



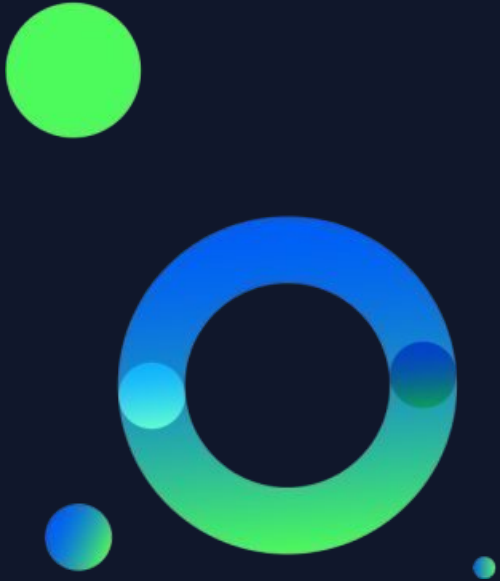
마인즈랩 기업설명회



2022. 12. 27



Make your own
AI service with maum.ai



인공지능 플랫폼의 상용화에 성공하다
마인즈랩의 AI Platform maum.ai

초개인화 맞춤형 인공지능 서비스 개발 도구
마음 오케스트라 (maum Orchestra™)

실제로 인간의 노동을 대체하는
마인즈랩의 AI Human 기술과 제품

ChatGPT의 기반기술 초거대AI 공급업체
초거대AI 언어모델과 챗봇을 활용한 서비스

과금형으로 인공지능 서비스를 제공한다
maum.ai 클라우드 AI API

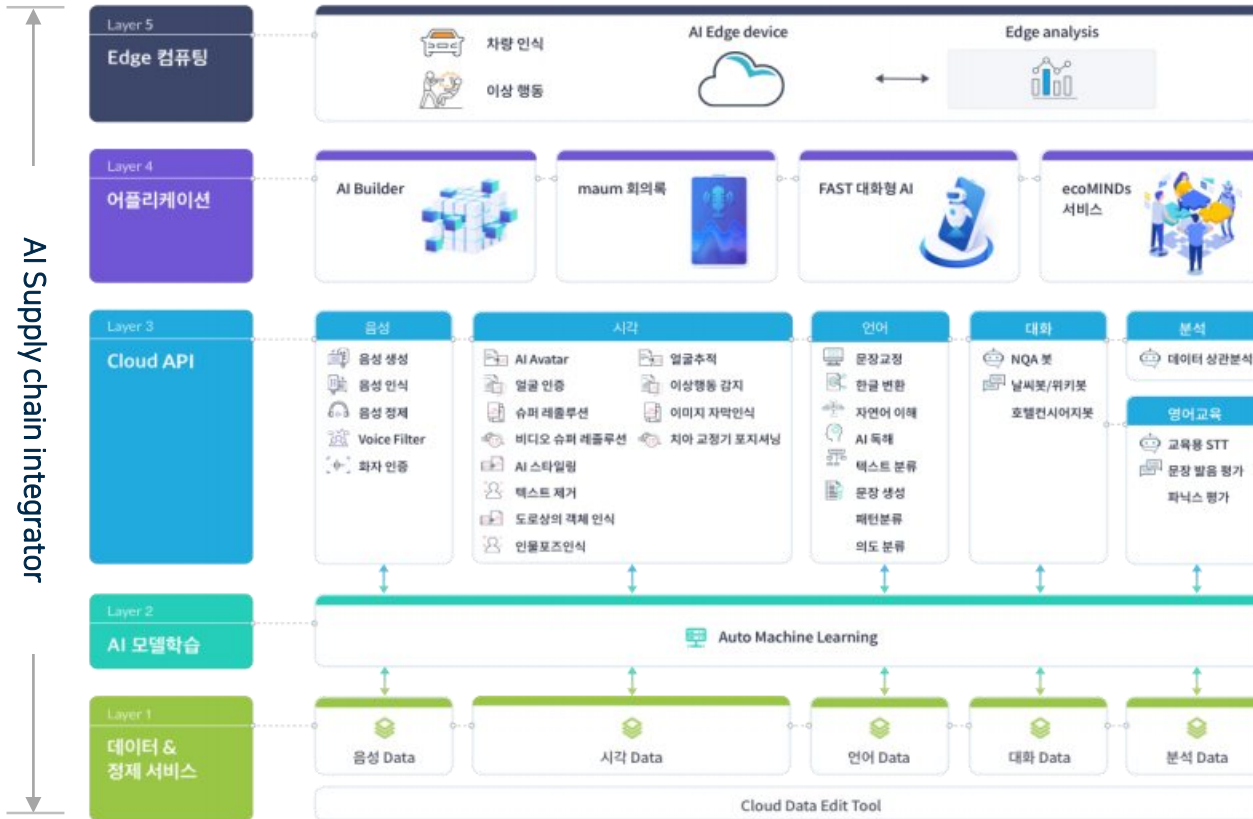
AI Transformation을 선도한다
Data, 전략적 파트너십, 컨설팅, 정부과제

마인즈랩은 종합 딥러닝 기술회사에서 휴먼AI 전문 기업으로, AI플랫폼을 선도하는 회사로 나아가고 있습니다. 기존에 개발한 인공지능 기술들과 최근에 완성한 플랫폼을 기반으로 다양한 수익을 창출하고 있습니다.

마인즈랩 일반



- 국내 유일 플랫폼 단위 AI 인프라
- AI 휴먼 시장 점유율 1위
- End-to-End AI 플랫폼
- AI 단일 매출 1위



주요 연혁

| | |
|------|---|
| 2022 | <ul style="list-style-type: none"> ■ AI Bulider 마음 오케스트라™ 런칭 ■ 대통령직 인수위원회 과학기술 부문 자문사 선정 ■ CES 2022 AI 은행원 공동 출품 (신한은행, 효성, 마인즈랩) ■ AI 휴먼 기술 논문 CVPR¹⁾ 2022 채택 ■ 음성 변환 모델 연구 논문 NeurIPS²⁾ 채택 |
| 2021 | <ul style="list-style-type: none"> ■ 11월 23일 기술특례상장 ■ 6월 S은행 디지털혁신점포 AI휴먼 구축 수주 ■ 3월 AI 휴먼 M1 (1세대) 최초 공개 (5월 런칭) ■ 한국지능정보사회진흥원, 국내 10대 AI 기업 선정 |
| 2020 | <ul style="list-style-type: none"> ■ 2020 대한민국 AI 소프트웨어 부문 대상 수상 ■ 대한민국 인터넷대상 인터넷 기술혁신 부문 과기정통부 장관상 ■ 4차 산업혁명 대상, 정보통신기획평가원장상 수상 |
| 2019 | <ul style="list-style-type: none"> ■ 2019 대한민국 컴퓨팅 부문 인공지능 대상 수상 ■ 세계최초 음성분리필터 (Voice Filter) 기술 구현 ■ 대한민국 SW 기업경쟁력 대상 과학기술정보통신부 장관상 ■ maum.ai AlaaS 구독형 인공지능 서비스 정식 출시 |
| 2018 | <ul style="list-style-type: none"> ■ 하이브리드 인공지능 고객센터 서비스 HAICC 출시 ■ 대한민국창업대상 산업통상자원부 장관상 수상 ■ 대한민국 임팩테크 대상 국무총리상 수상 |
| 2017 | <ul style="list-style-type: none"> ■ 대한민국 ICT 대상 지능정보부문 대상 수상 ■ 포브스, 주목할 만한 10대 스타트업 AI부문 선정 ■ 2016 기술 사업화대상 미래창조과학부 장관상 수상 |
| 2016 | <ul style="list-style-type: none"> ■ 딥러닝 기반 인공지능 제품 출시(실시간 음성 상담) |
| 2015 | <ul style="list-style-type: none"> ■ 통합 콜센터 Big Data 솔루션 출시(Minds VOC / RS / QC) |

2022.9.29에 maum.ai 시스템 V1.0 개발을 완성하고, 12.1에 오케스트라를 오픈했습니다.
그 이후로, 클라우드 기반 인공지능 시스템 전체를 제공하는 사업을 시작하였습니다.

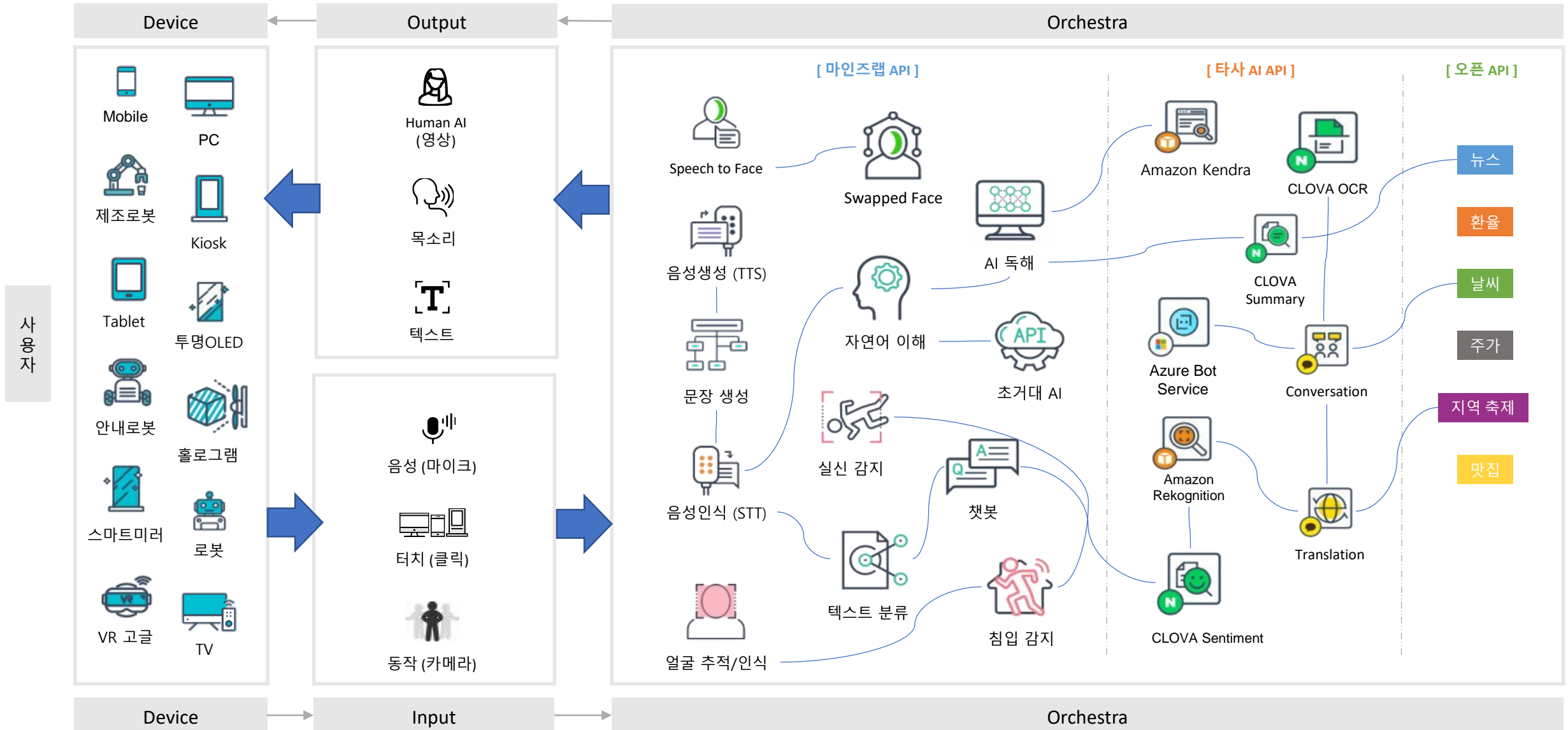


23 modules

Support Cloud & On-Premise

다양한 인공지능엔진들을 자유자재로 조합하여 서비스를 만들 수 있는 AI플랫폼을 세계 최초로 오픈했습니다.

maum.ai the AI service platform for the unmanned business



maum ai 플랫폼 현황

(2022.12.20) 현황

maum.ai world

다양한 AI 서비스들을
카테고리별로 모아 놓은 곳

카테고리

6개

서비스그룹

23개

AI Service

Orchestra Workflow 기반으로 만든
다양한 AI 서비스들

144개

AI
상담사

AI
속기사

IR
담당 AI

AI
세일즈맨

maum Orchestra Workflow

Cloud API의 다양한 AI와 서비스들을 연결해서 다양한 AI서비스를 만드는 틀

362개

전체 430 / 가용 362 / 미가용 68

Cloud API

다양한 AI와 서비스 API 등록/관리

54개

AI API
(MINDsLab)

엔진(Engine)

5개

모델(Model)

14개

AI API
(파트너사)

37개

AI API
(타사)

