

음성인식 AI 비서 시장의 현황과 시사점

■ 최지혜* · 이선희**

스마트폰에 시작된 음성인식 기술 경쟁은 사물인터넷(IoT)의 본격 확산과 맞물려 이제 집 안에서 본격적으로 불붙을 전망이다. 특히 주목할 만 한 점은 그 기기가 음성을 매개로 명령을 내리고 대화를 나눌 수도 있는 인공지능(AI) 스피커라는 점이다. 2014년 Amazon이 최초의 음성인식 AI 스피커 Echo를 출시한 이후 Google, Apple, Microsoft 등 글로벌 IT 공룡들이 제각기 자사 소프트웨어를 탑재한 스피커를 출시를 결정했으며, 국내 시장도 경쟁이 가속화되고 있다. 본고에서는 음성기반 AI 서비스와 스마트 기기의 기술적 특징 및 시장 전망을 살펴보고, 그 중에서도 AI 스피커를 출시하며 관련 시장을 선점한 Amazon과 Google을 비롯하여 국내외 주요 기업의 현황과 제품 전략을 정리하였다. 또한 음성기반 서비스의 본격 확산에 앞서 예상되는 주요 이슈를 검토하고 이에 대한 시사점을 도출하고자 한다.

목 차

I. 서 론 / 2

II. 음성인식 기반 AI 기술 개요 / 4

1. 음성인식 기술의 발전 과정 / 4
2. 음성기반 사용자 인터페이스의 특징 / 7
3. 음성인식 AI 스피커의 차별화 요소 / 10

III. 주요 기업의 음성인식 AI 스피커 전략 / 11

1. 해외: Amazon, Google, Apple, Microsoft, 기타 / 12
2. 국내: 삼성, SKT, KT, 네이버, 카카오 / 25

IV. 결론 및 시사점 / 31

* 정보통신정책연구원 ICT통계정보연구실 연구원, (043)531-4407, jihchoi@kisdi.re.kr

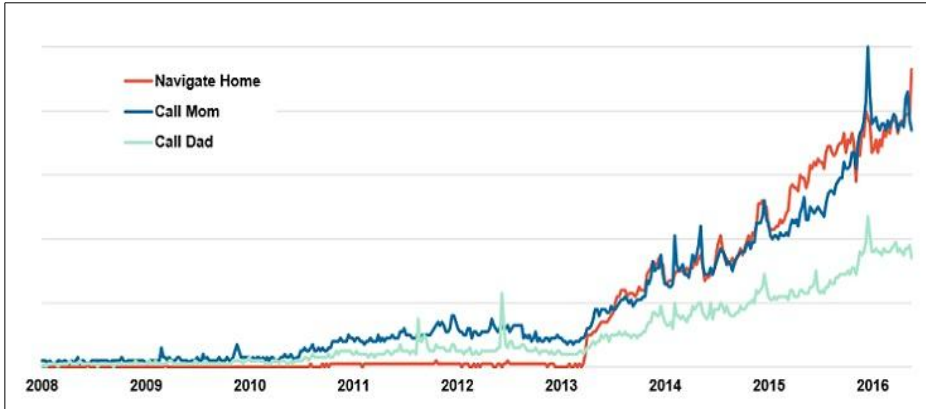
** 정보통신정책연구원 ICT통계정보연구실 연구원, (043)531-4088, imediagod@kisdi.re.kr

I. 서 론

1968년 개봉한 스탠리 큐브릭 감독의 “2001: 스페이스 오딧세이”에 등장하는 인공지능 슈퍼컴퓨터의 기종명은 ‘발견법적으로 프로그램 된 연산 컴퓨터’(Heuristically Programmed ALgorithmic computer), 줄여서 HAL(또는 HAL9000)이다. 임무 수행을 위해 목적으로 향하는 유인탐사선 디스커버리호에 탑승한 승무원은 자연어로 HAL과 대화를 나누거나 명령을 내릴 수 있다. 70년대 후반 미국의 SF 시리즈 “스타트렉”에 등장하는 우주함선 U.S.S. 엔터프라이즈호의 슈퍼컴퓨터도 키보드나 마우스 같은 입력장치를 필요로 하지 않는다. 커크 선장과 엔터프라이즈호 선원들이 “Computer”라고 음성으로 호출한 후 질문이나 명령을 하면 컴퓨터는 이에 대응한다.

“음성”을 이용하는 것은 인간과 컴퓨터의 상호작용에 가장 이상적인 인터페이스로 여겨져 왔다. 음성이야말로 인간이 가진 가장 본질적인 커뮤니케이션 도구이기 때문이다. 음성인식(Speech recognition) 관련 연구의 시작점은 1954년 IBM과 조지타운 대학이 공동으로 참여한 기계번역(Machine-translation) 기술 개발 프로젝트로 거슬러 올라갈 수 있다(The Economist, 2017. 5. 1). 그러나 방대한 음성데이터 처리를 위한 컴퓨터의 연산 능력 부족과 낮은 인식률로 말미암아 음성인식 기술은 2000년 중반까지도 일반인에 널리 사용되지 못하였다. 이 기술이 본격적으로 상용화되기 시작한 시점은 “Siri”를 탑재한 2011년 10월 아이폰4S 출시 이후로 보아야 할 것이다. Google도 2013년 출시한 안드로이드4.4 KitKat OS 부터 Google Now 애플리케이션을 통해 음성검색 기능인 “OK Google” 서비스 지원을 시작했으며, 이 서비스는 최근 인공지능 기반 비서 애플리케이션인 “Google Assistant”로 업그레이드 되었다. 한국어 지원은 아직 이루어지지 않고 있으나 구글은 최근 2017년 말까지 한국어를 포함한 4개 언어의 추가 지원 계획을 발표하였다. 정체되어있던 Google의 음성검색(Voice search) 이용 빈도는 “OK Google”기능이 도입된 2013년 이후 빠른 속도로 늘어나고 있다([그림 1]참조).

〔그림 1〕 2008~2016 음성검색 Google Trend(전 세계)



자료: Kleiner Perkins Caufield & Byers(2016. 5), Google Trend 재인용

스마트폰에 시작된 음성인식 기술 경쟁은 사물인터넷(이하 “IoT”)의 본격 확산과 맞물려 이제 우리 집 안에서 본격적으로 불붙을 전망이다. 지난 1월 미국 라스베이저스에서 열린 CES 2017¹⁾에서 발표된 “Tech Trends to Watch”에서도 올해의 핵심 기술 중 하나로 음성인식·제어(The New Voice of Computing)를 꼽았으며, 세계적 IT 뉴스 제공업체인 ZDNet역시 “음성은 차세대 컴퓨터 인터페이스”라는 분석 기사를 내놓았다(ZDNet, 2017. 1. 4). 또한 이번 전시회의 가장 큰 화젯거리는 미국의 전자상거래 업체인 Amazon이 개발한 음성기반 인공지능(이하 “AI”) 소프트웨어 “Alexa”와 자체 스피커 단말기인 “Echo”였다. 이를 두고 IT전문 온라인매체인 TechCrunch는 “Amazon과 Nvidia는 어떻게 올해 CES에서 승리했는가”²⁾라는 제목의 기사를 싣기도 했다.

본고에서는 음성기반 스마트기기의 기술적 특징 및 시장 전망을 살펴보고, AI 스피

1) CES(The International Consumer Electronics Show)는 미국 600여개의 소비재 전자산업 종사자 사업 연합체인 CEA(Consumer Technology Association)이 주관하는 세계 최대의 가전·IT제품 전시회로 독일에서 열리는 IFA(Internationale Funkausstellung), 스페인 MWC(Mobile World Congress)와 함께 전 세계 가전업계의 트렌드를 한눈에 파악할 수 있는 세계 3대 IT 전시회로 꼽힌다.

2) 원 제목은 “How Amazon and Nvidia won CES this year”,(TechCrunch, 2017. 1. 8)이다.

커를 출시하며 관련 시장을 선점한 Amazon과 Google을 비롯하여 국내외 주요 기업의 현황과 제품 전략을 정리할 것이다. 또한 음성기반 기기와 서비스의 본격 확산에 앞서 예상되는 주요 이슈를 검토하고 시사점을 도출할 것이다.

Ⅱ. 음성인식 기반 AI 기술 개요

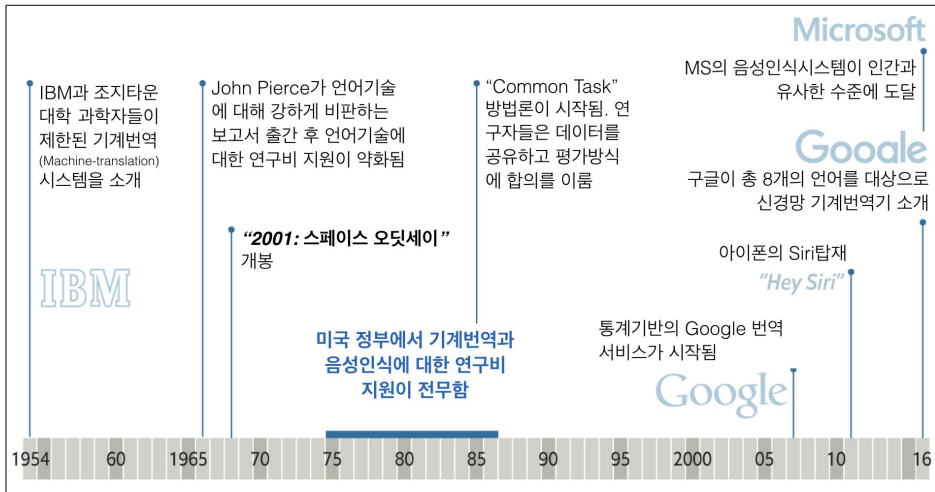
1. 음성인식 기술의 발전 과정

언어와 음성에 대한 연구는 앞서 언급한 바와 같이 1950년대로 거슬러 올라갈 수 있다. 초기에는 개별 숫자, 음절, 모음을 인식하는 시스템을 만들기 위한 노력이 이루어졌으며, 1963년 IBM이 총 16개의 단어를 인식할 수 있는 “Shoebbox”라는 기기를 소개하였다. 그러나 초기 음성인식 기술 연구를 주도해 온 AT&T Bell Lab의 John Pierce가 “자동 음성인식 시스템은 인공지능을 필요로 하기 때문에 향후 수십 년 간 실체가 없을 것”이라 비판한 이후 음성인식 기술 연구 개발에 대한 정부 지원이 급격히 약화되기도 하였다([그림 2] 참조). 그럼에도 불구하고 음성인식 기술은 1971년부터 시작된 미국 국방부 산하 국방첨단연구사업국(DARPA)의 음성이해연구 프로그램(Speech Understanding Research)을 통해 비약적으로 발전하는 계기를 얻었다(한국 콘텐츠진흥원, 2011. 11). 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 늘어났다(Furui, 2005). 또한 미국의 카네기멜론대학(CMU)이 불특정화자의 목소리를 인식할 수 있는 음성인식 소프트웨어 “SPHINX”를 소개하고, 은닉 마르코프 모델(Hidden Markov Model)³⁾로 대표되는 통계적 프레임워크가 이용되기 시작했으나 컴퓨터의 처리 속도 및 저장 공간 문제로

3) 은닉 마르코프 모델: 발화된 음성이 어떤 문자열로부터 도출되었는지 판단하기 위해 관찰된 음성 신호를 가장 잘 설명하는 최적해(문자열)를 찾는 2차 통계모델로 음성인식 분야에서는 은닉상태의 문자열을 찾기 위해 음성신호 변동이 확률 변수로 취급된다.

