

접근성 개선을 위한 AI음성인식 디바이스 활용

(주)팀랩스
박선호(shpark@teamlabs.co.kr)

요약

음성인식 기반 인공지능 서비스는

'14년 Amazon Alexa Echo 서비스가 상용화된 이후로

구글(구글 홈), 애플(시리), 마이크로소프트(코타나) 등이 합세하였고,

중국(바이두, 징둥, 샤오미, 알리바바)에 이어

국내에서도 SKT(Nugu), KT(기가 지니), Naver(클로바) 등의 플랫폼이 서비스를 하고 있음

키보드, 마우스와 터치스크린에 이어 음성인터페이스가 융합되어

빠른 입력과 함께 풍부한 감성과 자연스러운 대화형태의 정보이용이라는 새로운 사용자 경험을 제공하기 시작함

이러한 사용자경험은 정보이용에 있어

등 공공서비스 수요자나 노약자나 장애인 등 도움이 필요한 계층에게 크게 도움이 되겠지만

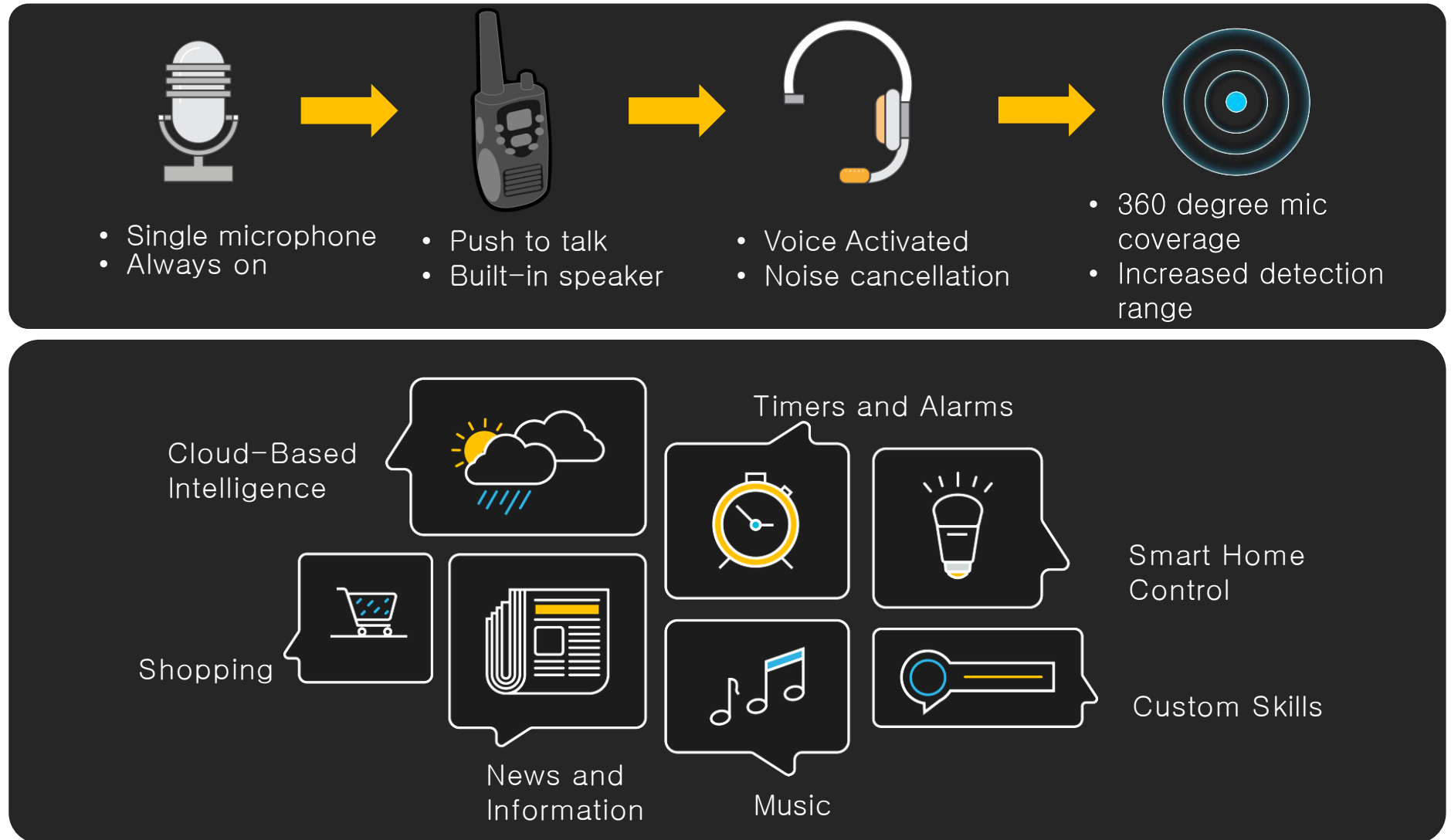
현재의 음성인식 기반 인공지능을 활용하는 서비스는 일상적인 비즈니스 영역의 서비스를 주로 제공하고 있음

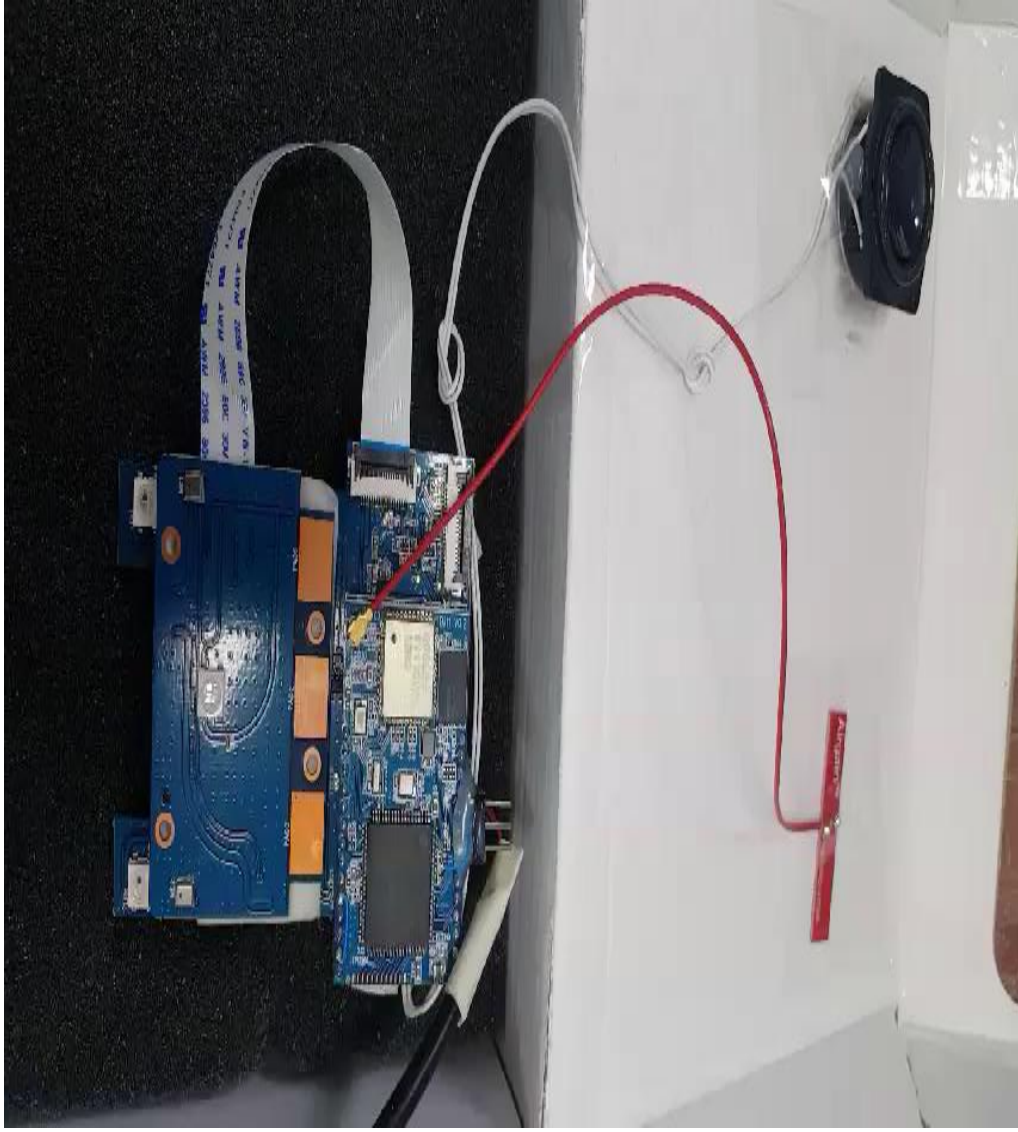
따라서 국내에서 개방되어 운영중인 플랫폼에 계층별 수요자를 위한 AI서비스를 체계적으로 구현하여 제공하는 것은 의미 있겠으며

