

2021. 09

유망시장 Issue Report

음성분석

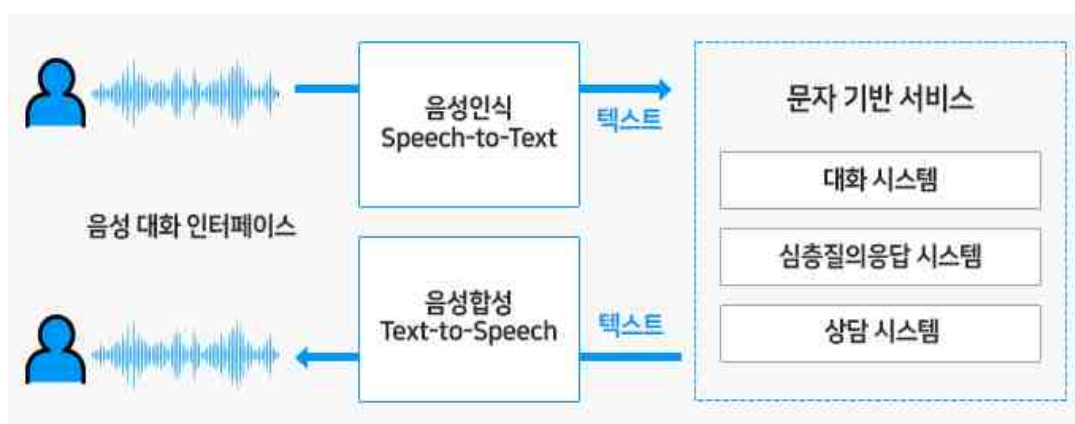
목 차

I . 개요	01
II . 정책동향	06
III . 기술동향	14
IV . 시장동향	23
V . 산업동향	30

I. 개요

1. 아이템 개요

- 음성분석은 상호작용의 음성 콘텐츠를 기반으로 녹음된 음성을 자동으로 검색, 식별, 분류 및 상호 참조할 수 있는 기능을 제공하는 기술을 뜻함
 - 음성분석은 인공지능 및 빅데이터 분야에 속하는 기술로, 사람이 말하는 음성언어를 컴퓨터가 해석해 그 내용을 문자 데이터로 전환하는 처리를 말하며 STT(Speech-to-Text)라고도 함
 - 음성분석은 인간의 자연어 발화를 컴퓨터가 자동으로 이해하고 처리하는 알고리즘을 연구하는 분야로, 자동 통번역, 음성 대화 시스템, 차세대 인터페이스, 질의응답 시스템 등 다양한 응용군을 포함



[그림] 음성분석 기술 개요

※ 출처 : 음성인식 기반 서비스에서의 개인정보처리 실태 분석 및 이용자 보호 방안 연구, 한국인터넷진흥원 2019

- 음성분석 기술은 친숙하고 편리하며 이동 및 작업 중에도 상시 입력이 가능하고, 화자의 고유 정보를 전달할 수 있으며, 입력 속도가 빠르다는 장점이 있음
 - 음성은 인간에게 가장 친숙한 정보 전달 방식으로, 별도의 학습이나 훈련 없이도 일상 제품의 복잡한 기능을 제어하는 데 용이
 - 음성분석 기술은 이동시 정보 입력에 유리하여 모바일 단말 이용이나 운전 중 기기 조작 시 안전성과 효율성 향상에 기여 가능
 - 화자의 음성을 통해 신원, 심리, 건강 상태, 언어 능력 등을 파악할 수 있어 개인별 서비스 제공에도 적합

- 보안 및 금융 분야에서 보안인증 및 신용평가 등에 활용될 수 있으며, 언어 교정 및 외국어 교육에도 사용될 수 있음

[표] 음성인식 기술의 특징

특징	활용 유형	활용 분야
편리성	음성인식은 인간에게 친숙한 정보 전달 방법이기 때문에 별도의 학습이나 훈련 없이도 기기를 손쉽게 사용할 수 있음	가전제품 및 게임 등
상시 입력 가능	손과 발이 자유롭지 못한 상황에서도 정보를 입력할 수 있기 때문에 이동시에도 기기 사용이 가능하며 안전성과 생산성을 제고할 수 있음	모바일 기기 사용 및 물류 등의 작업 수행
화자의 고유 정보 전달	화자의 음성을 통해 신원, 심리, 건강 상태, 언어능력 등을 파악할 수 있어 개인별 서비스 제공 가능	보안 금융 의료 교육 등
고속 입력 가능	입력 속도가 타자에 비해 빠르기 때문에 고속 또는 실시간으로 정보 처리 가능	콜센터, 방송, 통번역

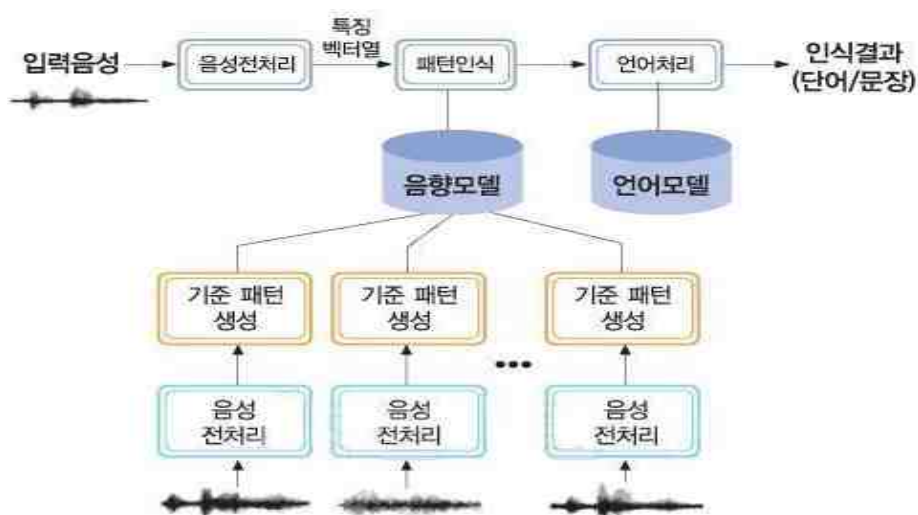
※ 출처 : 인공지능 기술 전망과 혁신정책 방향, 과학기술정책연구원, 2018

○ 음성분석 기술은 음성명령 컴퓨터는 물론 무인 전화번호 안내, 음성명령 주문형 비디오, 각종 음성안내 시스템, 가전제품을 비롯해 자동차 항법장치(Car Navigation System), 홈 오토메이션(Home Automation), 음성 다이얼링 등 이용 영역이 무한

- 음성분석 기술은 이동통신 기기는 물론 TV·냉장고를 비롯한 가정용 전자제품과 로봇 등에도 접목 가능

○ 음성분석의 과정은 크게 전처리부와 인식부로 구분되며, 두가지 처리 과정을 통해 인간과 컴퓨터 간의 자연스러운 의사소통을 가능하게 할 수 있다는 점에서 가치를 지님

- 전처리부는 사용자가 발성한 음성으로부터 인식대상이 되는 구간을 찾아 잡음을 제거하고 인식 과정을 위한 특징을 추출
- 인식부는 입력된 음성의 패턴을 인식하여 음성 데이터베이스와의 비교를 통해 가장 가능성 있는 단어를 인식 결과로 출력하거나 비교 단어를 제한하여 문장을 인식

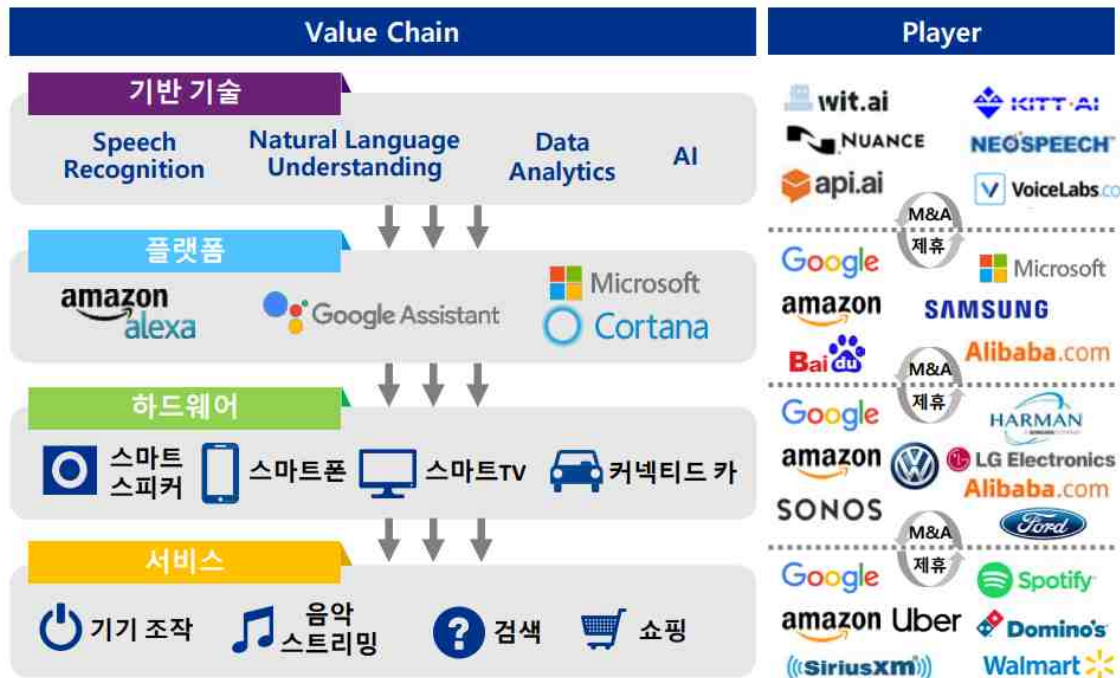


[그림] 음성분석의 과정

※ 출처 : 인공지능 산업동향, 경북테크노파크, 2019

- 딥러닝 및 잡음 처리 기술의 발전으로 음성분석 기술의 한계를 극복하기 위해 사람 간의 자연스러운 대화 음성을 분석할 수 있는 기술 고도화 진행 중
 - 음성신호를 입력받게 되면 실제 화자가 발성한 음성 부분만을 검출하여야 하는데 이 음성 검출 부분은 인식성능에 매우 큰 영향을 미침
 - 음성분석 기술은 소음이 있는 공간에서는 인식률이 급격히 떨어지는 문제점에 노출되어 있으며, 여러 단계에 걸쳐 수행해야 하는 정교한 명령에는 적합하지 않음
 - 특징 추출 과정에서는 불필요하게 중복되는 음성 정보를 없애고 동일 음성 신호들 간의 일관성을 높임과 동시에 다른 음성 신호와는 변별력을 높일 수 있는 기술 고도화 추세

2. Value Chain



[그림] 음성분석 밸류체인

※ 출처 : 음성 AI 시장의 동향과 비즈니스 기회, 삼정KPMG, 2020

○ 음성분석 시장 내 기업들의 가치사슬 구조는 기반 기술, 플랫폼, 하드웨어, 서비스 분야의 형태로 나누어짐

- 각 가치사슬의 구현을 위한 기술은 음성 인식, 자연어 처리, 음성 합성 등이 있으며, 딥러닝, 빅데이터, 클라우드가 기반 기술로 활용

○ (기반 기술 분야) 기반 기술은 음성인식, 자연어 처리와 데이터 분석, 인공지능 등을 말하며, 음성 기술 솔루션이나 원천기술을 보유한 업체들이 주요 플레이어로 활약

- 기반 기술 분야에 해당하는 업체들은 주로 기술을 B2B 라이선스 형태로 판매하기도 하나, 자체적인 음성분석 서비스나 애플리케이션을 보유

○ (플랫폼 분야) 플랫폼 분야의 기업들은 음성인식, 자연어 처리 기술 등을 자체적으로 개발하기도 하지만, 기술 보유기업으로부터 솔루션을 라이선스 형태로 계약해 사용

- 플랫폼 분야에는 아마존 알렉사(Amazon Alexa)나 구글 어시스턴트(Google Assistant)와 같은 음성 AI 비서 플랫폼 기업들이 포진