3개

▶ 추가 연동 가능한 모델(Model) : 50개

maum.ai 플랫폼



AI Service Users

채널별

유형별



AI Engine Providers

마인즈랩 AI 엔진

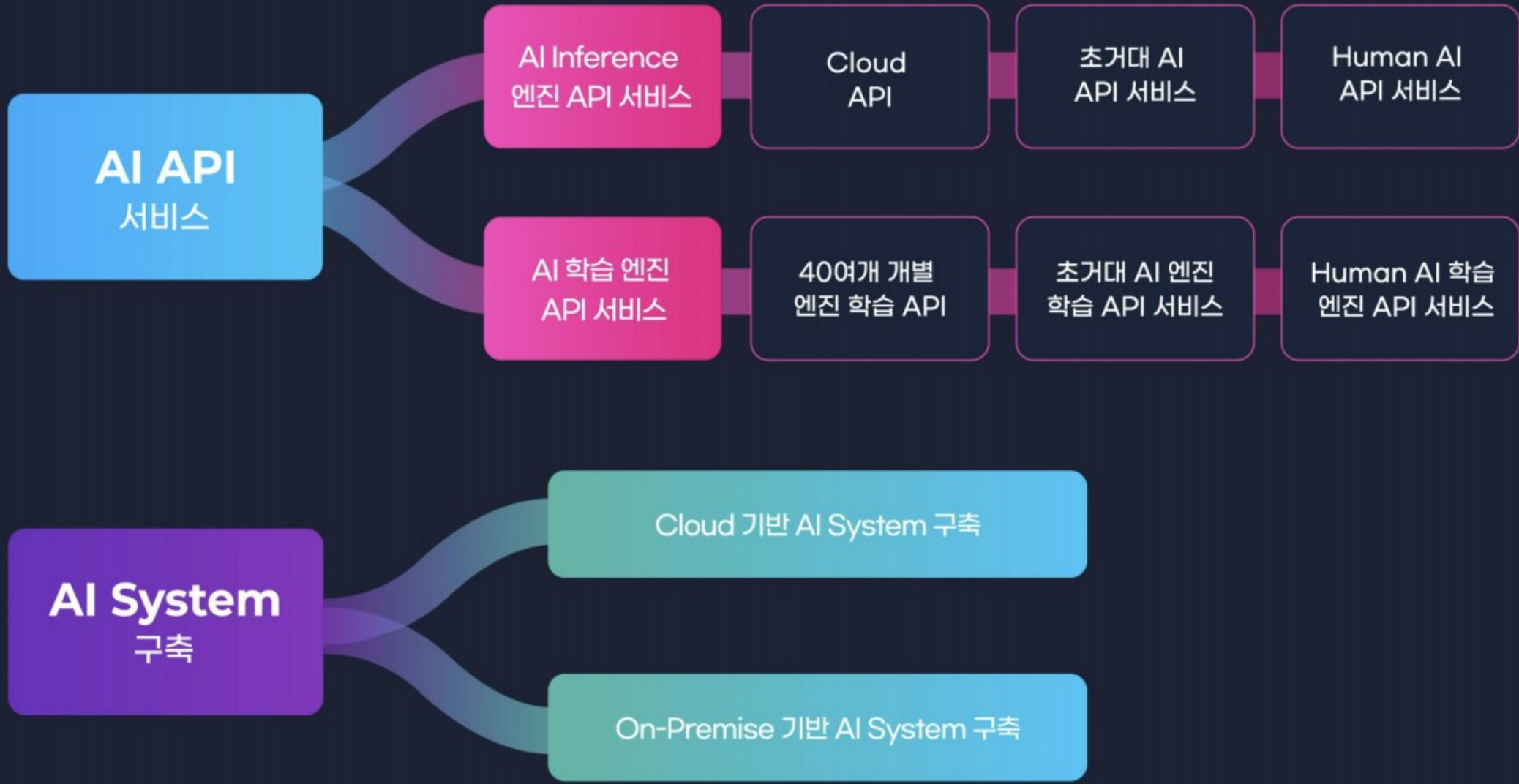
파트너사 AI 엔진

94개



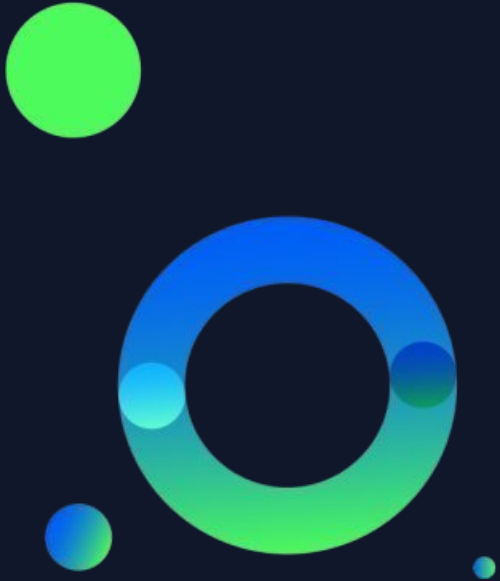
AI Service Providers

maum.ai 플랫폼의 완성에 따라서, AI시스템 전체를 고객사에 구축해주는 수익모델과 다양한 AI API를 제공하는 수익모델 등 크게 두가지 수익모델로 명확하게 방향을 설정했습니다.





Make your own AI service with maum.ai



인공지능 플랫폼의 상용화에 성공하다
마인즈랩의 AI Platform maum.ai

초개인화 맞춤형 인공지능 서비스 개발 도구
마음 오케스트라 (maum Orchestra™)

실제로 인간의 노동을 대체하는
마인즈랩의 AI Human 기술과 제품

ChatGPT의 기반기술 초거대AI 공급업체
초거대AI 언어모델과 챗봇을 활용한 서비스

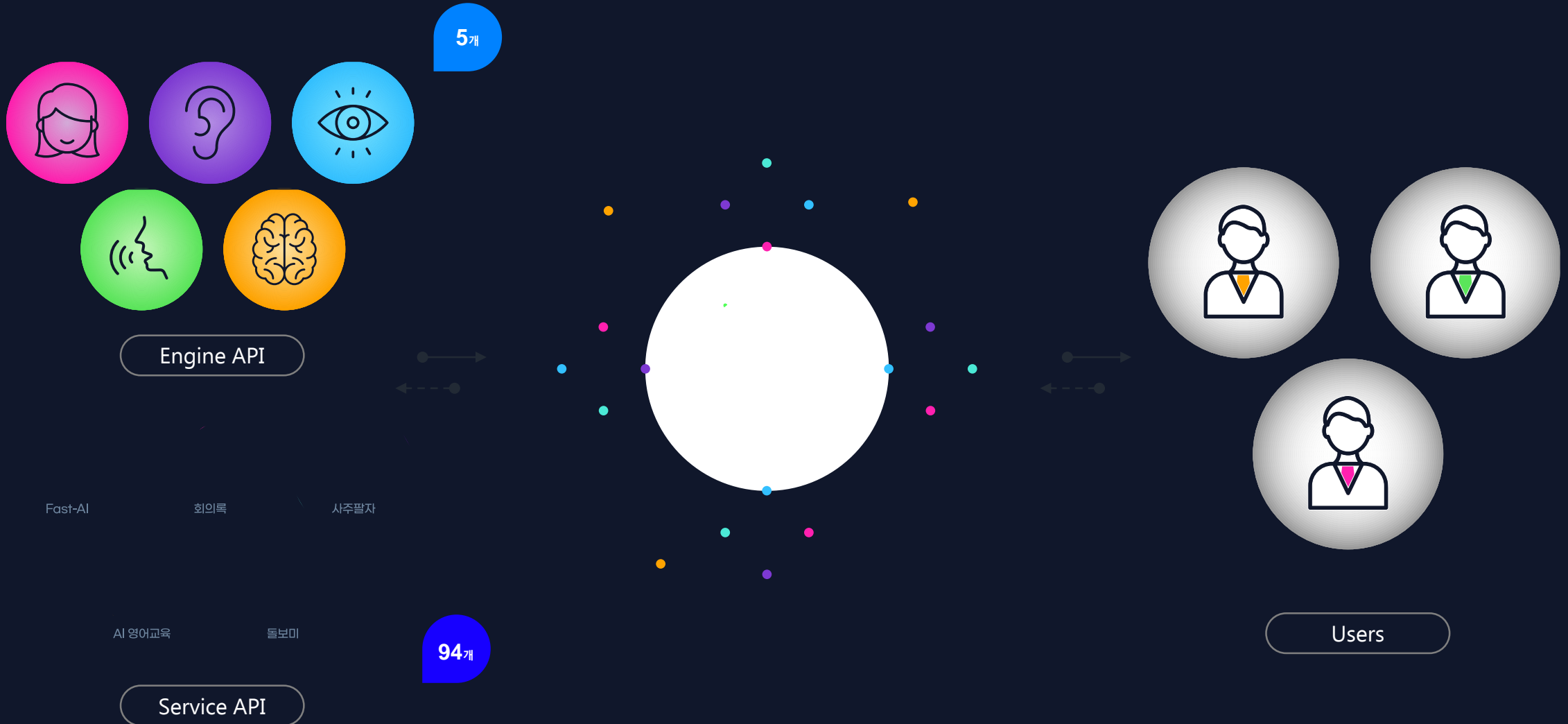
과금형으로 인공지능 서비스를 제공한다
maum.ai 클라우드 AI API

AI Transformation을 선도한다
Data, 전략적 파트너십, 컨설팅, 정부과제

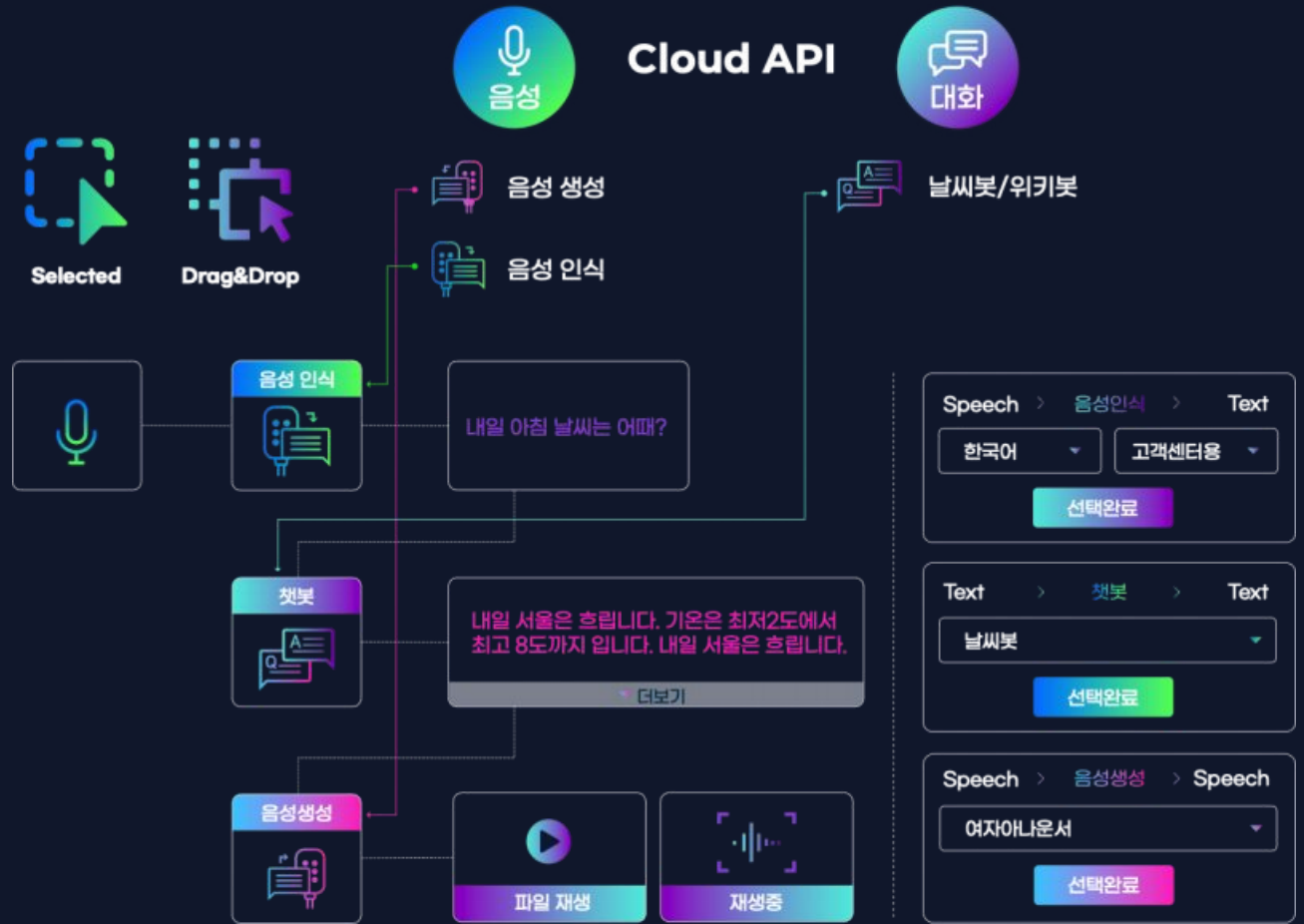
초개인화 맞춤형 인공지능 서비스 개발 도구
maum Orchestra



AI 엔진과 서비스를 연결해주는 maum Orchestra를 12/1 상용 오픈했습니다.



maum Orchestra AI Orchestration



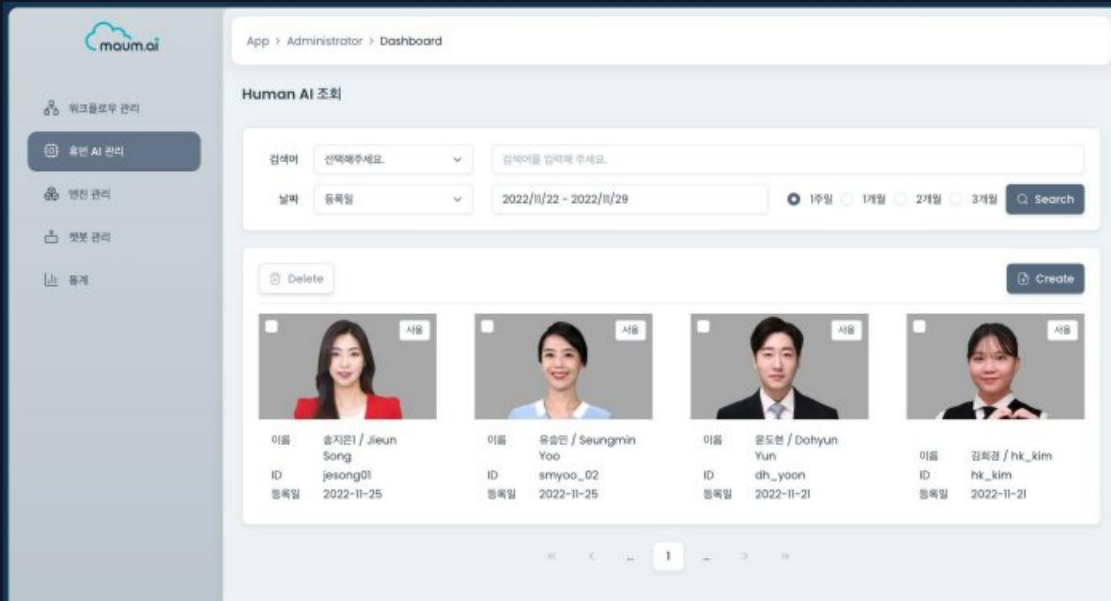
maum.ai 플랫폼 접속

maum Orchestra

엔진별 조건입력

원하는 시서비스 생성

Orchestra as an AI Human Builder



전신 행동 영상

행동 표현이 동반한 실사 인물 발화



상황에 맞는 의상, 인물

응답마다 인물 설정 가능.
문장에 따라 감정 표현 설정 가능



Any Device

휴먼시는 모바일, 태블릿, PC,
키오스크, 로봇 등 어떤
디바이스에도 탑재됩니다.

