

# Company Introduction

Seamless AI Conversations – From Calls to Robots

(주)인터로이드

Copyright© 2025 InterRoid. All Rights Reserved  
Company Introduction Presentation  
[www.interroid.com](http://www.interroid.com)

Apr. 2025



# Agenda

---

What is InterRoid ?

1. Overview
2. Our Business
3. Solution & Technology



**Revolutionizing Voice Call Automation  
with state-of-the-art AI technology**

Transforming Conversations with AI-Driven Voice Automation,  
Seamless AI-Powered Voice & Chat Solutions

# 01

## 인공지능(A.I)으로 혁신과 변화를 만드는 회사

---

## Overview

**InterRoid**는 빅데이터, 인공지능, 분석 플랫폼 및  
챗봇 분야의 전문성을 활용하여 다양한  
산업분야에서 AI기술 기반 혁신과 변화를 이끌어 냅니다.

지난 7년간

AI, 디지털 영역에서  
100개 고객사와 215개 이상의 성공적인 프로젝트 완수를 통해  
풍부한 경험치가 축적된 기업입니다.

업력

7년

고객사

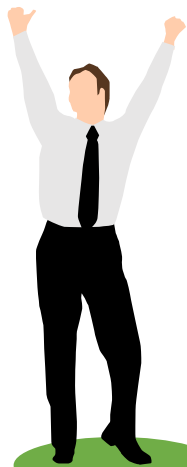
104

수행 프로젝트

215

누적매출(억)

406



## Our Journey for advanced AI Technologies

2018 ~

2020

### MLOps

- 머신러닝 분석화 자동화 솔루션 'Smart AutoML' 출시
- 세라믹 산업 빅데이터 `AI기반 세라믹 제조혁신 플랫폼 'Cerabig' 개발
- AI이미지 분석 및 OCR 솔루션 'Smart Vision' 개발
- 온라인 쇼핑 서비스를 위한 AI 실착 이미지 카탈로그 API 서비스 개발
- AI전문기업 '에이아이더'와 Bio전문기업 '뉴트리진' 합병



2021

### AI Vision

- 핵의학 영상 이미지 데이터(EDG PT-CT) 구축 및 응용서비스 개발
- AI골프 스윙영상 분석 앱 "나인포즈(9Poses)" 출시
- 〈음성단위 운율을 반영한 음성 합성 방법 및 장치〉 특허출원 획득



2022

### AI Voice

- AI기반 통화음성 특징분석 서비스 '굿 보이스(Good Voice)' 출시
- 감정분석 가능한 음성복제 STT엔진, TA분석 솔루션 'SPEECH-to-i' 출시
- 정보통신산업진흥원 SW품질 우수상 수상 (12.07.)
- 에이아이더뉴트리진 코넥스(311960) 상장 (11.29.)
- 글로벌 분석플랫폼 '데이터이쿠(Dataiku)'와 AI 플랫폼 사업제휴



2023

### LLM

- GPT기반 고객 의도분류
- GPT 적용 마이데이터 서비스 고도화
- LLM기반 Text Analytics 솔루션 AICC
- 지오코딩 솔루션 X-The 솔루션 개발