특히 풍부한 공공DB를 활용하여 동적인 정보형식으로 정보이용 취약자들에게 음성인식 기반 인공지능 서비스 확대를 제언하고자 함

Voice Interaction의 전환 과정

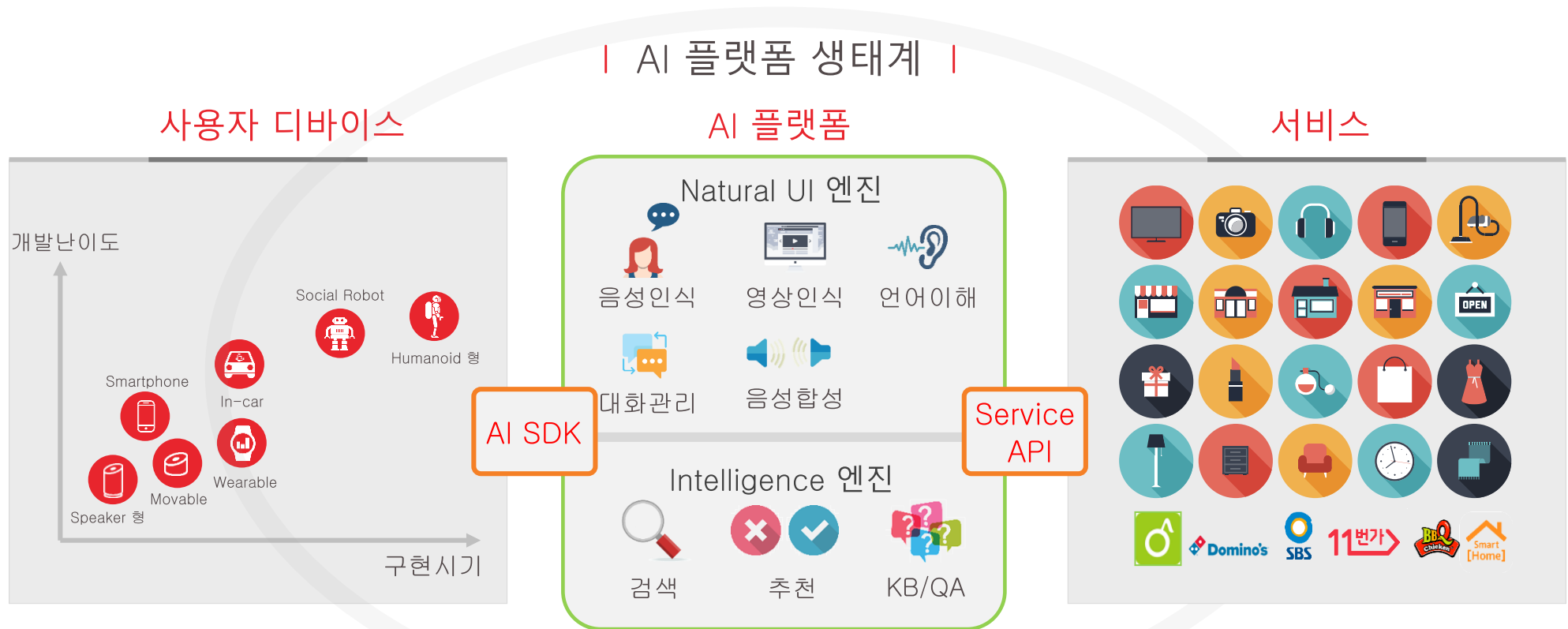
음성처리의 발전과 현재 모습





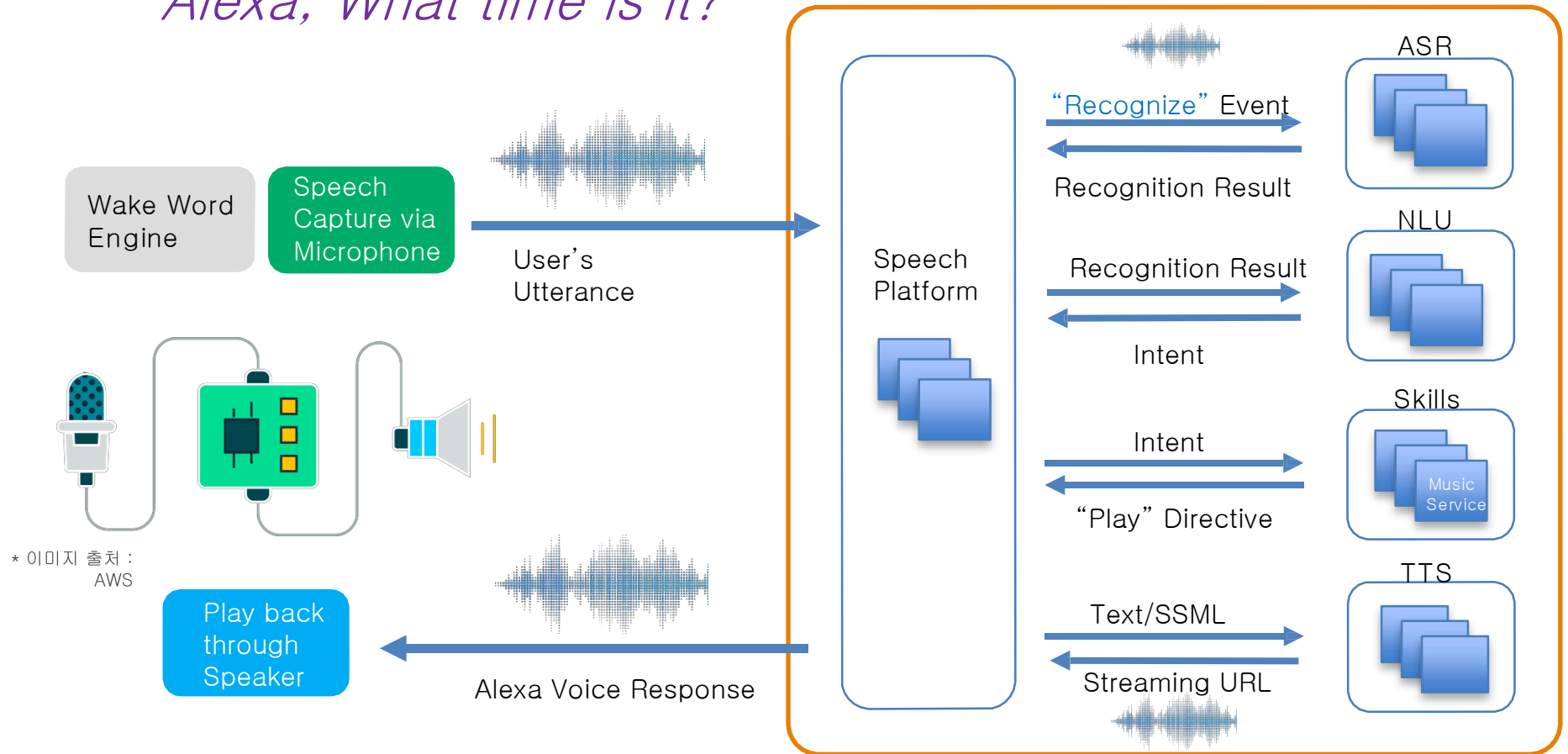
AI Ecosystem

AI 플랫폼 사업자들은 자사 플랫폼과 연동하는 환경을 폭넓게 제공하기 위하여 API/SDK 공개정책을 강화하면서 다수의 3rd Party가 참여하는 AI플랫폼 생태계를 