- **(하드웨어 분야)** 음성분석 하드웨어 업체로는 스마트 스피커 제조업체가 대부분으로, 스마트 스피커 제품은 원통형 기본 모델에서 미니형, 디스플레이 장착형 등 용도별로 광범위하게 포진되어 여러 제품에 장착됨
- 스마트TV는 음성 AI 플랫폼을 탑재하며 냉장고, 에어컨, 스마트 스피커 등을 컨트롤 하는 스마트홈 허브로 발돋움하기 위해 경쟁 가속화
- 운전 중 자동차에서 사용하기 편리하고 안전하다는 장점에 힘입어 커넥티드 카 시스템에도 음성분석 시스템이 활발하게 탑재되고 있음

- **(서비스 분야)** 음성분석 서비스로는 새로운 커머스 플랫폼으로서의 잠재력을 보유한 보이스 커머스, 자동차에서 메인 인터페이스로 활용 가능한 차량 내 음성 서비스, 여러 3차 협력 업체들의 참여로 다양한 서비스 출시가 가능한 확장 서비스 등으로 구분
- 음성분석 기능을 이용하는 서비스 기업은 기본적인 기기 조작 이외에도 음악 스트리밍, 온라인 라디오, 검색, 쇼핑 등 여러 영역에 걸쳐 분포
- 서비스 분야의 주요 플레이어는 아마존, 구글 등 플랫폼 기업부터 월마트와 같은 유통 업체, 피자 체인인 도미노, 음악 스트리밍 업체인 스포티파이(Spotify), 온라인 라디오 서비스 업체인 시리우스XM(Sirius XM), 차량공유 업체 우버(Uber)까지 매우 다양하게 분포

- 음성분석 밸류체인 내 플랫폼 기업과 하드웨어, 서비스 영역 간에는 M&A보다는 제휴와 협력이 활성화 추세
- 스마트 스피커와 같은 단말은 비교적 개발과 생산이 쉽지만, 스마트TV, 커넥티드 키는 시장 진입이 무척 어렵기 때문에 파트너십을 통한 생태계 확대가 활발
- 글로벌 업체의 경우 서비스 전 영역에서 사업을 진행하며 수직 계열화를 통한 시너지 창출에 힘쓰고 있지만, 한편으로는 오픈 플랫폼 전략을 구사하며 3차 협력 업체의 참여를 통한 생태계 확대 주력

II. 정책동향

1. 해외 정책동향

- 음성분석 기술 선도 국가의 경우 자생적 데이터 생성 및 유통구조를 확보하며 네트워크 효과를 달성, 인공지능 기술의 발전과 함께 음성분석의 정교화를 촉진하기 위한 정책 집행 진행
- 최근 미국, 중국, 캐나다, 호주 등 인공지능 기술 관련 앞선 행보를 보이는 주요국들은 경제·산업·사회적 영향을 고려해 각기 대표적인 인공지능 정책을 제시하면서, 인공지능 기술의 R&D 우선순위 상향 조정, 대폭적 예산 확충 등 국가적 역량을 결집
- 국가적 특성에 맞는 정책 거버넌스 체계를 마련하고 정부-민간의 공조를 통해 인공지능 생태계 조성의 효율성·효과성을 제고

[표] 국가별 인공지능(AI) 전략 특징

국가	주요 특징
미 국	▪ 연방정부 R&D에서 AI 우선순위를 높이고 예산을 확대, 적극적인 기초장기 R&D 투자 강화 ▪ 민간시장 주도적 AI 산업이 발전이 가능하도록 관련 부처들이 규제, 표준, 인력 대책 마련 중 ▪ '19년도 AI R&D계획에서 민간-공공 협력 체계 강화 추진
캐 나 다	▪ 글로벌 AI 연구혁신 허브로서의 위상이 높아지는 가운데 '범캐나다 AI 전략(17,3)'을 발표하고 AI 선도지역의 3대 권역의 연구 우수성 센터를 중심으로 한 생태계 확장에 집중 ▪ 세계적인 인재 확보, 국민적 인식 제고, 글로벌 기업의 R&D 활동 유인 등 AI 연구상용화 역량 강화를 위한 기반 마련에 노력
EU	▪ 화자의 음성을 통해 신원, 심리, 건강상태, 언어능력 등을 파악할 수 있어 개인별 서비스 제공 가능 ▪ Horizon 2020¹⁾ 등 AI 관련 R&D 지출을 확대하고, 민간투자가 촉진 될 수 있도록 노력
영 국	▪ '산업전략(17.11)'의 민간협력 방식인 '섹터달' 정책을 AI 분야에도 도입함으로써 AI 분야 민간협력을 최대화하고, 국가적 역량을 결집하는데 주력
독 일	▪ AI 연구개발 투자를 확대하고, 연구성과의 경제적 이전을 강조하는 'AI 국가 전략(18.11)'을 발표하여 인재 육성 및 투자유치와 함께 스타트업, 중소기업 중심의 AI 지원 강화 ▪ EU 및 프랑스 회원국과의 공조를 통해 유럽 전반의 AI 활성화 도모
중 국	▪ 경제의 질적 전환을 위해 국가 전반의 AI 융합 가속화를 가능하도록 차세대 AI 발전규칙('19.05)을 발표하고 AI 혁신 플랫폼, 혁신기지, 전문인력 등에 대한 대규모 투자를 단행 ▪ AI 분야에서 시장원리가 작동할 수 있도록 금융, 표준, 지적권, 법규 정비를 촉진

※ 출처 : 주요국 인공지능 정책 동향과 시사점, 정보통신기획평가원, 2019

1) Horizon Europe 2020은 EU의 과학기술 기반 및 산업경쟁력 강화를 촉진하고, 지속가능한 개발목표(SDGs)를 포함한 글로벌 사회과제를 해결하기 위한 정책을 말함

가. 미국

- 미국은 2011년도부터 NSTIC(National Strategy for Trusted Identities in Cyberspace) 계획에 의해 정부와 민간분야에 보안, 개인정보보호, 사용성, 상호운용성 등을 만족하는 *강한 인증(Strong Authentication) 요구

*강한 인증 : 다수의 인증요소나 요청방법을 사용하여 사용자나 기기를 인증 하는 방법으로 FIDO, OpenID Connect 등이 있음

- 지문, 음성, 얼굴 형상 스캔과 같은 특정 데이터를 수집하기 전에 정보주체로부터 서면 동의 자료를 확보하도록 요구

- 미국 정부는 지속적인 장기 투자를 통해 음성분석, 뇌과학 등 기초원천기술을 확보하고, 인공지능 핵심 기술개발과 인재 양성에 주력

- 미국 국방부는 CALO 프로젝트(2003~2008, 2억 달러 투자) 중 ‘음성 개인비서 연구 부문’을 독립시켜 벤처기업 Siri 설립('07.12)하였으며, 애플이 2억 달러 인수 후 애플의 음성분석 서비스로 탑재('11.1)

- 미국 방위고등연구계획국(DARPA)은 2018년부터는 차세대 인공지능(AI Next) 프로그램을 가동하면서, 현재 20여 개의 핵심적인 인공지능 프로젝트에 20억 달러 이상을 투입

- 트럼프 정부에서는 인공지능 관련 R&D 우선 지원, 인공지능 관련 교육 강화 등 인공지능 기술 발전을 위한 정책추진 방향을 제시하며 미국 *American AI Initiative(2019) 행정 명령으로 통합조정심의 정책 시행 거버넌스를 정비

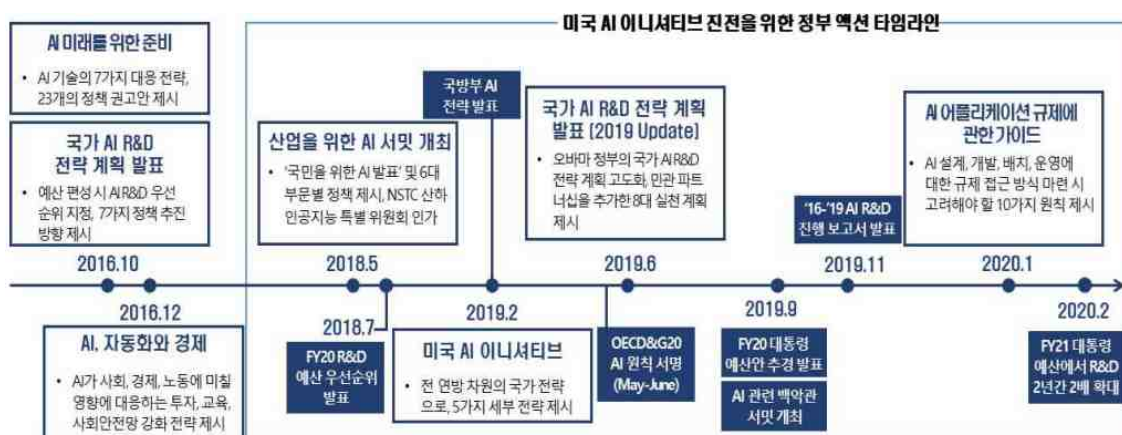
*American AI Initiative : 미국 자체 및 대외 산업 협력을 강화하기 위한 정책으로 인공지능 산업 관련해서 연구개발 투자, 자료 공개, 거버넌스 설정, 인력 양성, 국제 환경 조성, 신뢰할 수 있는 인공지능 수용 등의 세부적인 내용을 포함

- 미국의 연방데이터 및 컴퓨팅 자원을 공공 데이터법(Government Data Act)에 의거하여 AI 연구자 및 산업계 종사자들에게 공개
- 다양한 종류의 기술 및 산업 분야에서 AI 시스템 개발에 대한 표준을 설정하여 AI와 관련된 신기술이 안전하고 신뢰받을 수 있는 방식으로 개발
- 연방정부는 미국의 국가 및 경제 안보 이익에 중요한 AI기술을 해외로부터 보호받을 수 있도록 다양한 보호조치를 마련

[표] 미국 AI 이니셔티브 주요 내용

전략	내용
연구개발 투자	▪ 연방기관들은 장기적인 관점으로 AI 연구개발 투자를 최우선 추진
인프라 개방	▪ 데이터, 모델, 컴퓨팅 리소스를 AI 연구자에게 개방 ▪ 정부 데이터 법(OPEN Government Data Act)을 시행
거버넌스 표준화	▪ 인공지능 시스템 기술표준 개발 ▪ 인공지능 시스템 개발 지침 수립
전문인력 확충	▪ 펠로우십 및 연수 프로그램 운영 ▪ 국민의 AI 잠재력 향상을 위한 교육 확대
국제협력	▪ AI R&D를 장려하고 시장 창출이 가능한 국제적 환경 조성 ▪ 미국의 이익을 보호하고 경제 안보를 지키기 위한 액션플랜 개발

※ 출처 : AI 해외 동향, 경제정보센터, 2020



[그림] 미국 AI 이니셔티브 타임라인

※ 출처 : Office of Science and Technology Policy(2020) 재구성

○ 미국은 국가 AI R&D 전략 계획(2019년)을 발표하여 범분야 기반과 R&D, 산업적용 등 총 3개 부문으로 구성된 전략 체계 추진

- (범분야 기반) 윤리적·법적·도덕적 함의, 안전·보안, 공공데이터 제반 환경, 기술표준·벤치마킹, AI 연구개발인력, 민관파트너십 등 총 8개 전략을 포함
- (R&D분야) AI 진전에 필요한 많은 연구개발 분야를 포함
- (응용분야) 농업, 정보통신, 국방, 교육 등에 걸쳐 영향력 있는 발전과 긍정적인 방향으로 사회에 환원
- AI R&D 전략 계획 연구 우선순위는 기업이 해결하기 어려운 분야에 초점을 맞추고 있으며, 미국의 주도권 유지를 위한 연구개발 투자를 강조



