



DEEPLY

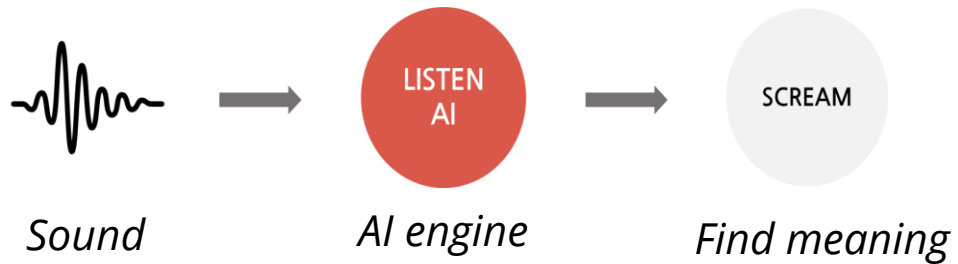
특별한 상황을 인식하는 소리기반 위험감지 AI

리슨. AI

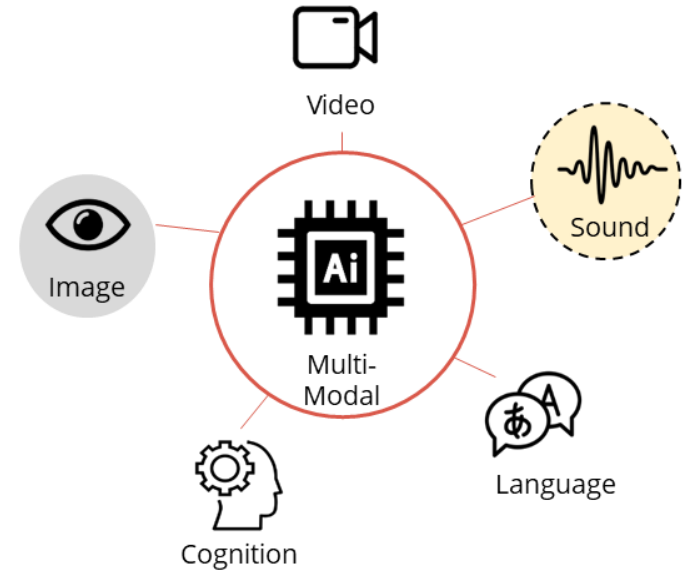
회사소개서
2023.11

“소리에 담긴 의미를 찾는다”

이제 열리고 있는 소리의 세상, 소리기반 인공지능 핵심 컨셉



- 인간의 귀처럼 다양한 음원을 분석
- 특히 훈련 받은 **전문가의 귀를 모방**
- 현재 발생하는 **소리가 어떤 의미인지 분석 및 탐지**



- 기존 영상 기반 시스템의 **부정확성** 및 **한계점** 존재
- **멀티 모달(Multi-Modal)** 인공지능로 나아가는 AGI

AI가 감지하는 무인자율환경

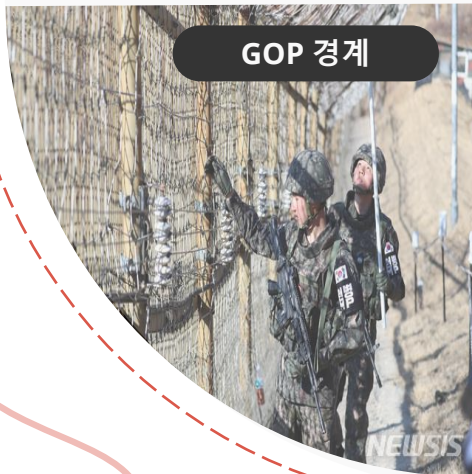
평소와 다른 특별한 상황을 인식하는 AI 엔진, 리슨 AI



대형 참사 사전 예방



산업 재해 예방



GOP 경계



사회 안전망

자율 환경 인식에 대한 속제

- 지능형 CCTV부터 라이더까지 유인보안 시장을 대체하기 위한 다양한 방법 화두
- 평소와 다른 상황에 대한 빠른 인식 요구
- 인공지능의 발전과 인력대체 요구증가로 무인지역의 자동화 속도 가팔라
- 인공지능의 발달과, 인구감소 및 인건비 증가로 인해 무인 산업 시장 규모가 급속하게 확대



99조 USD (25')

지능형CCTV

150만대 (2021년), 매년 10% 증가



1024조 USD (26')

스마트 홈, 시티

2050년 세계 인구 68%가 (65억명) 스마트 도시에 거주

사운드 인공지능 원천기술보유 딥테크 기업

- 2017. 07 (주) 디플리 설립
- 07 중소기업청, 청년창업사관학교
- 2018. 08 Newflight 투자 유치
- 08 벤처기업인증 및 기업부설연구소 설립
- 09 Facebook & 아산나눔재단 아시아 스타트업 참여, 서울대학교 협력 산학연 R&D
- 2019. 02 Google for startups 참여, Google 기술관련 공동 R&D, AI with google 2019
- 05 KOICA CTS, 데이터바우처 우수기업, SW 고성장기업 선정
- 11 LG CNS 비언어적 음성 공동개발 프로젝트
- 2020. 01 모바일앱 WAAH 출시
- 01 대통령 신년간담회 과학기술인 12인 참석, 데이터 바우처 2020 (2년차), SW 고성장기업 선정 (2년차)
- 03 CES 2020
- 09 아마존(AWS) 협력 정글 프로젝트
- 2021. 03 대통령직속 4차산업혁명 위원회 참여, 국무총리 데이터 & AI 기업 시연회
- 04 초보 부모 플랫폼 모바일앱 babba 출시
- 10 LG CNS AI 연구소 화자분리, 감정인식, LG CNS 군수관련 기술개발 공동 사업
- 2022. 01 보건복지부 고려대학교병원 노인건강돌봄 실증사업 (지자체 3000가구 실증)
- 03 KT Partnership, 따뜻한기술더하기, 음성관련 노인 실증 및 기가지니 프로젝트
- 05 KB그룹 국민은행 불완전판매 음성프로젝트 (KB스타터스, IITP)
- 07 육군본부 무인헬기 사운드 감지 프로젝트