조성하고 있으며, 동시에 자사의 AI플랫폼 고도화를 통하여 시장과 서비스가 더욱 활성화될 수 있도록 노력하고 있습니다.



음성인식 구조

“Alexa, What time is it?”

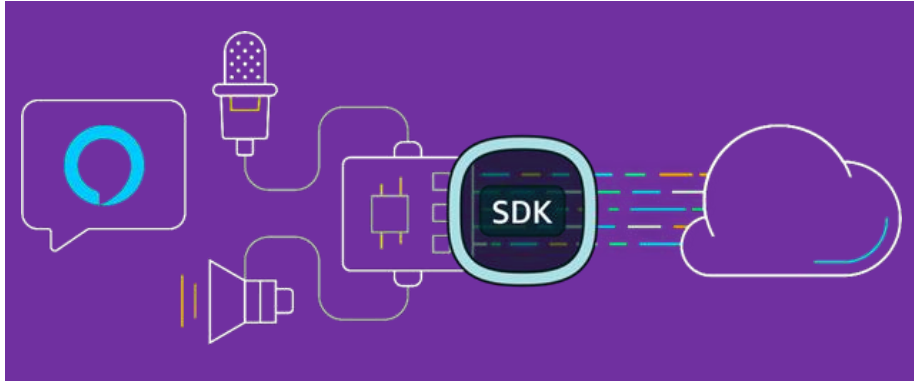


- * Automatic Speech Recognition (ASR)
- * Natural Language Understanding (NLU)
- * Text To Speech (TTS)

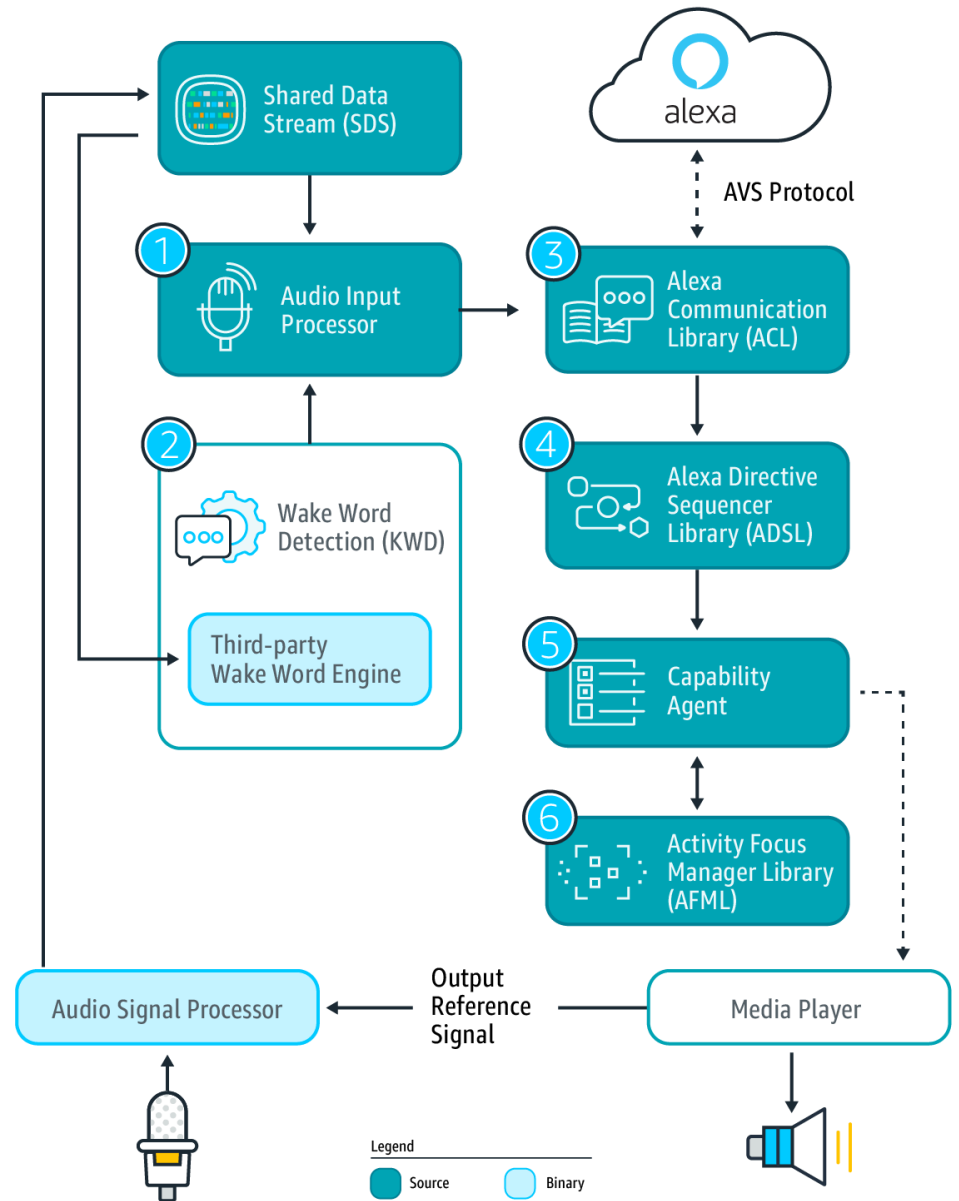
“It's 4:30 PM”