[그림] 미국의 국가 AI R&D 전략계획 체계

※ 출처 : National Science and Technology Council(2019) 재구성

- 바이든 정부는 2021년 연구개발과 혁신적 기술에 3천억 달러(340조 규모)의 추가적인 투자를 통해 전기자동차, 경량화 소재, 5G, 인공지능, 음성분석 기술 등 고부가가치 제조업과 과학기술 분야에서 양질의 일자리 창출 목표 발표
 - 3천억 달러의 연구개발 투자 분야는 첨단소재, 보건·의약, 생명공학, 청정에너지, 자동차, 항공·우주, 인공지능, 통신, 음성분석 등의 산업으로 미국 산업에서 일자리를 창출하는데 기여 예상
 - 중소기업 기술혁신사업(Small Business Innovation Research: SBIR)의 확대된 버전인 ‘미국시드펀드(America’s seed fund)’와 같이 중소기업의 첨단기술 사업화를 촉진하기 위한 경쟁형 자금 지원

나. 중국

- 중국은 “빅데이터 발전 촉진을 위한 행동 지침”, “차세대 인공지능 발전 계획(新一代人工智能发展规划)” 등 다양한 정책을 통해 인공지능 기술의 발전을 촉진하고, 정부가 인공지능 기술 적극 지원
 - 2015년 7월 중국 정부는 “인터넷 플러스 적극 추진”에 관한 행동 지도의견을 통해 인공지능 등 새로운 산업 모델 창출이 가능한 11개 중점 분야 제시
 - 2017년 중국 국무원은 “차세대 인공지능 발전계획(新一代人工智能发展规划)”을 발표하며 인공지능 산업을 국가 전략산업으로 육성하기 시작하였으며, 인공지능 산업과 세부 영역의 발전도 적극 추진

- 목표 달성을 위한 행동계획으로는 음성, 영상, 지도 등의 데이터를 포함한 대규모의 훈련 자원 라이브러리 구축, 스마트 음성처리, 바이오 인식, 신형 휴먼컴퓨터 등 인공지능 핵심기술에 대한 연구개발 및 산업화 촉진 진행

다. 호주

- 호주에서는 방대한 양의 AI 기술 개발을 진행 중이며 주요 AI 연구 분야에서 세계 최고 수준의 역량을 보유
- 호주는 정부 인공지능 준비지수 세계 12위 국가로 블록체인 기반의 핀테크, 음성분석, 몰입형 시뮬레이션 기술을 활용한 에듀테크, 로봇을 활용한 의료 테크, IoT 기반 농업테크 등 ICT 기술력을 보유
- 호주 내 AI 기술이 확산·적용되고 있으며, 전국적으로 AI에 대한 지속적 투자 지원을 진행
- 2018년 5월 호주 정부는 AI 기술 로드맵, 표준 프레임워크, 국가 윤리 프레임워크 개발, 기계학습 발전을 위해 2,990만 달러의 예산을 AI 분야에 할당
- 2020년 10월 남호주는 Adelaide 대학의 ‘AI·머신러닝 연구소’에 710만 달러를 투자하고, 퀸즐랜드는 브리즈번 혁신센터의 AI hub에 300만 달러 투자

다. 캐나다

- 캐나다 정부는 2017년 3월 전세계 최초로 국가 AI 발전전략을 발표하였으며, 각 도시의 스타트업 생태계를 확산시키고, 도시간 산·학·연 네트워크 연계를 강화하기 위해 중점 분야별 ‘슈퍼 클러스터’ 구축 사업(‘Innovation Superclusters Initiative’)을 추진
- 2017년 캐나다 혁신과학경제개발부는 2018년부터 2023년까지 5년간 AI, 농업, 생명과학, 첨단제조업, 해양 등 5개 분야의 연구개발 프로젝트인 Innovation Superclusters Initiative(ISI)를 집중 지원

라. 유럽

- 유럽은 각 회원국의 산업·기술 경쟁력을 제고하고 유럽적 가치에 기반한 신뢰를 담보할 수 있도록 ‘유럽 AI 전략을 마련(2018.04)’
- AI 기술·산업 역량 제고 및 도입 확대, AI로 초래될 사회·경제 변화 대응, 적절한 윤리·법적 체계 확립 등에 중점
- 유럽 전역의 AI에 대한 민·관 투자를 활성화하기 위해 EU 차원에서 AI 연구 지출을 확대하는 한편, 연구 허브도 강화



[그림] EU의 인공지능 전략 추진 경과

※ 출처 : ARTIFICIAL INTELLIGENCE FOR EUROPE FACT SHEET, European Commission, 2019

- 유럽은 ‘AI 윤리 지침 초안(2019.4)’ 을 발표하여 AI 윤리지침 초안 작성, 보호, 학습용 데이터의 품질 확보와 문서화, 개인정보 처리시 데이터 접근 규정을 마련
- EU의 법 규범, 디지털 단일 시장 구현이라는 공동의 목표 하에 AI 산업 활성화를 추진하되 신뢰 기반으로 AI 규범 체계를 우선적으로 정립 하는데 목적
- 책임성, 데이터 거버넌스, 모두를 위한 설계, AI 자율성 거버넌스 비차별, 인간자율성 존중, 개인정보보호 준수, 기술적 엄격성, 안정성 및 투명성 등 인공지능 설계 초기 단계부터 신뢰할 수 있는 요건 구체화

2. 국내 정책동향

- 정부는 2019년 ‘IT 강국을 넘어 AI 강국으로’를 비전으로, 2030년까지 범정부 역량을 결집하여 AI 시대 미래 비전과 전략을 담은 ‘AI 국가 전략’을 발표
 - 데이터·컴퓨팅자원 등 AI 산업의 핵심 인프라 확충을 위해 양질의 데이터 자원 확충을 위한 공공데이터 전면 개방(~2021)
 - 데이터 생산·유통·활용을 지원할 공공-민간 데이터 지도의 연계(~2021)
 - 민간의 AI 개발 지원을 위한 AI 허브의 컴퓨팅 자원 맞춤형 지원
 - 지역 산업과 AI 융합의 거점, 광주 AI 집적단지 조성(~2024)
 - 글로벌 선도국과 대등한 기술산업 경쟁력 확보, 누구나 창업하고 성장해 나갈 수 있는 혁신 생태계 구축



[그림] AI 국가 전략

※ 출처 : 인공지능 국가전략, 정부, 2019

- 정부는 2021년 ‘대한민국 데이터 119 프로젝트’를 추진하여 민간 수요가 높은 핵심 데이터를 다양한 방식으로 제공하며, 대규모 음성·자연어 데이터 구축 및 서비스 기반 제공
 - 일상생활, 상거래, 업무, 생산활동 등에서 한국인의 감성을 이해할 수 있도록 광범위한 대규모 음성·자연어 데이터를 제공
 - 데이터 유형별(이미지, 영상, 음성, 텍스트 등)로 활용이 용이하도록 고품질의 표준화된 형태로 정비·개방
 - 인공지능 기반 문서검색 서비스, 음성인식 비서, 민원 대응 챗봇 등

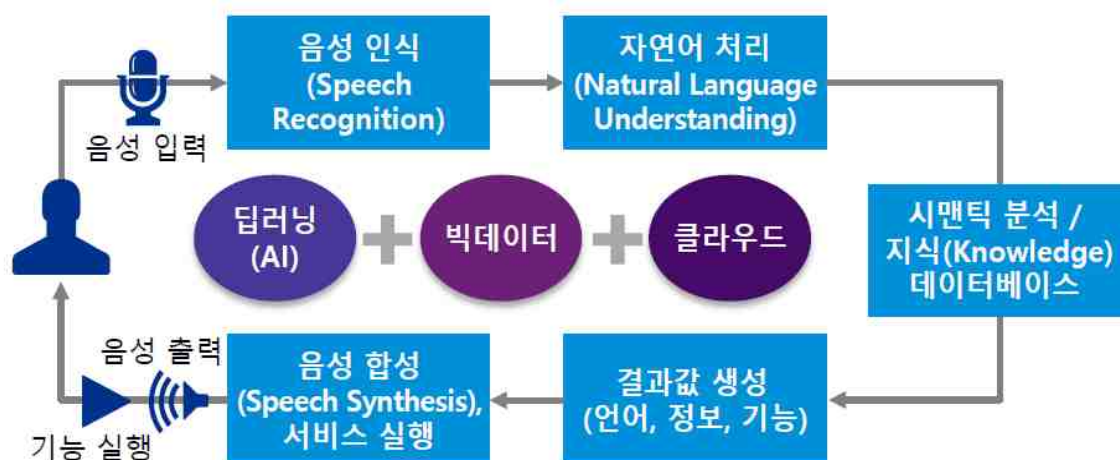
개별 단위 서비스 고도화를 위한 기반으로 활용 가능

- 정부는 음성 정보 및 바이오 정보를 포함한 개인정보를 보호해야 하는 의무 수행과 관련 산업이나 서비스의 발전에 따라 이를 활용할 수 있도록 하는 법제도 개정을 진행
 - 「개인정보 보호법」, 「정보통신망 이용촉진 및 정보보호 등에 관한 법률(약칭 : 정보통신망법)」, 「신용정보의 이용 및 보호에 관한 법률(약칭 : 신용정보법)」 등 3가지 법률을 통칭하는 국내 데이터 3법 개정(2020년 1월 9일)
 - 주요 개정 내용으로는 ①데이터 이용을 활성화하는 가명 정보 개념의 도입, ②유사하거나 중복된 규정을 정리하여 개인정보보호 거버넌스를 일원화, ③데이터 활용에 있어서 개인정보 처리자의 책임을 더욱 강조, ④모호한 개인정보의 판단 기준의 명확화
 - 법 개정으로 인한 민간기업들이 잇따라 간소화된 개인인증 절차로 음성분석 기술 등을 활용하면서 기술 활성화 기대 가능

Ⅲ. 기술동향

1. 기술범위 및 특징

- 음성분석은 크게 언어 이해, 대화 처리, 응답 생성의 순서로 이루어지며, 인간의 언어로 된 사용자 발화를 처리하는 전 과정을 NLP(Natural Language Processing) 또는 자연어 처리라고 함
- (언어이해) 언어 이해는 사용자의 발화를 분석하여 기계가 이해할 수 있는 의미 구조로 변환하는 과정을 말함
- (대화처리) 의미 구조를 적합한 대화 모델을 적용하여 시스템 응답의 의미 구조를 생성하는 것을 말함
- (응답 생성) 기계에서 사용자에게 전달하는 응답 발화나 사용자의 요청을 수행하는 행위를 생성



[그림] 자연어 기반 대화처리 기술 구조도

※ 출처 : 음성 AI 시장의 동향과 비즈니스 기회, KPMG, 2020

- 음성분석 기술은 크게 음성인식, 자연어처리, 시맨틱 분석 기술을 통해 입력받은 음성을 컴퓨터가 분석하고 특징을 추출한 다음, 미리 수집된 음성모델 데이터베이스와의 유사도를 측정해 가장 유사한 것을 문자 혹은 명령어로 변환
- 음성인식 기술은 일종의 패턴 인식 과정으로, 사람마다 목소리와 발음, 억양 등이 다르기 때문에 최대한 많은 사람들로부터 음성 데이터를 수집해 이로부터 공통된 특성을 추출, 기준 패턴을 생성

가. 음성 인식

- 음성분석의 가장 기본적인 기술로 사람의 음성을 텍스트 데이터 형태로 바꿔주는 기술임
 - 음성 입력과정에서 음성인식 정확도를 높이기 위해 녹음 장치로 수집한 데이터에서 잡음과 불필요한 데이터를 걸러주는 전처리 기술도 함께 활용
 - 음성인식 기술은 음성 신호로부터 문자 정보를 출력하는데, 상용 서비스에 적용되는 음향 정보 해독은 확률 통계방식을 사용

나. 자연어 처리 (Natural Language Understanding)

- 음성을 일단 텍스트 데이터로 올바르게 변환한 이후에는, 텍스트를 컴퓨터가 이해할 수 있는 형태로 변환하는 기술
 - 자연어 이해과정은 크게 형태소 분석, 구문분석, 개체명 인식, 화행, 의도분석, 감정 분석, 정보 추출 등 6단계로 이루어져 있음
 - 어절을 최소의 의미 단위인 ‘형태소’로 추출하거나 실제 언어 샘플 데이터인 말뭉치를 활용해 구절과 구문을 분석 진행
 - 자연어 이해과정은 기계와 인간의 상호 작용 과정에서 최우선적으로 이루어져야 하는 부분이기 때문에 자연어 이해과정에 대한 연구들이 활발하게 이루어지고 있으며, 그와 발맞추어 자연어 생성과정에 대한 연구들도 함께 진행