Logging Visualization

- AI엔진과 서비스 파트너에 대한 정확한 과금체계 완성

Data

Time

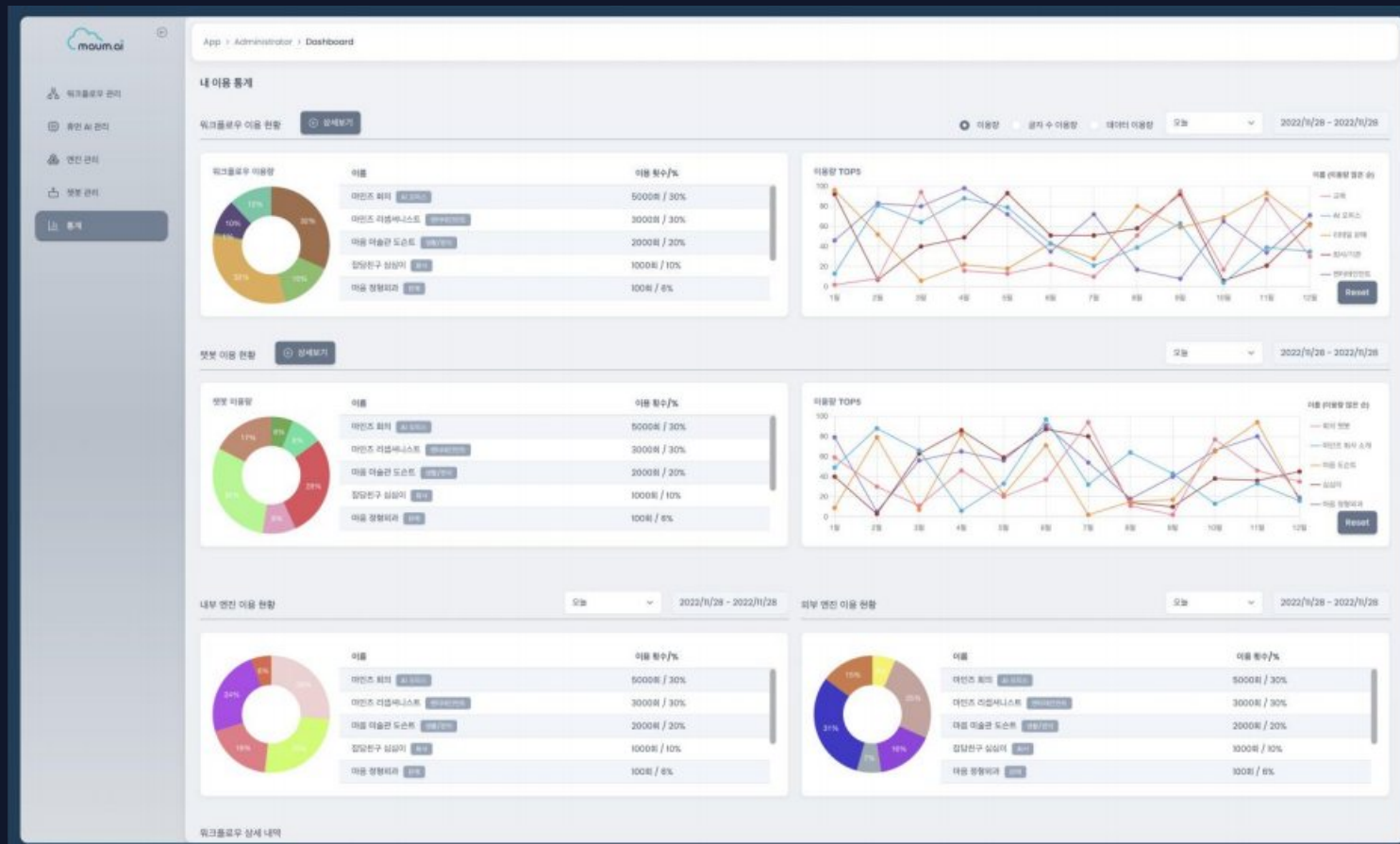
Text

STT, TTS,
STF, API,
ChatBot,
VSP, ISP,
LM Vision

API, Socket,
gRPC

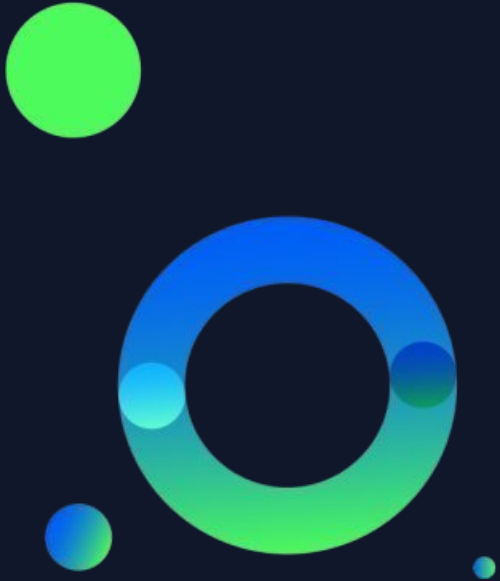
ChatBot,
Speaking
module

9개 종류 10개 항목 측정 및 Gauge 제공





Make your own AI service with maum.ai



- 인공지능 플랫폼의 상용화에 성공하다
마인즈랩의 AI Platform maum.ai
- 초개인화 맞춤형 인공지능 서비스 개발 도구
마음 오케스트라 (maum Orchestra™)
- 실제로 인간의 노동을 대체하는
마인즈랩의 AI Human 기술과 제품
- ChatGPT의 기반기술 초거대AI 공급업체
초거대AI 언어모델과 챗봇을 활용한 서비스
- 과금형으로 인공지능 서비스를 제공한다
maum.ai 클라우드 AI API
- AI Transformation을 선도한다
Data, 전략적 파트너십, 컨설팅, 정부과제

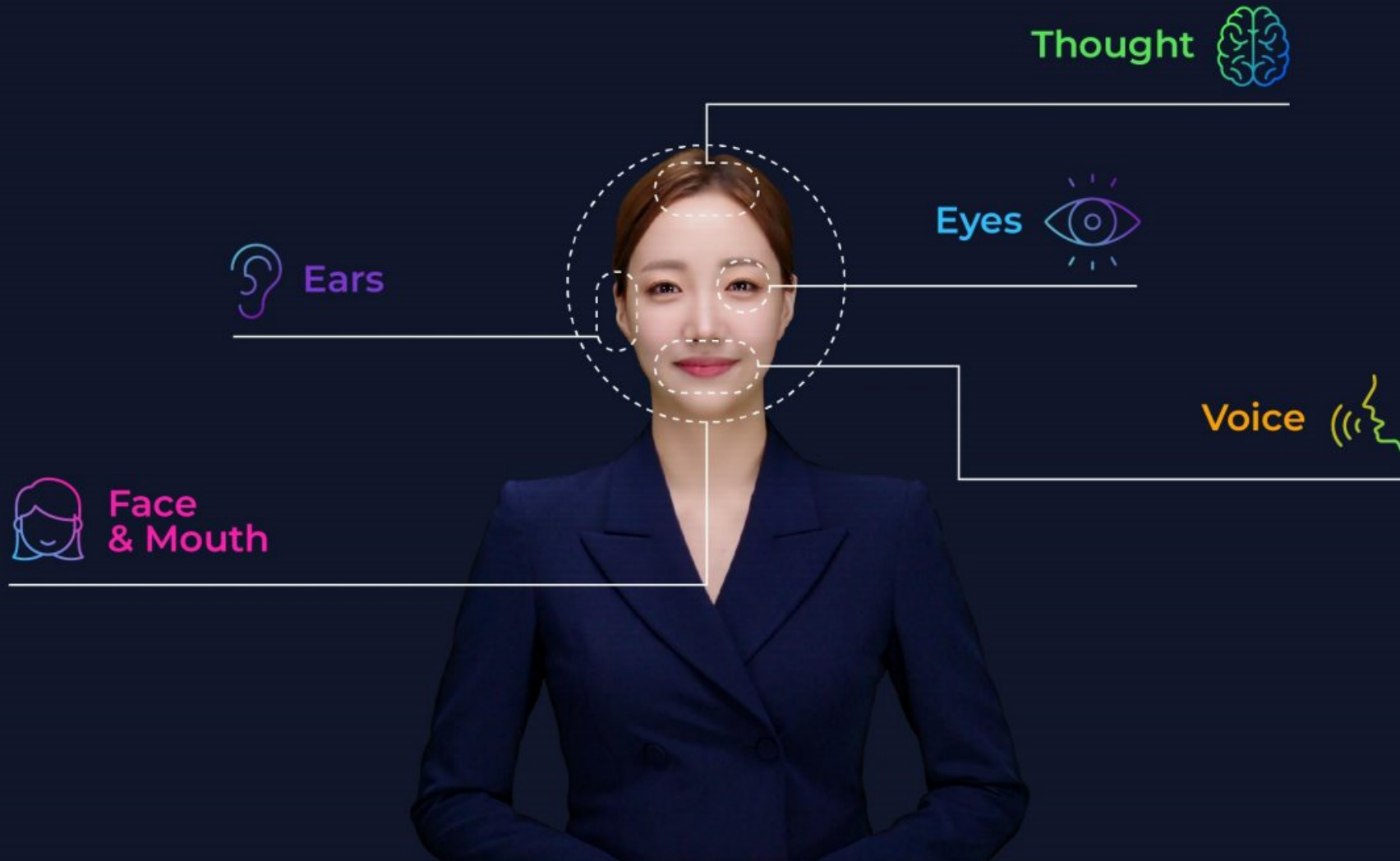
시휴먼



Human
AI



마인즈랩의 AI휴먼은 단지 목소리와 얼굴을 만들어내는 기술이 아닙니다. 인간의 노동을 대체할 수 있는 수준의 **Integrated Human AI Technology**를 구현하는 것이 마인즈랩의 기술입니다. **AI Human**의 제작, API 제공, 디바이스를 다양한 산업과 도메인으로 확산하고 있습니다



Human AI 클라우드 플랫폼 maum.ai World 서비스



Any Device

휴먼시는 모바일, 태블릿, PC, 키오스크, 로봇 등 어떤 디바이스에도 탑재됩니다.



User Experience

단순한 사용절차를 통해서 최고의 사용자 경험을 제공합니다.

Know-Who AI휴먼

maum.ai를 통해서 수많은 휴먼AI 서비스를 사용할 수 있습니다.

One App

maum.ai 앱 다운로드 하나로 모든 휴먼시 서비스를 사용할 수 있습니다. 사용자는 개인별로 서비스를 커스터마이징 할 수 있습니다.



maum.ai World

이것이 마인즈랩의 maum.ai World입니다.