- * Speech Synthesis Markup Language (SSML)

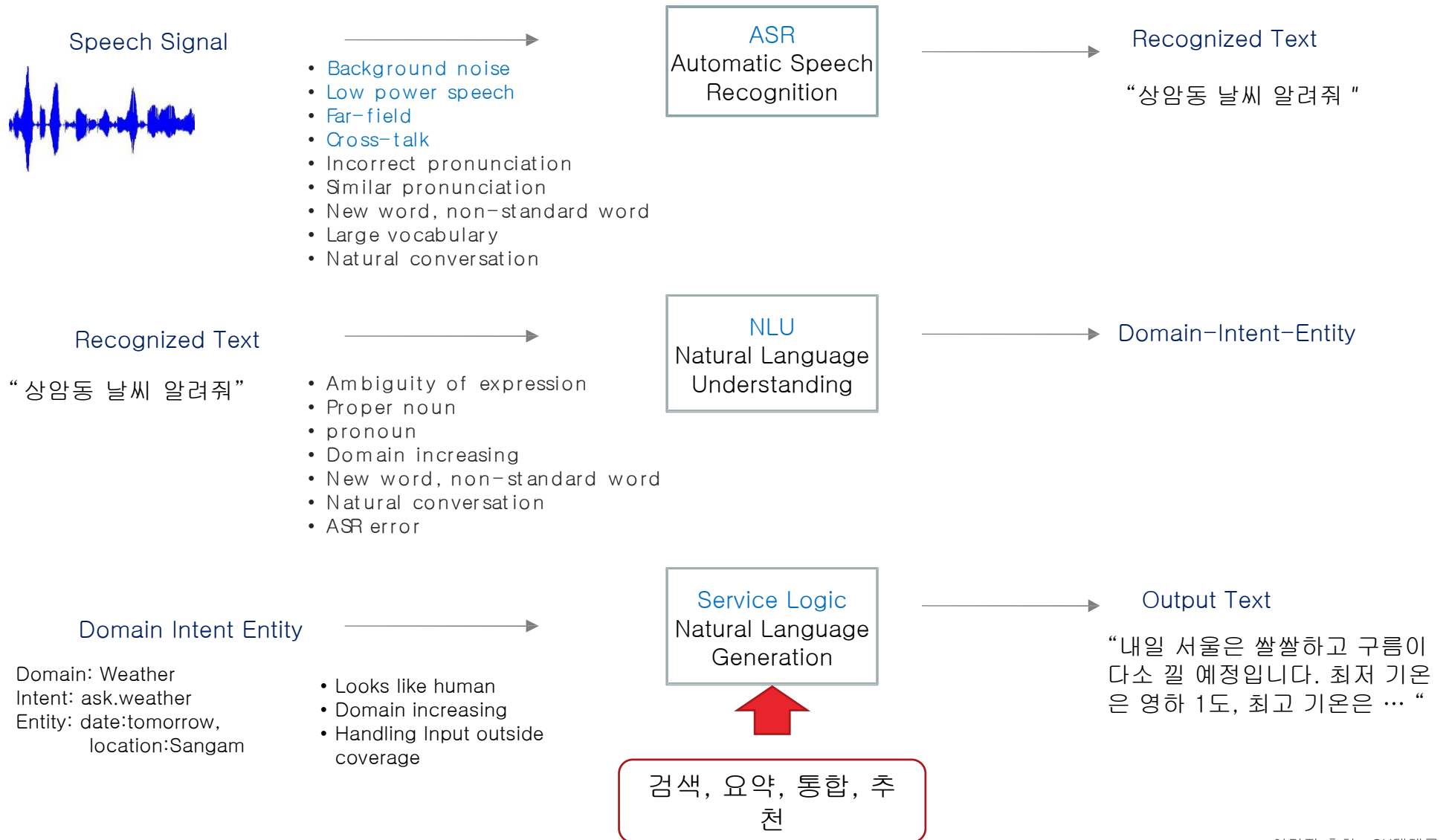
AVS SDK for Device software



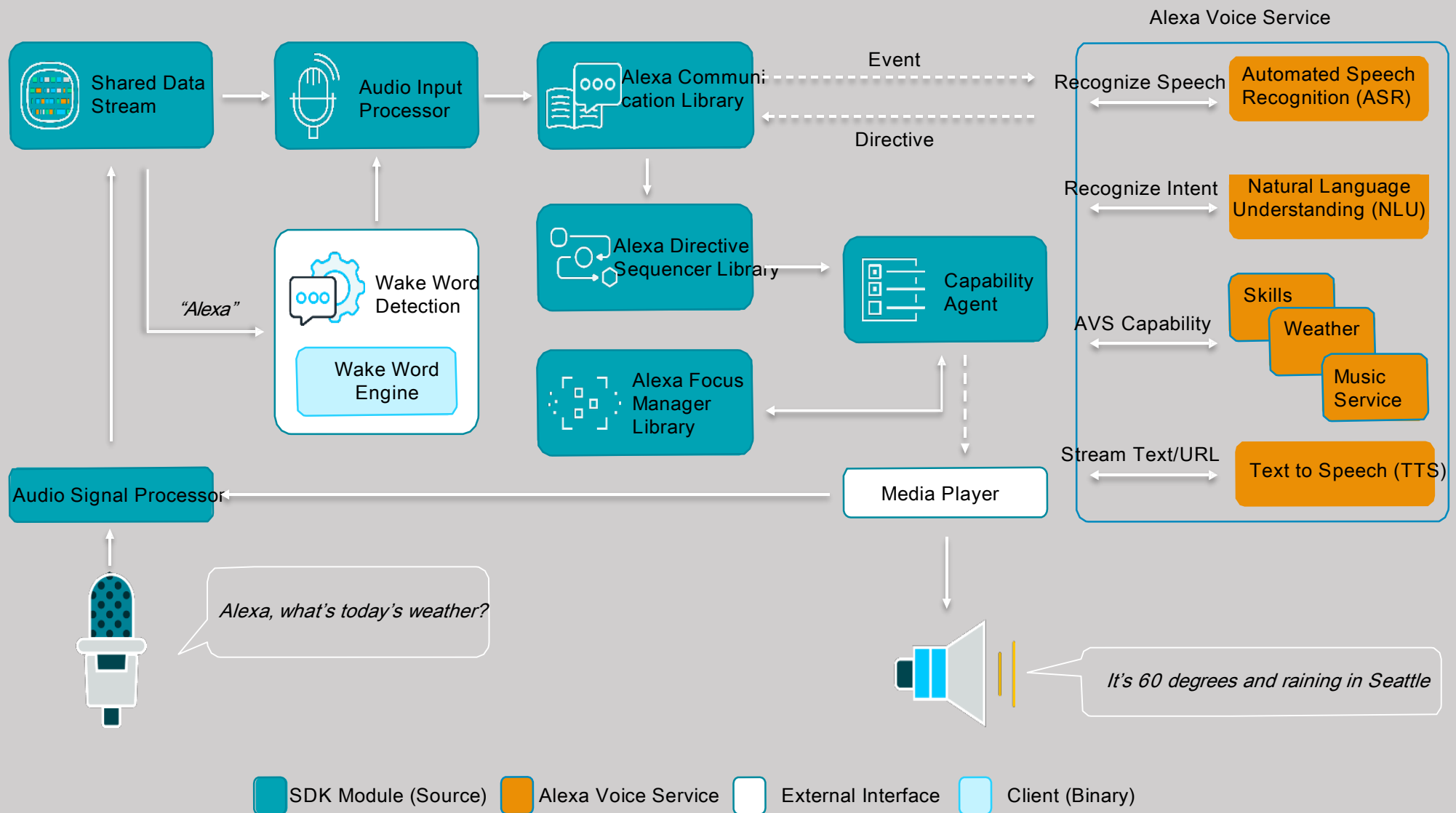
- Develop commercial-grade device software for AVS products
- Provides C++-based libraries to:
 - Process audio inputs
 - Establish persistent connections with AVS
 - Handle all Alexa interactions on the device
- SDK Architecture
 - Audio Input Processor(AIP)
 - Wake Word Detection(WWD)
 - Alexa Communications Library(ACL)
 - Alexa Directive Sequencer Library(ADSL)
 - Capability Agent
 - Activity Focus Manager Library(AFML)