2024

### AI ChatBot

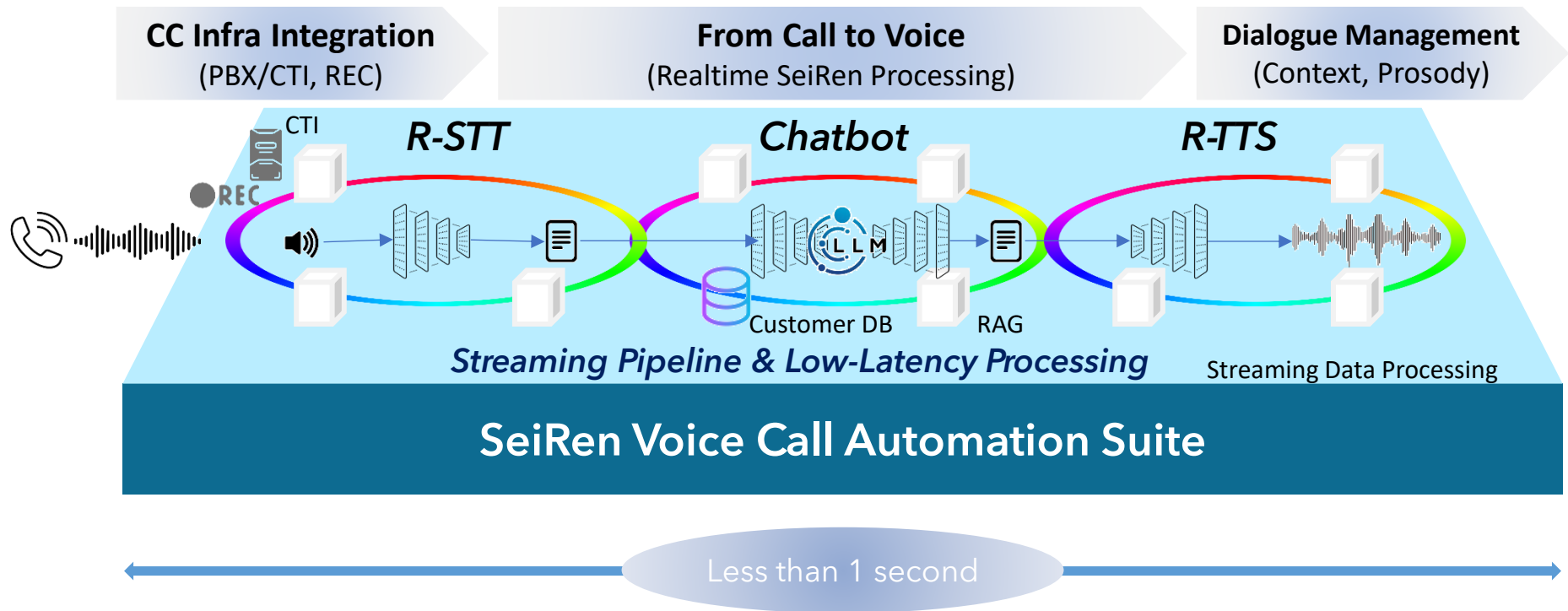
- LLM Platform, "AIThe LLMOps" 출시
- 생성형 Text2SQL 솔루션 개발
- AI Chatbot 서비스 SeiRen 개발