[표] 자연어 이해 과정: 언어 이해와 기계 학습 방법

언어 이해과정	기계 학습 방법
형태소 분석	플랫폼을 통해 연결된 데이터 또는 분석을 통해 획득한 데이터를 활용하여 제공되는 서비스 (스마트시티, 스마트공장, 스마트 의료, 스마트카 등)
구문 분석	문장을 구성 성분으로 분해하여 각 구성 성분 간의 의존 관계와 역할을 분석하며, 현재 딥러닝을 적용한 연구가 활발히 이루어지는 중
개체명 인식	인명, 지명, 기관명 등 의미를 문장 구성 성분에 부여하며 정보 추출, 정보 검색, 질의 응답 등의 목적을 위해 주로 사용
화행, 의도 분석	화행 분석은 사용자의 발화에 나타난 범용적인 의도를 파악하는 것을 말하며 이에 영역 지식이 더해진 구체적인 의도를 파악하는 것을 의도 분석이라 구분
감정 분석	문장에 나타난 긍·부정 표현을 분석하여 사용자의 감성을 파악
정보 추출	텍스트 형태로 변환된 데이터로부터 정형화된 정보를 추출하는 기술로 데이터베이스에 저장하거나 시맨틱 저장소에 저장하여, 빅데이터 분석과 시맨틱 검색 등의 용도로 사용

※ 출처 : 융합연구리뷰, 융합연구정책센터, 2020

○ 자연어 생성과정은 콘텐츠 결정과 텍스트 계획, 문장 계획, 문장 생성의 3단계로 구성

- 자연어 이해를 거쳐 얻은 정보를 기반으로, 어떤 콘텐츠를 어떤 문장을 통해서 사람에게 전달할 것인지 결정

[표] 자연어 생성 과정: 언어 생성과 기계 학습 방법

언어 이해과정	기계 학습 방법
콘텐츠 결정, 텍스트 계획	▪ 사용자에게 어떤 정보를 전달해야 할지(콘텐츠 결정), 정보를 수사적으로 구성하는 방법(텍스트 계획)을 결정하는 것
문장 계획	▪ 각 문장과 문단으로 정보를 나누는 방법과 텍스트를 매끄럽게 만들기 위해 어떤 문장 구성요소들로만들 것인지 결정하는 것
문장 생성(Realized)	▪ 문법적으로 올바른 방식으로 개별 문장을 생성하는 것

※ 출처 : 융합연구리뷰, 융합연구정책센터, 2020

다. 시멘틱 분석

○ 데이터의 의미를 해석하고 검색하는 기술로, 기존에 축적된 지식(Knowledge) 데이터베이스에서 유용한 정보를 검색·선별함으로써 결과값을 생성할 수 있는 기술

- 결과값은 질문에 대한 답변, 요청한 정보나 기능으로 나타날 수 있고, 결과값이 언어적 형태일 경우에는 음성으로 출력되며, 기능일 경우에는 해당 서비스나 조작을 실행

2. 국내/외 기술 Trend

- 음성분석 기술은 1950년대 초기부터 개별 숫자, 음절, 모음을 인식하는 시스템을 만들기 위한 노력이 이루어졌으며, 1963년 IBM이 총 16개의 단어를 인식할 수 있는 “Shoebox” 라는 기기를 소개
 - 음성인식 기술은 1971년부터 시작된 미국 국방부 산하 국방첨단연구사업국(DARPA)의 음성이해연구 프로그램(Speech Understanding Research)을 통해 비약적으로 발전
 - 고립단어(Isolated word)를 인식하는데 한정되었던 음성인식 기술은 1980년대에 들어서야 연결단어(Connected word)를 처리할 수 있게 되었으며 인식할 수 있는 단어도 10,000개 수준으로 크게 늘어남
 - 1990년대는 음성인식 기술이 최초로 상용화된 시기로, 드래곤(Dragon)社は 낱말 단위의 인식과 표현이 가능한 소비자용 음성인식 제품 ‘드래곤 딕테이트(Dragon Dictate)’ 를 출시

[표] 음성분석 기술의 발전 과정

세대	내용
1세대(1952~1968)	- 숫자, 음절, 모음, 음소 인식시스템 개발 시도
2세대(1968~1980)	- 동적시간워핑(Dynamic Time Warping) 기술 개발 - 고립단어(Isolated word) 인식 시스템 개발
3세대(1980~1990)	- 연결단어(Connected word)인식 - 통계학적 프레임워크 적용
3.5세대(1990~2000)	- 노이즈, 마이크로폰, 전송채널, 반향 등에 취약한 음성인식 - 음성인식 오류 최소화(discriminative)를 위한 접근방식 시도
4세대(2000~현재)	- 중요정보의 추적/추출/요약/번역 가능 - 즉흥적 음성인식(일본 CSJ프로젝트) 개발 - 복합(Multimodal)음성인식: 영상 - 음성 정보 동시 인식

※ 출처 : 음성인식 AI 비서 시장의 현황과 시사점, 정보통신정책연구원, 2019

- 현재 사용되고 있는 AI 음성인식 서비스는 대부분 인공신경망(ANN)* 기반 딥러닝 알고리즘이 탑재되어 있으며, 딥러닝 알고리즘에 자연어 처리 등의 기술을 결합

*ANN(Artificial Neural Network): 인공신경망으로, 인간의 뇌를 모방해 분류

또는 예측을 위해 만들어진 머신러닝 알고리즘

- 음성분석이 기존에는 단어 단위로 제한돼 있었다면, 문장 전체를 분석하고 의미를 추론할 수 있는 정확도를 향상 단계에 이름
- 바이두, iFLYTEK 등 중국 음성인식 선도업체들은 많은 사용자의 확보, 방대한 양의 데이터 축적, 그리고 고품질의 알고리즘 기술 등을 바탕으로 2019년에 99%의 음성분석 정확도를 기록

가. 해외 기술 동향

○ 해외 음성분석 연구는 플랫폼 기업 중심으로 이루어지고 있으며, 음성분석 시장 확대를 위한 보안과 프라이버시 문제 해결, 음성분석 성능 향상, 다양한 응용 서비스 개발이 활발

- 음성분석 유망 서비스로는 새로운 커머스 플랫폼으로서의 잠재력을 보유한 보이스 커머스, 자동차에서 메인 인터페이스로 활용 가능한 차량 내 음성 서비스, 3차 협력 업체들의 참여로 다양한 서비스 확장 진행 중

[표] 해외 기업의 주요 음성 분석 AI 에이전트

서비스	출시일	주요 탑재 기기	특징
애플 '시리'	2011년 10월	- 아이폰, 아이패드, 맥PC, - 애플워치 - 애플 스마트 스피커 '홈팟'	- 애플 아이폰 4S에 최초 탑재 후 아이패드, 맥OS로 확대 - 애플 기기에만 탑재되는 폐쇄형 서비스
아마존 '알렉사'	2014년 11월	- 아마존 스마트 스피커 '에코' 시리즈 - 카 인포테인먼트 - 스마트TV	- 초기에 아마존 에코 탑재를 시작으로 적용 단말 확대 중 - 협력 기업의 서비스를 간편하게 추가
구글 '어시스턴트'	2016년 5월	- 안드로이드 스마트폰, 태블릿 '구글 홈' 스마트 스피커 - 스마트TV - 카 인포테인먼트	- 알렉사 스킴과 유사한 확장 기능 '액션' 보유 - 방대한 사용자 데이터 확보로 타사 대비 뛰어난 응답 성능 확보
MS '코타나'	2015년 12월	- 윈도PC - 안드로이드 기기(앱)	- 윈도PC용 음성 AI 비서 서비스로 출발 - 안드로이드 OS 탑재 기기에서 호출 불가하며 앱 설치 방식으로만 작동
바이두 '두어'	2017년 11월	- 바이두 스마트 스피커 - 카 인포테인먼트	- 인공지능 기반의 음성 인식과 검색 기술을 결합한 두어OS를 개방형 플랫폼으로 공개 97% 이상의 정확한 음성 인식 성능

※ 출처 : 음성 AI 시장의 동향과 비즈니스 기회, KPMG, 2020

- 애플은 2010년 시리를 인수한 이후 음성인식 기술 기반의 음성분석 서비스를 제공
 - 2011년 ‘시리’를 출시하여 스마트폰 기반의 서비스로 시작하여 스마트폰 명령 수행에서는 가장 앞서있으나, 쇼핑이나 안내에 대한 응답은 알렉사나 구글 어시스턴트보다 다소 부족함을 보임
 - 2013년 개인비서 서비스 제공 앱 ‘Cue’를 인수함으로써 고도화된 가상 비서 서비스의 기반을 다짐
 - 2015년 10월에 영국 캠브리지에 있는 인공지능 스타트업인 보컬 ‘IQ(Vocal IQ)’를 인수하여 Carplay를 통해 아이폰에서 제공하던 시리를 핸들에서 호출할 수 있는 형태로 제공

- 아마존은 2014년 ‘알렉사’를 출시하여 다른 업체의 서비스를 손쉽게 알렉사의 기능으로 추가할 수 있는 ‘알렉사 스킬’ 보유
 - 알렉사(Alexa)의 어플리케이션을 전면 개편해 2020년 7월 부터 보급을 시작하였으며, 클라우드 학습 기반으로 작동하여 음악재생, 알람설정, 날씨 정보 제공, 교통정보 제공 등 음성인식을 통해 다양한 기능을 제공
 - 다양한 스마트TV, 스마트 스피커, 자동차 인포테인먼트 시스템 등을 확장하고 있는 추세이며, 현재 전세계적으로 AI 플랫폼 시장을 압도적인 점유율로 선도 중

- 마이크로소프트는 자사의 클라우드 ‘애저’에 AI ‘코타나’를 결합하여, 기존 오피스 및 운영체제 제품군을 강화 중이며, ‘애저’를 넷산, 도요타에 자율주행 플랫폼으로 공급
 - 마이크로소프트는 인간 언어를 이해하는 대화형 차세대 컴퓨팅 플랫폼 개발을 발표하여, 인공지능이 단순히 어휘나 문장의 의미를 인식할 뿐 아니라, 맥락과 상황까지 감안해 반응할 수 있도록 개발
 - 마이크로소프트는 아마존과 연합하여 알렉사 검색에 ‘Microsoft Bing’ 검색 엔진을 활용, 알렉사를 통한 Microsoft 오피스 접근과 동시에 MS 코타나를 통한 Amazon 전자상거래 플랫폼 오픈
 - 마이크로소프트의 AI 팻봇인 ‘헬스케어 봇’은 음성분석 기술을 적용하여 2020년 3월 중순부터 덴마크 인구 3분의 1을 책임지는 응급의료 기관

코펜하겐 EMS(Emergency Medical Services)에서 사람 대신 전화응대

- 아마존과 구글과 같은 IT 기업뿐만 아니라, 완성차 업체(OEM)도 자체 브랜드의 인공지능 음성분석 기술인 음성 비서 기능을 탑재 진행 중
 - BMW나 아우디, 포드와 같은 OEM은 자사 브랜드 비서를 아마존 알렉사, 구글 어시스턴트와 함께 탑재하여 운용
 - GM은 IBM의 인공지능 비서인 왓슨을 적용한 ‘OnStar Go’ 를 도입해 2017년부터 탑재된 차량을 출시
 - 벤츠는 ‘MBUX’ 라는 이름으로 차량 인포테인먼트 시스템을 도입하였고, 주로 스마트폰과의 연계, 내비게이션 길 안내나 라디오 등 음원 설정, 문자 확인 등을 제공 가능
- 알리바바는 자회사인 알리바바 클라우드에서 음성분석 기술을 개발 중이며, 음성 외에도 자연어처리(NLP) 기반 챗봇 등 다양한 대화형 AI를 각 산업분야에 접목할 수 있는 기술 개발 중
 - 인터넷 쇼핑 외에도 금융, 행정, 법률 분야에 업체의 음성인식 시스템을 도입되고 있으며, 2020년 공항에 음성분석 서비스를 출시하며 항공표 예약 및 구매에 적용