용어 참조



플로우 참조



AI 서비스 디바이스 제품 동향

기존 디바이스 제품 동향

음성인식 AI 스피커

- Harman Kardon 오디오 튜닝, 캐릭터 디자인 등으로 이용자의 선호도와 취향을 만족시키며 상품성을 높이는데 주력하고 있음

IPTV STB결합형

- 거실 TV를 위한 셋탑박스 결합형으로 TV채널 제어, VOD 주문 등 사용자의 AI서비스 이용성을 높이는 것과 함께
- IPTV 셋탑박스 교체 TV가입자를 대상으로 하는 디바이스 보급을 확대하기 위함

디스플레이 형

- Amazon Echo 라인업으로 비디오 정보를 추가함으로써 정보제공의 완성도를 높이고 있으며,
- Amazon쇼핑몰의 보다 자세한 상품 정보를 제공하는 것을 타겟으로 함

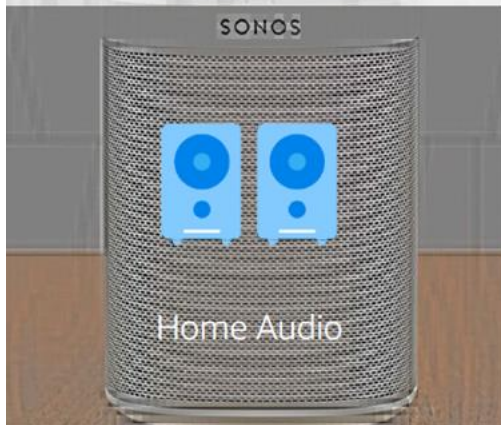
BT연결형

- 스마트폰 App과 연동하여 Far-field를 BT device에 구현하는 형태
- 헤드셋(InnoDesign), 에프터마켓 차량(Anker), Cane(WE-WALK)



AI 서비스 디바이스 제품 동향

Sonos Integrates Alexa for Voice Control



LG's InstaView Puts Alexa in a Fridge



DISH Brings Alexa voice control to Hopper 3 DVR



Martian mVoice Smartwatch with Alexa



Ford Brings Amazon Alexa to its Sync 3 Vehicles



Anker Unveils \$35 Eufy Genie with Alexa



Ecobee4 Upgrades the Smart Thermostat with Alexa



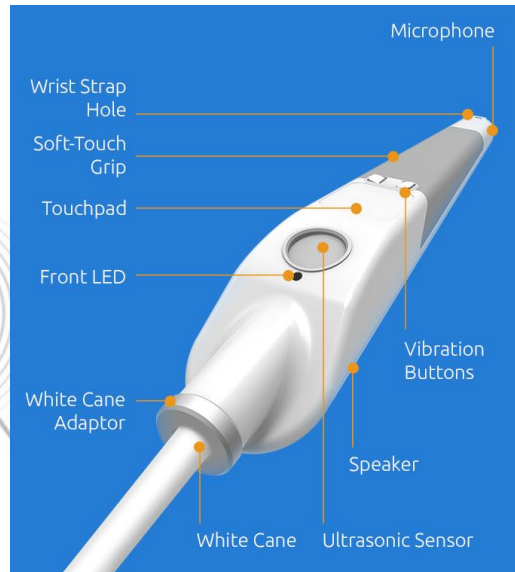
HTC U11 Brings Alexa to Android Phones



시각 장애, 청각 장애관련 시디바이스



“우선 시각 장애인들이 길을 걸을 때 가장 필요한 기능이죠. 장애물 감지 기능입니다. WE-WALK는 초음파 센서를 내장하고 있어 전방에 장애물이 있는 것을 알 수 있도록 해줍니다. 또한 진동과 사운드를 통해 경고를 해 줍니다.”



“음성인식은 청각 장애인 사용자의 리듬을 인식하기가 어렵다. 수화 인식 장갑부터 수화 아바타까지 많은 유사 프로젝트가 있지만, 수화는 손의 움직임 뿐만 아니라 얼굴 움직임에 따라서도 의미가 달라지기 때문에 수화 문법을 인식하는 데에 어려움이 있다.” – Abhishek Singh

시각 장애, 청각 장애관련 AR, AI 기술이용 디바이스

Vision Enhancement Using Audio



- 최근 인텔이 153 억 달러에 인수한 MobileEye에 이어서 2010년 설립된 이스라엘의 스타트업
- 개발 제품 MyEye는 시각장애인 또는 난독증 환자용 (3,500 달러)으로 안경에 부착해서 사용, 휴대 가능하며 직관적인 UX 를 제공함
- 텍스트를 읽거나 물체를 식별하는 일, 상대방 얼굴 또는 제품을 인식할 수 있게 해주는 장치로서, 영어와 히브리어를 지원
- 안경 프레임에 부착된 소형 카메라로 이미지를 스캔하여, 텍스트 또는 오디오로 변환