어떤 디바이스에도 Any Device



Mobile



Laptop



PC

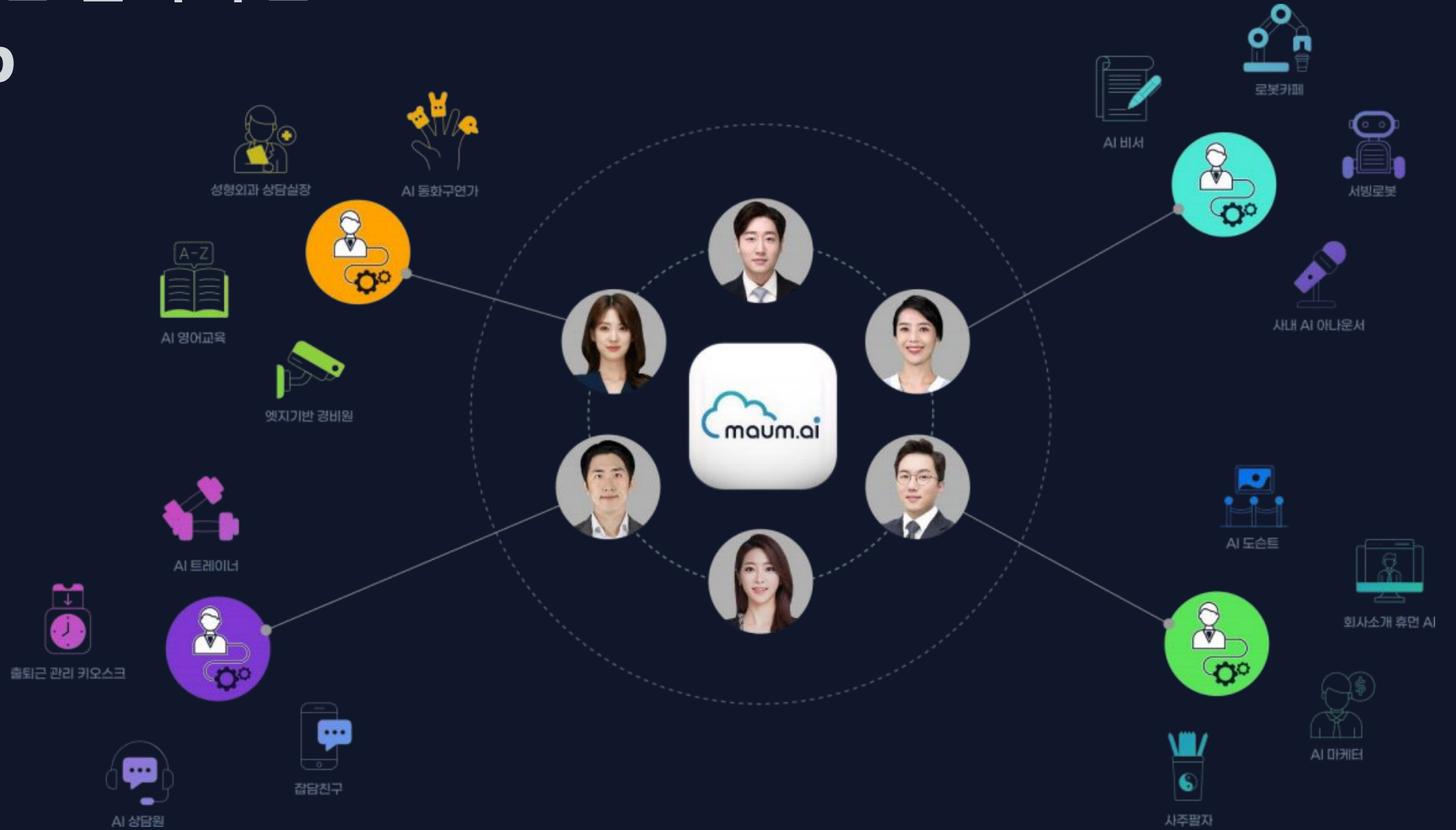


Tablet

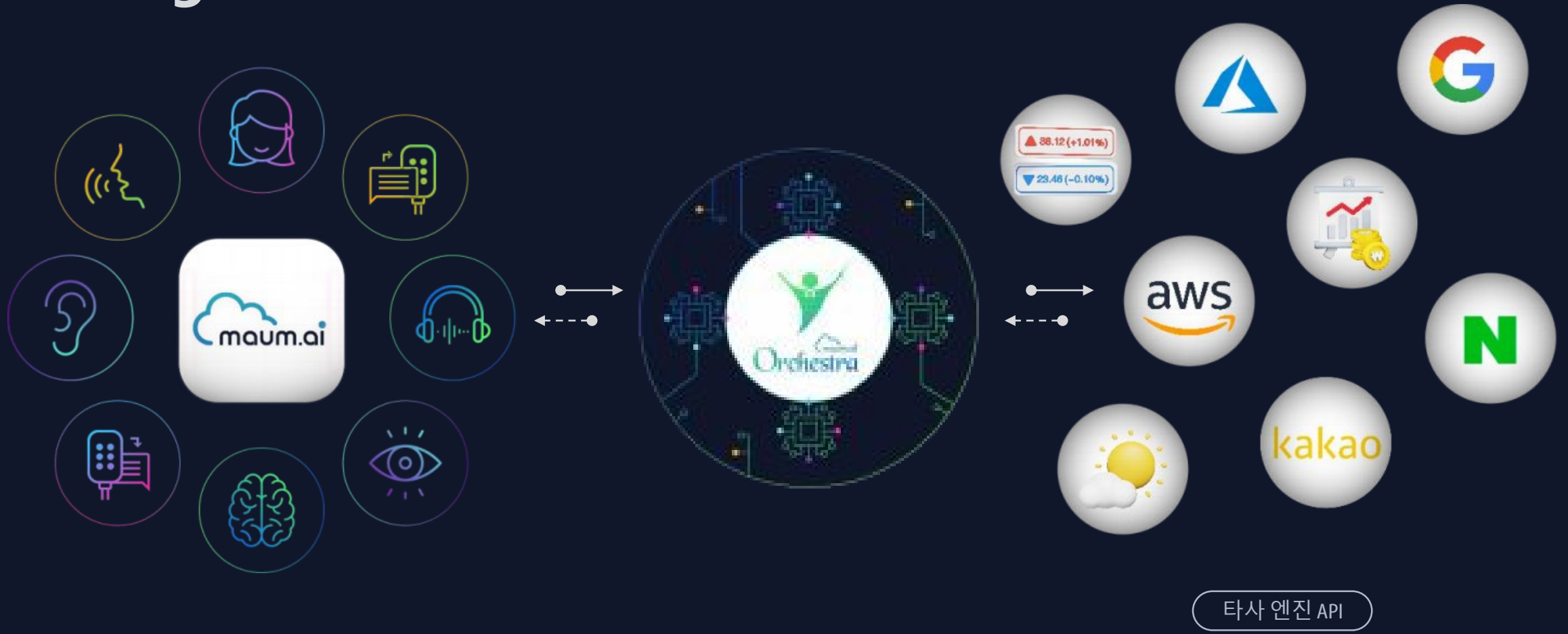


Kiosk

하나의 앱만 설치하면 One App



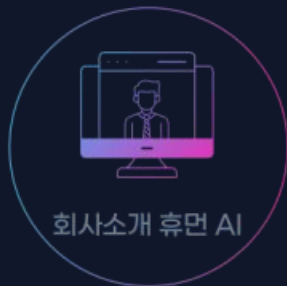
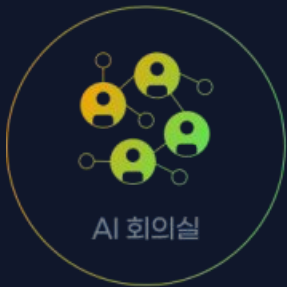
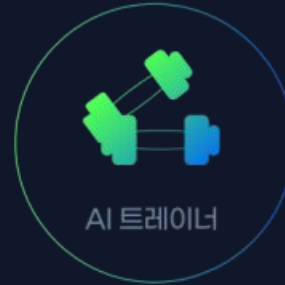
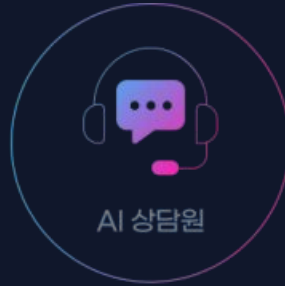
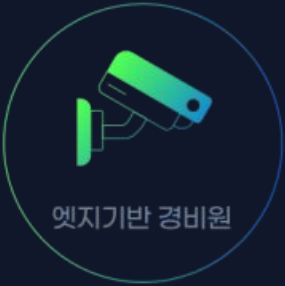
다양한 인공지능 엔진들과 AI Engine API



타사 엔진 API

다양한 인공지능 서비스들

AI Service API



최고의 사용자 경험을 누릴 수 있습니다

The Best User Experience



Multimodal

Interface

[정의]

사용자의 제스처, 손의 움직임, 음성, 패턴 등
사용자가 컴퓨터에 전달할 수 있는 다양한 입출력을 컴퓨터가
동시에 받아들이고 조합하여 이해할 수 있는 인터페이스 시스템



maum.ai에 깃든

최상의 사용자 경험

maum.ai!

가장 중요한 정보를 강조하라!

시선을 유도하라!

Call To Action!



Any Device  experience.

just **One app** download, 

Customize  your user interface.

활용사례

실제로 인간의 업무를 대체한 사례



여수MBC

활용사례

실제로 인간의 업무를 대체한 사례



L'ORÉAL



kotra

대한무역투자진흥공사

ADTech

광고에 활용되는 휴먼 AI



Metaverse

여행 안내



원하는 여행지를 선택해주세요.



AI Human 모듈은 딥러닝으로 학습시킨 인공지능이 은행원, 리셉셔니스트, 도슨트, 강사 등 실제 사람의 업무를 대신 수행하며, 초기에는 은행 등 금융업에서 시작하여 다양한 산업으로 확산되고 있음.

우리가 만드는 은행의 미래



금융의 물리적 환경과 디지털 환경의 혼합

금융기관은 AI로부터 비용 절감의 혜택을 누릴 수 있습니다. 자산운용사, 은행, 보험사는 대화형 AI, 로봇틱프로세스 자동화, 광학 문자 인식, 기타 머신러닝, 딥러닝 애플리케이션 등의 기술을 활용해 일상 업무에서 효율성을 창출할 수 있습니다.

- 1 고객 입장
- 2 AI 컨시어지
고객 응대 및 창구 안내 (번호표 발행)
- 3 AI 컨시어지 로봇
고객 응대 및 점포 안내
- 4 디지털 데스크
은행 업무
- 5 대면 창구
은행 업무
- 6 스마트 ATM
창구업무
체크카드 발급, 통장정리, 보안카드 / OPT 재발급, 바이오 등록 / 재등록, 온라인 뱅킹 신청 및 변경, 개인 정보 변경
ATM 업무
입출금, 계좌이체, 지로 / 공과금 납부
- 7 스마트 카드업무 키오스크
카드 신청, 교체 및 발급, 카드 대출 및 관리 등
- 8 고객 퇴장

미래의 리셉셔니스트



메타버스 가상공간 안에도 적용 가능한 인공지능 리셉셔니스트

회사 로비와 세미나장 등에서 인공지능 리셉셔니스트를 통해 방문 고객 응대 및 정보 제공 등 접객 업무를 수행할 수 있다. 메타버스 가상 공간에도 적용가능 함은 물론 키오스크, 스마트폰, 로봇 등 다양한 디바이스와 결합할 수 있다.

- 1 고객 입장**
고객이 입장 후 키오스크 인지
- 2 AI 리셉셔니스트**
고객을 인식하고, 고객의 의도를 파악해 정보 제공
- 3 직원 호출**
방문고객이 찾는 직원을 호출
- 4 메타버스와 결합**
가상공간 안에서도 메타버스 디지털트윈을 구현하며, 다양한 기기와 결합 가능

미래의 도슨트



코로나 19로 인한 비대면 수요 증가, 도입속도 높아지는 인공지능 도슨트

코로나 19로 인한 비대면 서비스 수요가 증가함에 따라 문화, 전시 공간에서는 인공지능 도슨트 도입에 속도를 내고 있다. 인공지능 도슨트를 통해 시 기반의 체험 콘텐츠 개발도 본격화하며, 색다른 체험을 선사하고 있다.

1 고객입장
고객이 입장 후
키오스크 인지

2 전시장 안내
관람객이 원하는 전시장의
시설 정보 안내

3 작품설명
관람객에게 작품을 설명하고
전시 가이드를 제공

4 관람객과의 대화
관람객에게 AI 체험 콘텐츠를
제공하고, 관람객의 자유 대화를
제공

미래의 경비원

CCTV 기능을 넘어 AI Edge로 일상의 새로운 변화를 가져오는 인공지능 경비원

카메라가 판단하는 Edge AI 기술을 적용하여 배회, 침입, 실신 등의 이상행동을 탐지하고, 주의 알림을 통해 전조단계에서의 빠른 대응을 하고 있다.



1 배회 인식
특정인이 공간을 배회하는 것을 인식하여 위험 상황을 신속하게 알림

2 침입 인식
임의로 설정한 영역 내에 사람이 침입한 것을 인식하여 위험 상황을 신속하게 알림

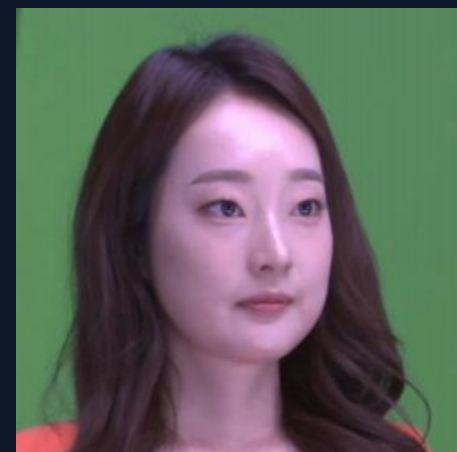
3 실신 인식
임의로 설정한 영역 내에 사람이 실신한 것을 인식하여 위험 상황을 신속하게 알림

최고 화질의 AI Human을 만들어내는 마인즈랩 "인공인간 제작소"입니다.

마인즈랩 인공인간 제작소 (압구정점)



촬영 원본



휴먼AI 영상 제작용 스튜디오 사용자

M 스튜디오를 사용해 영상이나 음성을 만들고자 한다면,
maum.ai 계정을 생성 후 한달간 무료로 사용 가능합니다.

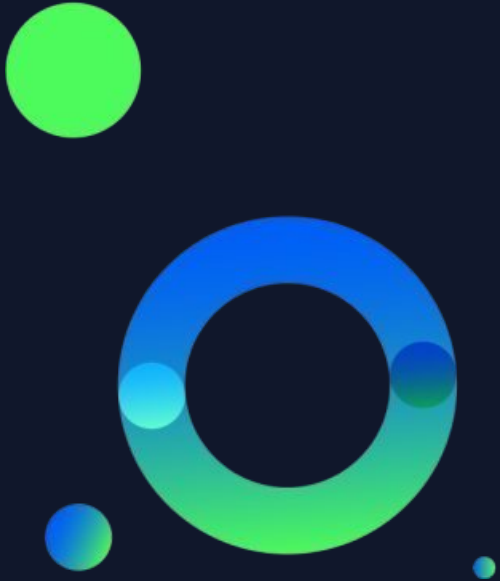


The screenshot displays the '영상 편집' (Video Editing) page of the M Studio web application. The interface includes a search bar, a table of video projects, and a navigation sidebar. The table lists various video projects with columns for No., Thumbnail, Title, Creation Date, Number of Characters, Resolution, Video Format, Duration, and Editing Status.

| No. | 영상 썸네일 | 제목 | 생성일 | 순다이트 개수 | 해상도 | 영상 용량 | 소요 시간 | 작업 상태 |
|------|--------|-----------|------------------|---------|-----|--------|--------|-------|
| 3046 | | New Title | 2022-09-22 14:07 | 1 순다이트 | UHD | 2.07MB | 22 sec | 지정완료 |
| 3044 | | New Title | 2022-09-22 13:35 | 1 순다이트 | UHD | 2.02MB | 35 sec | 지정완료 |
| 3043 | | New Title | 2022-09-22 13:28 | 1 순다이트 | HD | 0.22MB | 24 sec | 지정완료 |
| 3037 | | New Title | 2022-09-22 11:34 | 1 순다이트 | UHD | 0.64MB | 15 sec | 지정완료 |
| 3036 | | New Title | 2022-09-22 11:24 | 1 순다이트 | UHD | 0.46MB | 12 sec | 지정완료 |
| 3035 | | New Title | 2022-09-22 11:34 | 1 순다이트 | UHD | 0.59MB | 11 sec | 지정완료 |
| 3034 | | New Title | 2022-09-22 11:24 | 1 순다이트 | UHD | 0.42MB | 10 sec | 지정완료 |
| 3033 | | New Title | 2022-09-22 11:34 | 1 순다이트 | UHD | 0.54MB | 11 sec | 지정완료 |
| 3032 | | New Title | 2022-09-22 11:24 | 1 순다이트 | UHD | 0.36MB | 9 sec | 지정완료 |
| 3031 | | New Title | 2022-09-22 11:34 | 1 순다이트 | UHD | 0.61MB | 16 sec | 지정완료 |



Make your own AI service with maum.ai



- 인공지능 플랫폼의 상용화에 성공하다
마인즈랩의 AI Platform maum.ai
- 초개인화 맞춤형 인공지능 서비스 개발 도구
마음 오케스트라 (maum Orchestra™)
- 실제로 인간의 노동을 대체하는
마인즈랩의 AI Human 기술과 제품
- ChatGPT의 기반기술 초거대AI 공급업체
초거대AI 언어모델과 챗봇을 활용한 서비스
- 과금형으로 인공지능 서비스를 제공한다
maum.ai 클라우드 AI API
- AI Transformation을 선도한다
Data, 전략적 파트너십, 컨설팅, 정부과제

초거대 AI 언어모델과 챗봇





(MS & NVIDIA, Megatron-Turing NLG 530B; arXiv 2201.11990)



초거대 AI

초거대 AI Pros & Cons



초거대 AI Sweet Spot

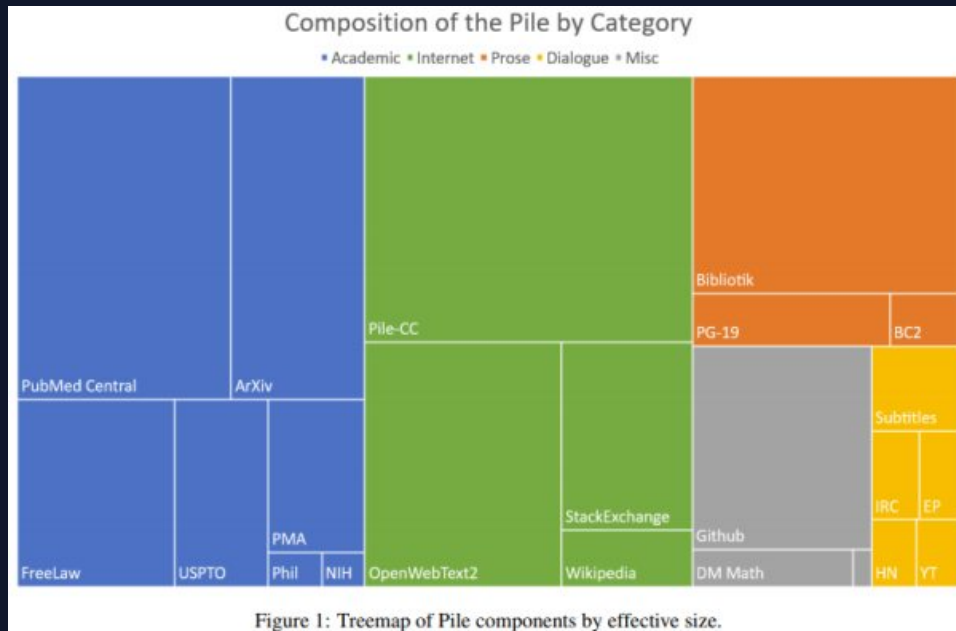
다양한 태스크
보통 성능
많은 비용
사전학습
데이터 필요

**GPU 한 장으로 추가 학습과
inference가 가능한 지점**

개별 태스크
좋은 성능
적은 비용
태스크
데이터 필요

초거대 AI Pretrain 학습데이터

다음과 같은 데이터들을
base model 학습데이터로 수집



EluetherAI, *The Pile*; arxiv 2101.00027

1. 논문

- a. 바이오/의학
- b. 컴퓨터공학
- c. 수학
- d. 물리학
- e. 법학
- f. 철학
- g. 특허 관련
- h. 기타 분야의 논문