CEO



이수지

Co-founder / CEO

2007-2011 KAIST 전기전자 / 경영 학사
2012-2014 서울대 뇌인지과학 석사
2014-2016 (주) ybrain 뇌파제품총괄개발

CTO



박한

Co-founder / CTO

2005-2009 연세대 컴퓨터과학 학사
2009-2017 연세대 컴퓨터과학 석박사수료
20xx-20xx 국가기관 보안 (6년)
2014-2016 큐브스튜디오 창업자/대표

CIO



류명훈

AI Engineer / CIO

2002-2010 고려대학교 뇌인지과학/심리학 학사
2012-2014 고려대학교 BCI 뇌공학 석사
2017-2022 서울대 오디오인공지능 (MARG lab) 박사수료



글로벌 Tech 기업을 고객사로 세계에서 인정받은 사운드 AI 기술력, 글로벌 기업 및 정부 참여

고객사·파트너사

- Google (DeepMind), Meta (Facebook), LG CNS 협업
- AI with Google 2019 연사 초청
- Facebook 개발자 컨퍼런스 F8 초청 (US)
- CES 2020 참가, 4차 산업혁명위원회 위원
- 고령가구 대상 위험감지 모델 실증 (광주 광산구, 3000가구)
- 스마트 도시 위험감지 모델 실증 (강원랜드&하이원리조트, 인천 IFEZ)
- 산업 현장 모니터링 모델 실증 (롯데건설, 한화비전 등)

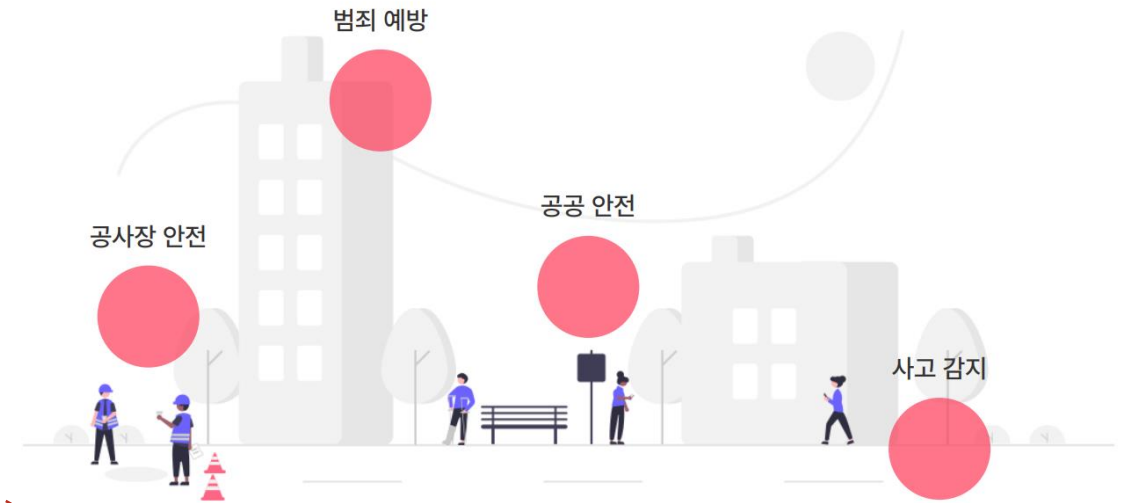


무인지역 보안을 위한 솔루션

리슨. AI Listen.AI

소리 기반 위험 상황 감지 인공지능 솔루션

- 비명, 신음, 혈떡임 등 **위급 상황 관련 사람 소리 감지**
- “살려주세요”, “119” 등 응급 구조 요청 소리 감지
- 차 사고, 폭발음, 알람음, 총 소리 등 **위험 상황 관련 환경음 감지**
- CCTV, 블랙박스 등 엣지 장치 (Edge Device)와 연동 가능



사용자
편의성



낮은 비용
하드웨어



사각지대와
어두운 환경



사생활 보호

비명	신음	혈떡임
구조요청언어	“살려줘”	“119”
총소리	사람 소리	군중 웅성거림
자동차 경적	브레이크	사이렌
발자국	유리 깨지는 소리	폭발음
이륜차	오토바이	트럭
기침	재채기	흠쩍임
하품	딸꾹질	코 푸는 소리
고함	혀차는 소리	박수 소리
울음	웃음	아기 소리
성별 구분	연령 분류	감정 분석
엔진 소리	환경음	+ More

Listen.AI

더 효율적으로, 더 안전하게 고객의 장소를 지킵니다

- Listen City 는 도시의 거리, 빌딩, 사람, 집, 공장 등에서 발생하는 다양한 소리를 실시간으로 감지합니다.
- 이를 통해 사고, 테러, 응급상황을 조기에 감지하고, 관내의 관제센터 및 응급지원센터와 연계하여 빠르게 대응 가능한 시스템을 구축합니다