## Focus

## Seamless AI Conversations – From Calls to Robots

## Intelligent Voice Call Automation



# 02

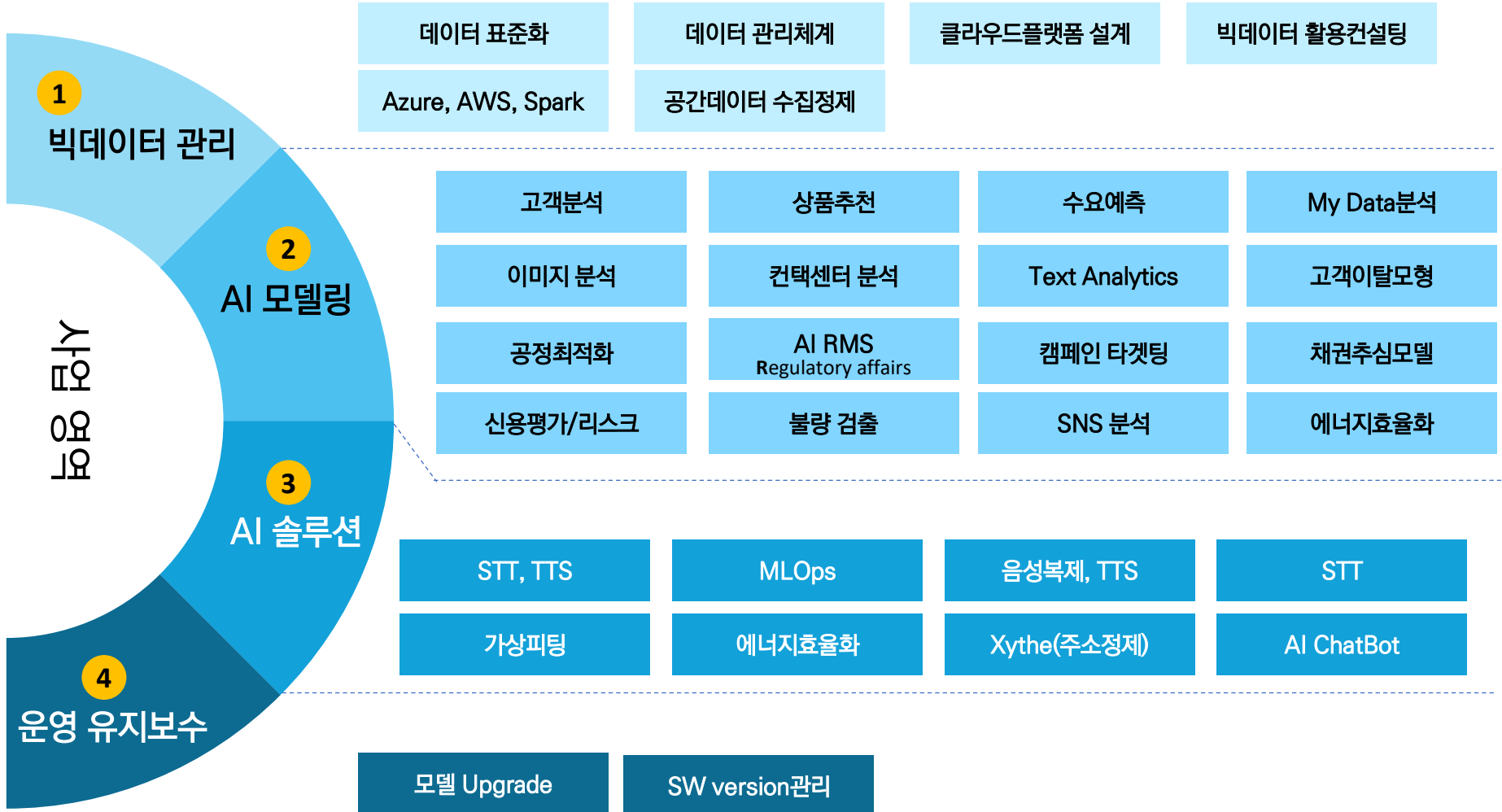
## 고객가치를 극대화할 수 있는 AI 기반 사업수행

---

## Our Business

InterRoid는 LLM, MLOps, Contact Center AI 분야에 자체 개발한 플랫폼을 바탕으로 고객의 니즈에 맞는 운영환경을 구축해 드립니다.

### Business Domain





## □ Key Customers

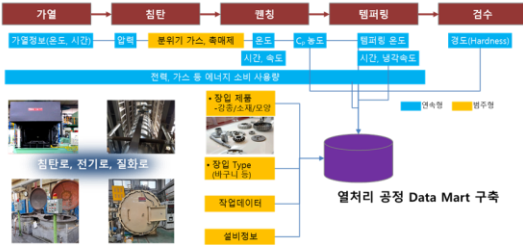
금융(Finance)	제조(Manufacturer)	유통/서비스(Dist. & Service)	공공(Public)
 KB 국민은행  우리금융그룹  AIA 생명  MERITZ 메리츠화재  KB 손해보험  KB 생명보험  라이나생명 <small>A Chubb Company</small>  삼성카드  비씨카드  Hyundai Capital  NH투자증권	 SAMSUNG 삼성전자  POSCO  HYUNDAI  LG화학  DOOSAN  SK chemicals  HK Kolmar  원일금속(주) WIM® <small>WONIL METAL CO., LTD.</small>  BUSUNG	 KEPCO  kt  SAMSUNG 삼성물산 패션부문  ENLIPLE  S-OIL 에쓰-오일  ORION  MetaM  UBIVELOX <small>유비벨록스주식회사</small>  성균관대학교 <small>SUNG KYUNKWAN UNIVERSITY(SKKU)</small>	 ETRI  보건복지부  서울특별시 <small>SEOUL METROPOLITAN GOVERNMENT</small>  NIA 한국정보화진흥원  IITP <small>정보통신기획평가원</small>  LH 한국토지주택공사  한국산업단지공단 <small>Korea Industrial Complex Corporation</small>  KAC 한국공항공사  국토안전관리원