2. 인터넷

- a. 커뮤니티 인기글
- b. 지식 질의응답
- c. 위키피디아
- d. 기타

3. 산문

- a. 픽션/논픽션
- b. 고전문학
- c. 기타

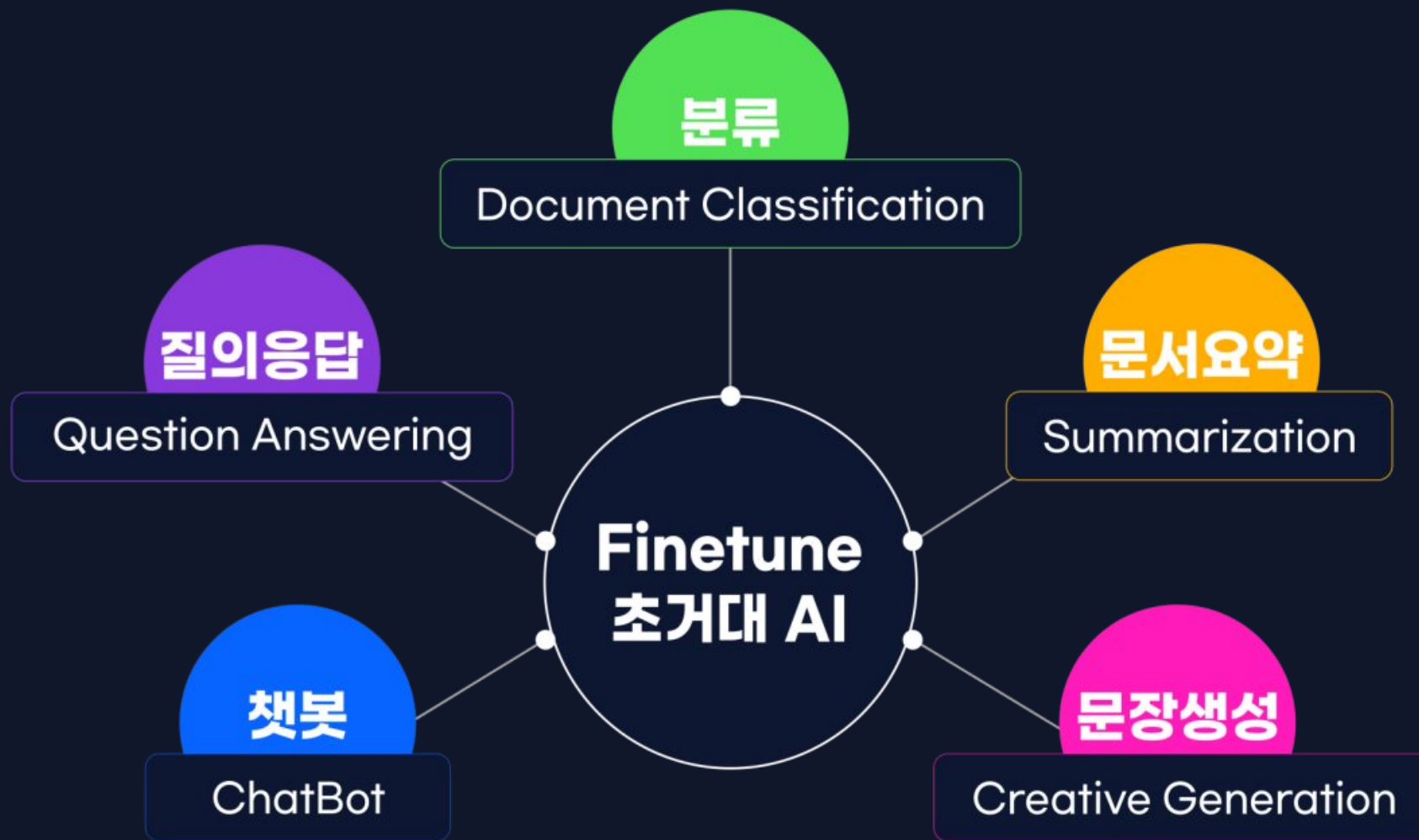
4. 대화

- a. 영화/TV쇼 자막
- b. 유튜브 자막
- c. 오픈채팅서버 대화로그
- d. 커뮤니티 등 지식/정보글 댓글 트리
- e. 기타 대화 데이터

5. 기타

- a. 깃허브 공개 코드
- b. 수학문제
- c. 이메일

기존 AI API 뿐만 아니라 초거대 AI를 개발하여, API로 제공하여 API 제공 수익 창출



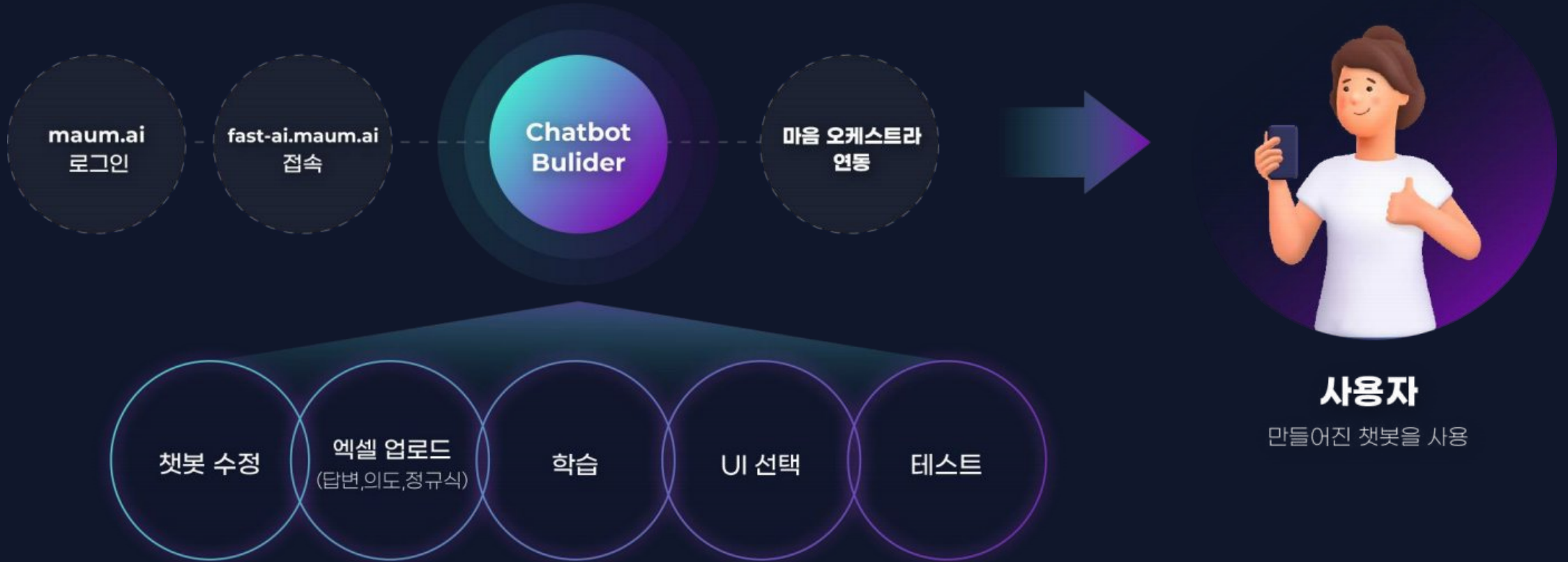
초거대 AI pretrain + Finetune API 서비스

(1.3B ~ 6B)



- 태스크별 layer 추가 (GPT는 그대로)
- 태스크별 base data로 학습

AI챗봇 모듈은 시나리오 기반 및 딥러닝 기반의 엔진 뿐만 아니라, 최근에 학습을 시킨 초거대AI를 적용하여, 고객의 문의에 대한 AI의 답변 수준과 정확도를 획기적으로 높였습니다.

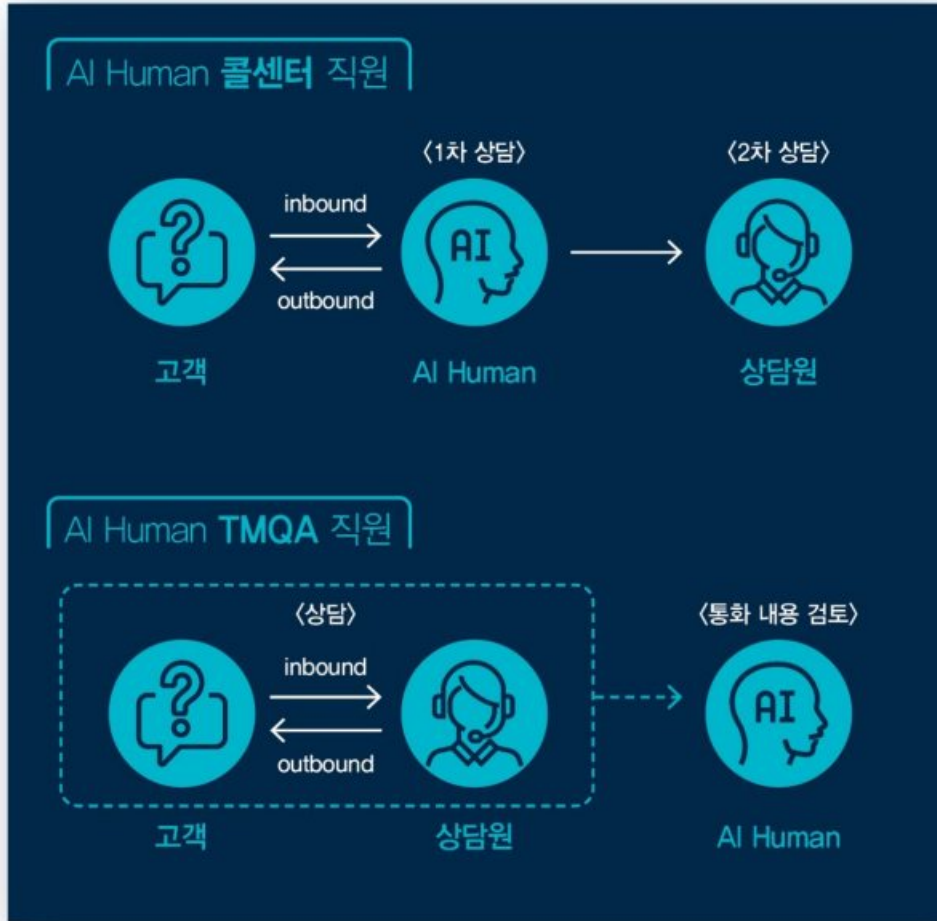


AI음성봇 모듈은 시나리오 기반 및 딥러닝 기반의 엔진으로 고객에게 AI가 전화를 걸어서 고객상담 업무를 진행하는 것으로서, 금융업에서 시작하여 다양한 산업의 고객센터로 확장되고 있음



TMQA 모듈은 고객센터의 텔레마케터의 콜을 AI가 듣고 품질검수(QA)를 하며, 주로 보험업 등 금융산업의 고객센터에 판매되어 지속적인 수익을 창출하고 있음

인공인간
고객상담원



음성봇
성공 콜수

연간
565,000 콜

마케팅 콜 검사(TMQA)
계약수

연간
2,300,000 콜

고객의 소리(voc)
분석 콜수

국내 연간
100,000,000 콜

미국 연간
10,000,000 콜



2016년부터 프로젝트 진행
(REVO 시스템 업그레이드)



현대해상

현대해상 AI 음성봇

VOC 모듈은 고객센터의 콜을 AI가 듣고 제품결함, 고객불만, 법적 위기사항 등을 감지하며, 다양한 산업의 고객센터에 제공되고 있음.

삼성전자 미국/캐나다 VOC 분석

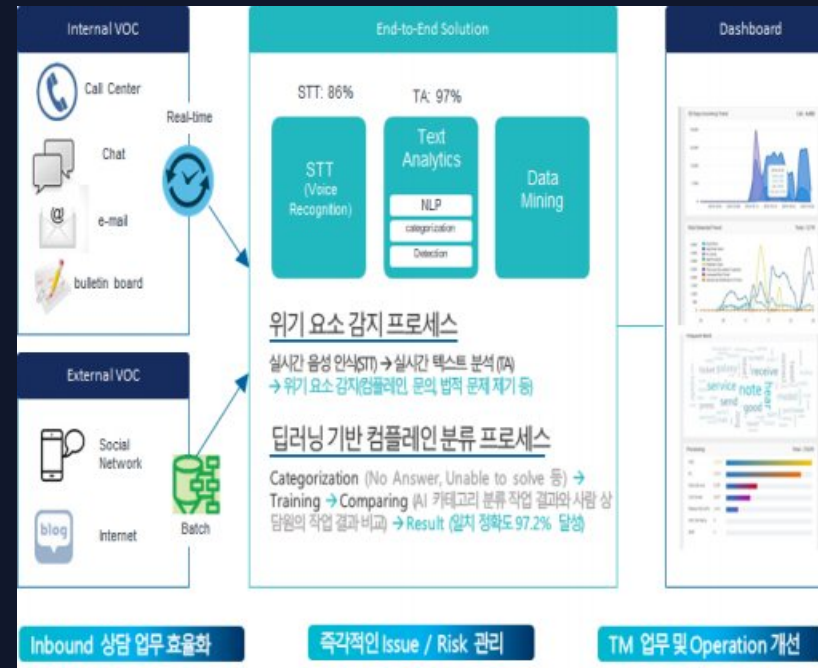
- 연간 10만건의 상담콜을 실시간 Active Recording하여 고품질로 음성인식하여 텍스트로 전환하고, Deep Learning 텍스트 분석
- 2016년 1차 프로젝트 이후 지속적으로 서비스 추가 및 품질 관리 서비스 제공 중

[음성인식 성공률]

- 2016년 : 84% (ETRI RNN 사용)
- 2018년 : 90% (ETRI LSTM 사용)
- 2020년 : 93% (마인즈랩 자체 CNN STT 사용)