소동 및 소란 감지 솔루션

클래스	소동	소동
적용 환경	실외, 건물 내, 주차장 등	
적용 분야	도시 방범/관제, 기업 및 사유지 보안	

위험 상황 감지 솔루션

클래스	비명	폭발음	총소리
적용 환경	실외, 골목 등		
적용 분야	도시 방범/관제, 기업 및 사유지 보안		

건강 이상 감지 솔루션

클래스	기침	재채기	혈떡임	신음
적용 환경	실내			
적용 분야	노인/1인 가구, 병원, 요양시설			

교통 사고 감지 솔루션

클래스	스키드마크	경적	알람	폭발음
적용 환경	도로, 교차로, 골목 등			
적용 분야	교통 안전, 주차장 관리, 어린이 구역			

공공안전, 산업안전, 군수분야까지 AGI - 멀티모달에 대한 관심 보안시장 진입 1년 내 10개 고객사에서 도입

경계 감시 분야



2023 - 현재

소리기반 위험감지 - 강원랜드 & 하이원 리조트

사건, 사고를 의미하는 다양한 소리 신호 탐지 (욕설/싸움 등 14종)
관제센터와 연계하여 사건 사고 조기대응
생명/재산 피해 규모 최소화 및 방문 고객 만족도 증가

계약



2023 - 현재

소리기반 위험 감지- 롯데건설 / 대림건설

건설 현장 내 안전 사고를 다양한 소리 신호로 탐지 (비명/ 사고 등 14종)
건설 부지 내에서 발생하는 규제 범위를 벗어나는 소음 감지
안전 사고 방지/조기대응 및 소음 관련 민원 최소화

계약

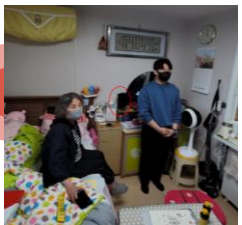


2023 - 현재

소리기반 위험감지- 인천TP (지하철, 버스정류장)

사건, 사고를 의미하는 다양한 소리 신호 탐지 (비명/ 사고 등 14종)
스마트 함체, 스마트 로봇 등 다양한 IoT 디바이스와 연계

계약



2021 - 현재

독거 노인 안심 솔루션 - 광주 광역시 광산구

+3,000 고령 인구 (호흡기/낙상 등 6종)
기침, 재채기, 호흡 등 안전관련 및 비명, 낙상, 부딪힘 통한 응급 상황 파악
모바일 앱 탑재, 안드로이드가 포함된 AI 스피커 확장

계약

군수 분야



2023 - 현재

소리기반 위험감지 - 한화테크윈((구), 한화비전) 외 5개사

이상 음원 감지 모듈을 카메라를 포함한 하드웨어에 연동
영상 기반 모니터링 시스템 연동 (VMS)

계약



2022 - 현재

대한민국 육군 - 포드림, 소리감지 무인헬기 탐지

무인헬기 (대형 드론) 소리감지 이상 탐지 기술개발 진행중
2인 탑승 가능한 대형 드론, 원격에서 이상상황 모니터링

계약



2021 - 현재

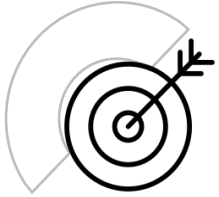
대한민국 해군 - LG CNS 기술개발파트너사 등록, 박람회등

LG CNS 기술파트너사 해군 함정 진입중
해군 관련 Closed 박람회 4회 참여 및 사업 수주 진행중
인공지능 세미나 등 진행

준비

세계 최대 규모 비언어 사운드 데이터셋

리슨 AI 엔진의 강점



세계 최대규모 비언어 데이터셋, 높은 정확도

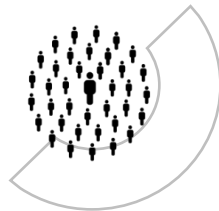
현존하는 가장 큰 비언어 데이터셋 (5만 시간 이상, 42종)을 기반으로 다양한 소리가 섞이는 현장에서도 안정적으로 작동



언어의 제한을 받지 않아 글로벌 확장 가능

음성과 음향은 언어의 제한을 받지 않음.