# Manufacturing Optimization, Optimal Stock Management, Quality Control

## 공정 최적화-1

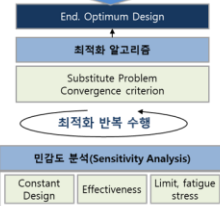
- 제조공정 최적조건(Gold Recipe)을 탐색하기 위해, ML 알고리즘을 통한 공정 분석
- 다양한 공정조건(Mix/Treater 재료, 온도, 압력, 프레스칭 시간 등) Input에 따른 품질 예측모형, 제조공정의 최적화 모형

### 열처리공정 및 수집 데이터



$$\text{Minimize } f(x) = \text{Temp\_열처리온도} + \text{EXP\_Time\_소송시간}$$

Subject to(공정 제약 조건) 경도(H) ≥ 50 HRC  
총 공정시간 ≤ 35 ...



### EDA & 열처리 공정 인자분석(Factor analysis)

- 공통 데이터 EDA(Exploratory data analysis) 분석 진행
- 이차원 이상을 공차치 처리
- 공통 변수간 연관성 분석

### M/L, D/L 예측 모델링

- 연속형 기반의 학습 및 분석 모델링 구축
- 크기 및 범위에 따라 데이터 종류에 영향을 주는 주요 요인 도출

### 핵심 공정 피쳐(Features) 도출



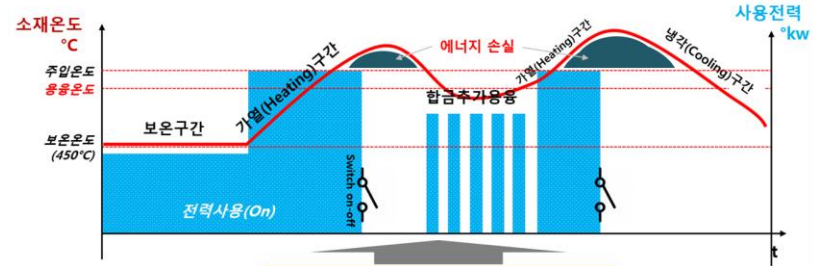
### ML/DL 예측 모델링



## AI에너지 효율화

- 예측/시뮬레이션 기반 최적 공정제어를 통한 에너지 손실 감소 알고리즘 및 솔루션 (주조, 열처리 등 에너지 다소비 기업 8개 기업 적용)

### 전기로내 소재 온도의 Heat Pattern과 소비 전력



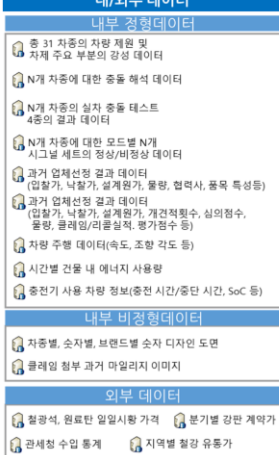
제품/원재료 구성	가열시간 (전력사용시간)	가열로 출력 (사용전력)	목표(용용) 온도	온도 손실
소재의 구성비, 부피 등에 따라 가열 시간 (에너지 소모)에 영향	완전 용용 온도까지 도달하기 위해 에너지 소모에 영향	출력이 높을 수록 에너지 소비가 많으며, 목표 온도 도달시간영향	가열 시간에 영향을 주기 때문에 에너지 소모에 영향	복사, 대류 등 소재 가열에 사용되지 않는 손실 에너지량