Vision Enhancement General



- eSight는 고속카메라와 2개의 OLED 스크린을 탑재
- 색상, 대비, 초점, 밝기 및 배율을 최대 24배까지 조절할 수 있는 카메라를 통해 사용자가 실시간으로 그들 앞에 위치한 모습을 순간 포착해 나중에 다시 보기도 가능
- 비디오 및 게임 스트리밍도 할 수 있는 시각장애인들을 위한 스마트안경(9,995달러)
- 사용자가 근거리와 원거리 사물을 쉽게 전환할 수 있도록 자동 포커스기능을 제공하며, 수동으로 사물을 줌인할 수도 있음. 블루투스 및 와이파이 기능 탑재

시각 장애, 청각 장애관련 AR, AI 기술이용 디바이스

Vision Enhancement General



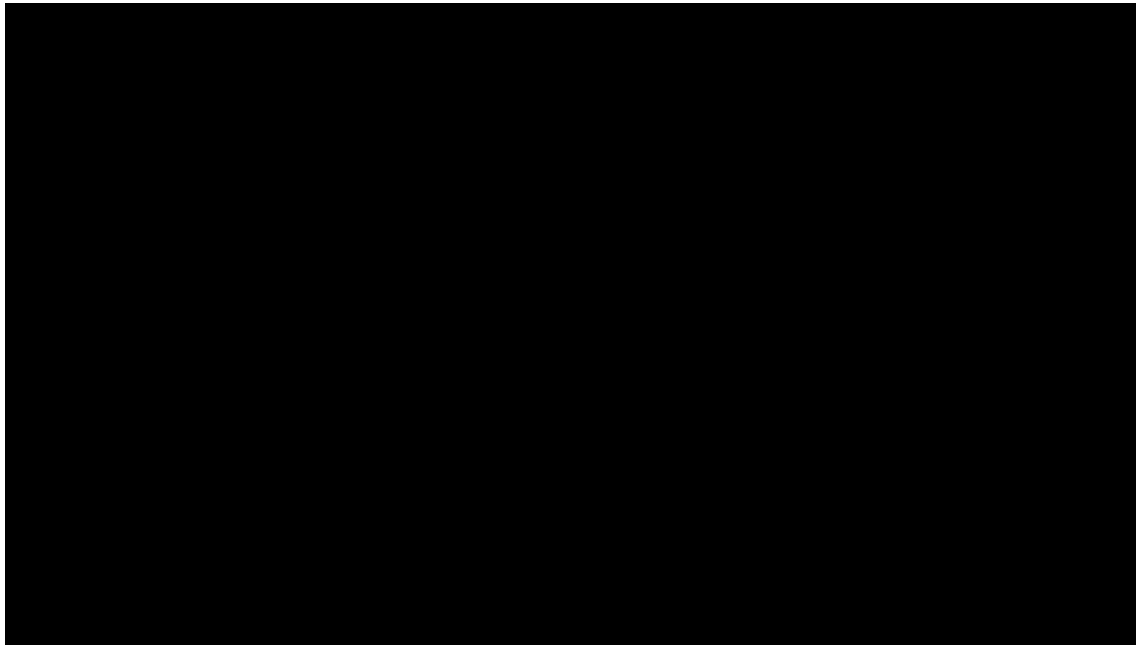
- 2014 년에 설립된 영국 스타트업으로 647,600 달러 투자로 출발함
- 시력강화 키트 SightPlus 를 출시(임상시험 완료), 증강현실을 이용한 시각보조장치
- 남아있는 시력을 향상시키고, 초점을 맞춰주는 역할을 함.
- 시력장애(특히 황반변성)의 영향을 받지 않은 눈에 빛을 집중시켜 줌
- SightPlus VR 은 전자 헤드셋 형태로 색상, 대비, 밝기, 배율 설정이 가능하며, 사용자가 보는 것을 라이브 스트리밍으로 전송 및 공유할 수 있음.



- 2016 년에 설립된 미국 기업으로, Osterhout Design Group(ODG)의 파트너사
- 대표제품 NuEyes Easy 는 시각장애인을 위한 웨어러블 디바이스(5,995 달러)로, 다양한 확대배율, 심도 있는 대비 및 색상 설정, 텍스트-음성 변환 및 미디어 스트리밍 기능 제공
- NuEyes Pro(6,195 달러)는 이메일, 소셜미디어 애플리케이션, 인터넷 브라우저, 그리고 안드로이드 태블릿의 모든 기능이 추가됨. 음성 또는 리모컨으로 작동 가능

시각 장애, 청각 장애관련 AR, AI 기술이용 디바이스

Vision Enhancement Using Humans



- 2015년에 설립된 미국기업으로 전맹을 대상으로 하는 원격 보조 기술임.
- Aira를 호출하면, visual interpreter가 원격으로 장애인과 통신하며 장애인이 착용한 스마트 글라스를 통해 해당자의 위치를 파악하고 주변에 보이는 것들을 즉시 알려줌.
- Aira의 플랫폼은 시각장애인 주변의 환경을 탐색하기 위해 증강현실 대시보드를 제공하며, 머신 러닝 기술을 이용하여 시각장애인이 볼 수 있는 것(또는 볼 수 없는 것)을 분류함.
- 월정액 방식의 Horizon 스마트안경을 착용하고 영상데이터가 전송되면, visual interpreter가 고객의 요청에 따라 길거리의 표지판, 슈퍼마켓 선반에 놓인 제품, 약통의 정보 등을 읽어 줌으로써 장보기, 병원 가기 등을 지원

Vision Enhancement Using Implants

아마존 인공지능 카메라 Deep Lens

- 인공 지능 개발자를 위한 세계 최초 딥러닝 기반 카메라
- 머신러닝 모델을 개발할 수 있는 플랫폼 세이지메이커 (Sagemaker)와 함께 '키네시스(Kinesis)', '그린그래스 (Green Grass)'등 AWS의 기존 플랫폼과 연동해 사용 가능
- 딥 렌즈(\$249)는 마이크로 HDMI(4), USB 포트(2), 인텔 아톰 프로세서, 8GB 램, 우분투, 4Mega픽셀, 1080p 카메라 장착으로 인공지능, 클라우드 컴퓨팅, IoT기술을 효과적으로 연결
- 인공지능을 활용해 동물은 물론 각종 사물과 문자를 인식
- Amazon Rekognition으로 이미지 인식처리로 촬영 후 10분 이내에 데이터 모델 활용 가능
 - '미국 표준 수화 번역 프로젝트'에서는 수화 번역 애플리케이션으로 사용자가 카메라에 대고 수화를 하면 그 의미를 말해주는 프로그램을 제작
 - 물류 창고 재고 품목을 자동으로 스캔하여 딥 렌즈 학습 데이터를 기반으로 품목 수량이나 변경 상태를 파악
 - 딥 렌즈에 음료수 한 병을 보여주면 다른 새 제품을 자동 주문할 수 있고
 - 애완동물을 인식하고 애완동물에게 사료를 공급할 수 있는 장치를 움직이게 할 수도 있으며