- RISK탐지, 상담원 QA, 상담유형 자동분류 등에 자연어 처리 기반 Text 분석 서비스 구축

VOC AI Analytics 시스템 (RISK탐지, 상담QA)

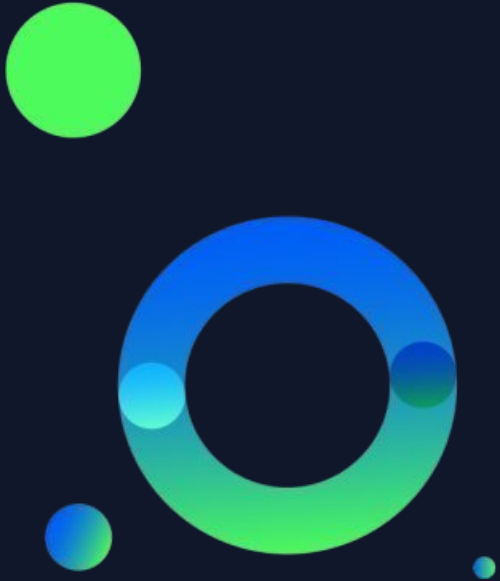


AI속기사 모듈은 회의 시에 회의 참가자의 발언을 음성인식하여 녹취록을 작성해주며,
클라우드 기반과 온디바이스 기반으로 판매되고 있음





Make your own AI service with maum.ai



- 인공지능 플랫폼의 상용화에 성공하다
마인즈랩의 AI Platform maum.ai
- 초개인화 맞춤형 인공지능 서비스 개발 도구
마음 오케스트라 (maum Orchestra™)
- 실제로 인간의 노동을 대체하는
마인즈랩의 AI Human 기술과 제품
- ChatGPT의 기반기술 초거대AI 공급업체
초거대AI 언어모델과 챗봇을 활용한 서비스
- 과금형으로 인공지능 서비스를 제공한다
maum.ai 클라우드 AI API
- AI Transformation을 선도한다
Data, 전략적 파트너십, 컨설팅, 정부과제

국내 최다 상용 인공지능 모델 운영
(Face, Voice, Ears, Eyes, Language, Edge AI)



다양한 인공지능 엔진과 서비스를 클라우드 기반 API로 제공하여 지속적인 수익 창출

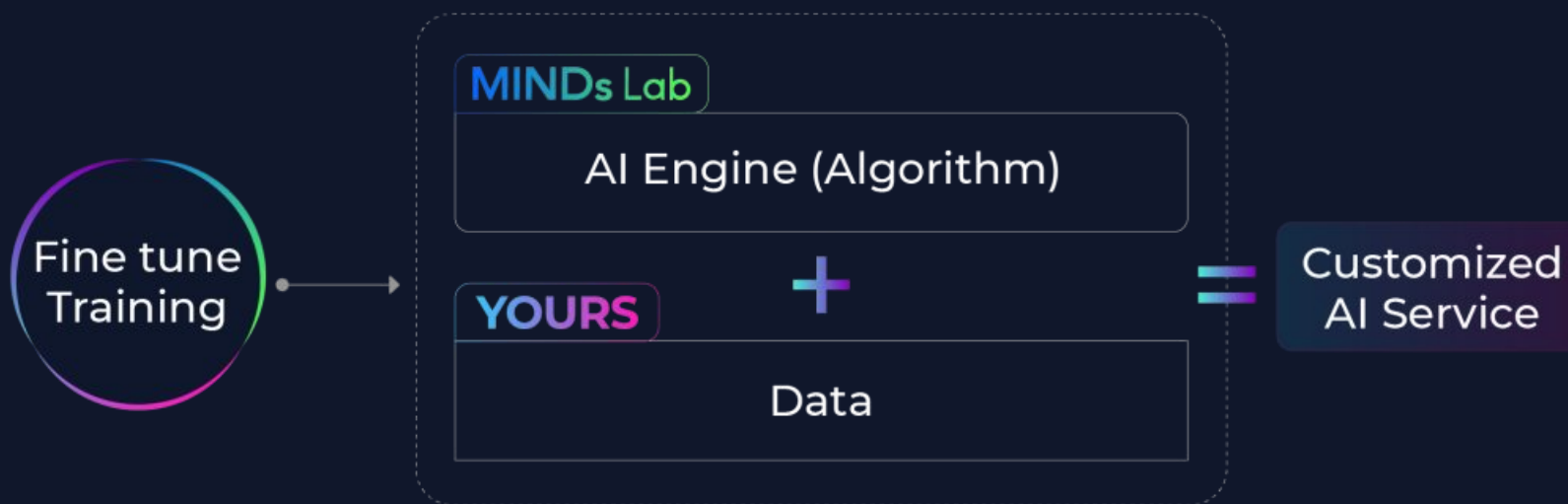
Pretrained AI API 서비스

Custom API Service



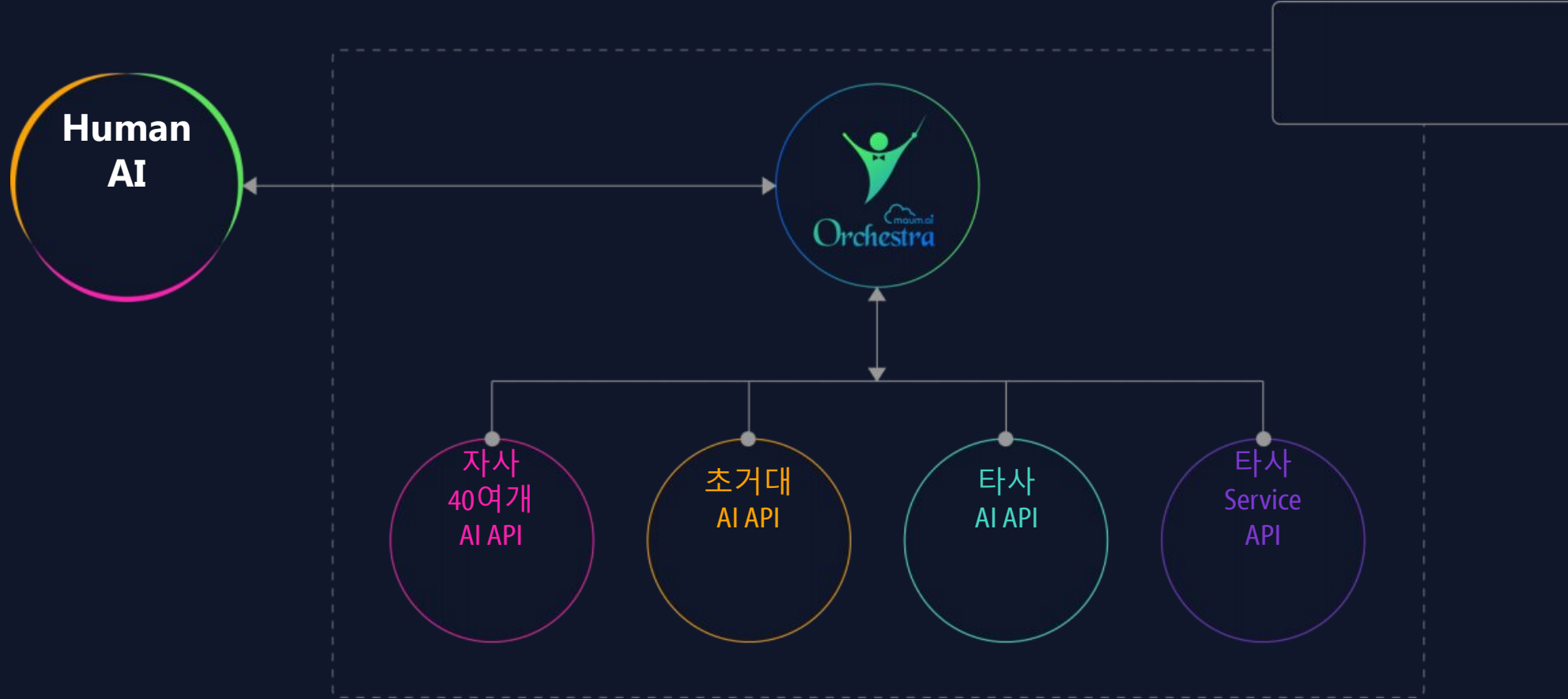
Finetune Train & Inference 서비스

Finetune Training을 통해 니즈에 부합하는 AI서비스를 제공



마인즈랩은 별도의 데이터 수집 및 학습을 통해 Customized AI Service를 제공할 수 있는 역량을 가지고 있습니다.
서버비, 연구비 등 별도의 금액이 발행할 수 있습니다.

세계 최고 수준의 최신 인공지능 중에서 당사의 사업 경쟁력과 직결되는 부문을 선택과 집중을 하여 연구개발을 하여, 탁월한 성과를 창출하였음. 특히, 2022년 12월1일에 마음오케스트라를 핵심으로 한 AI플랫폼을 완성하는데에 약 37억원의 연구개발비가 집중 투입되어 성공적으로 플랫폼을 오픈함.



단순히 목소리와 얼굴을 재현하는 타사의 기술과는
완전히 다른 초격차 기술!