음성인식 기술의 상용화는 요원했다.

[그림 2] 언어기술(Language Technologies)의 발전과정



자료: The Economist(2016. 5. 1), 저자 정리

30년 전 John Pierce의 분석대로 음성인식 기술 기술은 딥러닝, 머신러닝과 같은 AI 기술이 활용되면서부터야 음성인식 오류율이 낮아지고([그림 3] 참조), 본격적인 상용화가 이루어지고 있다. 미국의 벤처캐피털 Kleiner Perkins Caufield Byers(KPCB)는 2016년 발간한 보고서에서 음성 데이터의 의미 이해 정확도가 99% 수준에 도달한다면 대부분의 사람들이 음성인식 서비스를 이용할 것이라 전망하였다(KPCB, 2016. 6). AI 기술을 활용하기 위해서는 대용량의 데이터를 저장하고 빠르게 처리할 수 있는 컴퓨팅 자원을 필요로 하는데, 클라우드 컴퓨팅 기술에 힘입어 이것이 가능해졌다. 종전의 OS기반 플랫폼에서는 정보의 생산, 처리, 저장 및 이용이 기기 자체에 한정되어 이루어졌지만, 이제는 이러한 과정에 클라우드에서 이루어지면서 ICT 사용 환경이 크게 변화하였다(김대호 외, 2014; 최은정, 2013). 특히 실시간으로 유사 음성을 검색하고 대용량의 음성 데이터를 빠른 속도로 분석할 수 있게 되면서 정확도가 획기적으로 높아졌다(김대호 외, 2014).

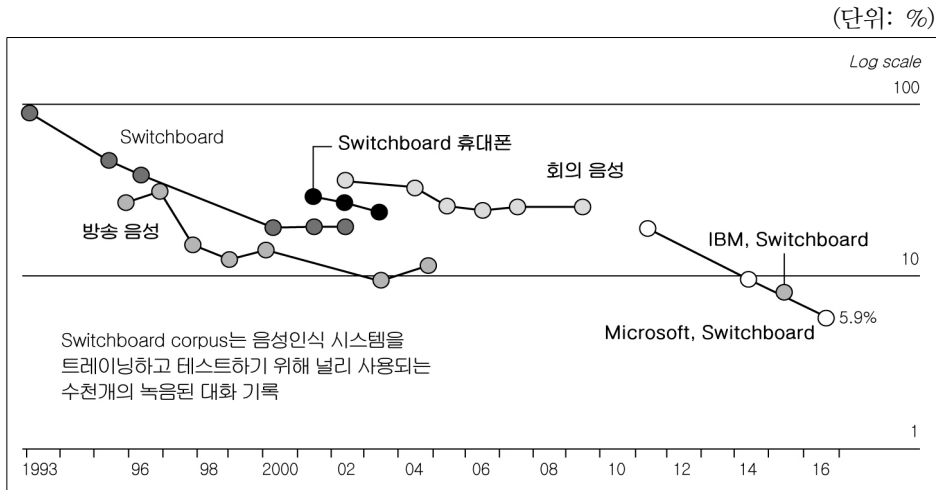
〈표 1〉 음성인식 기술의 세대별 구분

세대	내용
1세대(1952~1968)	- 숫자, 음절, 모음, 음소 인식시스템 개발 시도
2세대(1968~1980)	- 동적시간워핑(Dynamic Time Warping) 기술 ⁴⁾ - 고립단어(Isolated word) 인식 시스템
3세대(1980~1990)	- 연결단어(Connected word)인식 - 통계학적 프레임워크 적용 - 은닉 마르코프 모델, n-gram ⁵⁾ , Neural net - DARPA 프로그램: SPHINX(CMU), BYBLOS(BBN), DECIPHER(SRI), Lincoln Labs, MIT, Bell Lab 등
3.5세대(1990~2000)	- 오류최소화(discriminative)를 위한 MCE(Minimum Classification Error), MMI(Maximum Mutual Information) 등의 접근방식 시도 - 노이즈, 마이크론, 전송채널, 반향 등에 Robust한 음성인식 - DAPRA 프로그램: ATIS task, Switchboard task
4세대(2000~현재)	- DAPRA 프로그램: EARS(Effective Affordable Reusable Speech-to-Text) 프로그램, 중요정보의 추적/추출/요약/번역 - 즉흥적 음성인식(일본 CSJ프로젝트) - 복합(Multimodal)음성인식: 영상-음성 정보 동시 인식

자료: Sadaoki Furui(2015), 저자 정리

4) 동적시간워핑(Dynamic Time Warping)은 발화된 음성 신호로부터 추출된 음성 특징 벡터열을 대상으로 두 개의 다른 벡터열 간 유사성을 측정하는 방식이다(출처: 한국정보통신기술협회 정보통신용어사전)

5) n-gram언어모델: 1개 단어나 음절이 서로 다른 단어 혹은 음절과 함께 발화될 확률을 계산하는 모델로 자연어 처리에 적합하며, Google도 자체 클러스터링 기술과 n-gram을 기반으로 한 언어 모델을 이용하고 있다(출처: 한국정보통신기술협회 정보통신용어사전)

〔그림 3〕 NIST Switchboard⁶⁾ 테스트 오류율

자료: The Economist(2016. 5. 1), 저자 정리

2. 음성기반 사용자 인터페이스의 특징

Apple의 음성인식 비서 서비스 Siri는 막대한 양의 데이터베이스를 활용하여 질문에 최적 답변을 선택하는 구조를 갖고 있으며, Google의 음성검색 기능 역시 입력된 음성데이터를 클라우드 서버에서 텍스트로 변환하여 분석하고, 그 결과에 따른 실시간 검색결과를 기기로 재전송하는 방식이다. 최은정(2013)에 따르면 Google은 약 23,000억 개의 영단어를 성별, 연령별, 억양별로 구분된 음성 데이터로 구분하여 클라우드에 저장하고 실시간으로 처리할 수 있는 컴퓨팅 능력을 보유하고 있다. 더 많은 음성데이터가 축적될수록 “인간과 동등한(Human parity)” 수준의 정확한 인식이 가능해 질 것이다.

음성인식 기술은 단독으로보다는 IT 기기나 서비스와 결합됐을 때 그 파급력이 한층 더 클 것으로 예상된다. 음성기반 인터페이스는 분당 약 150단어 수준의 입력이 가능하기 때문에 분당 평균 40단어 가량 입력이 가능한 터치나 타이핑 방식에 비해 그

6) NIST Switchboard는 음성인식 기술 벤치마크 톨로 사용되는 전화 통화 모음으로 현재 사용되는 테스트의 샘플레이트는 8KHz 수준으로 2000년도 통화품질 수준에 상응한다.

속도가 3배 이상 빠르다(〈표 2〉참조). 또한 필요한 정보검색이나 기능 구동을 위한 랜덤액세스가 가능하며, 단방향이 아닌 컨텍스트 기반 문제해결방식이기 때문에 텍스트 입력이나 터치보다 훨씬 인간의 자연스러운 커뮤니케이션 형태에 가깝다. 이러한 기술적 특성을 고려하였을 때 음성인식 기술은 IoT와 인간을 매개하는데 가장 적합한 인터페이스로서의 잠재력을 갖고 있다.

아직까지 스마트홈이나 홈 오토메이션은 컴퓨터나 스마트폰과 같이 터치나 텍스트 입력을 필요로 하며, 출력 역시 GUI(컴퓨터그래픽사용자인터페이스) 형태의 통합 컨트롤 시스템으로 제공되고 있다. 그러나 음성인식 인터페이스를 활용할 경우 직접 접촉하지 않고도 IoT 기기나 서비스를 제어할 수 있을 뿐만 아니라, 조작방식을 학습할 필요가 없기 때문에 편의성 측면에서 터치나 텍스트 기반보다 우수하다.

〈표 2〉 음성과 터치/텍스트기반 사용자 인터페이스 비교

	음성기반	터치/텍스트기반
입력속도	150단어/분	40단어/분
입력방식	Random Access	Hierarchical GUI
정보 입출력	Microphone, Speaker	Keyboard, Display
문제해결방식	컨텍스트 기반 문제해결	단방향 문제해결
멀티태스킹	Hands-free 기능으로 멀티태스킹 (운전 등) 가능	터치나 텍스트입력 시 멀티태스킹 불가

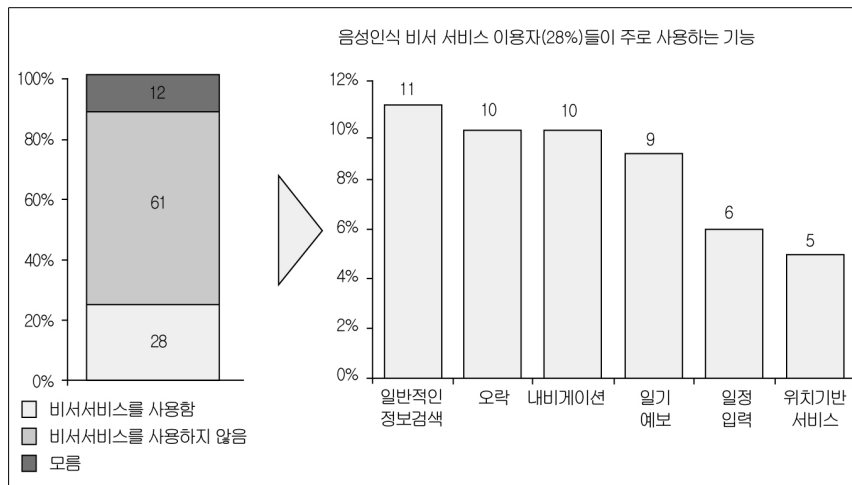
자료: 언론자료 취합

음성인식 기술 경쟁의 포문은 Apple의 Siri와 Google Now 같은 음성인식 AI 비서 서비스가 스마트폰에서 이용 가능해지며 열렸지만, 이것이 격화된 시점은 2014년 Amazon이 인공지능 스피커를 표방하는 Echo를 출시한 이후로 보아야 할 것이다. 2014년 11월 출시 당시에는 단순한 가정용 홈 오디오 스피커 정도로 대중에 인지되었으나, 음성인식 소프트웨어인 Alexa의 개발자 키트를 외부에 오픈하면서부터 Echo를 통해 활용할 수 있는 Skill(일종의 애플리케이션)이 빠르게 늘어나고, 이에 상응하여

Echo 판매량이 증가하는 선순환 구조가 창출되었다.