비명, 자동차 사고, 유리 깨짐 등의 전세계에서 공통적으로 발생하는 다양한 소리를 분석



자체 개발된 사운드 AI 엔진, 20만명 이상 사용

단순한 오픈소스 모델이 아닌 자체 6년 이상 연구 개발된 AI 모델을 사용

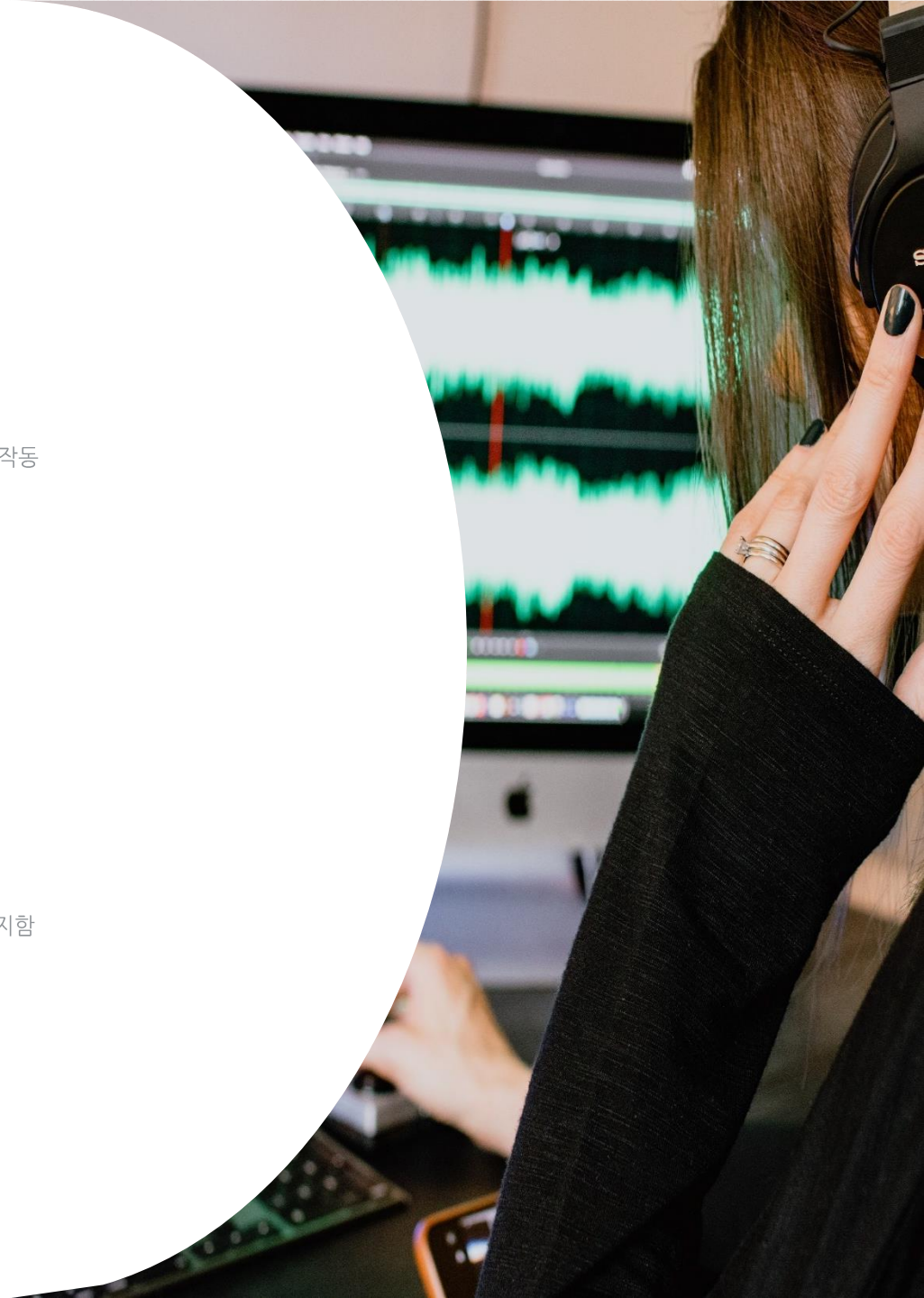
전 세계 20만명 이상의 유저로부터 검증된 Sound AI 엔진으로 어떤 기기에서든 실시간으로 원하는 소리를 즉시 감지함



빠르고 간편한 연동과 디바이스 대응

SDK, Server SDK, API, HW Chip과 module을 통해 다양한 제품에 연동 가능

직접 보유한 디바이스에서 실증, 현장 배포하여 정확도를 개선함



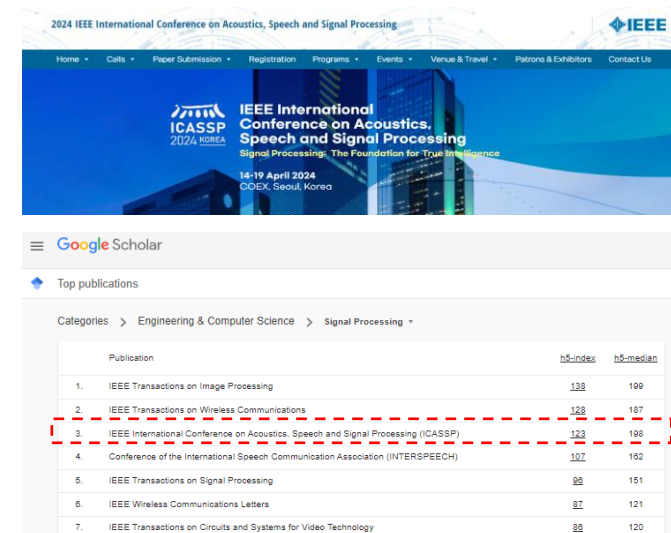
검증된 사운드 인공지능 개발 기술력

특허, 지적재산권, 검증시험, 논문

- 음성의 특징을 실시간으로 구분하여 상태를 나타내는 자사 사운드 AI 핵심 기술 특허 보유
- 국내 및 해외 특허 등록 3개, 출원 7개, PCT 및 해외특허, SW 인증시험 등 보유
- 세계 최정상급 신호처리 학회 ICASSP 2024 통한 기술력 검증



연구개발기관 (소유권자)	지식재산권명	국가명	출원·등록번호 /출원·등록일
(주)디플리	실시간 소리 분석 방법 및 시스템	대한민국	한국/10-2018-0075331 제10-2238307호
(주)디플리	실시간 소리 분석 방법 및 장치	대한민국	한국/10-2018-0075332 제10-2155380호
(주)디플리	자동 레이블링이 가능한 머신 러닝 방법 및 장치	대한민국	한국/10-2018-0075333 제10-2189362호
(주)디플리	데이터압축 및 암호화방법 그리고 장치	대한민국	한국/10-2018-0140185
(주)디플리	이상탐지방법	대한민국	한국/10-2018-0140186
(주)디플리	딥러닝 모델의 분리 수행을 위한 방법, 장치 및 시스템	대한민국	한국/10-2019-0129241
(주)디플리	대용량 오디오데이터의 효율적인 딥러닝 학습 방법 및 시스템	대한민국	한국/10-2019-0129242
(주)디플리	오디오 데이터의 공간 독립적인 분석을 위한 장치, 방법 및 컴퓨터 프로그램	대한민국	한국/10-2020-0065199
(주)디플리	자동 레이블링이 가능한 머신 러닝 방법 및 장치	PCT	OP19-0073PCT
(주)디플리	Method and device for analyzing real-time sound	미국	미국/36DEEP-BA10002PA
(주)디플리	Baby sound analysis v1.0	대한민국	한국/BT-B-19-0167
(주)디플리	Audio event analysis v1.0	대한민국	한국/BT-B-21-0080