당사는 현재 88개의 최신 인공지능 상용모델을 제공 중임



Eyes

6 개



Voice

26 개



Face & Mouth

48 개



Thought

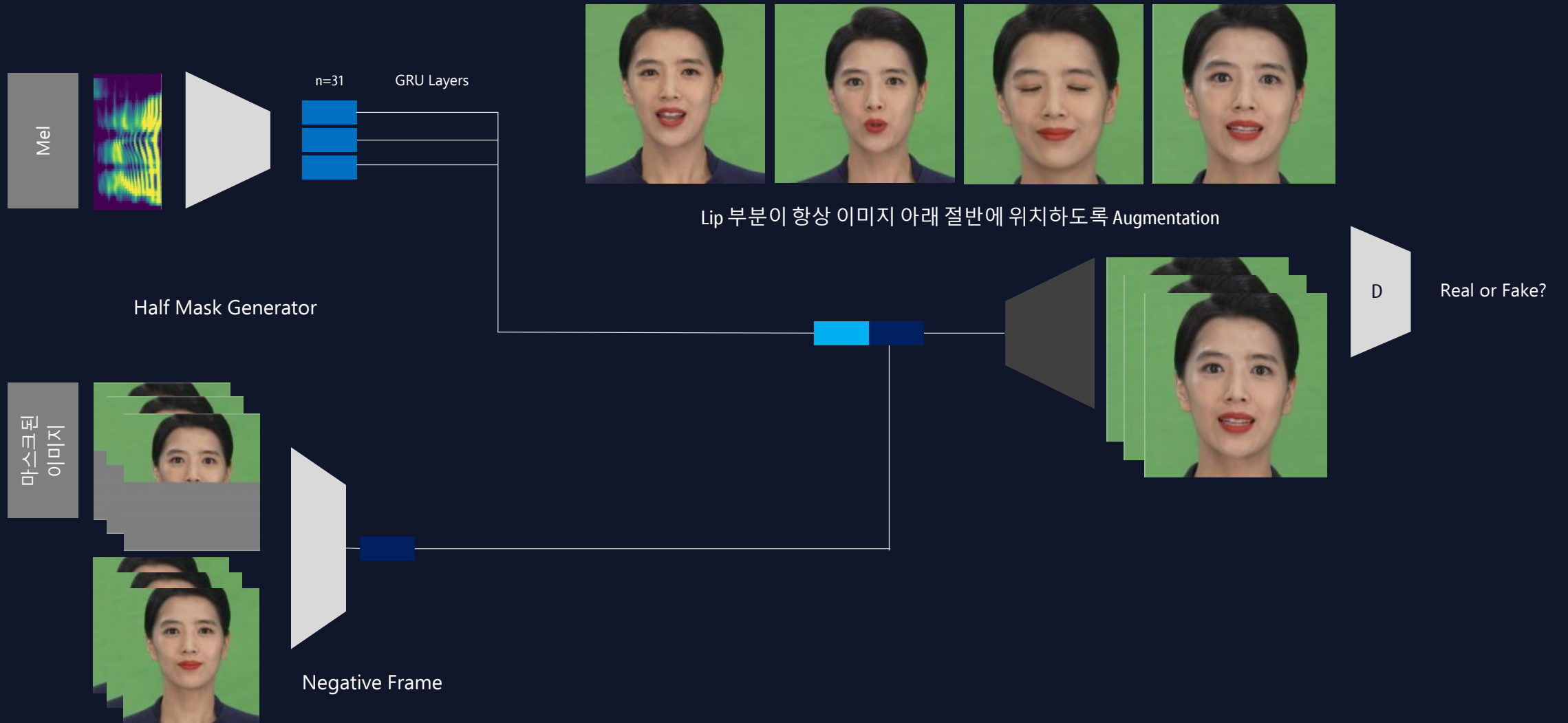
6 개



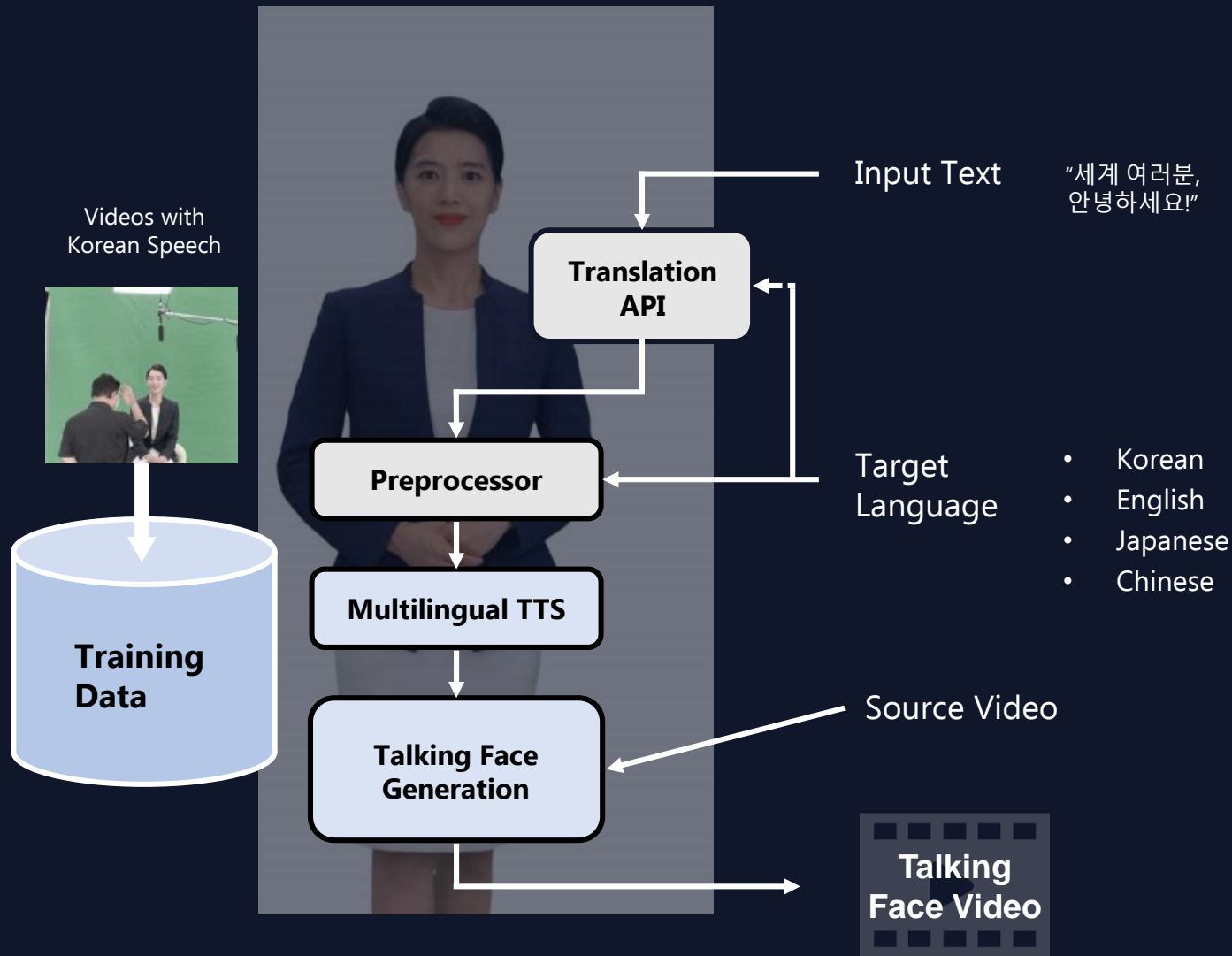
Ears

2 개

STF 엔진 – Face



Multi-lingual STF 엔진



StyleGAN 으로 만든 가상 얼굴로 진행한

Face Swapping



Interspeech 2022

NU-Wave 2 : A General Neural Audio Upsampling Model for Various Sampling Rates

Seungu Han^{*} and Junhyeok Lee

Conventionally, audio super-resolution models fixed the initial and the target sampling rates, which necessitate the model to be trained for each pair of sampling rates. We introduce NU-Wave 2, a diffusion model for neural audio upsampling that enables the generation of 48 kHz audio signals from inputs of various sampling rates with a single model.

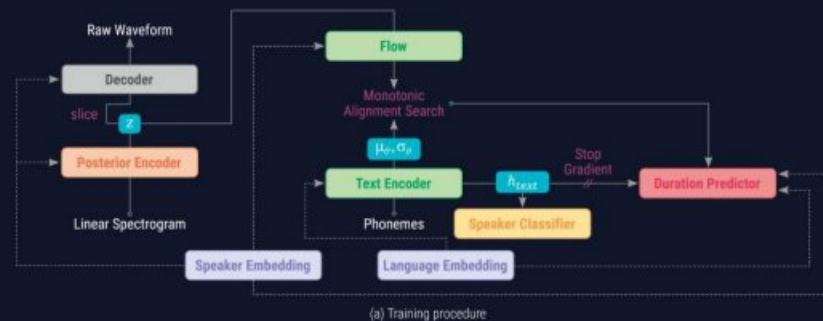


Interspeech 2022

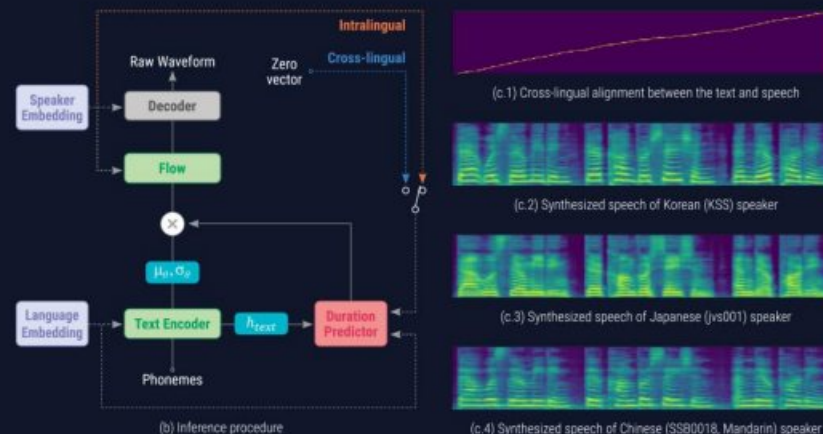
SANE-TTS : Stable And Natural End-to-End Multilingual Text-to-Speech

Hyunjae Cho^{*}, Wonbin Jung, Junhyeok Lee, and Sang Hoon Woo

In this paper, we present SANE-TTS, a stable and natural end-to-end multilingual TTS model. By the difficulty of obtaining multilingual corpus for given speaker, training multilingual TTS model with monolingual corpora is unavoidable. We introduce speaker regularization loss that improves speech naturalness during cross-lingual synthesis as well as domain adversarial training, which is applied in other multilingual TTS models. Furthermore, by adding speaker regularization loss, replacing speaker embedding with zero vector in duration predictor stabilizes cross-lingual inference.



(a) Training procedure



(b) Inference procedure

(c.1) Cross-lingual alignment between the text and speech
(c.2) Synthesized speech of Korean (KSS) speaker
(c.3) Synthesized speech of Japanese (jvs001) speaker
(c.4) Synthesized speech of Chinese (SSB0018, Mandarin) speaker

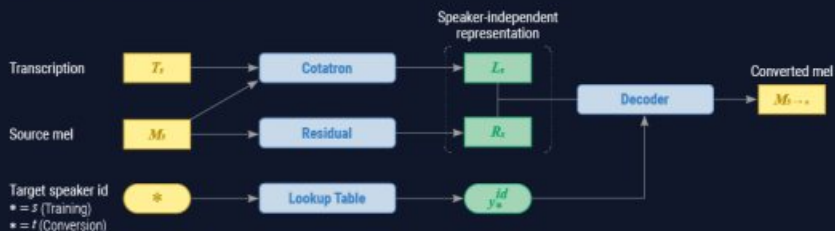
Interspeech 2020, 2021

Interspeech 2020

Cotatron : Transcription-Guided Speech Encoder for Any-to-Many Voice Conversion without Parallel Data

Seung-won Park*, Doo-young Kim, and Myun-chul Joe

We propose Cotatron, a transcription-guided speech encoder for speaker-independent linguistic representation. Cotatron is based on the multispeaker TTS architecture and can be trained with conventional TTS datasets. We train a voice conversion system to reconstruct speech with Cotatron features, which is similar to the previous methods based on Phonetic Posteriorgram (PPG). Our system can also convert speech from speakers that are unseen during training, and utilize ASR to automate the transcription with minimal reduction of the performance.

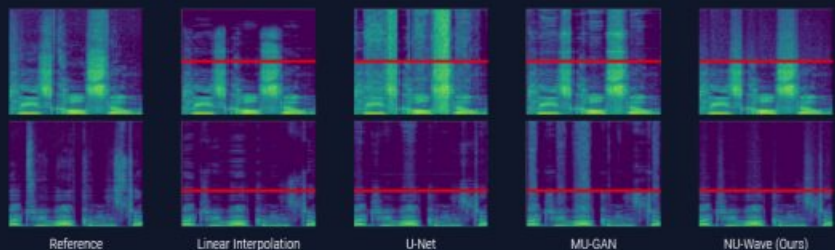


Interspeech 2021

NU-Wave : A Diffusion Probabilistic Model for Neural Audio Upsampling

Junhyeok Lee* and Seungu Han

In this work, we introduce NU-Wave, the first neural audio upsampling model to produce waveforms of sampling rate 48kHz from coarse 16kHz or 24kHz inputs, while prior works could generate only up to 16kHz. NU-Wave is the first diffusion probabilistic model for audio super-resolution which is engineered based on neural vocoders.

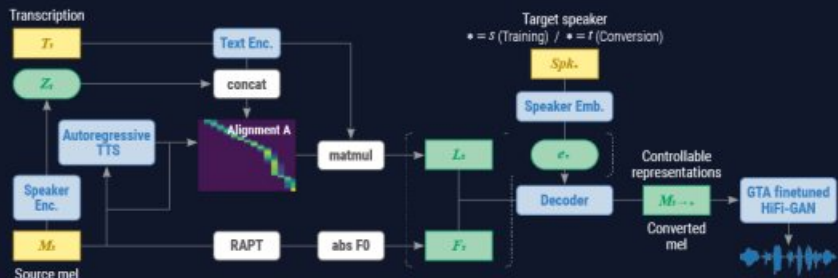


NeurIPS Workshop 2021

Controllable and Interpretable Singing Voice Decomposition via Assem-VC

Kang-wook Kim* and Junhyeok Lee

We propose a singing decomposition system that encodes time-aligned linguistic content, pitch, and source speaker identity via Assem-VC. With decomposed speaker-independent information and the target speaker's embedding, we could synthesize the singing voice of the target speaker. In conclusion, we made a perfectly synced duet with the user's singing voice and the target singer's converted singing voice.

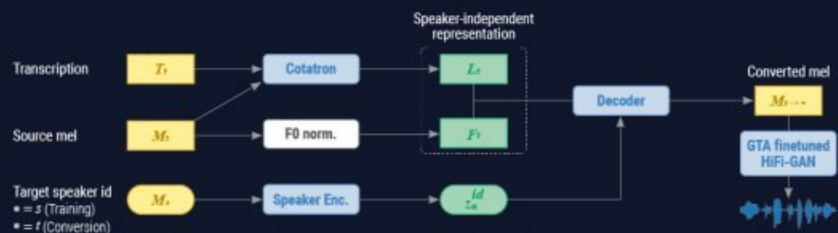


ICASSP 2022

Assem-VC : Realistic Voice Conversion by Assembling Modern Speech Synthesis Techniques

Kang-wook Kim*, Seung-won Park, Junhyeok Lee, and Myun-chul Joe

Recent works on voice conversion (VC) focus on preserving the rhythm and the intonation as well as the linguistic content. To preserve these features from the source, we decompose current non-parallel VC systems into two encoders and one decoder. We analyze each module with several experiments and reassemble the best components to propose Assem-VC, a new state-of-the-art any-to-many non-parallel VC system.

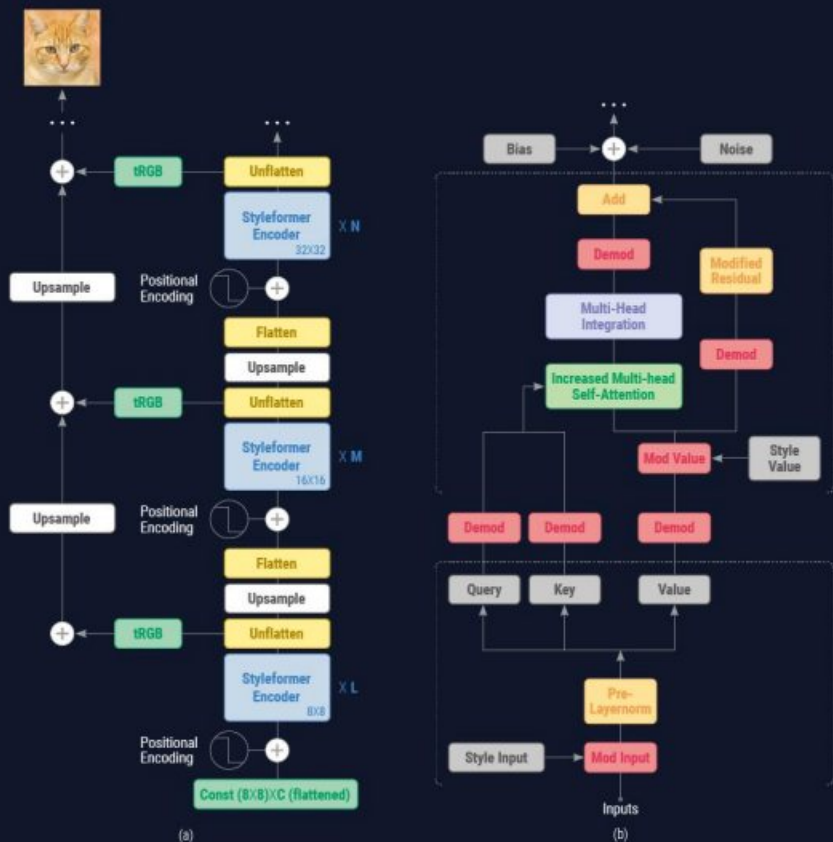


CVPR 2022

Styleformer : Transformer based Generative Adversarial Networks with Style Vector

Jeeseung Park* and Younggeun Kim*

We propose Styleformer, which is a style-based generator for GAN architecture, but a convolution-free transformer-based generator. We explain how a transformer can generate high-quality images, overcoming the disadvantage that convolution operations are difficult to capture global features in an image. Furthermore, we change the demodulation of StyleGAN2 and modify the existing transformer structure (e.g., residual connection, layer normalization) to create a strong style-based generator with a convolution-free structure. We also make Styleformer lighter by applying Linformer, enabling Styleformer to generate higher resolution images and result in improvements in terms of speed and memory.