Apple와 Google 모두 일찍이 음성기반 인터페이스가 차세대 스마트기기의 중요한 제어기술이 될 것이라 예상은 하였지만, 스마트폰의 음성인식 비서 서비스의 활용은 예상보다 부진했던 것으로 보인다. 2015년 실시된 스마트폰 이용자 조사에 따르면, 18세 이상 미국인 스마트폰 이용자 중 36.7%는 음성인식 비서 서비스를 전혀 이용해 본 적이 없으며(Statista, 2015. 10), 그 다음해 영국에서 실시된 조사에서는 약 61%의 스마트폰 이용자가 관련 서비스를 이용해 본 적이 없다고 응답하였다(Statista, 2016. 6). 그 이용자들도 대부분 ‘일반적인 정보 검색(11%)’, ‘오락(10%)’, ‘내비게이션(10%)’, ‘일기예보 확인(9%)’, ‘일정입력(6%)’, 그리고 ‘위치기반 서비스(5%)’([그림 4] 참조) 등의 단순기능을 중심으로 사용하고 있는 것으로 조사되었다(Statista, 2016. 6).

[그림 4] 음성인식 비서 서비스의 주요 사용 목적



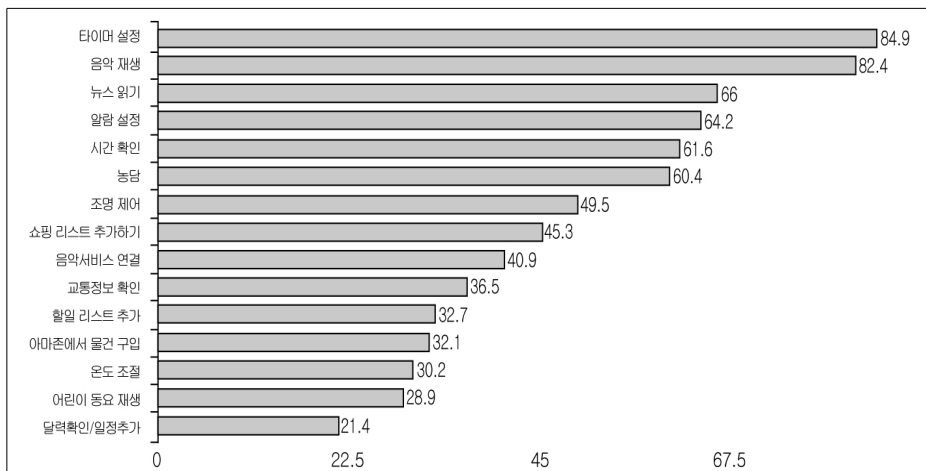
주: 스마트폰 사용자 3,251명을 대상으로 조사한 결과이며, 주로 사용하는 기능은 빈도가 높은 상위 6개만 위의 표에 포함됨

자료: Statista, Deloitte(2016. 6) Statista 재인용

3. 음성인식 AI 스피커의 차별화 요소

Echo와 Google Home으로 대표되는 음성인식 기반 AI 스피커에서 주로 이용되는 비서 서비스 기능은 스마트폰의 그것과 큰 차이가 없다([그림 4], [그림 5]). 그림에도 불구하고 스피커형태의 형태의 기기가 스마트폰이나 음성인식 기능을 탑재한 TV나 셋톱박스를 제치고 IoT 시대의 차세대 핵심 기기로 주목받는 이유는 무엇일까.

[그림 5] Amazon Echo 이용자가 1회 이상 사용해본 기능



주: Amazon Echo 사용자 1,300명 대상으로 설문 실시

자료: Experian(2016)

소리나 음성을 주고받는데 가장 일반적으로 쓰이는 입출력 장치는 마이크와 스피커이다. 따라서 음성인식 기술 구현에 가장 핵심적인 입출력 장치만으로 이루어진 스피커는 데이터 수집 측면에서도 스마트폰이나 TV보다 훨씬 효율적이다. Siri 애플리케이션이나 OK Google 기능은 활성화 되어 있을 때만 음성데이터를 녹음하고 처리할 수 있다. 이에 반해 스피커는 Mute기능을 사용 시를 제외하고는 상시적으로 음성데이터를 녹음하고 클라우드 서버에 축적할 수 있으므로 음성인식 시스템 개선에 필요한 중요한 자원을 훨씬 빠른 속도로 수집한다.

사용 환경도 중요한 요인이다. 인간은 하루 종일 주변에서 들려오는 다양한 소리에 둘러싸이지만 모든 소리를 동일한 수준으로 인지하지 않는다. 인지과학자 Cherry Colin(1953)은 특정한 소리나 자극에 대한 인간의 선택적 지각(Selective perception) 현상을 “칵테일파티효과(Cocktail Party Effect)”라고 이름 붙였다. 음성인식 기기나 소프트웨어는 아직 인간의 발화와 그 외의 소음을 인간만큼 완벽히 구분해내지 못한다. 따라서 실외에서 주로 사용되는 스마트폰의 음성인식은 외부 잡음으로 인한 오류율이 높을 수밖에 없다. 이에 비교하여 상대적으로 잡음이 적은 실내에서 사용되는 스피커는 음성인식 기술 구현에 적합하다.

접촉할 필요도 없고 확인할 스크린도 없는 원통형의 스피커가 스마트폰을 이어갈 차세대 스마트 기기로 부상하고 있는 현 상황에 대해 벤처캐피탈 Accel의 Sameer Ghandi는 “이것을 진화라고 단언해서 말 할 수 있는 사람은 없겠지만, 매우 자연스러운 방식임은 분명하다”고 평가했다. 다시 말해 음성인식 인터페이스는 “가장 자연스러운 인간과 기기 간의 상호작용 방식”이며(Wired, 2016. 6. 24), 향후 2~3년 이내에 스마트폰, 스마트워치, Hearable기기 등(〈표 3〉참조)에 도입될 것으로 전망된다(Gartner, 2017. 3).

〈표 3〉 디바이스·기능별 음성제어 기능의 효용수준 비교

	데스크탑/ 노트북	태블릿	스마트폰	HMDs	스마트워치 /밴드	AI 스피커	Hearable 기기
효용수준	매우낮음	낮음	높음	높음	매우높음	매우높음	매우높음

자료: Gartner(2017. 3)

Ⅲ. 주요 기업의 음성인식 AI 스피커 전략

음성인식 AI 스피커는 시작 단계에 있는 신생 시장이다. 시장조사업체에 따르면 시장은 연평균 43%씩 성장하여 2020년에는 약 21억불 규모에 이르고, 전 세계 가구 중 3.3%가

해당 기기를 사용할 것으로 예측된다(Gartner, 2016. 9). 이번 장에서는 국내외 주요 기업의 음성인식 기반 AI 비서 서비스 개발 및 AI 스피커 제품 전략을 살펴본다.

1. 해 외

해외 주요 IT기업의 스타트업 인수 실적을 살펴보면 2010년부터 음성인식(STT: Speech-to-Text), 문자음성(TTS: Text-to-Speech), 시맨틱검색, 실시간 음성번역 등의 기술 확보를 위한 노력이 이루어져 왔음을 확인할 수 있다(〈표 4〉참조). 최초로 스피커 형태의 기기에 음성인식 AI 비서 서비스를 탑재한 제품을 출시하여 새로운 시장의 문을 연 Amazon의 전용 스피커 단말기는 2016년 한 해 동안 전 세계적으로 약 510만대 가량 판매된 것으로 추정된다(Statista, 2016. 12. 22). 그러나 Echo의 독주가 2017년에도 계속될 수 있을지는 불확실하다. 작년 11월 출시된 Google Home이 강력한 경쟁 상대로 떠올랐으며, 2017년에는 Microsoft와 Harman/Kardon이 공동 개발한 AI 스피커 Invoke도 경쟁에 가세할 예정이다(USA Today, 2017. 5. 15).

〈표 4〉 해외 주요 IT 기업의 음성제어 스타트업 인수 실적

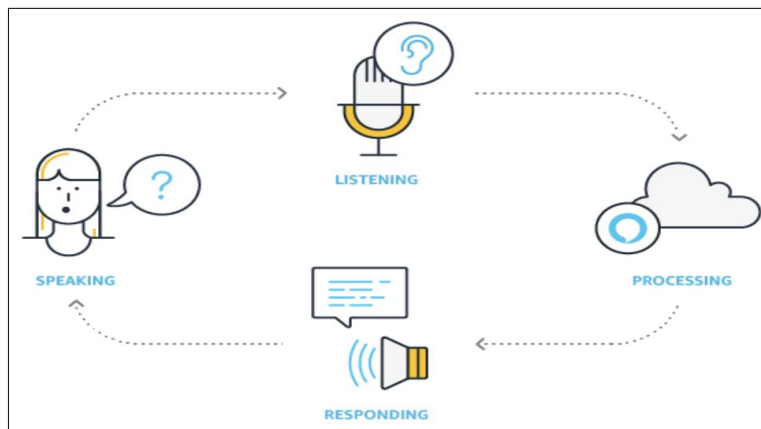
기업명	내용
Apple	- 2010년 'Siri'(음성인식 소프트웨어) 인수 - 2016년 VocalIQ(음성 인터페이스)인수
Amazon	- 2011년 'Yap'(STT: Speech-to-Text) - 2014년 'Evi'(지식기반 Semantic Search Engine), 'INOVA'(TTS: Text-to-Speech) 인수
Facebook	- 2013년 'Jibbigo'(모바일 다국어 음성번역) - 2015년 'Wit.at'(음성인식), Two Big Ears(VR오디오) 인수
Google	- 2014년 'DeepMind Technologies'(AI 머신러닝) - 2016년 API.AI(음성인식, 자연어 처리) - Limes Audio(음성처리 소프트웨어) 인수

자료: 언론자료 취합

(1) Amazon

2014년 11월 출시된 Amazon Echo의 성공은 해외 언론에서도 Sleeper hit⁷⁾이라 칭해질 만큼 예상치 못했던 것이다(USA Today, 2016. 3. 28). Echo 이전에 Amazon이 출시한 기기로는 이북리더기 Kindle과 태블릿PC인 Fire, 그리고 스마트폰인 Fire Phone이 있었다. 2014년 7월 기대 속에서 출시된 Fire Phone이 소비자로부터 외면 받으며 5.44억 달러의 영업 손실⁸⁾을 기록하였기 때문에 Echo에 대한 기대는 자연스레 낮아졌던 것으로 보인다.

[그림 6] Alexa Voice Service(AVS) 개요



자료: Amazon Developer 홈페이지(<https://developer.amazon.com>)

Echo는 “Alexa”나 “Amazon”, “Echo” 등의 명령어를 호출 시 구동되며, 총 7개의 내장 마이크로 녹음된 음성데이터는 무선인터넷을 통해 Amazon의 클라우드 시스템인 Amazon Web Service(AWS)로 전송되어 분석이 이루어진다. 그 결과를 바탕으로 Echo는 사용자의 명령을 수행하고, 사용자는 자신의 명령에 대한 피드백을 Alexa의 음성으로 들을 수

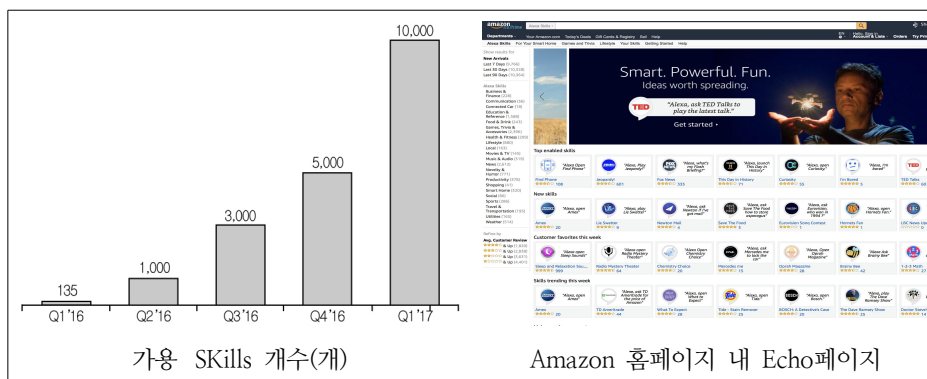
7) 엔터테인먼트 산업에서 사용되는 관용어로 흥행에 대해 큰 기대가 없었으나 예상을 깨고 크게 성공한 작품을 지칭한다(Wikipedia, https://en.wikipedia.org/wiki/Sleeper_hit).

