일레븐랩스 내부 기준 및 안전 정책이 있다면요?

안전은 혁신과 분리할 수 없는 가치입니다. 일레븐랩스는 다층 방어 체계로 운영하고 있습니다. 첫째, ‘보이스CAPTCHA’로 본인 음성만 고정밀 복제가 가능하도록 검증합니다. 둘째, 정치인 등 특정 인물에 대한 ‘No-Go Voices’ 차단 목록을 운영합니다. 셋째, ‘인공지능 스피치 클래스피파

이어(Classifier)’로 일레븐랩스에서 생성된 음성을 누구나 검증할 수 있도록 무료 공개하고 있습니다. 넷째, ‘금지된 사용 정책(Prohibited Usage Policy)’ 위반 시 영구 차단과 수사기관 협조를 진행합니다. ‘선거 안전에 대한 기술 협약(Tech Accord on Election Safety)’ 창립 서명사로 글로벌 선거 안전 협력에도 참여하고 있습니다.

Q. 일레븐랩스는 최고 수준의 음성 인공지능 엔진 기업을 지향하고 있는지, 아니면 플랫폼 기업으로 확장을 목표로 하고 있는지 궁금합니다.

두 정체성을 동시에 추구합니다. 출발점은 세계 최고 수준의 음성 인공지능 모델이지만, 이것만으로는 부족합니다. 그래서 TTS, STT, 음성 복제, 음악, 에이전트까지 음성 전 영역을 동일한 연구 기반 위에서 제품화하고 있습니다. 이미 200만 개 이상의 음성 에이전트가 저희 플랫폼 위에서 운영 중이고요. 독일 도이치텔레콤(Deutsche Telekom) 같은 통신사와는 네트워크 임베드 형태로의 인프라화도 진행 중입니다. 엔진의 깊이를 유지하면서 플랫폼으로 확장하는 ‘양수검장’ 전략입니다.

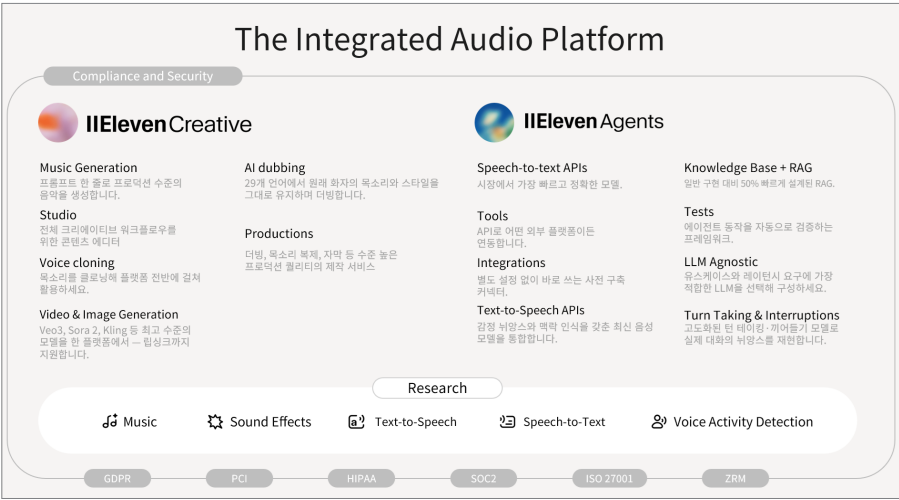
Q. 오픈AI(OpenAI), 구글, 앤트로픽(Anthropic) 등 주요 빅테크(Big Tech) 기업이 거대언어모델과 인공지능 추론 시장에서 점유율을 빠르게 확장하고 있고 음성 사업도 시작하는데요. 일레븐랩스의 차별화된 경쟁력은 무엇입니까?

빅테크는 거대언어모델과 추론 전반을 다루기에 음성은 그들의 부분 기능에 가깝습니다. 일레븐랩스는 음성 한 영역에 모든 연구 자원과 제품 역량을 집중합니다. 결과적으로 표현력, 다국어 자연스러움, 라이브 대화 응답 속도, 음성 복제 정확도 등 핵심 지표에서 일관되게 앞서고 있습니다. 또한, 연구 조직과 제품 조직이 한 팀으로 움직이기에 모델 개선이 곧바로 고객 경험으로 반영됩니다. 무엇보다 거대언어모델이 중립적으로 설계되어 있어서 오픈AI, 앤트로픽, 구글 등 어떤 인공지능 기업 모델과도 결합 가능하다는 장점이 있습니다.

Q. 마지막으로 아시아 시장에서 일레븐랩스 중장기 목표, 올해 한국과 일본 시장에서 달성하고자 하는 구체적인 목표는 무엇입니까?

아시아는 음성 인공지능에서 가장 빠르게 성장하는 권역입니다. 중장기적으로는 한국, 일본, 동남아 주요 시장에서 산업 분야 표준 음성 인프라로 자리 잡는 것이 목표입니다. 올해 한국에서

<그림 6> 일레븐랩스 AI 오디오 플랫폼 개요



제공: 일레븐랩스

는 통신, 미디어, 게임 분야 대형 참고 자료 확보와 함께 5월에 개최된 ‘AI 엑스포 코리아’를 시작으로 시장 인지도를 본격 확장시키고 있습니다. 일본에서는 도쿄해상, NTT 도코모를 중심으로 금융과 통신 분야 엔터프라이즈 사례를 가시화하고 콜센터 BPO 산업에서 공동 솔루션을 만들어가는 것을 핵심 성과 목표로 삼고 있습니다. 한국과 일본 시장에서 음성 인공지능 분야를 이끄는 기업으로 자리매김할 계획입니다.

5 인터뷰이: 홍상원 지사장 소개

GTM(Go-To-Market) 매니저로 일레븐랩스에 합류했다. 한국과 일본을 통틀어 일레븐랩스의 첫 임직원이며 양국 지사장을 겸하고 있다. 리츠메이칸(立命館) 아시아태평양 대학 경영·국제경영학 학사학위를 받았다. 일레븐랩스 합류 전에는 미로(Miro) 일본지사의 첫 임직원으로 합류해 엔터프라이즈 영업 등을 담당했으며, 슬랙(Slack) 일본지사 설립 초기 멤버로 엔터프라이즈营业을 맡았다. 일본 시장 내 글로벌 기업 현지 지사 세팅(setting) 전문가로 자리매김한 것이다. IBM(International Business Machines Corp.) 일본지사에서는 영업 대표로 재직하기도 했다.