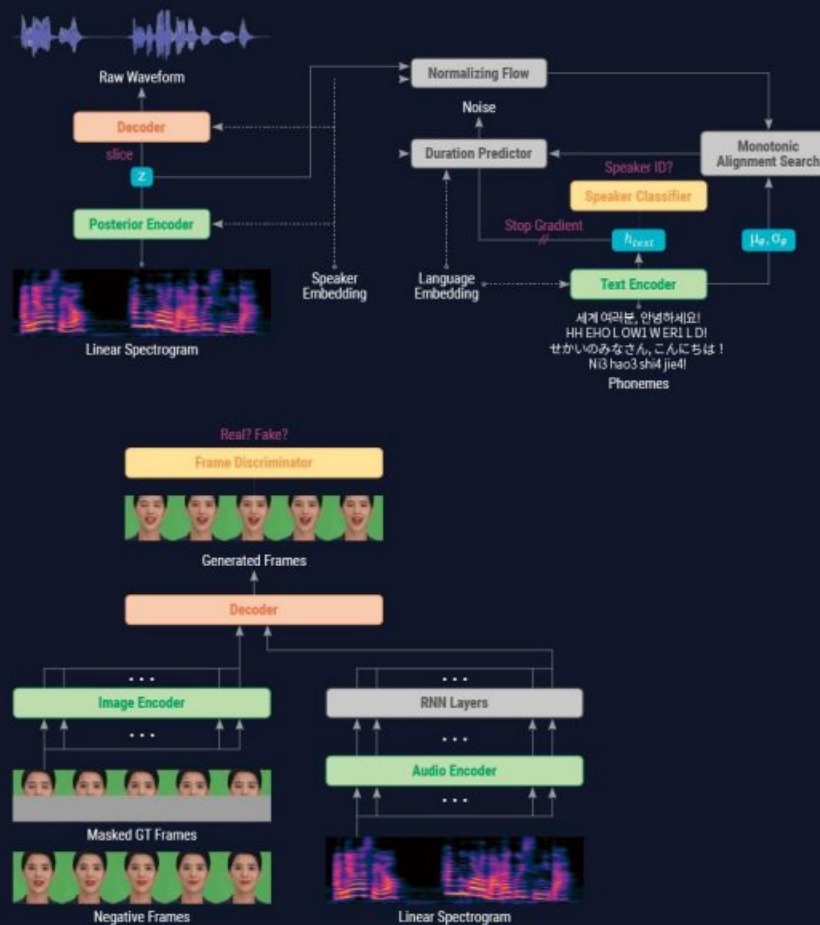


CVPR Demo 2022

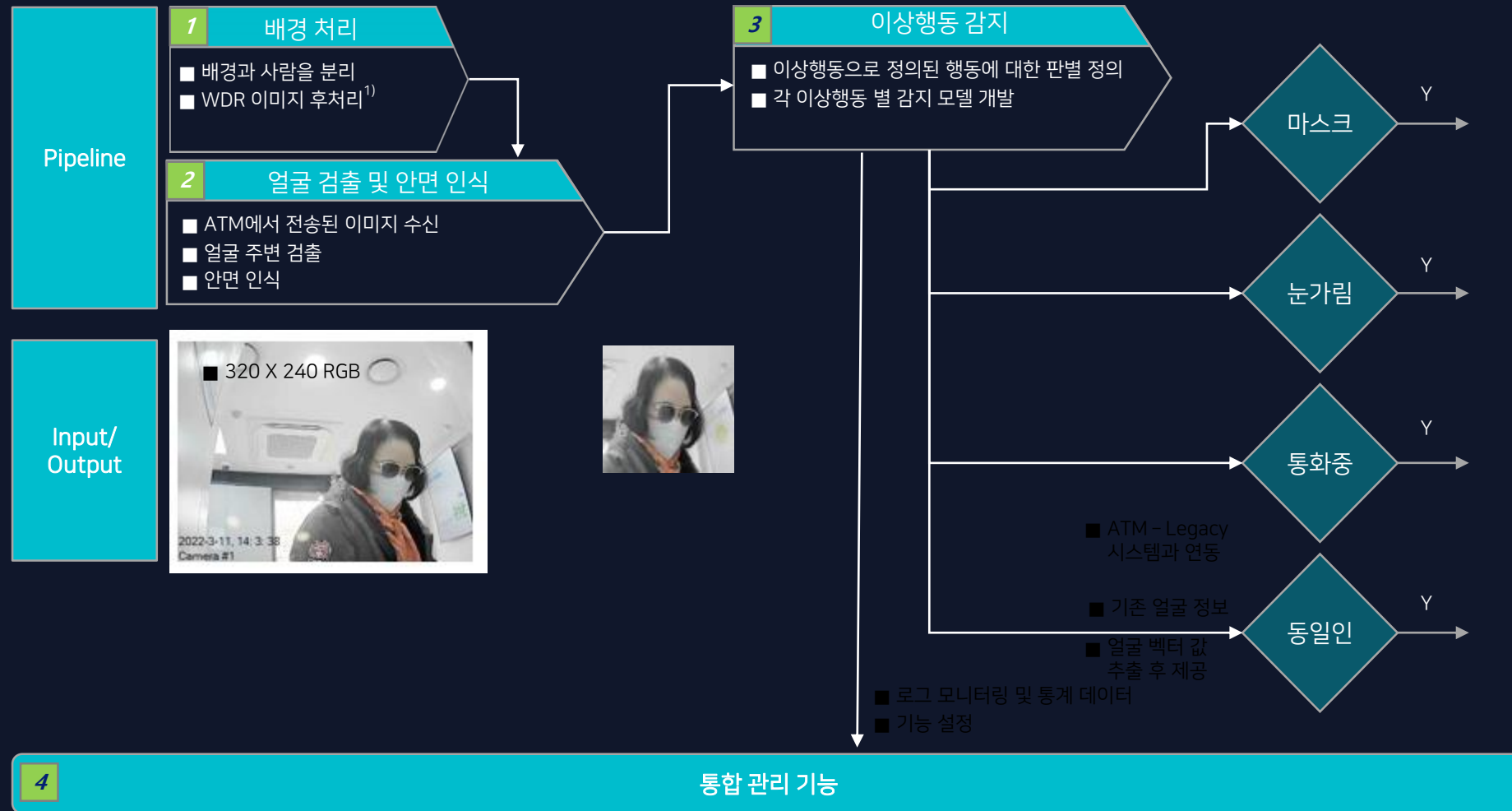
Talking Face Generation with Multilingual TTS

Hyoung-Kyu Song*, Sang Hoon Woo*, Junhyeok Lee, Seungmin Yang, Hyunjae Cho, Dongho Choi, Kang-wook Kim, and Youseong Lee

In this work, we propose a joint system combining a talking face generation system with a text-to-speech system that can generate multilingual talking face videos from only the text input. Our system can synthesize natural multilingual speeches while maintaining the vocal identity of the speaker, as well as lip movements synchronized to the synthesized speech. We demonstrate the generalization capabilities of our system by selecting four languages (Korean, English, Japanese, and Chinese) each from a different language family.



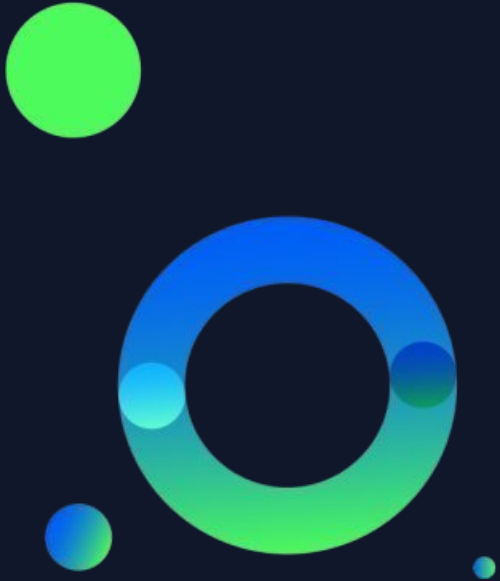
시각 모듈은 다양한 AI 시각엔진기술을 기반으로 금융사 보이스피싱 방지 등 다양한 산업, 다양한 용도로 매출이 발생함



1) Histogram Equalization & Adaptive Gamma Correction ;

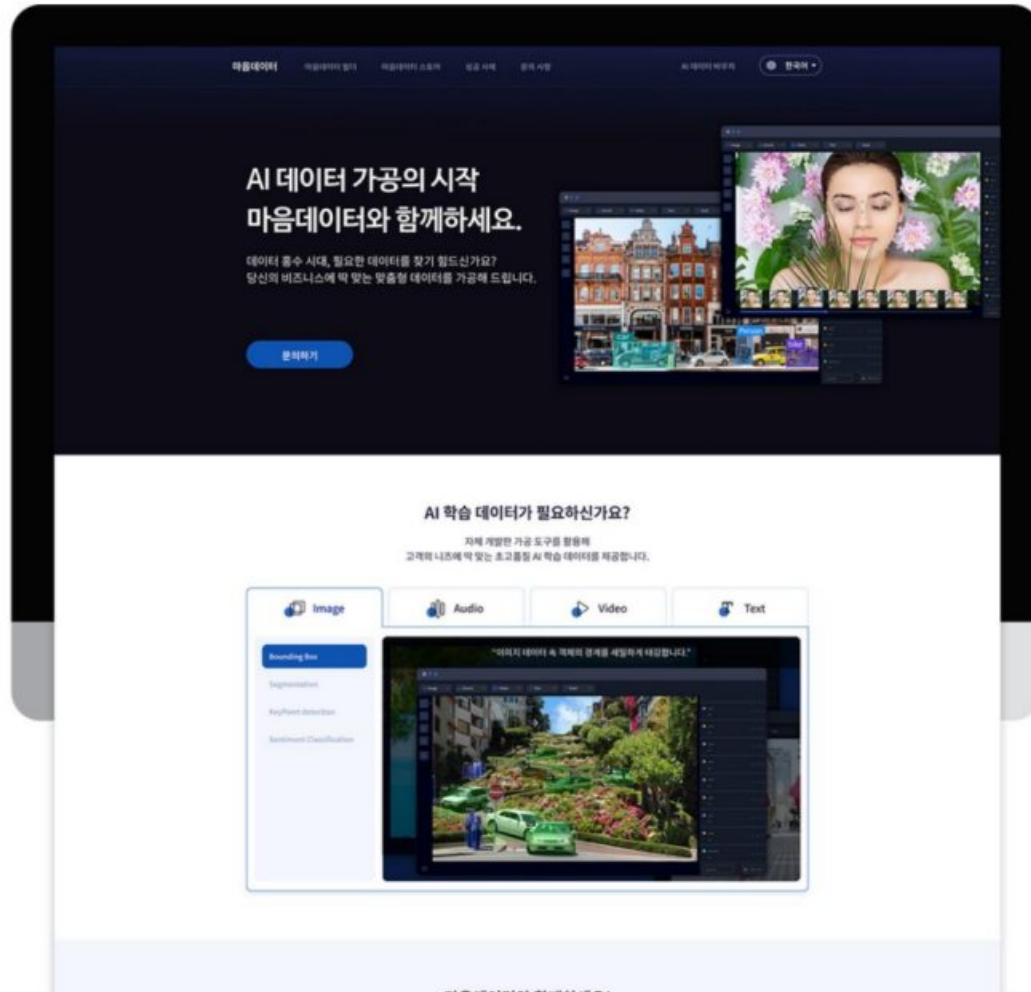


Make your own AI service with maum.ai



- 인공지능 플랫폼의 상용화에 성공하다
마인즈랩의 AI Platform maum.ai
- 초개인화 맞춤형 인공지능 서비스 개발 도구
마음 오케스트라 (maum Orchestra™)
- 실제로 인간의 노동을 대체하는
마인즈랩의 AI Human 기술과 제품
- ChatGPT의 기반기술 초거대AI 공급업체
초거대AI 언어모델과 챗봇을 활용한 서비스
- 과금형으로 인공지능 서비스를 제공한다
maum.ai 클라우드 AI API
- AI Transformation을 선도한다
Data, 전략적 파트너십, 컨설팅, 정부과제

AI 학습데이터를 정제/라벨링하는 마음데이터 Tool을 기반으로, 데이터 라벨링 서비스와 데이터 Tool 임대사업을 제공하여 지속적인 수익 창출



마음데이터 소개

마인즈랩의 데이터 서비스 플랫폼 **마음데이터**는 인공지능 전문 지식을 바탕으로 AI 데이터 가공, 데이터 판매, 그리고 데이터 구축을 원하는 기업을 대상으로 가공 도구 구축 및 임대 서비스를 제공합니다.

마음데이터 홈페이지 : <https://data.maum.ai/>

글로벌 화장품회사의 Beauty AI Platform 구축을 전략적 파트너로서 컨설팅사업을 수행하고 있으며, 다른 기업으로 확산할 수 있는 중요한 Reference를 쌓았음.

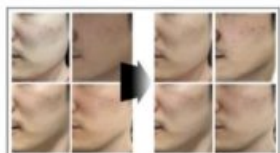
1 피부 측정 엔진 고도화

- 측정 항목 확대
As-Is : 4종 (모공, 주름, 홍반, 색소침착)
To-Be : 추가 4종 + 알파 (홍조, 잡티, 눈주름, 피부나이...)
- 엔진 성능 고도화
엔진 개선
촬영 사진 품질 개선
데이터 품질 개선

- Pose estimation 기반 구도 보정 및 최적 구도 자동촬영 기능



- 색감 보정 기능



2 피부 진단/분석 엔진 개발 및 고도화

- 고객 문진 & AI 측정결과 진단 표준 수립
AI 피부 측정 결과와 고객 문진을 활용한 진단 표준 수립
- 진단 표준에 기반한 AI 분석 시스템
표준 항목 결과로 분석 결과를 제공
- 피부 변화의 시스템 관리
피부 상태 변화를 시스템에서 관리
- 직관적인 피부 상태 확인/개선 방안 제시

| | |
|-------------------------|---|
| 사진 측정 값 | <ul style="list-style-type: none"> • 모공, 잡티, 트러블, 주름, 잔주름, 홍조, 홍반, 피부나이, 피부건강 등 |
| 원본 사진 AI 분석 | <ul style="list-style-type: none"> • 5축 점수 항목 • 피부 질환 분류 -민감 관련 피부질환: 여드름, 습진, 주사 -기타 피부질환: 피부암 등 |
| 문진 빅데이터 분석 | <ul style="list-style-type: none"> • 유분, 수분: 사진 측정의 제약을 보완 • 민감: 가려움, 따가움, 화끈 거림 등 주관적 증상 • 색소, 주름: 약화인자 확인 (흡연, 자외선 노출 등) |
| 생활데이터 시계열 분석 (항후 추가 계획) | <ul style="list-style-type: none"> • 외부 영향: 날씨, 계절, 미세 먼지 등 • 개인 생활패턴: 생리, 수면, 스트레스 등 |



3 추천 엔진 개발 및 고도화

- 피부진단 결과를 복합적으로 이해
- 기성제품/제형 개인화 추천 API 개발
- Near Real Time 개인화 추천
추천 모델의 최신 피쳐 반영한 지속적 모델 버저닝



AI를 도입하고자 하는 다양한 기업들에게 maum.ai 플랫폼을 기반으로,
AI Transformation을 컨설팅을 통해서 maum.ai 시스템과 API를 제공하여 수익 창출



인공지능 관련 정부과제를 약 25건 수행하였음.
대한민국 정부의 인공지능 산업 진흥정책에 적극적으로 부응



과학기술정보통신부



산업통상자원부



NIA 한국정보화진흥원



대한민국해군
REPUBLIC OF KOREA NAVY



대구광역시
DAEGU METROPOLITAN CITY

ex 한국도로공사



Incheon Airport



한국저작권보호원
KOREA COPYRIGHT PROTECTION AGENCY



건강보험심사평가원
HEALTH INSURANCE REVIEW & ASSESSMENT SERVICE



강원창조경제혁신센터
GANGWON-CHUNGCHONG ECONOMIC REVITALIZATION CENTER



KLID 한국지역정보개발원
KOREA LOCAL INFORMATION RESEARCH & DEVELOPMENT INSTITUTE

대법원

2023년 마인즈랩의 Vision



마음오케스트라를 기반으로 한 세계 최초 AI플랫폼 성공적 구축을 새로운 출발점으로 하여



인공지능 플랫폼 회사로 우뚝서서, 유니콘을 향해 전진해 나가겠습니다.

사용자