8) 더 자세한 내용은 “Amazon’s losses balloon to \$544 million, driven by weak Fire Phone sales”, (The Verge, 2014. 10. 23) 참조.

있는데, 이러한 시스템을 Alexa Voice Service(AVS)라고 한다([그림 6] 참조).

출시 초반 Amazon의 Prime 회원(유료이용자)과 사전 예약자에게만 판매되었던 Echo는 2015년 6월 일반인을 대상으로 판매가 시작됨과 동시에 Alexa 소프트웨어와 연동할 수 있는 “Alexa Skills Kit(ASK)”이 공개되었다. 이때부터 외부 개발자들은 ASK를 이용하여 Echo의 음성제어 기능을 활용할 수 있는 서비스나 연동 애플리케이션을 자유롭게 추가할 수 있게 되었다. 이러한 스마트 기능은 “Skills”⁹⁾라고 불리는데, 2년이 지난 2017년 1분기 기준 이용 가능한 Skills 개수가 10,000개를 넘어섰다(Statista, 2017. 2. 27). 대표적인 Skills로는 Uber 호출, Amazon 쇼핑 플랫폼을 통한 물건 구입 등이 있으며, 최근에는 스타벅스 커피 주문도 가능해졌다.¹⁰⁾

[그림 7] Alexa Skills 운영 현황



자료:(좌) Amazon, Statista 재인용, (우) Amazon 홈페이지(<http://www.amazon.com>)

Skills의 개수보다 Amazon에게 더 중요한 것은 Alexa를 통한 새로운 오픈 생태계의 구축이다. AVS는 반드시 Echo를 통해 작동될 필요가 없기 때문에 원하는 기업은 자유로이 Alexa의 음성인식 및 피드백 시스템을 활용할 수 있다. Alexa 연동제품

9) Skills는 스마트폰의 애플리케이션과 유사한 기능을 하나, 기기에 직접 설치되는 애플리케이션과는 달리 Amazon 클라우드 상 존재하며 Echo에 명령을 내린다.

10) 더 자세한 내용은 “Make Alexa order your iced grande caramel macchiato”,(CNET, 2017. 2. 6)을 참조

(Alexa-enabled Product)으로는 “Phillips Hue(스마트조명)”, “Belkin WeMo(스마트스위치)”, “Ecobee(스마트온도조절기)” 등의 스마트홈 기기가 있다(〈표 5〉 참조). AVS를 사용하는 가정용 기기 및 전자제품이 늘어남에 따라 장기적으로는 Echo가 홈오토메이션 시스템을 컨트롤하는 허브로 발전해 나갈 가능성이 제기되고 있다(박병근, 2016)

2016년 3월 Amazon은 “Amazon Tap”과 “Echo Dot”을 출시하며 라인업 확대를 시작했다. Echo는 상시 전원에 연결되어 있어야 하는 반면, Tap은 충전식 스피커이기 때문에 휴대가 용이다. Echo Dot 1세대는 스피커 기능이 포함되어 있었으나 지난해 10월 출시된 2세대에서는 스피커가 없는 형태의 음성인식 특화 기기로 출시되었다. 이는 Echo Dot을 집안 곳곳에 설치하여 음성인식 기능 활용을 극대화 하고자 하는 Amazon의 전략을 반영한 것이다. 2017년에는 카메라를 탑재하여 사용자의 옷매무새나 의상을 평가해주는 기능에 특화된 Echo Look과 7인치 터치스크린을 통해 영상 통화, 영상과 사진재생 기능 등 을 지원하는 Echo Show가 출시되었다.

〈표 5〉 Alexa연동기기(Alexa-enabled products)

구분	주요 제품			
가정용				
	조명 〈Phillips Hue〉	온도조절기 〈Ecobee〉	비디오 도어벨 〈Ring〉	스마트플러스 〈TP-Link〉
				
	천장 팬 〈Haiku Home L Series〉	라우터 〈Luma〉	시큐리티시스템 〈Blink〉	스프링클러 제어기 〈Rachio〉

2017/5/16

구분	주요 제품	
자동차	 <p>Ford 자동차 <SYNC></p>	 <p>현대자동차 <BlueLink Skills></p>
스마트폰	 <p>아이폰 <Lexi App></p>	

자료: Amazon 홈페이지, TechCrunch, 언론자료 취합

(2) Google

Echo의 유력한 대항마로 떠오른 음성인식 AI 스피커 Google Home은 2016년 11월 출시되었다. 그 이름에서 예상할 수 있듯이 홈 오토메이션을 주력 기능으로 내세우고 있다. Google Home의 운영체제는 “Chromecast”의 OS를 변형한 것이며, 음성인식 소프트웨어 Google Assistant는 기존에 Android OS 스마트폰에 도입되었던 Google Now에서 AI 기능이 강화된 버전이다. Google Now와 동일하게 “OK Google”이라는 호출어로 구동되는데, 이용자는 Google Home으로 음악 재생, 일기예보 확인, 물건 주문 등의 기능은 홈IoT를 제어할 수도 있다. Google Home으로 제어할 수 있는 대표적 IoT로는 “Chromecast(멀티미디어 스트리밍 어댑터)”, “Google Nest”의 스마트홈 기기(표 6) 참조), “Phillips Hue(스마트조명)”, “삼성 SmartThings(IoT)”가 있다. 또한 스마트기기 자동화 애플리케이션인 “IF”¹¹⁾의 레시피를 직접 생성하고 명령을 수행

11) 회사명과 동일하게 IFTTT(If This Then That)라는 이름으로 처음 출시된 이 애플리케이션에서 Recipe라고 불리는 주문서를 미리 만들어 놓으면, 그 조건이 충족되었을 때 자동으로 저장된

하게 만들 수도 있다(Techradar, 2017. 1. 18)

〈표 6〉 Nest에서 출시된 스마트홈 기기

제품명	내용
Nest Learning Thermostat	<ul style="list-style-type: none"> - 사용패턴에 따라 온도 설정 Self-scheduling 가능 - 동작인식 센서를 통해 사용자 외출 시 자동 온도 조절 - 실내온도, 습도, 외부날씨를 분석하여 온도 설정
Nest Protect	<ul style="list-style-type: none"> - 화재경보기, 연기 및 일산화탄소 감지 기능
Nest Cam Indoor	<ul style="list-style-type: none"> - 실내 영상촬영, 스트리밍 가능한 보안카메라
Nest Cam Outdoor	<ul style="list-style-type: none"> - 실외 영상촬영 - 움직임을 포착하여 스마트폰으로 경보 전송 기능

자료: Wikipedia, 언론자료 취합

Google은 지난 2014년 인수한 Nest Lab의 스마트 온도조절장치를 중심으로 Google Nest라는 플랫폼을 구축하고 IoT 전용 OS인 “Brillo”를 출시하는 등 스마트홈 시장 선점을 위해 노력해 왔는데, Google Home은 이런 기존 플랫폼을 연결하는 인터페이스의 역할을 할 수 있을 것으로 보인다(Kotra해외시장뉴스, 2016. 6. 29). Chromecast를 보유하고 있는 가정에서는 Google Home으로 TV나 홈엔터테인먼트 기기 제어¹²⁾도 가능해지기 때문에 Google이 구축해 온 미디어 플랫폼도 AI 스피커를 매개로 통합될 수 있다.

선발주자인 Alexa와 차별화 되는 Assistant 소프트웨어의 경쟁력은 기존에 보유하고 있는 강력한 검색엔진과 오랜 기간에 걸쳐 축적된 사용자 데이터(예: 개인이력, 활동정보, Google 서비스 이용정보 등)에 있다. 사용자의 질문이나 음성명령에 대한

명령을 수행한다. 예를 들어 “특정 장소에 도착하면 스마트폰을 진동모드로 전환해라”나 “새로운 트위터 팔로워가 생겼을 때 리스트를 자동으로 저장해라”같은 기능을 설정할 수 있다(동아일보, 2015. 1. 23, <http://it.donga.com/20272/>)

12) Chromecast가 설치되어 있을 경우 음성명령으로 YouTube 비디오를 검색하고, 검색결과를 TV에서 재생되게 할 수 있다.

적절한 답을 Google 데이터베이스에서 찾지 못했을 때는 위키피디아 등의 인터넷 소스를 검색하여 피드백 한다. 음성명령을 통해 다른 방에 있는 복수의 스피커에서 동일한 음악이나 팟캐스트를 동시에 재생할 수 있는 통합 멀티룸오디오기능도 Google Home에서만 가다. 전화통화(Hands-free calling) 기능과 영상응답(Visual response) 기능을 추가하는 업데이트 계획¹³⁾이 발표됐다.

Alexa의 Skill처럼 AI 스피커에서 구동되는 기능은 Google Home과 Assistant 생태계에서 “Action”으로 불린다. “Actions on Google”로 명명된 Google Assistant 개발자용 플랫폼은 2016년 12월 공개되었으며, 외부 개발자들은 이 플랫폼을 활용하여 Google Home에 써드파티 서비스를 연동시킬 수 있다. Alexa Skills와 Google Action의 가장 큰 차이는 구동 방식에 있다. Echo 사용자는 특정한 Skill을 사용하기에 앞서 해당 기능을 직접 활성화시켜야 하지만, Google Home에서는 Google Home App의 서비스섹션에 등록된 Action이라면 별도의 활성화 과정 없이 음성명령만으로 구동시킬 수 있다(CNET, 2017. 1. 3).

Google Home은 Echo와 비교했을 때 간단한 음성명령의 수행이나 검색에서는 두 기기 간 큰 차이가 없으나, 무작위 질문에 대한 응답에서 Echo보다 우수한 성능을 갖고 있는 것으로 보인다(WSJ, 2016. 11. 7). 그러나 두 기기 모두 스피커의 본질적인 기능인 음악 감상에는 적절하지 않은 것으로 보인다. 이에 대해 CNET(2016. 11. 3)에서는 Google Home과 Echo의 음질을 분석한 기사의 말미에 “만약 음악용 스피커를 찾고 있다면 둘 다 적절한 선택은 아니다(If you want a musical speaker, the choice is simple: get neither).”라는 혹평을 내리기도 했다.