나. 국내 기술 동향

- (한국전자통신연구원) 한국전자통신연구원은 2016년 한국어 음성인식 기술을 개발하여, 지니톡 대국민 시범서비스 및 2018년 평창 올림픽 서비스 지원을 통해 실증 검증
 - 자동통역서비스, 인공지능 비서, 군사용 음성 인식 기술 등 특정 응용 분야에 대해서는 상용화 수준
 - 한국어에 대하여 자연스러운 발성을 텍스트 문장으로 자동변환하는 ‘한국어 대어휘 자연어 음성인식’ 기술 개발
 - 대표적인 활용분야로써 자동통역 서비스, 인공지능 비서, 딕테이션, 대화형 지능형로봇, 군사용 응용 등에 활용

○ (KAIST) KAIST는 2021년 초고감도 인공지능(AI) 기반 화자 식별 및 음성 보안기술을 만들 수 있는 공진형 유연 압전 음성 센서를 개발

- 원거리에서 스마트 기기들을 정확하게 제어하는 미래 사물인터넷(IoT) 기술과 음성을 암호화하는 보안기술을 연결 가능
- 다수 채널을 보유하기 때문에 AI 음성 서비스에 적은 데이터량으로도 화자 식별 정확도를 높일 수 있어 조건에 따라 오류율을 60%에서 95%까지 줄일 수 있음

○ 국내 음성분석 연구는 대기업 중심으로 이루어지고 있으며, 가전 업체인 삼성전자, LG전자, 통신회사인 SK텔레콤, KT, LG유플러스, 인터넷 플랫폼 기업인 네이버, 카카오가 음성분석 플랫폼 연구 활발

- 기업별로 오픈 플랫폼을 공개 및 클라우드 방식의 AI 플랫폼 개발 도구를 공유하여 개방형 생태계 조성을 선도하려는 추세

[표] 국내 대기업의 주요 음성 분석 AI 에이전트

서비스	주요 기능
LG전자 'Q보이스'	• 자체 AI 브랜드인 씽큐(ThingQ)를 스마트 기기와 가전제품에 탑재하여 출시 • 다른 음성 인식 플랫폼에 비해 기술적으로 부족한 부분은 구글 어시스턴트나 아마존 알렉사 등 다른 음성 AI 플랫폼을 함께 탑재하여 상호 보완적으로 사용하도록 하는 전략을 추진
SK텔레콤 'NUGU'	• 통신업체에서 자체 개발한 '누구'를 스마트 스피커와 함께 출시 • 2019년 오픈 플랫폼을 공개하며 협력을 통한 개방형 생태계 조성을 선도
KT '기가지니'	• 2015년 IPTV 셋톱박스에 스마트 스피커를 결합한 형태로 음성 AI 플랫폼을 처음 선보임 • 2019년 클라우드 방식의 AI 플랫폼과 개발 도구를 개방하여 개방형 플랫폼 생태계 추진
네이버 '클로바'	• 네이버 인터넷 서비스와 함께 지니, 벅스 등과 연동한 복수의 음원 서비스를 이용할 수 있음 • 2020년 AI 고객센터 데모 버전을 출시하여, 성남시 보건소의 코로나 능동감시 대상자를 관찰 하는데 활용
카카오 '카카오'	• 스마트 스피커인 카카오톡에 탑재되고 있으며, 자동차용 내비게이션 애플리케이션인 '카카오내비'와 현대자동차 인포테인먼트 시스템에도 적용 • 카카오는 다음 카카오 계열의 서비스와 연계된 것이 특징 • 2019년 '서버형 음성 인식 서비스'를 신형 소나타에 적용
삼성전자 '빅스비'	• 가전업체인 삼성전자에서 개발한 음성 인식 AI 에이전트 플랫폼 • 스마트폰을 시작으로 삼성에서 출시하는 모든 스마트 기기와 가전제품에 탑재하고 있는 추세로, 대화형 인터페이스와 함께 혼합해 사용 가능 • 2019년 빅스비 캡슐을 개발할 수 있는 플랫폼을 오픈하여 이용 가능한 서비스 확장

※ 출처 : 음성 AI 시장의 동향과 비즈니스 기회, KPMG, 2020

- 국내 자동차 업체인 현대자동차와 기아자동차도 음성인식 기술을 꾸준히 개발하여 현재 상용화 진행중
 - 사용자의 음성을 주변 노이즈와 상관없이 정확하게 인식하도록 하는 Barge-in 기술과 기존의 2~3단계에 걸쳐 조작했던 것을 한 단계로 줄이는 지능형 음성 인식 기술을 발전
 - 자사의 기술력을 바탕으로 홈-투-카(Home-to-Car)와 같은 차량 제어 기능은 물론 카카오의 인공지능 플랫폼 ‘카카오 i’ 와 협업하여 2019년 차량 내에서 사용자가 사용 가능한 음성분석 서비스인 ‘서버형 음성 인식 서비스’ 를 신형 소나타에 적용

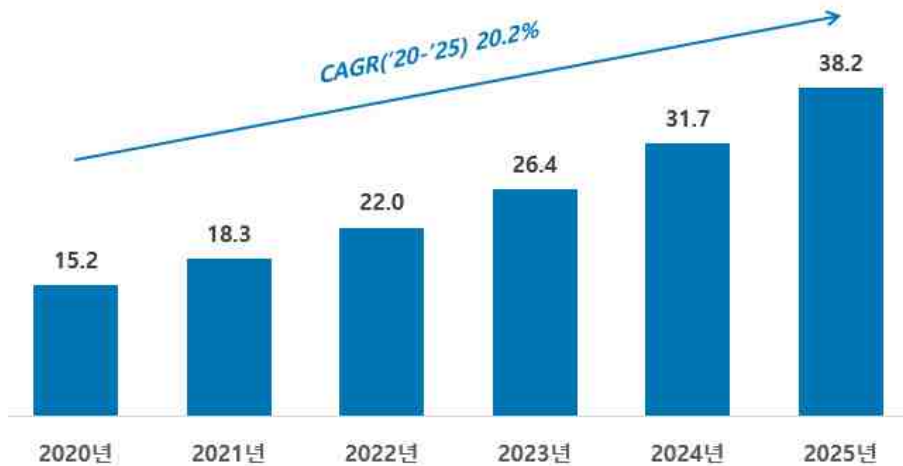
IV. 시장동향

- AI 자동화 시스템 채택이 증가하면서 비즈니스 서비스 분야에 지출이 늘어나고 있으며, AI 애플리케이션을 개발·구현·관리하는 기술 전문성에 대한 고객의 수요가 음성분석 서비스의 확장을 촉진
 - 국가적 차원에서 각국 시민들의 음성분석 데이터를 데이터베이스로 만드는 대규모 프로젝트와 같은 공공 서비스 및 안전을 위한 투자가 증가
- 금융, 리테일 및 HR, 마케팅 등 다양한 분야에서 음성분석 기술 수요의 증가 추세
 - 은행, 보험사, 핀테크(Fintech)와 같은 금융 분야에서 AI가 빠른 속도·효율성·정확성을 제공함으로써 음성기반 बैं킹이 고객 서비스를 향상
 - AI 기반 음성분석 기술을 이용하여 빠른 신원 확인 및 문제 해결을 통해 고객 만족 및 유지를 증대시키고 있으며, 고객에게 계좌 정보 제공·결제·환전 서비스도 가능
- 포스트 코로나 시대 이후 음악 스트리밍, 검색, 온라인 쇼핑 등의 산업 성장 및 딥러닝 기술 발달, 다양한 고객 점점 단말 확대, 유저의 인식 및 행동 변화에 따라 관련 시장 부상
 - 인공 지능(AI)이 음성 및 음성 인식의 정확성에 미치는 영향이 커짐에 따라 소비자 및 엔터프라이즈 부문의 음성 제어 기반 스마트 보조 장치의 증가
- 음성분석 시장 생태계 구축과 선점을 위한 기술, 플랫폼, 하드웨어 업체간 제휴와 경쟁이 심화되면서 음성 인식 기술 기업에 대한 M&A가 활발히 진행
 - 국내외 업체의 향후 스마트홈·IoT 허브 역할을 장악하기 위한 음성 AI 경쟁이 치열해질 것으로 전망

가. 해외 시장 동향

- 전 세계 음성분석 시장은 2020년 15억 2,500만 달러에서 연평균 성장률 20.2%로 증가하여, 2025년에는 38억 2,200만 달러에 이를 것으로 전망됨

단위 : 억 달러

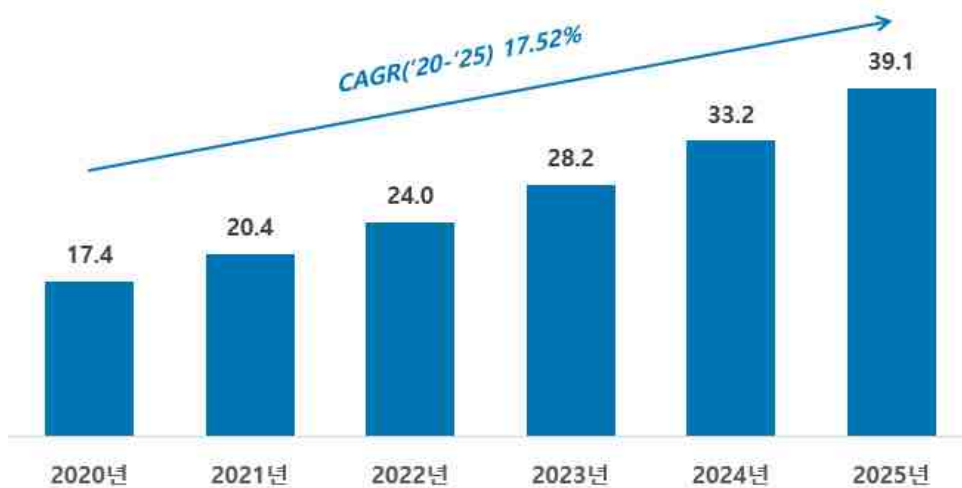


[그림] 글로벌 음성 분석 시장 규모 및 전망

※ 출처 : Speech Analytics Market, MarketsandMarkets, 2020

- 전 세계 음성 회화 분석 시장은 2018년 17억 4,444만 달러에서 연평균 성장률 17.52%로 증가하여, 2023년에는 39억 1,070만 달러에 이를 것으로 전망됨

단위 : 억 달러



[그림] 글로벌 음성 및 회화 분석 시장 규모 및 전망

※ 출처 : Global Voice and Speech Analytics Market, MarketsandMarkets, 2019

○ (세부 항목별 시장규모) 전 세계 음성분석 시장은 구성 요소에 따라 솔루션 및 서비스로 분류됨

- 솔루션은 2020년 9억 9,000만 달러에서 연평균 성장률 19.2%로 증가하여, 2025년에는 23억 8,700만 달러에 이를 것으로 전망됨
- 서비스는 2020년 5억 3,500만 달러에서 연평균 성장률 21.8%로 증가하여, 2025년에는 14억 3,500만 달러에 이를 것으로 전망됨