"Microphone conversion: mitigating device variability in sound event classification"
: 다양한 디바이스(마이크)에서 수신되는 음향의 차이에 대해 소리 분류 성능을 높임

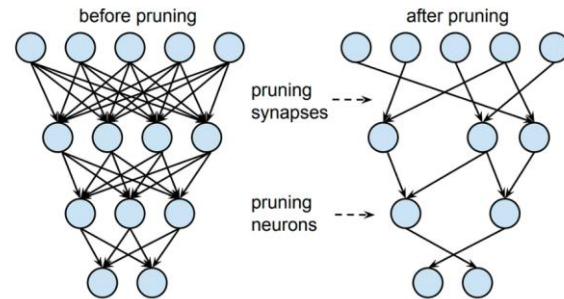
** 세계 최정상급 신호처리 (오디오부문) 학회/컨퍼런스
Interspeech, ICASSP, WASPAA, DCASE

유연한 연동을 통한 빠른 AI 모델 최적화 가능

- 인공지능의 대중화, 그러나 성능이 좋은 제품은 무거운 경우가 많음
- **모델 경량화**를 통한 CPU에서 작동하는 **엣지 모델** / 무거운 **서버 모델** 동시 보유
- 보장된 AI 모델 성능을 위한 마이크 최적화 프로세스 보유

모델 경량화 및 추론 가속화 적용

- Knowledge Distillation
 - Pruning
 - Model Quantization
 - Weight Quantization
 - Model Compression
-
- GPU, ASIC
 - TensorRT Library
 - Batch Inference



실제 연동 완료한 엣지 디바이스 및 서버 목록

- 삼성 갤럭시 S 시리즈 이상
- 삼성 갤럭시탭 S시리즈
- 한화비전 임베디드 카메라
- 고성능 CPU 카메라 / AI 카메라
- Raspberry pi, Orange Pi
- CPU, GPU 모두 지원 (저사양 가능)



API / SDK를 통한 편리한 연동 지원

API

REST API
RTSP / gRPC
WebSocket / ONVIF

SDK

Android - Device/Cloud SDK
Linux/Windows - C++ SDK
Cloud Server - Container (Docker)
Edge Server - Container (Docker)

마이크 최적 성능 보장

- 일반 스마트폰, 스마트태블릿 탑재 마이크
- 내장 MEMS MIC / 외장 USB MIC (ex. SAMSON)
- PoE MIC SPON



고객사 및 제품 다양화 전략 - 다양한 산업군에 소리 AI 엔진 제공

고객사
확보

유인 경계/감시 시장의 대체

소리기반의 안전 모니터링 제공 (공공보안, 물/리조트, 건설현장 외)
* 유인보안 인건비 절감 연간 보안 예산 중 인건비 예산 60%이상
공간사업자, 대·중견기업 보안솔루션 적용 (2023' 6개사업장 확보)
글로벌 고객사 확대 (일본 TB그룹외 협의중)

~
2024

제품
다양성
확보

리슨AI - 소리기반 응급상황 감지 AI 엔진 솔루션

응급상황별 - 방법, 사고, 시니어, 기업보안, 교통용
MLOps 기반의 AI 엔진 및 솔루션

~
2025

적용
산업
확장

군산업 본격 적용




경계감시 지역 GOP 보안 적용
방위청 사업 수주, 총소리 감지 등 군수물품 상용화

~
2027



인력을 대체하는 보안, 자율주행, 드론 상황까지

적용 가능 산업군 확장 및 향후 산업 추진 계획

	 지능형 CCTV	 스마트 홈/시티	 스마트 공장	 커넥티드 로봇	 미래형 드론	 자율 주행차
국내 (원) 시장규모*	2조 (22')	26조 (23')	16억 (22')	17조 (25')	4조 (26')	26조(35')
글로벌 (원) 시장규모**	99조 (25')	1040조 (26')	554조 (25')	203조 (25')	90조 (26')	1243조 (35')
Sound AI	<ul style="list-style-type: none"> - 비명, 신음 등 응급 상황 감지 - 구조 요청 언어 감지 - 총 소리, 폭발음 등 환경음 감지 	<ul style="list-style-type: none"> - 위험 시설 모니터링 - 방범 강화 및 범죄 예방 - 자연 재해 조기 감지 	<ul style="list-style-type: none"> - 엔진 오작동, 과부하 등 정상 상태와 다른 엔진 소리 감지 - 고속 회전체 또는 스피들의 이상 작동 감지 - 팬 또는 냉각수 이상 작동 감지 		<ul style="list-style-type: none"> - 운전자 건강 상태 감시 - 엔진 이상 및 오작동 감지 - 음성 감지를 통한 UI/UX 개선 	
고객사 분류	영상관리시스템(VMS) /CCTV 업체	지자체, 건설업체, 보안업체	제조/설비업	제조업, 서비스업	군/군수업체	Auto사 플랫폼 기업, 자동차 제조사, 전장업체
고객사 파트너사	C사*, 에스원, 세콤	IBM, Cisco, 광주시	Subaru, Panasonic	KT, 구글	해군, 육군, 포드림	현대자동차, Vector