13) 더 자세한 내용은 “Google’s Home Speaker can now make phone call”.(The Verge, 2017. 5. 17)에서 확인 가능하다.

〈표 7〉 출시된 AI 스피커(Amazon Echo와 Google Home) 비교

	Amazon Echo	Google Home
가격	\$179.99(Echo Dot: \$49.99)	\$129
크기(mm)/무게(g)	235×84×84mm/1,064g	96.4×142.8mm/477g
음성인식 마이크 개수	7개(빔형성 기술 지원)	2개
스피커 시스템	2.5 Woofer와 2인치 Tweeter	2인치 Driver와 2인치 Passive Radiator 한 쌍
상시녹음 여부	상시녹음	상시녹음
스테레오시스템 외부 출력	블루투스나 Echo dot	Chromecast
구동단어	“Alexa” 또는 “Echo”, “Amazon”	“OK Google”, “Hey Google”
음악 스트리밍 서비스	Amazon Prime Music, Spotify, Pandora, iHeartRadio, TuneIn	Google Play Music, YouTube Music, Spotify, Pandora, TuneIn
소프트웨어개발자도구	Alexa Skills Kit(ASK)	Actions on Google
특징	<ul style="list-style-type: none"> - 최근 카메라 기능에 중점을 두고 있는 Echo Look과, 터치스크린 추가된 Echo Show 출시 	<ul style="list-style-type: none"> - Google 자체 검색엔진 활용 - 통합 멀티룸 오디오 기능 지원 - Chromecast로 외부 출력가능 - 복수 사용자의 음성 구분
2016년 누적판매대수	약 510만대	확인 불가

주: Amazon Echo 판매대수는 Consumer Intelligence Research Partners(2016. 11)의 추정치임
 자료: Valuewalk, Statista(Consumer Intelligence Research Partners 재인용), 저자 재정리

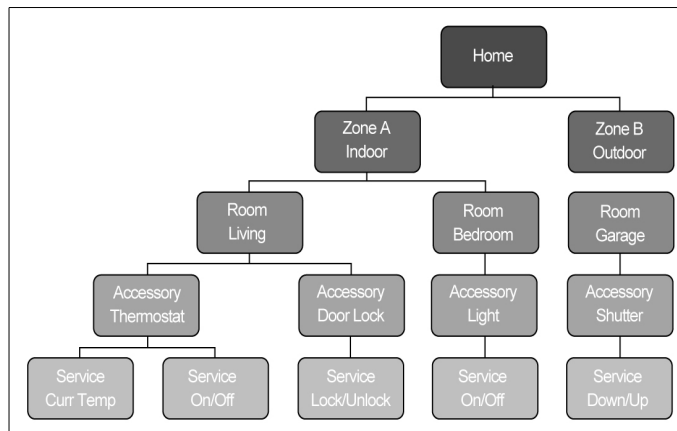
(3) Apple

음성인식 기술 경쟁이 스피커 형태로 본격화됨에 따라 Siri를 통해 독자적 음성인식 시스템을 갖고 자사 제품군에 적용하고 있는 Apple의 AI 스피커 출시 여부에 많은 관심이 쏟아지고 있다. 공식적으로 발표된 내용은 없으나 Apple은 Echo나 Google Home의 강력한 경쟁 제품을 단기간 내에 출시할 만한 역량을 보유하고 있는 것으로 평가된다. Apple은 2014년부터 “HomeKit”¹⁴⁾을 통해 스마트홈 플랫폼을 구축해왔는데,

14) HomeKit이란 Apple의 기기에 설치된 ‘홈’앱을 통해 다양한 스마트홈 기기나 액세서리를 조작할 수 있도록 해주는 사물인터넷(IoT) 플랫폼이다(<https://www.apple.com/kr/ios/home/>).

이는 AI 스피커의 홈 오토메이션 시스템과 동일한 방식으로 구동된다. Siri를 통해 음성명령을 내리면 HomeKit에 연결된 IoT를 제어할 수 있다. 또한 2016년 추가된 “Scene”기능을 이용하면 복수의 기기들을 동시에 제어할 수 있으며(아이뉴스24, 2015. 6. 30), 공간 별¹⁵⁾ 관리도 가능하다([그림 8] 참조).

[그림 8] HomeKit이 작동하는 계층 구조



자료: IoTenableDevice.com(<http://www.iotenableddevices.com/>)

Siri의 차별화 요소 중 하나는 데이터 처리 방식이다. Alexa와 Assistant는 방대한 음성데이터 처리를 위해 클라우드 시스템을 이용한다. 그러나 Siri는 딥러닝과 AI 기능 대부분이 기기 내부에서 이루어지기 때문에 인터넷에 연결되어 있지 않을 때도 사용 가능하며 프라이버시 유출에 대한 우려도 적다. 이는 지난해 테러용의자의 아이폰 잠금 해제를 두고 벌어진 FBI와의 법정공방¹⁶⁾과정에서 확인할 수 있었던 Apple의 프라이버

15) HomeKit사용자는 Home 애플리케이션에서 선택적으로 공간을 지정할 수 있다. 각 방에 명칭을 정한 후 “Siri, turn off the bedroom lights(시리, 침실 불을 꺼줘)”라고 음성으로 명령하면 침실의 조명만 소등할 수 있다. 여기서 Accessories는 개별 사물 인터넷 기기이며, Service는 기기에서 구동할 수 있는 기능을 말한다(자세한 내용은 <http://www.iotenableddevices.com/apple-homekit-automation-protocol-framework/>).

16) FBI는 2015년 12월 발생한 미국 캘리포니아주 San Bernardino 총기테러사건의 용의자가 소지하

시 보호에 대한 입장이 투영된 것으로 보인다.

최근 Apple은 SiriKit¹⁷⁾이라 명명된 소프트웨어개발자도구를 외부개발자에 개방하였다. SiriKit이 지원하는 분야는 VoIP Calling, 메시지, 지불 서비스, 사진, Workouts, Ride Booking, CarPlay, 레스토랑 예약 등 AI 스피커를 통해 구현할 수 있는 것과 유사한 것이 대부분이다. Apple의 음성인식 비서 서비스의 외연이 AI 스피커까지 확대될지 여부에 대해서는 더 지켜보아야 할 것이다.

(4) Microsoft

Microsoft는 1993년 카네기멜론대학 컴퓨터공학과와 Xuedong Huang을 영입하면서부터 내부적으로 음성인식 상용화를 위한 연구 개발에 착수했다. 1994년에는 Speech API를 소개했으며(Brown, 2008), 2000년부터 Office2003 등 일부 소프트웨어에 음성인식 기능이 적용되었다([그림 9] 좌측 참조). 2000년대 중반까지 Windows에서 지원되는 음성인식 소프트웨어는 간단한 단어¹⁸⁾도 제대로 구분 못할 정도로 인식률이 낮았다. Microsoft의 Chief Scientist인 Rico Malvar는 2009년 등장한 딥러닝 기술이 음성인식 개선에 활용되면서 음성인식 서비스 상용화 작업이 본격화되었다고 설명한다(동아일보, 2017. 5. 16). 5년이 지난 2014년 Microsoft는 음성인식 비서 서비스 “Cortana”를 출시하고 Windows10 운영체제¹⁹⁾에 우선적으로 적용했다([그림 9] 우측 참조). 2015년부터는 iOS와 Android OS에서도 Cortana를 사용할

고 있던 아이폰의 사용기록을 확인하기 위해 Apple에 잠금해제를 요구했으나, Apple은 고객의 프라이버시 보호를 위해 지원이 불가능하다는 입장을 유지하였다. 이 갈등은 소송전으로 이어졌으나, FBI가 자체적으로 잠금해제에 성공하며 소송을 취하였다.

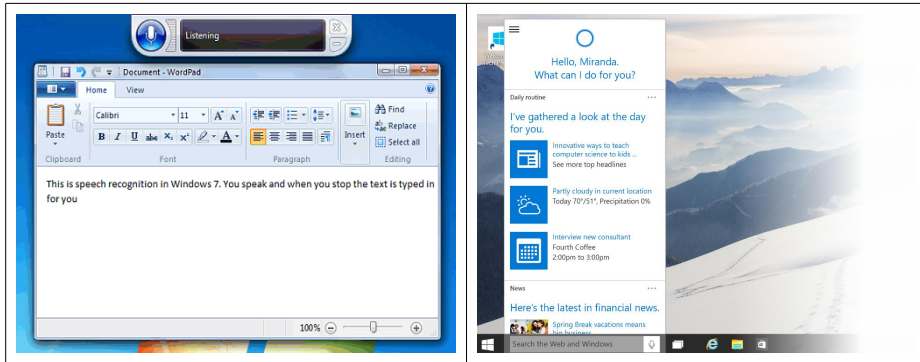
17) SiriKit Programming Guide 홈페이지(https://developer.apple.com/library/content/documentation/Intents/Conceptual/SiriIntegrationGuide/index.html#//apple_ref/doc/uid/TP40016875-CH11-SW1)

18) 2006년 Windows Speech Recognition Program은 엄마(Mom)과 숙모(Aunt)를 구분하기 힘든 수준이었다(더 자세한 내용은 <http://it.chosun.com/news/article.html?no=2828084>).

19) Windows 10, Windows 10 Mobile, Windows Phone 8.1, Microsoft Band, Xbox One, iOS, Android, Windows Mixed Reality에서 적용되었다(출처: <https://www.microsoft.com/en/mobile/experiences/cortana/>)

수 있게 되었는데, 타사 OS에서 시스템영역까지 제어하는 것은 불가능하므로 “Hey Cortana”같은 구동명령어 사용 등에 제약이 따른다.

[그림 9] Windows7의 음성인식 소프트웨어(좌)와 Windows10의 Cortana(우)



자료: Raw Computing(<http://rc.rawinfopages.com>), VentureBeat(<https://venturebeat.com>)

Cortana는 사용 패턴을 학습하고 적응할 수 있기 때문에 개별 이용자에게 최적화 된 서비스를 제공하고 능동적으로 Action을 제안할 수 있다는 점에서 차별성을 갖는다. 별도의 음성명령어 없더라도 Cortana는 이용자의 상황인식(Contextual awareness) 정보를 분석하여 특정한 애플리케이션이나 웹사이트 방문을 권할 수 있다(Microsoft Build 2016, 2016. 3. 25). 눈에 띄는 다른 기능은 음성과 시각적 요소의 결합이다. Cortana는 푸른색의 움직이는 후광을 이용하여 18가지 감정을 표현할 수 있기 때문에 이용자들은 컴퓨터나 기계가 아니라 실재하는 상대와 대화하는 것 같은 경험을 할 수 있다.

앞서 살펴본 기업들과 마찬가지로 Microsoft 역시 2016년 말부터 Cortana를 중심으로 한 음성인식 서비스 생태계 조성을 위한 다양한 전략을 취하고 있다. 첫 번째는 기기의 확장과 관련된 계획이다. 음향기기전문업체인 “Harman/Kardon”과의 협력 하에 Cortana 플랫폼을 탑재한 전용 AI 스피커 “Invoke”가 올해 안에 출시될 예정이며, 컴퓨터 제조업체인 “HP” 기기에도 Cortana를 적용하기 위한 협약이 이루어졌다(The Verge, 2017. 5. 10). 소프트웨어 측면에서는 음성인식 비서 서비스를 스마트홈의 인터

페이스로 활용하고자 하는 전략이 구체화 되고 있다. 또한, 외부개발자들이 기기나 OS 관계없이 Cortana 연동 서비스를 개발할 수 있게 해주는 “Cortana Skill Kit”도 공개될 예정이며(Microsoft News Center, 2016. 12. 14), 2017년 9월 업데이트에서는 Windows10 홈버브와 Cortana를 이용한 홈IoT 제어 등의 기능이 추가될 전망이다.