AI 플랫폼 연동 서비스 동향



기존 서비스 동향

Amazon Alexa SKILL

- Amazon Alexa Now Has 50,000 Skills Worldwide, works with 20,000 Devices, Used by 3,500 Brands, on September 2, 2018
- Amazon 자체 서비스 외에 음악스트리밍과 홈IOT에서 항공편 예약까지 50,000개 이상의 SKILL 제공



국내 서비스

- SKT 누구 서비스는 서비스 2년 동안 36개 자체 서비스, 102개 제휴 서비스 제공(누구 PLAY)
- 네이버, KT, 카카오 등도 제휴서비스를 위해 자사 플랫폼 연동을 촉진하기 위해 개방형 API정책을 갖추어 가고 있음
- IOT연동과 음악스트리밍이 주종을 이루고 있으며, 라디오팟캐스트, 오늘의 운세, 음식배달 등으로 확대하고 있음

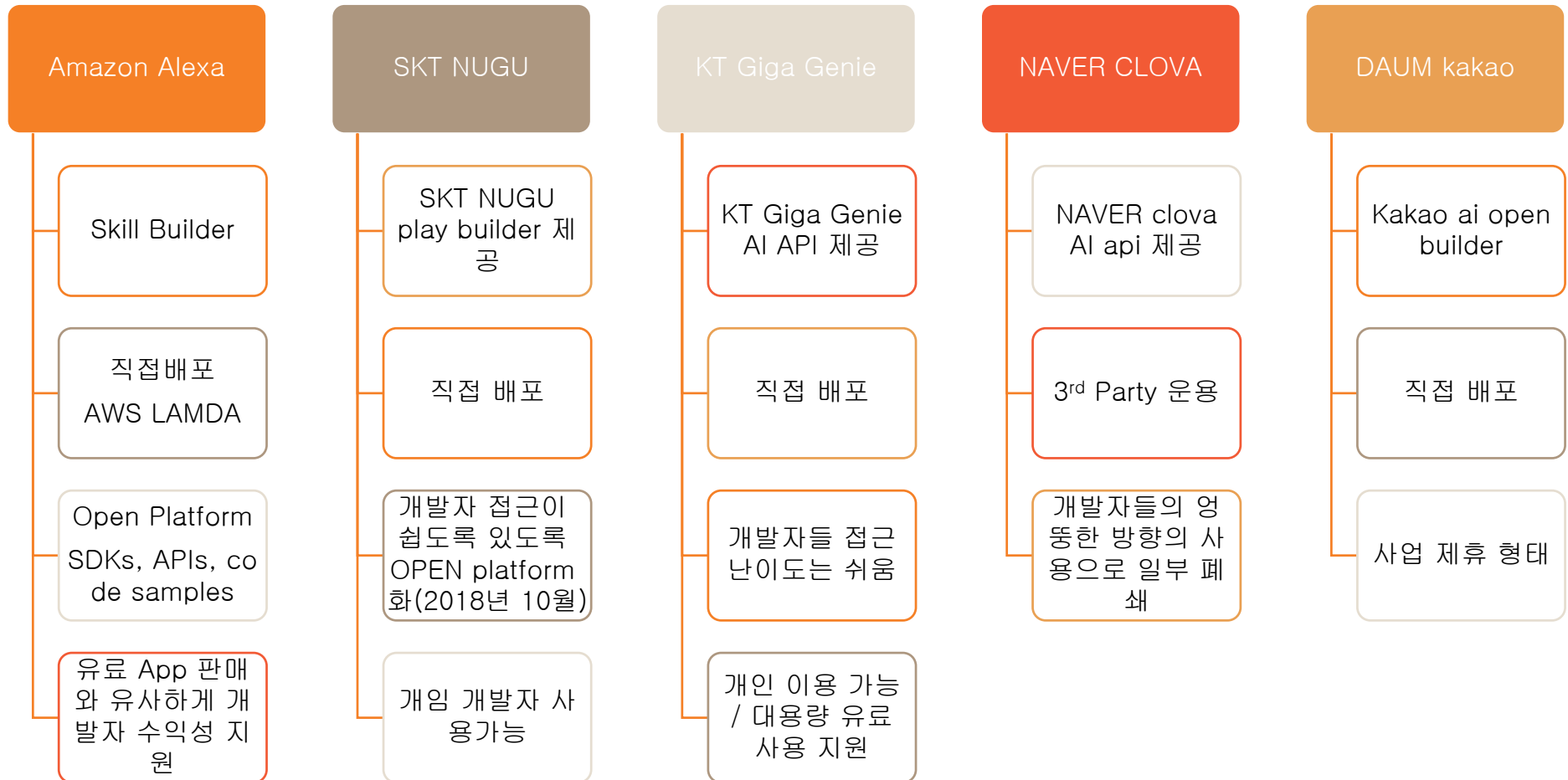
공공서비스

- 각 플랫폼 사의 자체서비스 로드맵에는 공공서비스 부분이 언급되어 있으나 구청업무시간이나 도서관 위치 등 긴요하지 않지만 빠르게 적용하여 구색을 갖추 수 있는 분야를 반영하고 있음
- 공공기관 주도의 AI음성기반 서비스는 아직 구체적으로 발표된 내용 없음

홈 시장에서 맞붙은 국내외 AI스피커

업체	AI스피커명	호출어	특장점
구글	구글홈	오케이 구글	- 구글 AI비서인 '구글 어시스턴트' 서비스 - 영어·한국어로 질문하면 각각의 언어로 답변 - 최대 6명의 목소리 인식, 개인별 다른 답변 제공
		헤이 구글	
네이버·LG유플러스	클로바	클로바 등 7종	- 택배 배송, 음식 배달, 국내외 시황 제공 - 오디오 클립 통화·동요 무료 제공
카카오	카카오톡 미니	헤이 카카오톡	- 음성으로 카톡 메시지 전송·보이스톡 무료 통화 - 택시 호출, 배달 음식 주문 등 서비스 제공
SK텔레콤	누구	아리아 등 4종	- 102개사와 제휴해 400여 종 기기 연동 - 호텔·편의점 등에서도 서비스
KT	기가지니	지니야 등 4종	- 기가지니 소리 통화, 핑크종 영어 따라하기 등 - 호텔·아파트·레지던스 등에서 서비스

AI 플랫폼 연동 서비스 동향



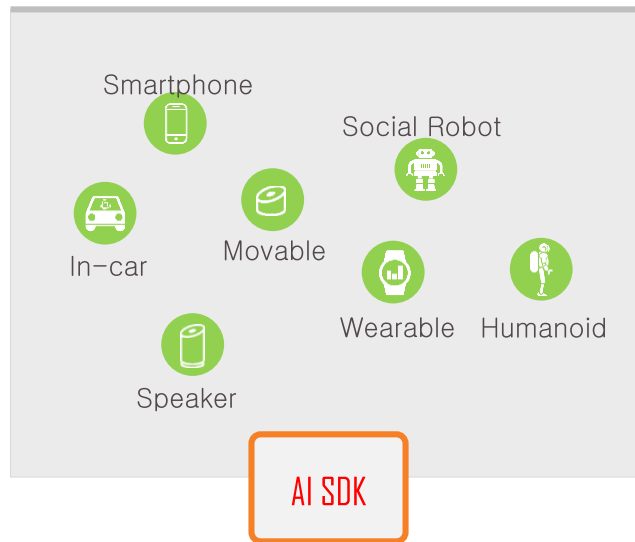
※ 코웨이는 세계 최초로 알렉사와 연동된 공기청정기 에어메가가 큰 인기를 끌면서 아마존 내 코웨이의 2017년 공기청정기 판매량은 전년 대비 3.3배 증가했다. (매경, 2018.3.23)