에너지 소비 절감을 위해 각 에너지 소비 요소의 최적 값 산출

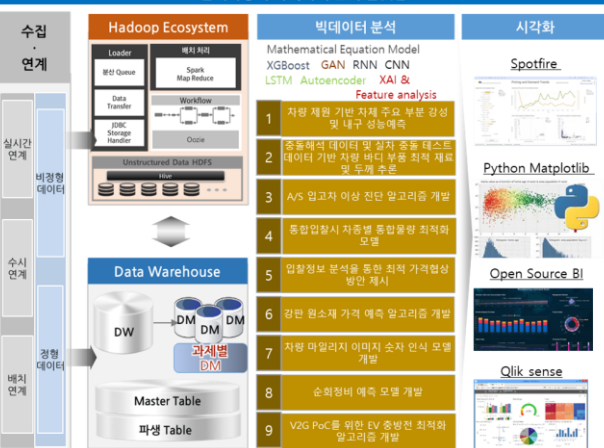
## 공정 최적화-2

- AI Smart Factory 아웃소싱 : 제조업 AI, 분석업무 아웃 테스킹(Out-tasking) 사업

### 내/외부 데이터



### 현대자동차 빅데이터 분석 플랫폼



## 재고관리 최적화 (수요예측)

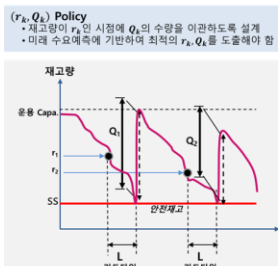
- 수요예측(Demand-forecast) 기반 최적 재고관리 시스템

### 이관/재고 최적화 4 Key Components

#### 수요량 예측(Demand Forecast) 기반

- 이관 시점 (Movement Point, ROP) 결정
- 이관량 (Quantity) 결정
- 안전재고 (Safety Stock) 관리수준 결정
- 운용 Capacity 관리수준 (Level) 결정

#### Demand Forecast-based Dynamic $(r_k, Q_k)$ Policy



#### 최적 이관/재고 운영 방안

- 최적 운용 전략 설계
- $F_k$ : 수요예측 모델 기반 k시점의 수요 예측치
- $L$ : 리드타임
- $z$ : k 시점의 재고
- CSL: Cyclic Service Level, stock-out 방지를 위한 주기
- $\phi$ : 표준 정규분포 누적 확률 함수
- CFU<sub>k</sub>: 예측 k기간의 수요의 불확실성(분산) (Cumulative Forecast Uncertainty)

$$r_k = \sum_{j=1}^{L+k} F_{k+j-1} + \phi^{-1}(CSL)\sigma_{CFU_{L+k}}$$

$$Q_k = \max\left[\sum_{j=1}^{L+N_k} F_{k+j-1} + \phi^{-1}(CSL)\sigma_{CFU_{L+N_k}} - Lr_{-1}, 0\right]$$

#### $\alpha$ 레벨의 안전재고 산출

- $\alpha$ : 오차한계 (예: 0.5%, 1.0%)
  - Z: safety Factor
  - $\alpha$ 별 수준
  - 예: 정규분포  $\alpha = 95, Z = 1.65$
  - $\sum_{j=1}^k D_j$  분포가 정규분포가 아닐 경우 Empirical Simulation
  - 최종 안전재고 수준 ( $\alpha$  레벨)
- $$\sum_{j=1}^k D_j + z \times \sqrt{\text{var}(\sum D)}$$

# AI Image Processing, Video Analytics, Motion Detection, Similar Object Detection

## 이미지 분석-1

다양한 산업 현장에서 필요한 이미지 분류, 객체인식, 정보추출(OCR)을 위한 이미지 처리 기술

**Hyundai Capital** | -2018~2024기간 동안 솔루션 공급 및 신차/차종 증가에 대한 유지보수, SW 업그레이드

### AI 기반 사진 판독을 통한 차량 정보 자동화 처리



## 이미지 분석-2

골프 스윙 자동 진단(Video Analytics, Pose Detection)