Microsoft는 음성인식 비서 서비스와 AI 스피커 시장에서 후발주자에 속하지만, 그 잠재력은 매우 크다. Microsoft에 따르면 현재 Windows10 이용자는 약 5억 명 규모이며 그 중 약 1억 4천명이 매달 Cortana를 이용하고 있다(CNBC, 2017. 5. 10). 이로 미루어볼 때 Microsoft의 독보적인 PC OS 시장지배력²⁰⁾이 Cortana 서비스 이용자 수 증가에 큰 영향을 미치고 있는 것으로 보이며, Windows10 OS 이용자가 늘어남에 따라 Cortana 도 꾸준히 시장 지배력을 넓혀갈 것으로 예상된다.

〈표 8〉 AI 비서 서비스(Apple Siri, Google Assistant, Microsoft Cortana) 비교

	Apple Siri	Google Assistant	Microsoft Cortana
구동 방식	물리적 키(홈버튼)로 구동 가능	구동명령어로만 가능	물리적 키로 구동가능
사용자 프로필	기기 내 저장	클라우드에 저장	클라우드에 저장
주 검색엔진	Wolfram Alpha	Google	Bing
주 웹브라우저	확인 불가	Chrome	Edge
연동기기	Apple에서 출시된 모든 기기(Apple TV, Mac PC 등)	스마트폰(픽셀), 스피커(Google Home) 등	Windows10 OS 기반 PC, Windows 모바일기기, 스피커(Invoke)

자료: Search Engine Land & Fllehippo, 저자 재정리

20) 시장조사업체인 Netmarketshare에 따르면 데스크탑PC OS시장에서 Windows 10의 점유율은 26.28%이며 가장 사용자가 많은 OS는 Windows 7(48.5%)이다. 비 Windows(예: Mac OS X, Linux 등) OS의 시장 점유율은 9.65%로 조사되었다(출처: <https://www.netmarketshare.com/>).

(5) 기타: 중국 · 일본 기업 동향

한자는 복잡한 문자 형태를 갖고 있고 글자 수만도 수천 개에 달하며, 병음 체계를 모르는 중국인들도 많기 때문에 중국은 음성인식 인터페이스가 활용되기에 최적의 조건을 갖고 있다. 대표적 검색 엔진 기업 Baidu는 이미 2015년 음성인식 AI 비서 서비스인 Duer를 모바일 애플리케이션으로 출시한 바 있다. Baidu의 음성인식 기술은 Apple이나 Google에 뒤지지 않는다는 평가를 받는데, 이러한 높은 평가의 뒤에는 Baidu Research가 개발한 전용 소프트웨어 Deep Speech가 있다. 2016년 9월 Baidu World Conference에서 발표된 내용에 따르면 이 소프트웨어의 음성인식 정확도는 97% 수준이며, 시끄러운 환경에서도 인식률이 매우 높다(한국경제, 2016. 9. 2). Duer의 음성인식 기술은 Little Fish라는 이름의 가정용 로봇에 적용되어 CES 2017에 소개되었다(ZDNet, 2017. 1. 6).

일본에서는 NICT(총무성 산하 정보통신연구기구)가 주도하는 가운데 NTT도코모, 후지쓰, 도요타자동차 등이 일본어 기반 음성인식 AI 서비스 개발을 위해 협력하고 있는 것으로 보도된다(뉴스1, 2017. 1. 9, 산케이 재인용). 이미 상용화에 성공한 기업도 있는데, 2017년 3월 라인(Line, 한국 네이버의 일본 자회사)에 인수된 IoT 스타트업 vinclu가 대표적이다. vinclu의 전용기기인 Gatebox는 내부가 빈 유리관 형태로 되어 있으며, 그 안에서 여성형 AI 캐릭터가 홀로그램으로 구현된다. Gatebox는 집 안의 스마트홈 기기를 제어하고 인터넷 검색을 통해 필요한 정보를 찾아주는 등 기능상으로는 Echo나 Google Home과 크게 다르지 않지만, 인간 형상으로 구현된 캐릭터와 상호 교감을 나눌 수 있다는 점에서 특이점을 갖는다. 2016년 12월부터 일본과 미국에서 Gatebox의 예약 판매가 시작되었으나, 지원언어가 일본어에 한정되어 있고 가격이 2,600불 가량으로 매우 높기 때문에 글로벌 시장에서 성공 가능성은 높지 않다.

〈표 9〉 해외 주요기업의 음성인식 AI 비서 서비스와 스피커 현황 요약

	Amazon	Google	Apple	Microsoft	Baidu	vinclu
						
SW (출시연도)	Alexa (2014)	Assistant (2016)	Siri (2011)	Cortana (2014)	Duer (2015)	—
전용기기 (출시연도)	Echo (2014)	Home (2016)	—	Invoke (2017)	Little Fish (2017)	Gatebox (2016)
지원언어 (한국어)	2개 (지원안함)	8개 (지원예정)	20여개 (지원함)	8개 (지원예정)	1개 (지원안함)	1개 (지원안함)
클라우드 사용여부	사용함	사용함	사용안함	사용함	확인불가	확인불가
강점	만개이상의 Skills 보유	구글 이용자의 빅데이터 보유	개인정보 보호 및 보안 우수	Windows 유저를 잠재적 이용자로 보유	해외기업의 진입장벽 높은 중국시장에서 유리	최초의 홀로그램 AI상용화

자료: 각 사 홈페이지 및 언론자료 취합

2. 국 내

국내에서는 2016년 9월 SKT “누구”를 처음으로 2017년 1월에 KT “기가지니”가 AI 스피커 시장에 진입하였다. 이동통신사의 AI 스피커 외에도 2017년 3월 삼성전자 “갤럭시s8”에 자사 AI 비서 서비스인 “빅스비”와 Google의 “어시스턴트”를 동시에 탑재하면서 스마트폰에서 음성인식 기반 AI 비서 서비스가 적용되었다. 포털에서도 자체 개발한 음성인식 AI 비서 기술을 상용화하려는 움직임을 보이고 있다. 네이버는 2017년 5월, AI 비서 앱인 “클로바”를 출시하였으며, 자사의 검색과 AI 기술을 바탕으로 AI 스피커와 디스플레이를 선보일 계획이다. 이 밖에 다양한 가전제품에 활용되는 것을 목표로 API 공개를 계획하고 있다. 카카오 또한 음성인식 AI 비서 앱과 스피커를 2017년 하반기에 선보일 것으로 계획하고 있다. 국내에는 방대한 검색 정보 데이터를

축적하고 있는 포털, 자사의 가전제품에 기술 적용이 용이한 가전사, 홈IoT 중심의 AI 스피커를 출시한 이동통신사 사이에서 AI 비서 기술과 이를 적용한 서비스 및 기기 시장을 선점하기 위한 경쟁이 치열해질 것으로 예상된다. 본 파트는 음성인식 AI 비서 서비스의 주요 기업별 현황을 살펴보고, 향후 기술 개발 및 서비스 전략을 정리하고자 한다.

(1) 삼성

삼성의 빅스비는 AI 비서 UI로, 2017년 3월말 출시된 갤럭시s8에 탑재되면서 서비스되었다. 2017년 5월에는 “보이스”라는 기능이 추가되었는데, 기존 삼성 스마트폰에 탑재된 “S보이스”나 Apple의 “Siri”와 유사하다. 사용자의 음성 인식 외에도 터치나 텍스트 입력 방식으로 보이스와의 대화, 정보 탐색, 메신저 문자 발신 등이 가능하다. 빅스비의 또 다른 기능인 “비전”은 GPS 설정시 랜드마크를 식별하여 해당 장소에 대한 정보 제공하고, 이미지 속 외국어의 번역이 가능하며, 궁금한 상품을 뷰파인더에 맞추면 해당 상품에 대한 쇼핑 정보를 제공한다.

[그림 10] 삼성 빅스비 연동 예시



자료: 삼성전자 홈페이지(<http://www.samsung.com/sec/apps/bixby/>)

삼성은 스마트폰 외에 가전제품으로 영역을 넓히는 움직임을 보이는데, 최근 “패밀리허브 2.0” 냉장고에 빅스비를 적용시켰다. 해당 상품은 장착된 디스플레이를 통해

음성인식, 다른 가전제품제어, 삼성페이 결제기능, 스마트폰과 연동 기능을 제공하는데, 2017년 5월에 빅스비가 자동으로 업데이트되어 AI 비서 기능을 더하게 되었다. 삼성은 향후 출시할 TV를 포함한 가전제품에 빅스비를 설치할 계획을 밝혔는데(매일경제, 2017. 5. 14), 이는 다양한 가전제품을 개발하는 가전사의 장점을 기반으로 홈IoT 시장을 공략하려는 전략으로 보인다.

(2) SKT

SK텔레콤(SKT)은 2016년 9월에 국내 최초로 AI 스피커인 “누구”를 출시하였으며, 가격은 약 25만 원 선으로 알려졌다. 출시 7개월 만에 10만대 이상을 판매하였으며, 대화량이 1억 건이 넘어(연합뉴스, 2017. 5. 3), 국내 음성인식 AI 스피커 시장의 발전 가능성을 확인하였다. “누구”는 음성인식을 통해 자체에 내장되어있는 스피커, 조명 뿐 아니라 연결된 앱과 기기, IPTV 등을 제어할 수 있으며, 라디오 재생, 날씨와 일정 등 확인, 알람기능 등을 요청할 수 있다. 음성인식 외에도 AI 스피커와 연결된 앱을 통해 추가적인 설정과 이용이 가능하다.

〔그림 11〕 국내 AI 스피커 형태



자료: 각 사 홈페이지

2017년 5월, SKT는 KEB하나은행과 음성인식 AI 금융 서비스의 제공 협약을 체결하여, 등록된 계좌의 잔액 및 거래내역 조회, 환율과 환전 조회, 지점 안내 등을 음성으로 문의하고 듣는 “음성 금융 서비스”를 계획하고 있으며(ZDnet, 2017. 5. 17), 향후 다양한 서비스와 업체 제휴를 통해 AI 비서 서비스 영역을 넓힐 것으로 보인다.

(3) KT

KT는 2017년 1월 UHD TV와 AI 스피커가 결합된 AI TV인 “기가지니(Gigagenie)”를 출시하였다. SKT “누구”가 AI 스피커 중심이었다면 기가지니는 TV를 중심으로 AI 스피커와 함께 홈 비서 역할을 수행하며, 별도 카메라 설치를 통해 홈캠으로 활용할 수 있다. 가격은 단품으로 29만원(카메라는 9만원 별도) 선이며, KT의 인터넷, UHD TV과의 결합 상품도 제공하고 있다. 기가지니는 리모콘 없이 TV를 제어하고, 음악 서비스(지니), 기가지니와 연결된 가전제품 및 홈시스템을 제어하는 홈IoT, 일정 관리 및 알람, 대화형 정보 제공 등의 기능을 수행한다.