단위 : 억 달러



[그림] 글로벌 음성 분석 시장의 세부 항목별 시장 규모 및 전망

※ 출처 : Speech Analytics Market, MarketsandMarkets, 2020

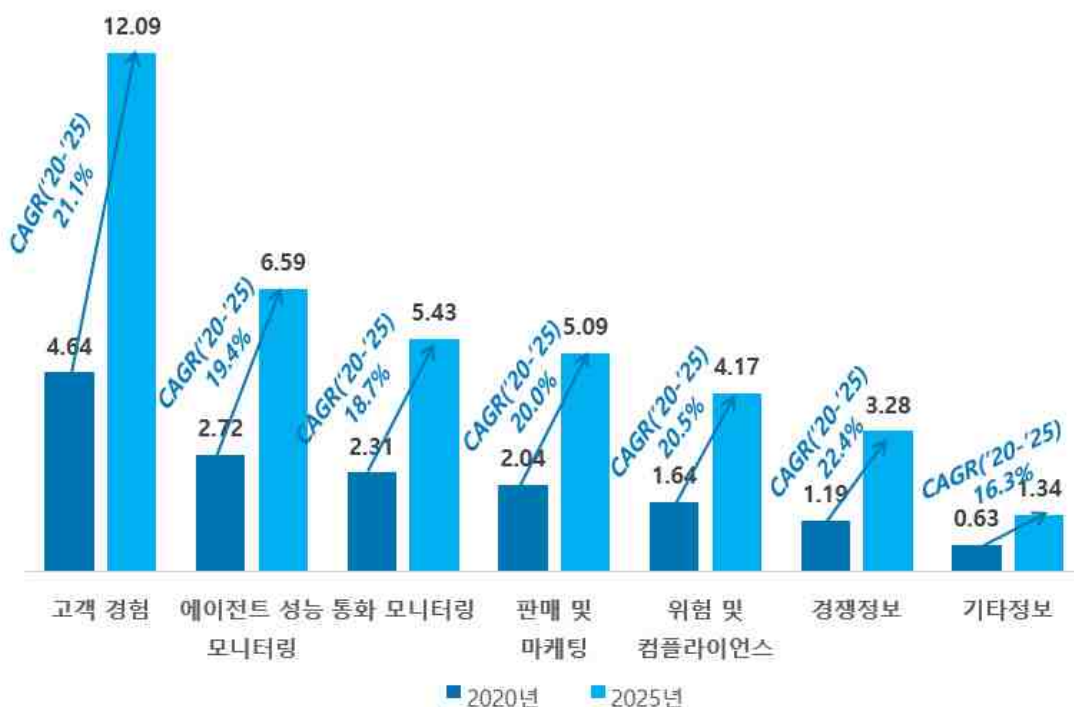
○ (용도별 시장규모) 전 세계 음성분석 시장은 용도에 따라 고객 경험 관리, 에이전트 성능 모니터링, 통화 모니터링, 판매 및 마케팅 관리, 위험 및 컴플라이언스 관리, 경쟁 정보, 기타 용도로 분류됨

- 고객 경험 관리는 2020년 4억 6,400만 달러에서 연평균 성장률 21.1%로 증가하여, 2025년에는 12억 900만 달러에 이를 것으로 전망됨
- 에이전트 성능 모니터링은 2020년 2억 7,200만 달러에서 연평균 성장률 19.4%로 증가하여, 2025년에는 6억 5,900만 달러에 이를 것으로 전망됨
- 통화 모니터링은 2020년 2억 3,100만 달러에서 연평균 성장률 18.7%로

증가하여, 2025년에는 5억 4,300만 달러에 이를 것으로 전망됨

- 판매 및 마케팅 관리는 2020년 2억 400만 달러에서 연평균 성장률 20.0%로 증가하여, 2025년에는 5억 900만 달러에 이를 것으로 전망됨
- 위험 및 컴플라이언스 관리는 2020년 1억 6,400만 달러에서 연평균 성장률 20.5%로 증가하여, 2025년에는 4억 1,700만 달러에 이를 것으로 전망됨
- 경쟁 정보는 2020년 1억 1,900만 달러에서 연평균 성장률 22.4%로 증가하여, 2025년에는 3억 2,800만 달러에 이를 것으로 전망됨
- 기타 용도는 2020년 6,300만 달러에서 연평균 성장률 16.3%로 증가하여, 2025년에는 1억 3,400만 달러에 이를 것으로 전망됨

단위 : 억 달러



[그림] 글로벌 음성 분석 시장의 용도별 시장 규모 및 전망

※ 출처 : Speech Analytics Market, MarketsandMarkets, 2020

○ (전개 형태별 시장규모) 전 세계 음성분석 시장은 전개 형태에 따라 온프레미스(On-premises) 및 클라우드(Cloud)로 분류됨

- 온프레미스(On-premises)는 2020년 11억 3,200만 달러에서 연평균 성장률 19.3%로 증가하여, 2025년에는 27억 3,400만 달러에 이를 것으로 전망됨

- 클라우드(Cloud)는 2020년 3억 9,400만 달러에서 연평균 성장률 22.5%로 증가하여, 2025년에는 10억 8,800만 달러에 이를 것으로 전망됨



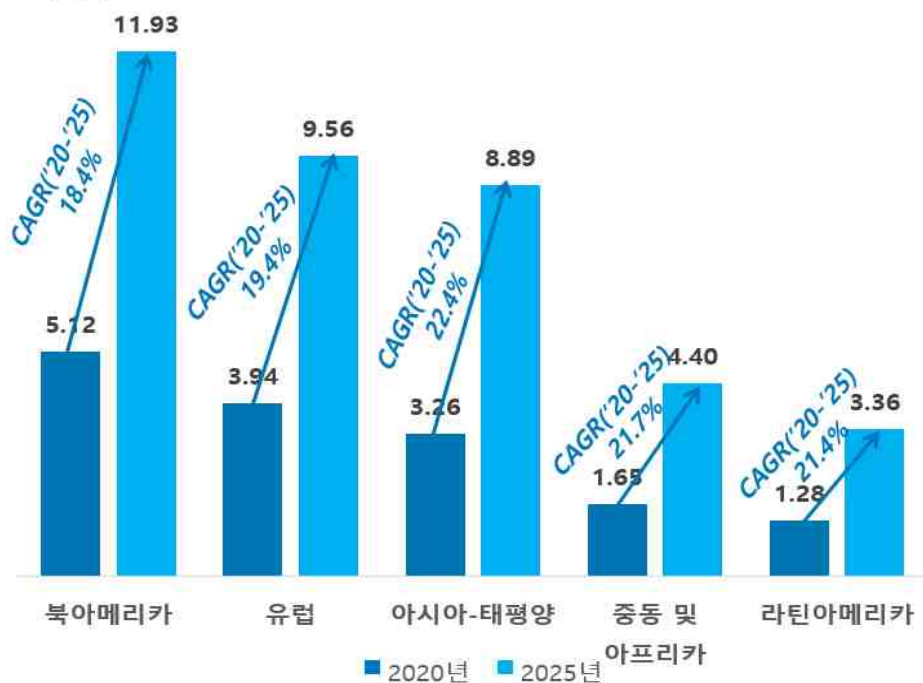
[그림] 글로벌 음성 분석 시장의 전개 형태별 시장 규모 및 전망

※ 출처 : Speech Analytics Market, MarketsandMarkets, 2020

- **(지역별 시장 규모)** 전 세계 음성분석 시장을 지역별로 살펴보면, 2020년을 기준으로 북아메리카 지역이 33.6%로 가장 높은 점유율을 나타내었음
 - **(북아메리카)** 북아메리카 지역은 2020년 5억 1,200만 달러에서 연평균 성장률 18.4%로 증가하여, 2025년에는 11억 9,300만 달러에 이를 것으로 전망됨
 - **(유럽)** 유럽 지역은 2020년 3억 9,400만 달러에서 연평균 성장률 19.4%로 증가하여, 2025년에는 9억 5,600만 달러에 이를 것으로 전망됨
 - **(아시아-태평양)** 아시아-태평양 지역은 2020년 3억 2,600만 달러에서 연평균 성장률 22.4%로 증가하여, 2025년에는 8억 9,800만 달러에 이를 것으로 전망됨
 - **(중동 및 아프리카)** 중동 및 아프리카 지역은 2020년 1억 6,500만 달러에서 연평균 성장률 21.7%로 증가하여, 2025년에는 4억 4,000만 달러에 이를 것으로 전망됨

- (라틴 아메리카) 라틴 아메리카 지역은 2020년 1억 2,800만 달러에서 연평균 성장률 21.2%로 증가하여, 2025년에는 3억 3,600만 달러에 이를 것으로 전망됨

단위 : 억 달러



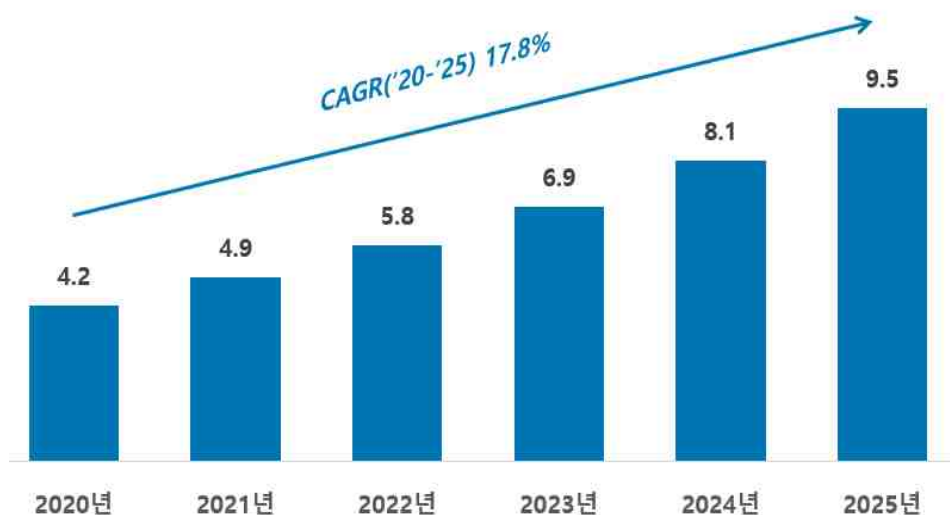
[그림] 글로벌 음성 분석 시장의 지역별 시장 규모 및 전망

※ 출처 : Speech Analytics Market, MarketsandMarkets, 2020

나. 국내 시장 동향

- 국내 음성분석 솔루션 시장은 2020년 4조 2,000억 원에서 연평균 성장률 17.8%로 증가하여, 2025년에는 9조 5,000억 원에 이를 것으로 전망됨

단위 : 조 원



[그림] 국내 음성분석 솔루션 시장 규모 및 전망

※ 출처 : 국내 인공지능 AI시장 전망, 한국IDC, 2020

- 국내 음성분석 시장은 기술 및 인력 측면에서 후발주자지만 우수한 정보통신기술(ICT) 인프라와 교육수준 및 신기술에 대한 빠른 수용성을 통한 빠른 성장 가능
 - 정부와 공공기관이 AI 인재양성 및 생태계 구축에 적극적으로 나서고 투자를 강화하며 디지털 트랜스포메이션을 가속화 추진
 - 현재 보안·교통관제·비즈니스, 인텔리전스 등 음성인식 및 처리가 필요한 분야에 지능형 시스템이 상용화되어 있으며, IP 카메라, 클라우드 기술의 접목으로 시장의 니즈에 대한 수요 대응 가능

V. 산업동향

1. 국내/외 산업 동향

- 음성분석 기술은 하드웨어를 넘어 서비스에 적용되며 급속히 확산되고 있으며, 4차 산업혁명의 핵심 요소인 AI와 결합한 음성분석 기술은 스마트폰 기반 대화형 개인비서, 스피커형 홈 허브, 커넥티드 카, 동시통역, 챗봇(Chatbot), 로봇 등 신산업에 확대 적용
- 음성분석 기술을 확보하기 위해 글로벌 기업들은 벤처기업을 인수하거나 기술 투자를 단행하고 있으며, 이를 통해 기존 사업을 넘어 새로운 사업으로 확장



[그림] 글로벌 인공지능 분야 스타트업 (음성분석 분야 포함)

※ 출처 : Venture Scanner 홈페이지(www.venturescanner.com)

- 음성분석 기술은 AI의 일부 영역인 딥러닝에 의해 비약적으로 진화하고 있어, 플랫폼 비즈니스는 더욱 강화
- 사람의 음성명령을 인식해서 작동하는 가상비서 기능이 다양한 제품에 본격적으로 도입되어 시장 창출과 성장을 견인하고 있으며, 최근 실제 음성인식 기술을 적용한 가전제품, 자동차 등이 연이어 출시
- 아마존, 구글 등 플랫폼 기업들의 지배력이 강화되고 있으며 다양한 제품 검색, 예약, 주문, 결제 등을 장악하여 새로운 비즈니스 진출 진행 중