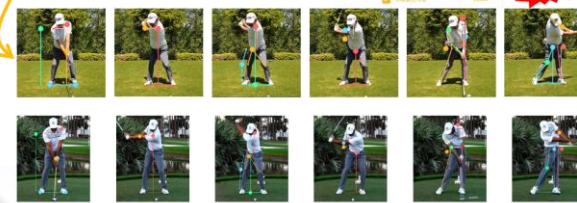


### 나인포즈 앱 출시

#### Video Analytics Tech.

- Video Pre-processing
- Pose Detection
- Swing Analysis

#### 1 내 스윙동작 자동 분석 & 점수



#### 2 타이거우즈의 스윙과 자동비교분석



## 이미지 분석-3

다양한 산업 현장에서 필요한 이미지 분류, 객체인식, 정보추출(OCR)을 위한 이미지 처리 기술

### AI 기반 사진 판독을 업무처리 자동화

#### 1 발효식품 Aging Stage 자동인식

발효식품의 발효 단계 및 품질 예측 제공 프로그램 개발

- 발효 단계와 발효 품질에 대한 예측 정보를 제공
- 적절한 유통 및 섭취 시기, 저장 기간 등의 정보를 제공
- 해당 프로그램은 발효식품의 이미지 및 입력 값에 따라 즉연에 AI 예측 결과를 간략히 표시



#### 2 한식 메뉴인식 및 나트륨 섭취 측정



#### 3 SK케미컬, 화장품용기 자동인식

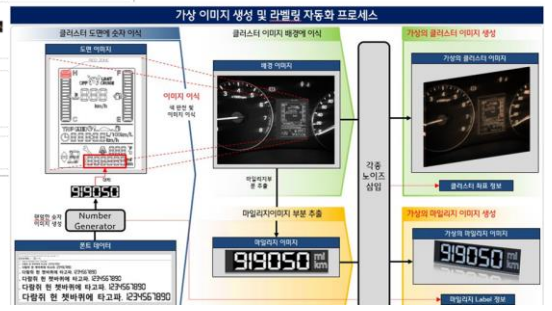


## 이미지 분석-4

제품 점유율 확인, 매대 정보 DB화, 차량 계기판 마일리지 숫자 인식(OCR)

### AI 기반 매대인식 AI 모델 자동화

#### 1 매대 내 제품 위치 탐지 2 각 제품별 이미지 크롭 3 제품 카테고리 분류





# Customer Analytics, Product Recommendation

## 고객분석

▪ Cross/Up-sell을 위한 고객 프로파일 기반 전략/전술 세그먼트를 개발 모델

### ML/DL기반 고객세분화 모델링

#### 1 군집화(Clustering) 및 분류(Classification) 적용 고객 세분화

① 기존고객 : 군집화/SOM 적용 고객 세분화

- K-means Clustering 적용
- 최적의 세그먼트 그룹 수를 결정
  - Elbow method, Silhouette method, Gap statistic 이용
- PCA, Auto-encoder 방법에 의한 세분화 Visualization
  - 군집 시각화를 위한 차원축소 적용(PCA, t-SNE, Auto-Encoder)
- SOM(Self-Organizing Map)
  - 인공신경망 SOM 기법을 활용한 기존 군집화 적용

② 신규고객 세분화 : 분류(Classification) 적용

- K-means에 의한 기존 고객 세분화 모델 수립
- 신규고객 가입시, 그룹 생성
  - K-Nearest Neighbors, Decision Tree, Random Forest, Support Vector Machines 적용 예정

#### 2 Joint Embedding 기법 적용 고객 세분화

- 고객 변수 중, 텍스트(Text) 변수 포함
- 텍스트형 변수의 경우 Word Embedding 을 통한 Vectorization 변환

Numeric, multinomial or text-string 변수 모두 포함 모델링 적용

수치형, 범주형 텍스트형

- 적용모델
  - DEC(Unsupervised Deep Embedding for Clustering Analysis)
  - Client2vec model(Word2vec 또는 mSDA를 통한 임베딩 기법)
  - K-means 분류 알고리즘 적용
  - 상기 Client2vec Joint embedding 모형을 고객세분화 그룹핑을 위해 K-means 분류 적용