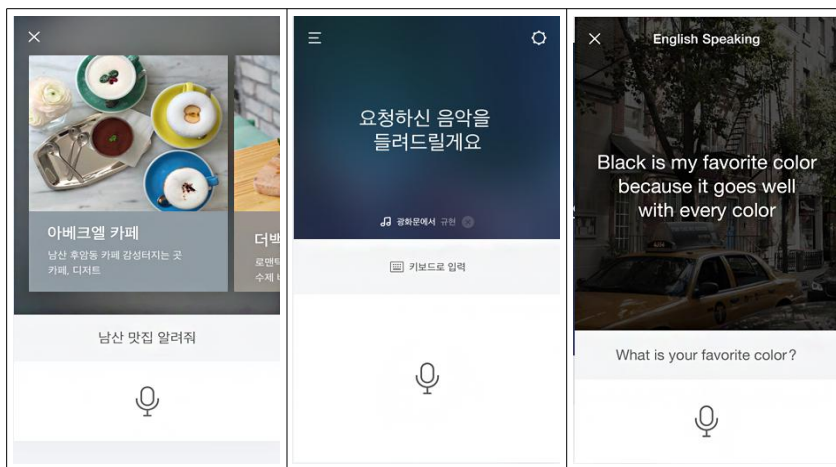
2017년 4월, 미래에셋대우와 MOU를 맺었으며, 주가 및 지수 확인, 시황정보, 종목 및 금융상품 추천 등 금융정보를 제공할 예정이다(매일경제, 2017. 5. 30) 이 밖에도 기가지니는 모터쇼에서 자동차와 연동해 원격시동과 위치안내 등의 차량제어를 시연하면서(아시아투데이, 2017. 5. 5), 다양한 영역에서 음성인식 AI 비서 서비스의 확장할 것으로 보인다

(4) 네이버

2017년 5월, 네이버는 자회사 라인(Line)과 함께 개발한 인공지능 비서 앱인 “클로바(Clova)”를 베타 버전으로 공개하였다. 클로바 공개 이전인 2017년 3월에 베타 버전으로 음성인식 AI 대화형 엔진인 “네이버i”을 선보였는데, 클로바에는 네이버i를 포함한 다양한 AI 알고리즘이 적용되었다(전자신문, 2017. 5. 18). 네이버i는 음성인식, 사용자 대화 이해 및 대화 관리, 자연어 생성, 음성합성 기술, 네이버가 보유하고 있는 방대한 검색DB와 연결하여 개발한 AI 대화형 엔진 서비스이다. 네이버i가 검색 결과를 정확하게 제공하는 것이 목표라면, 클로바는 스스로 판단하는 능력에 초점을 맞췄다

(전자신문, 2017. 5. 18). 사용자의 대화를 누적해 취향을 고려한 콘텐츠를 추천하는 기능이 그 예이다. 이 밖에 클로바와의 대화, 필요한 정보 제공(음성 검색), 일정 관리, 콘텐츠와 음악 추천 등의 기능을 제공한다. 네이버는 AI 서비스에 검색 정보 데이터를 적극 활용하고 있는데, 이는 음성인식 AI의 대화 데이터뿐 아니라 포털에 축적된 방대한 검색 데이터로 사용자가 원하는 정보의 정확성을 높여 AI 비서 기능의 고도화 하려는 전략으로 보인다.

[그림 12] 네이버 클로바 연동 예시



자료: 네이버 클로바 홈페이지(<https://clova.ai/ko>)

AI 비서 서비스를 제공하는 다른 업체와 유사하게 네이버 또한 개발툴을 써드파티와 제조사들에게 오픈할 계획을 밝혔다. 네이버는 자사의 음성인식 AI 기술이 스피커, 자동차, 냉장고 등의 다양한 기기에서 활용되는 것을 목표로 API를 연내 개방하여 네이버 클라우드 플랫폼(NCP)를 통해 서비스를 개발 가능케 할 계획이다(아이뉴스24, 2017. 5. 22). 하드웨어 제조사, IoT 업체를 포함한 개발사들을 위해 클로바의 API, SDK, 개발 문서를 제공하는 플랫폼(Clova Interface Connect: CIC)를 제공할 계획이다(클로바 홈페이지 참고).

(5) 카카오

카카오는 음성 인식에 관심을 갖고 2012년 음성인식 전문 기업인 “다이알로이드”를 인수하여 음성 언어를 컴퓨터가 해석하여 문자로 변환해주는 시스템인 “뉴톤(Newton)”과 문자를 음성 언어로 변환해주는 시스템인 “뉴톤 톡(Newton talk)”을 개발하였다(아이뉴스24, 2014. 6. 26). 뉴톤과 뉴톤 톡의 API를 2014년 6월부터 공개하여(아이뉴스24, 2014. 6. 26), 음성인식 서비스의 확장을 위해 써드파티에게 개발툴을 공유하였다. 최근 카카오앱과 카카오내비에도 “뉴톤”이 탑재되었으며, 향후 카카오지하철과 카카오버스 등의 앱에도 음성 검색 기능을 도입할 계획이다(머니투데이, 2017. 2. 17).

카카오는 기존의 음성인식 서비스 확장 외에도 AI 비서 서비스를 추진하고 있다. AI 기반의 메시저 로봇인 “챗봇”과 AI 스피커 등을 포함한 음성 프로젝트를 진행하고 있으며, 2017년 하반기에 AI 비서 앱과 AI 스피커를 출시할 예정이다(전자신문, 2017. 5. 11). 이용자가 많은 카카오톡, 멜론 등 카카오의 다양한 서비스들이 AI 비서 앱이나 스피커와 연동될 경우, AI 서비스 이용자 유입에 도움이 될 것으로 보인다.

〈표 10〉 국내 주요 기업의 AI 비서 서비스 현황

기업	서비스(출시일)	종류	특징
SKT	누구 (NUGU)	AI 스피커	<ul style="list-style-type: none"> - 국내 최초 음성인식 AI 비서 스피커 - 조명, 연결된 앱과 기기, IPTV 제어 - 개인 스케줄 관리 및 알람, - KEB하나은행과의 협약으로 계좌, 주식, 환율 정보 제공 계획
KT	기가지니 (GiGA Genie)	AI TV/스피커	<ul style="list-style-type: none"> - AI TV와 스피커 결합으로 홈IoT 환경 제공 - 기가지니 카메라 별도 설치를 통해 홈캠 활용 가능 - TV 및 음악 감상, 일정관리, 홈IoT(홈 가전제품, 홈 시스템 제어), 주식 정보 제공

기업	서비스(출시일)	종류	특징
삼성전자	빅스비 (Bixby) (2017. 3, “보이스”는 2017. 5)	AI 비서 UI	<ul style="list-style-type: none"> - 등록된 사용자 음성 인식 - 스케줄 확인, 뉴스 정보 제공, 앱 이용, 콘텐츠 추천 - 카메라로 사물, 이미지, 텍스트, QR 코드 등을 인식해 정보(번역, 쇼핑 등) 제공 - 현재 스마트폰 “갤럭시s8”과 냉장고 “패밀리 허브2.0”에 탑재
네이버	클로바(Clova) 베타 (2017. 5) 웨이브(AI 스피커), 2017. 하반기 출시예정	AI 비서 앱	<ul style="list-style-type: none"> - 음성인식, 번역, 콘텐츠(음악, 영화 등)추천 - AI 대화형 엔진 네이버아이(i)와 다양한 AI 기술을 접목 - 포털에 축적된 방대한 검색 데이터와 딥러닝 기술을 더해 사용자가 원하는 정보의 정확성 향상 가능 - 페이스(AI 비서 디스플레이) 출시 계획
카카오	AI 비서 앱 2017. 7 출시예정 AI 스피커 2017 하반기 출시예정	AI 비서 앱	<ul style="list-style-type: none"> - 음성 언어를 컴퓨터가 해석하여 문자로 변환해주는 시스템인 “뉴톤”과 문자를 음성 언어로 변환해주는 시스템인 “뉴톤 톡”의 기술 기반 - 2017년 하반기 AI 비서 앱과 AI 스피커 출시 예정

자료: 각 사 홈페이지 및 언론자료 취합

IV. 결론 및 시사점

2008년부터 2015년까지 매년 두 자리 수의 높은 성장세를 유지하던 스마트폰 시장은 그 속도가 점차 둔화되고 있는 것으로 보인다(Gartner, 2016). 중국, 멕시코 등 신흥 시장에서도 시장 성숙에 따라 신규 수요는 줄어들고 있으며, 대부분의 스마트폰 제조사들이 교체 수요에 의존해야하는 상황으로 분석된다. 스마트폰 시장 포화 상황에서 새로이 주목받고 있는 새로운 시장 중 하나는 사물인터넷(IoT)이다. 무선인터넷을 통해 집안의 각종 시설과 가전, 전자 제품이 서로 통신을 주고 받을 수 있게 되면서, 이를 컨트롤 하는 허브 역할을 두고 TV, 셋톱박스, 냉장고, 스마트폰 등이 경쟁해왔다. 그러나 2017년 현재 스마트홈을 통제할 핵심 기기로 가장 주목받는 것은 음성인식 및 AI

기능이 결합된 “스피커”이다. Google, Apple, Microsoft 등 글로벌 IT 공룡들이 제각기 자사 소프트웨어를 탑재한 스피커를 출시하며 Echo가 독식하던 시장에 진입했으며, 국내 시장도 경쟁이 가속화되고 있다. 2016년 SKT “누구”를 시작으로 2017년 KT, 네이버, 카카오 등 진입(혹은 예정) 사업자가 많아 경쟁이 치열해질 것으로 보인다.

각 사는 스피커뿐만 아니라 음성인식 AI 기술력을 확장하여 AI 시장 우위를 점하는 것을 목표로 삼고 있다. 앞서 살펴본 바와 같이 Amazon은 Echo의 라인업을 다양화하고 있으며, 현 시점에서는 가장 높은 시장점유율로 유리한 고지에 있다. 기존에 충성도 높은 고객층을 보유하고 있는 Apple, Microsoft와 삼성은 AI 비서 서비스를 향후 자사 제품에 적용할 것을 계획하고 있는데, 이를 통해 고객의 편의성을 향상시키고 이용 데이터 축적을 통해 자사 AI 비서 서비스의 품질을 안정화 시킬 수 있을 것으로 보인다. Google과 국내 포털사는 다수의 경쟁력 있는 온라인서비스, 축적된 정보량, 검색 기술 등을 기반으로 AI 비서 서비스에서 이동통신사의 홈IoT 중심의 AI 전략과 차별화될 것으로 보인다. 마지막으로 이동통신사의 경우, 음성인식 AI 서비스를 통해 홈 시스템과 가전제품을 제어하는 홈IoT 중심으로 확장될 것으로 보이며, 온라인 서비스 자회사(인터넷 쇼핑몰, 음악제공서비스 등)의 콘텐츠를 활용하여 온라인상에서도 경쟁력을 확보할 것으로 보인다.