접근성 관점의 AI 디바이스 검토

공공정보 및 생활필수 정보 음성서비스를 위해서는 공공성격의 AI음성정보 서비스가 필요하겠으며, 사용 환경에 맞는 품팩터의 디바이스가 필요

- 각종 지원정책 정보 (임신출산/영유아/노년/장애인/서민금융/창업지원/귀농지원/양육지원/의료지원 등)
- 공간위치 정보(개방화장실/재난대피소/감염병발생 등)
- 재난 구난 정보, 알림 등

사용자 디바이스



다만 신체적 기능이나 건강에 이상이 있는 계층의 경우, 보다 편리하게 정보접근을 할 수 있는 특화된 기능을 갖춘 디바이스가 필요함

시력/청력이 약하거나, 재난상태 등으로 방송공지가 필요한 계층

- 바코드리더, 이미지 인식 등 인식장치가 있는 디바이스
 - 가전제품 등 제품이용방법 음성검색
 - 식음료 및 의약품 등의 제조유통기한 안내와 설명
 - 의약품의 성분, 기능, 복용법 안내설명 등에
 - 제품이용(사용법, 복용법, 각종 사용안전 매뉴얼 등) 등의 분야
- 제스처인식 지원장치가 있는 디바이스
 - 수어인식과 자막처리 지원
- 긴급방송을 지원하는 오디오/비디오 장치가 있는 장치
 - 대피가 필요한 재난지역 인근에 있는 경우
- 보행내비게이션 등 이동을 지원하는 디바이스
 - 버스정류장에서 탑승할 버스로 안내 등

일반적인 디바이스 개발 프로세스

COMPANION APP DEVELOPMENT

App UI/UX design and developing, Companion App test and certification

SOFTWARE DEVELOPMENT

OS and Application Software, Software OTA update Services, Security

HARDWARE DEVELOPMENT

Review Physical UI, Schematics Design & Layout, Final PCB Board, ID/Mechanical/Acoustic Designs, Review design checklist

PRODUCTION

EVT or Engineering Validation Testing(H/W bring up)
DVT or Design Validation Testing(FCC, BQB, CE, Wi-Fi., etc.)
PVT or Production Validation Testing(Define and Optimize Standard
Operation Procedure, Provide automated test tools)
MVT or Manufacturing Valuation Testing

CERTIFICATION

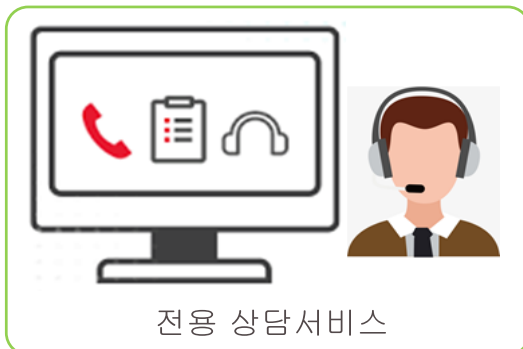
LAUNCHING

On-going OTA updates
Instant in-app feedback
Efficient issue tracking system
Penetration tests

DEVELOPMENT PHASE

서비스 측면 검토

서비스



전용 상담서비스

공공서비스 전반에 대한 시음성지원서비스 우선순위 기반 순차적 접근

- 장애인/노인
- 재난/구난/안전/공공보건의료
- 사회서비스/복지/아동보육 등

정적인 공공데이터 DB를 활용, 동적인 정보로 전환하여 제공

- 재난관리정보/지진대피소정보/공공시설물안전정보
- 민방위대피시설/재해위험지구/지진해일대피소
- 공중화장실/무더위쉼터/교통사고다발지역
- 복지서비스 정보/장애인편의 시설/장애인콜택시/
- 식품의약품안전처 식품, 의약품 정보/건강보험심사평가원 의약품 성분 약효 정보 등

전용 상담 서비스

- 장애인 등 취약계층의 정보지원을 위한 전용상담서비스
- 콜센터 챗봇 이상 단계의 보다 강화된 수준의 AI서비스로 이행되는 과정의 간극을 메우기 위해 필요

유관 협회단체의 역할이 중요

- 시각 장애인 등 특정 계층이 필요로 하는 실질적 서비스를 발굴하는 과정이 필요

서비스 - 디바이스 개발 프로세스



“이 약은 언제 조제했지?”
“어떨 때 먹는 약인가?”
“언제까지 먹을 수 있나?”
“병원에 언제 또 가야 되나?”
??



의약품 AI SKILL



- 도메인 : 약품
- 2nd Party : 처방전

Voice UX 기획

발화문

처방전 언제 조제?

무슨 약이지?

병원에 언제 또 가야되나?

⋮

Intent

inform.prescription.openaday

inform. medicine.efficacy

inform.clinic.schedule

(time:9 p.m, title : hospital)

Yes

No

상위기획

상세기획
(VUX)

기술검토

시나리오 Fix
(User Test)

개발/QA

서비스 출시

- 서비스 확정
- 시나리오 Ideation

- 개발 시나리오
- VUX 기획

- Feasibility검토 (ASR/NLU/TTS 등)
- 서비스로직 검토
- App검토

- 사용자 검증
- 사용자 테스트
- 시나리오 Fix

- 상세개발요건 산출
- AI 엔진/서비스 로직/App 개발
- 상용 QA/Closed Beta Test

효용성과 실용성 그리고 실현가능성이 높은 분야

음성인식 기반 인공지능 서비스는

국내외 다수의 플랫폼이 경쟁적으로 고도화를 하면서 서비스를 하고 있으며, 첨단 기술과 융합하여

매우 자연스럽게 편리한 사용자경험을 창조하고 있음.

한편으로 사용자별 계층과 이용 유형에 맞추어 다양하고도 새로운 유형의 서비스시나리오와 서비스 디바이스가 창출되고 있음.

이렇게 축적되는 기술과 노하우를 노약자와 장애인의 보조서비스로 활용하게 되면 실질적 편의성과 실용성을 높일 수 있으며,

궁극적으로는 UI/UX 혁신으로 일반 비즈니스 서비스 영역에서도 크게 도움이 될 것임.

도래하는 20배 빠른 초고속(20Gbps), 10배 많은 초연결($1/m^2$), 10배 짧은 저지연(1ms)기반의 5G 통신환경을 맞이하면서, 기술 접근성이 높아진 AI/VR/Voice Assistance/Image Processing/Sensing의 영역에서 접근성 관점 솔루션을 찾는 계기가 되기를 희망함.