- 글로벌 기업들은 콜센터 업무의 효율 개선을 위해 음성인식 관련 솔루션을 활용사례 및 음성인식 기반 안내 로봇, 의무기록지 음성인식, 스마트 진료 보조 로봇의 상용화
 - 차이나 모바일(中國電信)은 ' 16년부터 스마트 음성인식 로봇 EVA(移娃)를 고객 콜센터에 도입
 - 음성 문자 변환(STT : Speech to Text), 문자 음성변환(TTS: Text to Speech), 자연언어 이해(NLU : Natural Language Understanding) 등 의 기능이 탑재
 - 고객이 실제 상담사와 통화를 원할 경우 필요한 업무영역의 전문 상담사를 연결해 빠르고 정확하게 서비스를 제공할 수 있도록 설치

- 아마존, 마이크로소프트 등 글로벌 기술 기업들은 음성 지원 제품에 대한 새로운 프로그램인 '음성 상호 운용성 이니셔티브(Voice Interoperability Initiative)' 를 구축(2019년 09월)
 - 음성 상호 운용성 이니셔티브는 ①고객의 개인 정보 및 보안을 보호하면서 다른 사람과 원활하게 작동할 수 있는 음성 서비스 개발 ②여러 개의 동시 발화 단어를 통해 선택과 유연성을 향상시키는 음성 지원 장치 구축 ③단일 제품에 여러 음성 서비스를 보다 쉽게 통합할 수 있는 기술 및 솔루션 출시 ④음성 서비스의 폭, 품질 및 상호 운용성을 향상시키기 위해 머신러닝 및 대화식 AI 연구 가속화 등을 목적으로 구축
 - 음성 상호 운용성 이니셔티브는 아마존, 마이크로소프트, BMW, 세일즈 포스(Salesforce), SFR, 소노스(Sonos), 하만(Harman), 스포티파이(Spotify), 사운드 유나이티드(Sound United), 세렌스(Cerence), 에코비(ecobee), 이노미디어(InnoMedia), 오렌지(Orange's), 텐센트, 바이두, 소니(Sony Audio Group), 버라이즌(Verizon), 컴스코프(CommScope), 마이박스(MyBox), 디스크 비전(DiscVision) 등 글로벌 기업 35개사가 동맹을 구축
 - 이들 기업들은 단일 장치에서 여러 개의 상호 운용 가능한 음성 서비스를 지원하며 한 대의 기기에서 개별 기업들이 개발한 각기 다른 음성인식 솔루션이 공통 지원돼 소비자들이 취향과 목적대로 솔루션 선택 가능



[그림] '음성 상호 운용성 이니셔티브' 가입 기업

※ 출처 : 서울경제, 2019

- 국내외 스타트업 현황의 경우 대부분의 기업들이 자연어 처리를 위해 서버 기반의 AI 음성인식 기술 보유하고 있으며, AI 스피커, IoT, 스마트홈, 챗봇, 로봇 등의 분야로 사업 진출
- 기업의 모든 비즈니스 분야에서 음성분석 기술이 활용되면서 내부 직원, 외부 파트너, 그리고 최종 소비자 단의 경험 향상에 적용되는 과정에서 전반적인 비즈니스 생태계 혁신을 이끄는 필수 기술로 대두

[표] 국내외 음성인식 관련 스타트업 현황

구분	Unisound	바블리	Conversation HEALTH	OTOAI	Talkatoo	마인즈랩
기업 로고						
본사	Beijing (중국)	Toronto (캐나다)	Toronto (캐나다)	New York (미국)	Nova Scotia (캐나다)	성남 (대한민국)
사업분야	Cloud Data Services/ Natural Language/ Robotics	AI/ Voice Assistant/ IoT	AI/ Speech Recognition	AI/ Natural Language/ Speech Recognition	AI/ Natural Language/ Speech Recognition	AI/ Deep learning/ Natural Language/ STT/TTS
설립년도	2012	2018	2019	2017	2019	2014

※ 출처 : 크런치베이스(www.crunchbase.com)

2. 국내/외 기업 동향

가. 해외 기업 동향

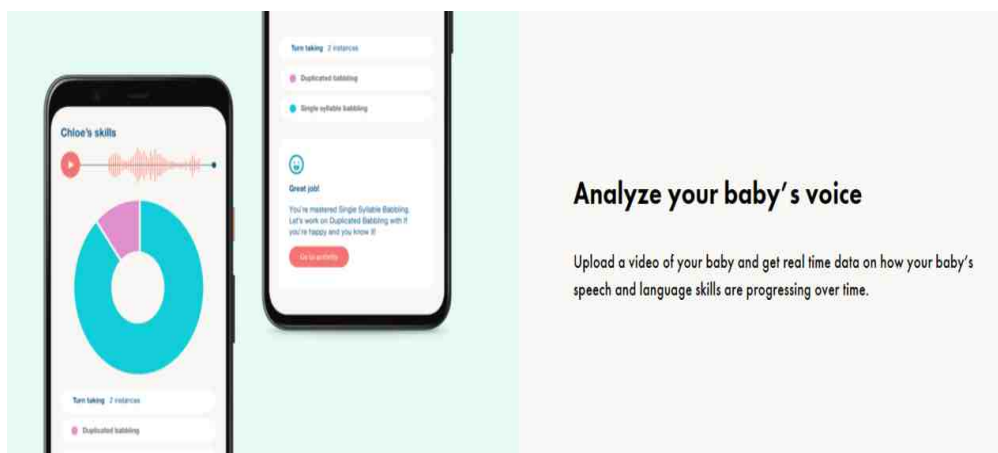
- (유니사운드(Unisound)) 2012년에 설립된 중국 음성인식 및 처리 기술 스타트업으로 클라우드 서비스, 자연어 해석, 로봇틱스 등의 사업을 진행하며 2018년 기준 기업 가치가 10억 달러를 돌파해 유니콘 기업으로 성장
 - 인공지능 음성인식 관련 독자적 기술(지식 재산권)을 보유하고 있으며, 중국 및 해외 100여 개의 특허 및 소프트웨어 저작권을 보유
 - 중국 및 해외 포함 총 2만여 개사와 협력 관계에 있으며, 매달 관련 사용자 2억 명 초과



[그림] 유니사운드의 음성분석 프로그램 구동 장면

※ 출처 : 유니사운드 홈페이지

- (바블리(Babbly)) 2018년에 설립된 캐나다 토론토에 기반을 둔 스타트업으로, 바블리는 영유아 언어 발달 과정을 분석하는 AI 서비스를 제공
 - 자녀가 웅얼이하거나 말하는 영상을 스마트폰 앱에 업로드하면 사용 어휘 수, 언어발달 추이 등 데이터를 확인 가능
 - 소아과 의사, 아동 언어 병리학자 등 전문가가 데이터를 기반으로 자녀 언어 발달에 필요한 조언 기능 보유



[그림] 바블리 서비스 'Babbly' 구동 장면

※ 출처 : 바블리 홈페이지

○ (컨버세이션 헬스(Conversation HEALTH)) 2019년에 설립된 캐나다 스타트업
컨버세이션 헬스는 임상 시험, 의료 업무 및 상업 분야에서 환자 및 의료
전문가를 위한 대화형 에이전트 지원

- 환자와 의료 전문가를 이어주는 대화형 AI 챗봇을 제작하여 환자가
약을 언제, 얼마큼 복용해야 하는지 질문하면 AI 알고리즘이 답변



[그림] 컨버세이션 헬스 대화형 에이전트 구동 장면

※ 출처 : 컨버세이션 헬스 홈페이지

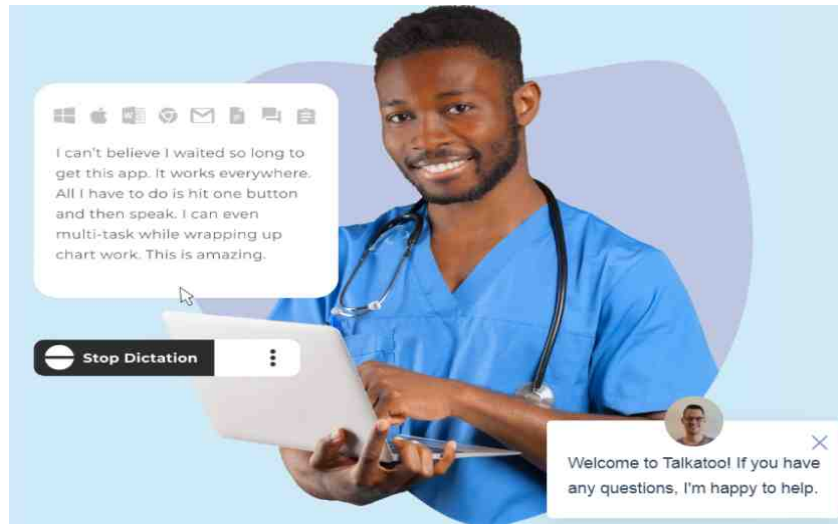
- (오티오.AI(OTO.AI)) 2017년에 설립된 미국 뉴욕에 있는 스타트업으로 오티오.AI는 음성 속 감정을 분석할 수 있는 서비스를 제공
 - 오티오.AI가 개발한 AI 프로그램 딥톤(DeepTone)으로 목소리에서 웃음, 짜증, 피곤함 등의 감정을 포착
 - 녹취 파일을 단순 문장으로만 기록하던 것을 넘어 억양과 뉘앙스 등 감정적 요소도 함께 기록할 수 있게 되어 보건의료, 로봇 공학, 콜센터 등 분야에 폭넓게 이용 가능



[그림] 오티오.AI 딥톤 동작 과정

※ 출처 : 오티오.AI 홈페이지

- (토크투(Talkatoo)) 2019년에 설립된 캐나다 스타트업 토크투는 수의학 분야에 쓰이는 TTS(Text To Speech) 서비스를 제공하며, 음성합성 시스템으로 AI가 사람의 음성을 분석해 텍스트 형식으로 변환
 - 난해한 전문 용어가 활용되는 수의학 분야에 특성화됐다는 장점을 지니며 의학 사전이 내장돼 수의학 전문 어휘를 분당 200개 이상 단어를 기록 가능
 - 인간에 비해 말을 듣고 받아 적는 속도가 5배나 빠르며, 국적마다 다른 억양을 고려하고, AI가 자체 분석으로 적절히 구두점을 추가하는 기능 보유

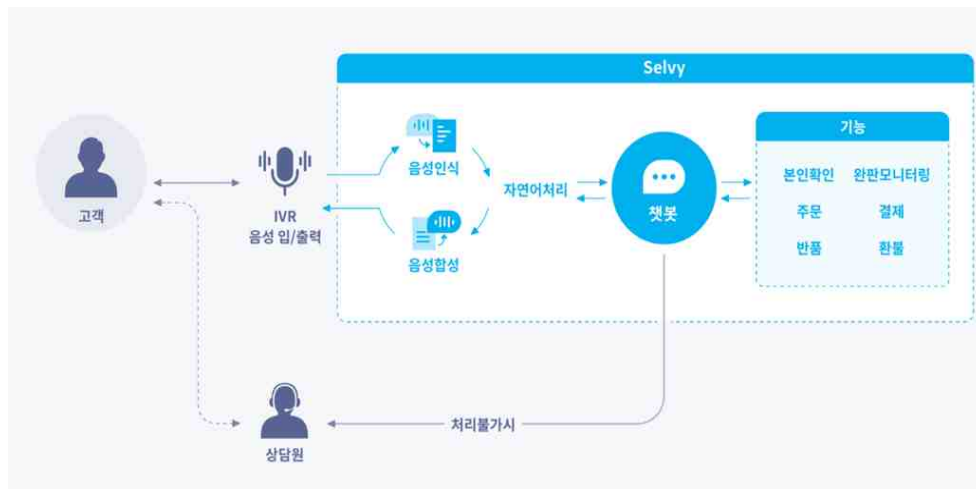


[그림] 토카투 음성인식 서비스 구동 장면

※ 출처 : 토카투 홈페이지

나. 국내 기업 동향

- (셀바스AI) 셀바스 AI는 1999년에 설립된 코스닥 상장 기업으로, 음성지능 (Selvy STT, Selvy TTS), 필기지능(Selvy Pen), 영상지능(Selvy OCR) 서비스 제공
 - 상담 이후 별도의 상담 기록을 남겨야 하는 법률 사무소 내 전문가들은 인공지능 상담, 조사 녹취록 솔루션 도입을 통해 실시간 AI 상담록을 작성 가능
 - 자동으로 회의록이 작성되어 워드, 한글, 텍스트 파일 등 다양한 포맷으로 자료를 관리 및 출력할 수 있어 법무법인 및 법률 사무소의 관련 업무효율이 개선
 - 셀바스 AI의 음성인식 기술은 고립어 및 연속어 음성 인식기술을 활용한 딥러닝 방식을 적용하여 전문언어를 쉽게 인식할 수 있도록 집중적으로 학습해 90% 후반에 이르는 국내 최고의 음성 인식률을 확보