이러한 시도는 기기를 더 많이 판매하기 위한 경쟁으로 이해하기보다는 향후 터치나 텍스트 입력 방식을 대체할 음성 인터페이스의 표준을 선점하기 위한 노력으로 바라보아야 할 것이다. 특정 음성인식 AI 소프트웨어나 스피커를 이용하여 제어할 수 있는 서비스나 기기가 많아질수록 이 경쟁의 승자가 될 가능성이 높아지기 때문이다. 기업들은 자사의 음성인식 기술을 외부에 개방함으로써 독자적인 오픈 생태계를 조성하려 노력하고 있다. 향후 음성인식 AI 개발툴 공개로 음성인식 AI 관련 응용 서비스들이 활발히 개발될 것으로 기대되며, 이는 국내 음성인식 AI 시장을 활성화되는데 기폭제가 될 것으로 예상된다. Amazon처럼 자사 이커머스 플랫폼과 시너지를 추구하는 것 외에는 아직 명확한 비즈니스 모델이 없지만 현 시점에서는 시장 규모를 키우기 위한 이용자 확보가 당면 과제로 생각된다.

국내 음성인식 AI 서비스들이 국내외 시장에서 글로벌 기업과 경쟁하는 것은 불가피한 상황이다. 국내 시장에서 Apple Siri와 같이 이미 한국어 지원을 하고 있는 글로벌 기업과 경쟁해야한다. 대부분의 국내 음성인식 AI 비서 서비스들은 한국어 중심으로 실행되고 영어를 포함한 외국어 지원은 차후에 계획되어 있어, 세계 시장에서의 경쟁력을 갖추기보다 우선은 국내 시장을 선점하는데 주력할 것으로 보인다.

AI 스피커 시장의 본격화에 앞서 깊고 넘어가야 할 중요한 이슈 중 하나는 이용자의 프라이버시 보호와 관련된 사항이다. AI 스피커를 통해 수집된 데이터는 무선인터넷을 통해 클라우드 서버로 전송되기 때문에 데이터 도난이나 해킹에 대한 우려도 존재한다. 실제 개인용 클라우드 백업 서비스인 iCloud나 Dropbox의 해킹이 꾸준히 증가하고 있으며, 그로 인한 프라이버시 침해 사례도 증가하고 있다. 국내 가정용 IP카메라로 촬영된 영상이 외부 인터넷망을 통해 전송되는 과정에서 유출되어 중국 사이트에 게시된 사례도 있었다(보안뉴스, 2017. 5. 4). 이용자의 집 안에서 발생하는 모든 소리를 녹음하고 클라우드로 전송하는 음성인식 AI 스피커 역시 해킹 위험에서 자유로울 수 없다.

또한, 앞서 살펴본 바와 같이 AI 스피커는 전원이 연결되어 있을 때에는 항상 이용자의 음성에 귀 기울이고 있다. 즉, 상시 녹음이 이루어지고 있으며 그 데이터가 Amazon이나 Google의 클라우드 서버에 저장되고 있는 것이다. 그 뿐만 아니라 음성인식 기반 AI 스피커는 이용자가 시청한 TV 쇼프로그램이나 인터넷 검색결과, 구매내역, 수면 시간 등의 행동 정보를 지속적으로 수집하고 있다. 물론 상시 녹음 기능은 음소거(Mute) 기능을 사용할 경우 끌 수 있지만, 음성명령을 내리기 위해서는 다시 녹음 기능을 활성화시켜야하기 때문에 불편함이 발생한다. 이러한 기능에 대해 어떤 이들은 “엿듣고 있다”라고 생각할 수도 있을 것이다. 또한 수집된 행동 정보를 바탕으로 특정한 제품 광고를 노출하거나, 구매를 제안하는 등의 상업적 이용이 이루어질 가능성도 있다.

AI 스피커를 출시한 기업들은 수집된 개인정보는 비밀이 유지될 것이라 약속하지만, 구체적으로 어떤 방식으로 이 개인정보가 보호되는지에 대해서는 알려져 있지 않다. 따라서 각 사는 이에 대해 구체적이고 투명한 대책을 마련하고 잠재적 이용자들의 신뢰를 확보하는 것이 우선이다.

참고문헌

- 권오욱·최승권·노윤희·김영길·박전규·이윤근 (2015. 8), “자유발화형 음성대화 처리 기술동향”. 전자통신동향분석 Vol.30, No.4, Aug. 2015, pp.26-35
- 김대호·최선규·이재신·신동희·안재현·전경란·이상우·김성철·김도연·심용운 (2014), “ICT 생태계”, 커뮤니케이션북스
- 박병근 (2016), “음성인식 스피커 Amazon Echo 생태계”. 디지에코 보고서(2016. 6. 9.) Issue & Trend
- 최은정 (2013), “클라우드 혁명이 바꾸는 미래”. SERI경영노트, 2013. 3. 28(제 181호)
- 한국콘텐츠진흥원 (2011. 11), “음성인식 기술의 동향과 전망”. 2011년 11월 문화기술 (CT) 심층리포트.
- Brown, R. (2008). “Exploring new speech recognition and synthesis APIs in Windows Vista. Talking Windows” MSDN Magazine. <http://msdn.microsoft.com/hi-in/magazine/cc163663.aspx>.
- Colin, C., (1953). “Some Experiments on the Recognition of Speech, with One and with Two Ears” The Journal of the Acoustical Society of America. 25(5): 975-79. doi:10.1121/1.1907229. ISSN 0001-4966.
- Furui, S. (2005). “50 years of progress in speech and speaker recognition research”. ECTI Transactions on Computer and Information Technology(ECTI-CIT), 1(2), 64-74.
- Gartner (2016. 9). “Forecast Snapshot: VPA-Enabled Wireless Speakers, Worldwide, 2016”.
- _____ (2017. 3). “Market Trends: Human-Machine Interactions — Role of Voice Across Devices and Ways to Enhance Customer Experience”
- KPCB (2016. 5). “Internet Trends Report 2016”.
- Saon, G., Kurata, G., Sercu, T., Audhkhasi, K., Thomas, S., Dimitriadis, D., Cui, X., Ramabhadran, B., Picheny, M., Lim, L.L. and Roomi, B. (2017).

“English conversational telephone speech recognition by humans and machines”. arXiv preprint arXiv:1703.02136.

[참고자료]

- 《동아일보》 (2017. 5. 16), “연구, 개발의 대가 MS의 4가지 혁신 비법”.
- _____ (2015. 1. 23), “IFTTT, 네가 다 알아서 해줘”.
- 《머니투데이》 (2017. 2. 17), ““너의 목소리가 보여” 음성기술 집중하는 네이버와 카카오”.
- 《매일경제》 (2017. 5. 14), “인공지능 빅스비, 삼성 가전에 첫 탑재”.
- _____ (2017. 5. 30), “KT, AI·5G상용화·스마트에너지…4차 산업혁명 이끈다”.
- 《아시아투데이》 (2017. 5. 5), ““누구” “기가지니” AI서비스 대중화…수익 모델 확보 방안은?”
- 《아이뉴스24》 (2017. 5. 22), ““클로바 API 곧 제공…”네이버 클라우드, 모든 시장 타깃”
- _____ (2015. 6. 30), “[정구민] 애플 홈킷이 제시하는 미래 스마트홈 서비스”.
- 《연합뉴스》 (2017. 5. 3), “SK텔레콤 “누구” 판매량 10만대 돌파”.
- 《전자신문》 (2017. 5. 11), “[이슈분석] AI 스피커 뛰어드는 네이버·카카오”
- _____ (2017. 5. 28), “[이주의해시태그-#네이버-클로바] 쑥쑥 크는 네이버 AI”.
- 《조선일보》 (2016. 12. 18), ““엄마”와 “숙모”도 구분 못하던 음성인식, AI달고 “헐헐””.
- 《한국경제》 (2016. 9. 2), “AI기술 공들인 중국 바이두, 자율주행차 개발 속도낸다”.
- 《Kotra해외시장뉴스》 (2016. 11. 21), ““구글 홈” 출시로 스마트 스피커 시장 경쟁 본격화”
- 《Microsoft News Center》 (2016. 12. 14), ““마이크로소프트, 모두를 위한 AI” 향한 대화형 컴퓨팅의 새로운 비전 공유”
- 《ZDNet》 (2017. 5. 17), “KEB하나은행-SKT, 국내 첫 인공지능 스피커 금융서비스”.
- _____ (2017. 1. 6), “바이두, 중국판 아마존 에코 만든다”.

- 《CNBC》(2017. 5. 17), “Microsoft is halfway to its goal of 1 billion Windows 10 devices”.
- 《CNET》(2017. 2. 6), “Make Alexa order your Iced Grande Caramel Macchiato”.
- 《Digital Trends》(2017. 4. 4), “A holographic virtual girl lives inside Japan’s answer to the Amazon Echo”.
- 《Experian》(2016. 9. 7), “[Infographic] Unpacking the Breakout Success of the Amazon Echo”.
- 《Filehippo》(2014. 2. 26), “Halo’s Cortana: Siri’s Nemesis?”.
- 《Mashable》(2016. 12. 17), “Hologram-in-a-jar is the perfect buddy for lonely people with no friends”.
- _____ (2017. 2. 3), “How to use third-party Actions on Google Home”.
- 《TechCrunch》(2017. 1. 8), “How Amazon and Nvidia won CES this year”.
- 《Techradar》(2016. 8. 20), “How to control your smart home with IFTTT”.
- 《The Economist》(2017. 5. 1), “Technology Quarterly: Finding a Voice”.
- 《The Verge》(2014. 10. 23), “Amazon’s losses balloon to \$544 million, driven by weak Fire Phone sales”.
- _____ (2017. 5. 10), “Microsoft shows how Cortana will work in speakers and cars”.
- _____ (2017. 1. 5), “Baidu’s “Little Fish” home robot could be China’s Echo”.
- 《USA Today》(2016. 3. 28), “Amazon Echo turns into a sleeper hit, offsetting Fire’s failure”.
- _____ (2017. 5. 15), “Google Home wants to be your assistant, heating up the rivalry with Amazon Echo and soon, Apple”.
- 《Wired》(2016. 6. 24), “The Amazon Echo is winning the race to a screenless future”.
- 《ZDNet》(2016. 9. 29), “Amazon, Google, Facebook, IBM, and Microsoft form AI non-profit”.
- _____ (2017. 1. 4), “CES 2017: Voice is the next computer interface”.

<https://www.amazon.com>
<https://www.apple.com/ios/home>
<https://www.apple.com/ios/siri>
<https://assistant.google.com>
<https://build.microsoft.com>
<https://clova.ai/ko>
<http://developers.daum.net/services>
<http://gatebox.ai/>
<http://gigagenie.olleh.com>
<http://www.iotenableddevices.com>
<https://www.microsoft.com/en-us/windows/cortana>
<https://www.netmarketshare.com>
<http://www.nugu.co.kr/product.jsp>
<http://www.samsung.com/sec/apps/bixby/>
<http://searchengineland.com/>
<https://www.statista.com>
<http://www.valuewalk.com>
<http://word.tta.or.kr>