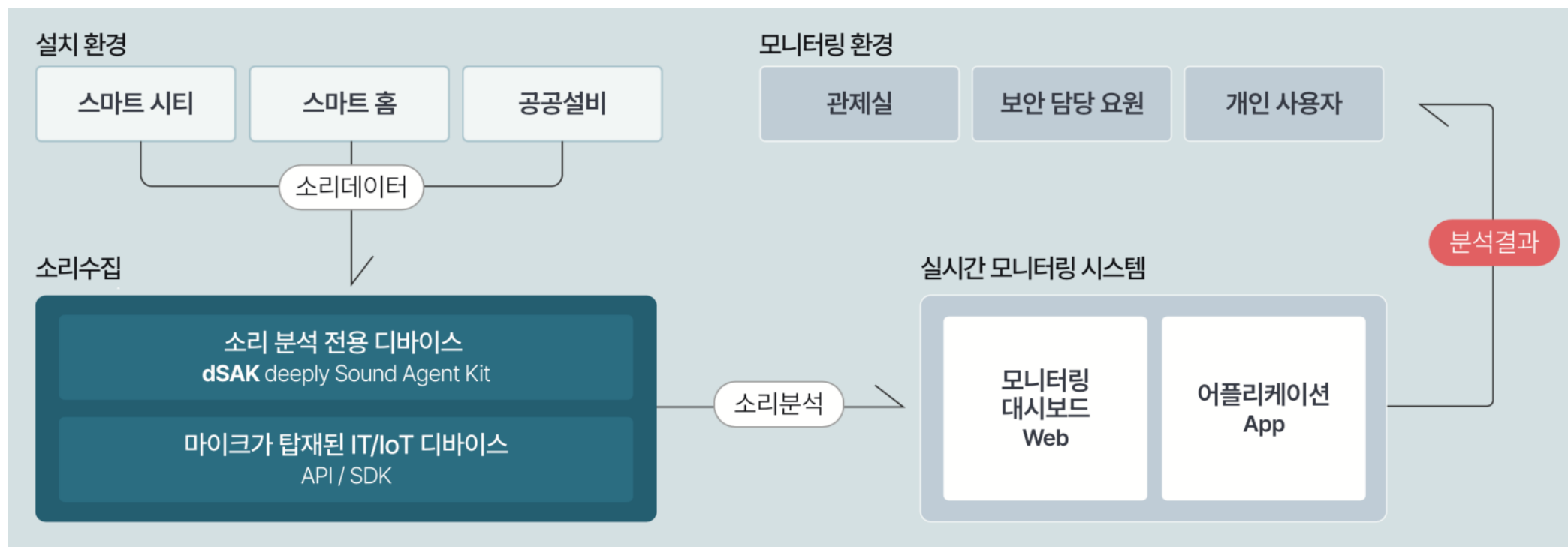
THANK YOU

Sound AI Emergency Solution



APPENDIX

인공 지능 소리 분석 시스템



타사 대비 강점

하드웨어 디바이스 현장 설치, HW+SW 통합 플랫폼 운영 (즉시 사용 가능)

- 개발 보드 및 하드웨어 마이크 특징을 반영하여 제품을 구성.
- 강원랜드, 인천IFEZ, 지하철/ 버스정류장을 비롯하여 직접 운영하는 디바이스 및 SW 플랫폼
- 최종 양산형 제품으로 테스트 및 제공 가능



dSAK v1.0



dSAK v2.0



편리한 사용자 이용 - 현장에 최적화된 높은 이상 음원 감지 정확도, VMS, 앱, 문자연동을 통한 편리한 알람

- 최고 성능의 사운드 인공지능 (딥러닝) 모델을 통한 높은 정확도, 특히 설치 후 2주간의 최적화 과정을 통한 높은 이상음원감지 성능 (MLOps)
- 기존 비디오 모니터링 관제시스템과의 연동 가능 (VMS)
- 이외 앱/ 문자/ 전화 등 다양한 방식의 알람 지원