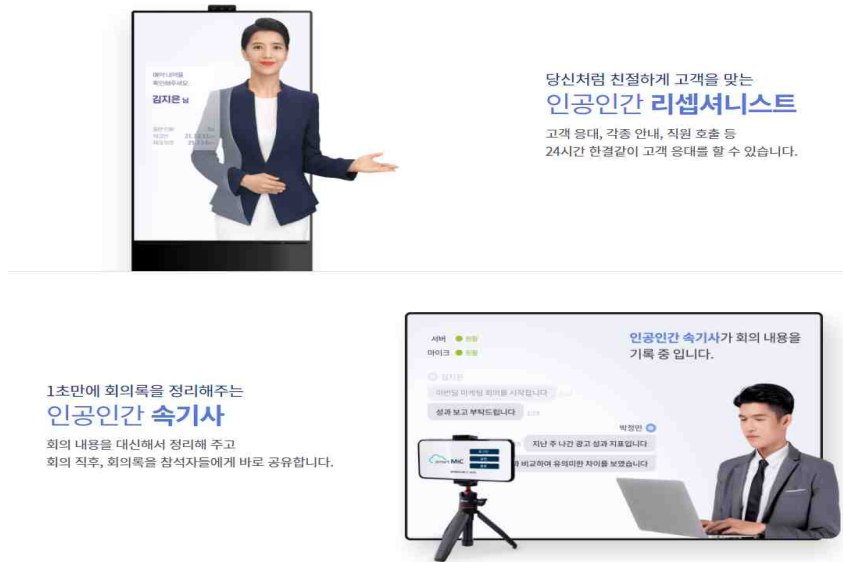


[그림] 셀바스AI 인공지능 음성분석 기술 개요도

※ 출처 : 셀바스AI 홈페이지

- (마인즈랩) 마인즈랩은 2014년 설립된 국내 스타트업으로 소셜 빅데이터 분석 플랫폼 및 AI 보편화를 목표로 출시된 구독형 AI 서비스 ‘마음AI(maum.ai)’는 음성, 시각, 언어, 대화 등 40여 가지의 AI 기술과 솔루션을 API방식으로 제공하여 언제 어디서든 이용 가능
 - 설립 이후 5년 만에 180억원 규모의 Pre-IPO 투자유치와 함께 750억원의 기업가치를 인정받으며 IPO 절차 개시
 - 마인즈랩 기술의 집약체라고 할 수 있는 구독형 인공지능 서비스(AIaaS) 플랫폼 ‘마음AI(maum.ai)’는 음성인식, 자연어 처리, 시각 지능과 같은 인공지능 기술과 솔루션을 모듈화해 광범위한 분야에 맞춤형으로 적용할 수

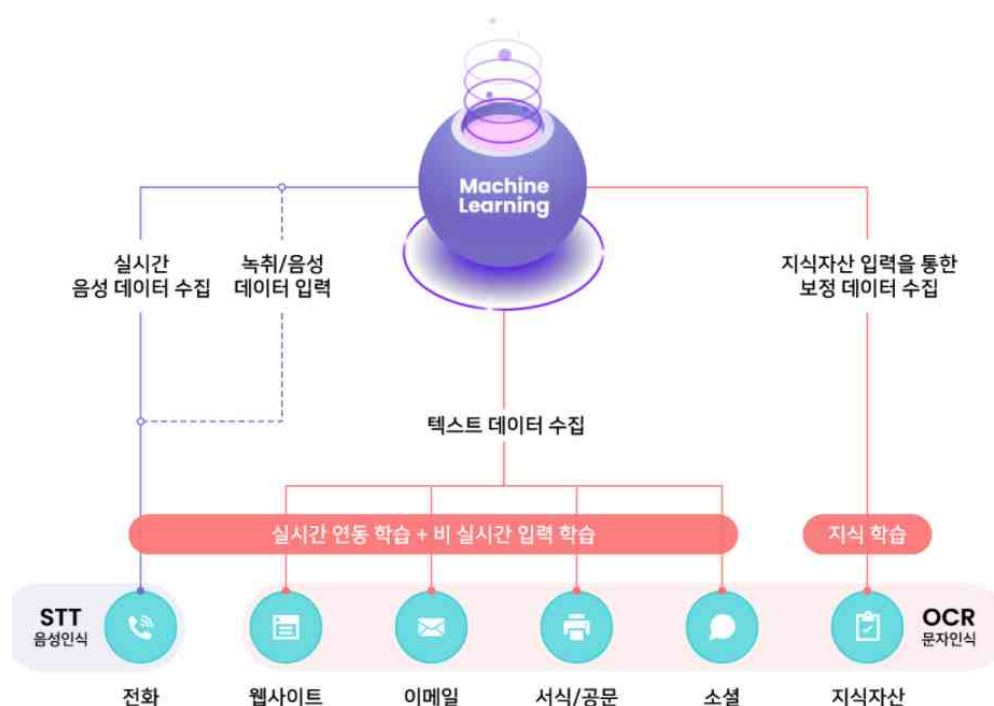
있도록 플랫폼화



[그림] 마인즈랩 마음AI(maum.ai)'서비스

※ 출처 : 마인즈랩 홈페이지

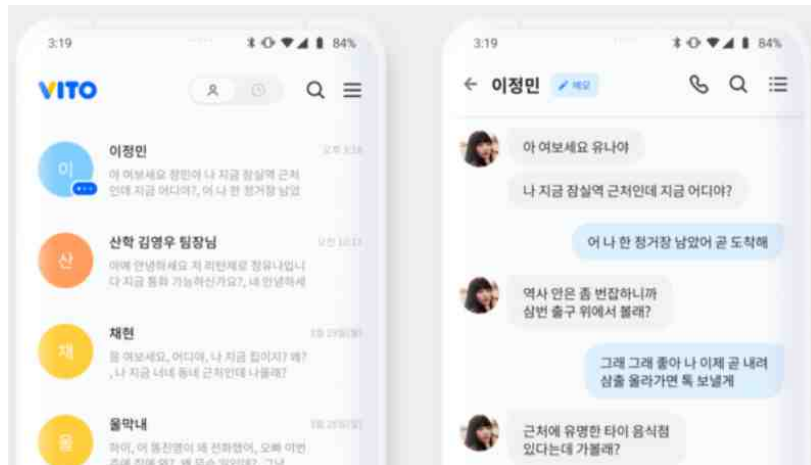
- (두유비) 두유비는 2015년 설립된 실시간 데이터 분석 플랫폼 스타트업으로, 2021년 의료 기록을 실시간으로 녹취하고, 음성을 텍스트로 변환해주는 자동 음성 인식 서비스 ‘스피치EMR(SpeechEMR)’을 출시
 - SpeechEMR은 의료 분야에 특화된 인공지능(AI) 음성 인식 기술을 사용하며, 진료 환경에 따라 최대 95%의 높은 인식률을 보장
 - 음성 인식 서비스는 텍스트로 바꾼 내용에서 오타자를 표시하고 50만 단어의 일상 용어, 10만 단어의 의료 용어 등 의료 전문 음향·언어 모델 및 용어 사전을 활용해 빠르게 내용을 수정하고 저장
 - 현재 두유비의 텍스트 분석 솔루션은 다양한 금융 분야에서 활용되고 있으며 국내 유명 보험사 미래에셋증권, 삼성화재해상보험을 포함한 많은 대형 금융기관 컨택 센터에 사용
 - 한양대병원은 두유비 의료 음성 인식 서비스의 첫 협약 기관이며, 두유비의 주요 분야인 대화 분석 서비스는 대형 컨택센터 시스템 구축 업체와 네이버 등 클라우드 서비스 업체와 협력을 통해 시장을 확대



[그림] 두유비 자동음성 인식 시스템 '스피치EMR(SpeechEMR)' 개요도

※ 출처 : 두유비 홈페이지

- **(리턴제로)** 리턴제로는 2018년 설립된 국내 스타트업으로 전화 목소리를 텍스트로 변환해 채팅처럼 보여주는 서비스 '비토'를 개발하여 통화녹음, 음성인식 및 화자분리, 메신저 형태의 말풍선 화면(UI/UX) 서비스를 한 번에 누를 수 있는 것은 물론, 부분 재생, 검색, 편집, 내보내기, 메모와 같은 부가 기능도 다채롭게 이용 가능
 - 리턴제로의 '비토' 서비스는 국내 모바일 이용자의 75%에 달하는 안드로이드 버전에서 구동되며 암호화, 비식별화 기술로 개인정보 및 프라이버시 보호를 강화
 - 통화 음성을 텍스트로 바꿔주는 STT(Speech To Text) 기술 기반 '소머즈 엔진', 사용자의 목소리를 분석해 화자를 분석하는 '모세 엔진' 등을 자체 개발해 원천 기술을 보유하고 있으며, 휴대폰 용량 부족으로 통화 녹음이 부담스러운 이용자들을 위해 '비토 클라우드' 서비스도 제공



[그림] 리턴제로의 '비토' 서비스(통화 녹음의 메신저화 장면)

※ 출처 : 리턴제로 홈페이지

○ (효성ITX) 효성그룹의 IT 전문 계열사인 효성 ITX는 자체 개발한 음성 분석 솔루션 ‘익스트림 VOC(Voice of Consumer)’와 가상 비서 솔루션 ‘익스트림 AI’를 도입하여 콜센터를 지능형 고객 상담 센터로 발전

- 어드바이저가 통화 내용을 문자로 변환하여 키워드에 맞는 예상 답변을 보여주고, 음성분석을 통해 고객의 감정 변화를 실시간으로 파악하여 관리자가 고객 불만 및 긴급 상황에 신속히 대응할 수 있도록 돕는 시스템을 제공
- 상담 내용을 실시간 분석하여 교육청, 학교, 학부모 등 서비스 이용자들의 상담 편의를 향상시키고, 서비스 품질을 개선 가능

[참고문헌]

- 음성 AI 시장의 동향과 비즈니스 기회, KPMG, 2020
- 바이든 신행정부의 주요 정책 전망과 시사점. 국회입법조사처, 2021
- 음성인식 AI, 한국전자통신산업진흥원, 2021
- 글로벌 산업기술 주간 브리프, 산업통상자원부, 2020
- 중국, 인공지능 기술 발전 요인 및 과제 분석, 하버드비즈니스리뷰, 2021
- AI 해외 동향, 경제정보센터, 2020
- 음성인식 기반 서비스에서의 개인정보처리 실태 분석 및 이용자 보호 방안 연구, 한국인터넷진흥원, 2019
- 음성인식 AI 비서 시장의 현황과 시사점, 정보통신정책연구원, 2019
- 인공지능 산업동향, 경북테크노파크, 2019
- 인공지능 기술 전망과 혁신정책 방향, 과학기술정책연구원, 2018
- 호주, AI 로드맵 주요 내용, 한국정보화진흥원, 2020
- 포스트 코로나 시대, 4차 산업혁명은 어디로?, 대한무역투자진흥공사, 2020
- KOSME 산업분석 Report, 융합금융처, 2019
- 혁신성장동력 시행계획, 관계부처 합동, 2018
- 인공지능 분야 국내외 동향 분석 및 투자전략 수립, KISTEP, 2018
- 인공지능 기술의 진화와 AI 반도체·컴퓨팅의 변화, 정보통신정책연구원, 2019
- Speech Analytics Market, MarketsandMarkets, 2020
- ARTIFICIAL INTELLIGENCE FOR EUROPE FACT SHEET, European Commission, 2019
- www.venturescanner.com
- www.cruchbase.com