무인 시장에서 보안카메라 기술의 변화

영상
녹화

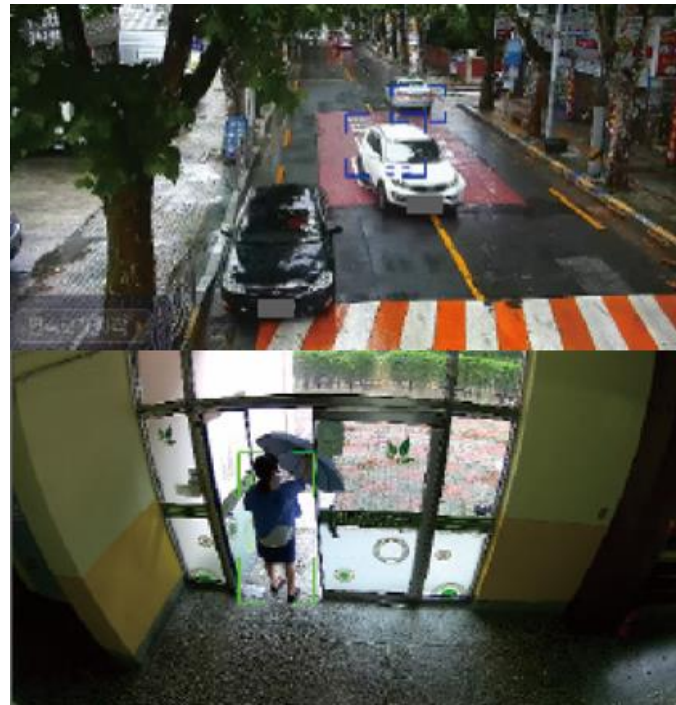
단순 영상 녹화
사후 확인 및 증거 확보

영상 AI
분석

영상 기반 분석 시스템
시야각/정확도 문제 존재
어두운 환경 조명 필요

영상 +
음성 AI
분석

영상과 음성이 상호 보완적으로 작동
시야각 및 정확도 문제 해결
사운드 분석 기반 새로운 시장 존재



42 Types of Sound

Scream

Moan

Pant

Human Voice

Crowd

Footstep

Explosion

Glass Brake

Rescue Language

Car Horn

Car Brake

Gunshot

Alarm

Siren

Whistle

Vehicle

Motorcycle

Truck

Cough

Sneeze

Nose Blowing

Yawn

Throat Clearing

Lip-Smacking

Sigh

Tongue Click

Clapping

Crying

Laughing

Yelling

Gender Classification

Age Classification

Emotion Analysis

Baby Sound

Home Sound

Noise

Speaker Identification

Speaker Verification

+ More



도시 내 위험 상황 관련 소리



건강 상태 관련 소리

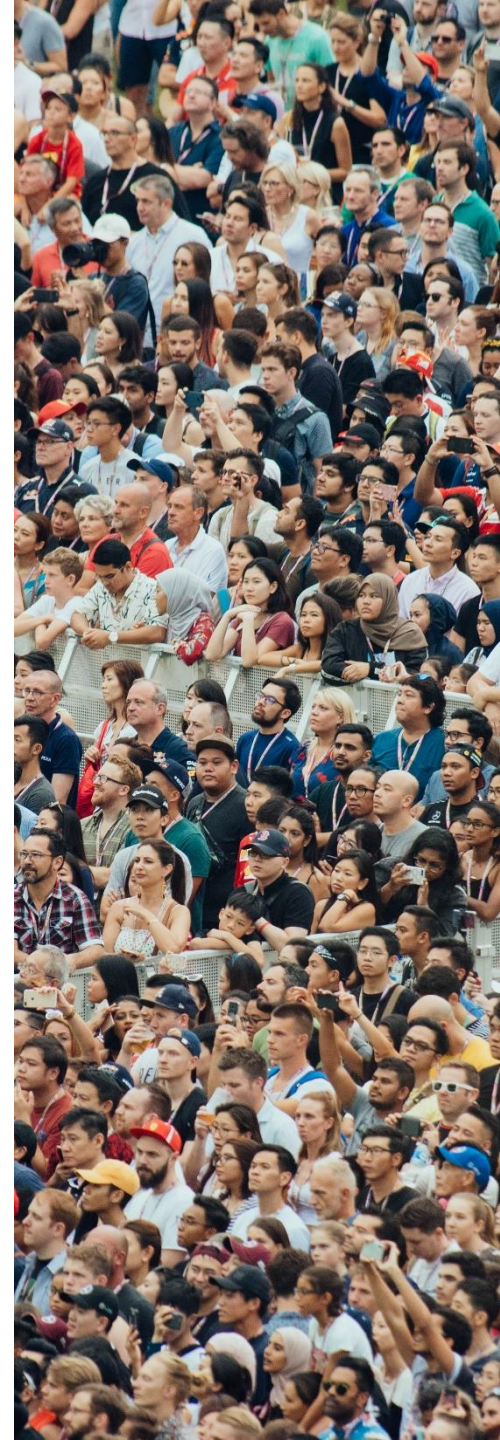


기타 화자 관련 기능

소음제거기능 Noise Reduction

- 화자의 대화만 남기고 소음을 제거
- 비/ 바람이 불거나, 지지직거리는 소리, 주변 음악소리등을 전부 제거
- 외부환경 10m 밖에서 발생하는 소음 제거
- 정확도 높음
- 라이브 데모 가능
- 샘플 파일 참고

	배경음	배경음2	고객사
원본 파일			
노이즈 처리 파일			





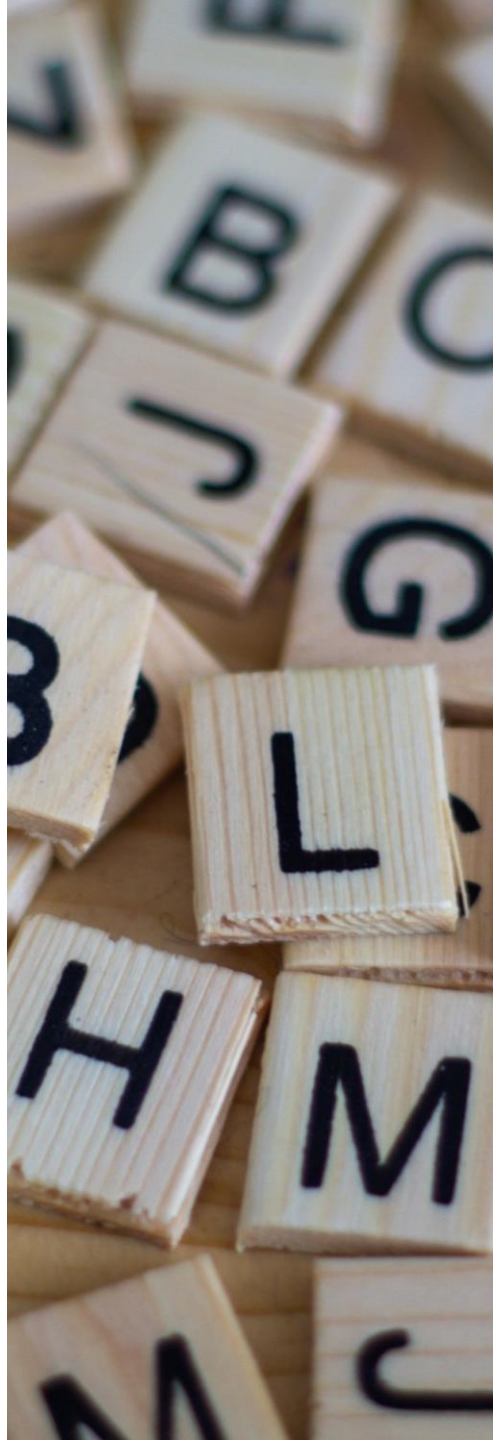
사람 목소리 구분 Voice activity detection

- 사람의 목소리 (Speech)인지 그 외에 다른 소리인지 구분
- 실시간/ 배치 분석 가능
- 예시
 - 녹음된 파일에 사람들의 대화가 포함되었는지 확인하기 위해 사용
 - 법률/규제 문제로 녹음이 불가능한 경우,
사람 목소리를 인식하여 저장하지 않고 기타 소리만 녹음하는 용도로 사용
(법률에 의해 사람들의 목소리를 녹음할 수 없는 경우)
- 사람들이 목소리를 내는 상황만 감지하여 제품이 작동하도록 동작 가능
- 라이브 데모 가능



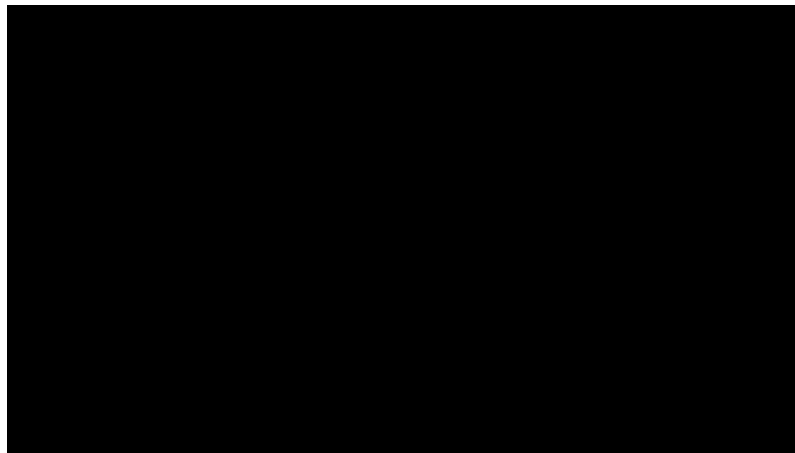
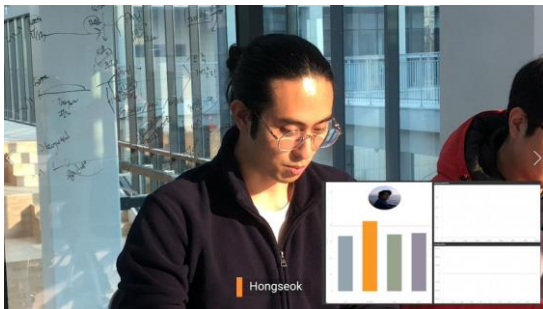


특정단어인식 Keyword Spotting

- 특정 단어에 대한 인식
 - 5개 단위로 단어에 대해 제작
 - 기보유 모델 - 디플리, 해지, 취소 등
 - 라이브 데모 가능
 - 작동하는 상황에 따른 (집안, 엘리베이터, 차안, 전화등) 최적화 및 기능향상
 - 예) 취소. 해지 알겠습니다. 인지했습니다. 등
- 

본인확인기능 Speaker Identification

- 현재 말하고 있는 화자가 본인인지 아닌지 실시간으로 확인
- 같은 사람 여부를 지속적으로 확인
- 초기 화자에 대한 데이터 필요
- 초기 데이터를 이용하여 재학습 시킨 모델을 배포하여 구성
- 관리자모듈 연동 별도
- AI 면접 / 인증 / 전화등 본인여부를 지속적으로 확인하기 위한 목적
- 하단 데모 동영상 참고



```
px;  
0px;  
: 10px;  
10px;  
: 10px;  
: 10px;  
t: 10px;  
: bold;  
auto;  
  
" id="login_form" >  
ler='red'><b>Authenticat  
ror!' class="dError1">P  
  
ml-auth-status>-1</sam  
click="window.top.local
```


감정분석

Emotion Classification

- 화자의 감정에 대해 분석
- 긍정/중립/부정의 3단계로 분류
- 짧고 명확한 음성일수록 추론이 용이함
- 분명한 음성에 대한 판별 추천
- 데이터 추가 학습 및 엔진 업데이트 필요





박수소리등 트리거 감지 Indicator detection

- 7종의 Indicator를 감지하여 시스템에 연동하여 사용

Clapping	Double clapping	Alarm sound	Doorbell
Whistle	Finger snap	Beep	

- 예시사용
박수를 쳐서 불을 끄거나, presentation 장표를 넘김
박수두번으로 세탁기를 작동
휘파람으로 에어컨을 작동
- 집안 환경과 연동하여 스마트 홈 제어용으로 사용 가능
- 라이브데모 가능