## AI 캠페인 타겟팅

▪ 정교한 타겟팅 알고리즘 기반의 캠페인 자동화 솔루션 (운영 노하우 사례 : 온/오프라인 : 약 600개 캠페인 수행, 15종의 캠페인 유형)

원천데이터: Online DW, SAS, etc.

AI Customer Target Platform: Target Extraction, Simulation Module, Prediction Module, Target AI Algorithm, Feature Extraction, Analysis Module, Visualization

AI 기반 Targeting Algorithm: XGBoost Model, Gradient Boosting Model, Random Forest, Rule Association

#### 캠페인 시뮬레이션 및 예약 스케줄 관리

#### 캠페인 실적 및 반응 정보

#### AI Feature 설정 및 관리

## 고객이탈 모형

▪ 고객이탈 모형에 따른 이탈 스코어 산출 및 원인파악, 원인별 이탈방어 전략 설계

### 고객이탈 모형 및 대응전략

#### 이탈 모형

- 개인적 상황 (Personal Circumstances)
- 보험상품/계약 특성 (Product & Policy)
- 고객서비스 경험 (Customer Experience)
- 경쟁사 이탈 (Competitor Churn)
- 시장/외부환경 (Market and Externals)

#### 이탈 스코어 모델

#### 이탈 원인 해석/대응

이탈 원인 파악용 기법 (개인 고객별)

- SHAP (Shapley Additive Explanations)
- LIME (Local Interpretable Model-agnostic Explanations)
- 트리 기반 알고리즘의 Feature Importance
- 요인 분해 (Factor Decomposition)

## 상품추천

▪ 기존 Cosine Similarity 기반 CF 알고리즘보다 성능이 우수한 행렬분해법/ML/DL 알고리즘 이용

### 추천 알고리즘 : 기존 CF → Matrix Factorization

#### 추천 시스템(CF 협업필터링)

##### User-based filtering

##### Item-based filtering

Cosine Similarity의 한계

#### Factorization Machine

Output: Factorization Machine

$$y^{hnt}(x) = w_0 + \sum_{i=1}^n w_i x_i + \sum_{i=1}^n \sum_{j=1}^n \langle v_i, v_j \rangle x_i x_j$$

Dense Embedding, Sparse Features

유저정보 | 아이템정보 | 고객 반응정보 | 고객 선호정보

---

#### Deep Component

Output: Deep Component

Hidden layer, Dense Embedding, Sparse Features

유저정보 | 아이템정보 | 고객 반응정보 | 고객 선호정보



# 03


## 기업 디지털 혁신을 구현하는 AI Edge- Solutions

---

## Solutions & Technology

InterRoid는 ML/DL분석 자동화, Image Processing, STT 및 음성복제 등 다양한 솔루션과 최근 주목받는 AI 챗봇 솔루션을 제공하여 고객사의 기업가치 향상에 노력하고 있습니다.

# AI Solutions




**Smart AutoML**

ML 모델개발  
No Coding 플랫폼




**Smart Vision**

이미지 전처리,  
OCR, 객체인식 등



**Virtual Fitting**


의류대상  
가상피팅 솔루션



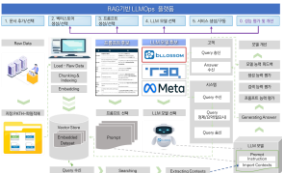
**SeiRen Suite**

<b>R-STT</b>	<b>Chatbot</b>	<b>R-TTS</b>	<b>Good Voice</b>
Real-time STT엔진 (Speech-to-Text)	LLM기반 챗봇 개발 Builder	개인화 음성복제 솔루션	음성 품질 측정 /분석

**LLMOps** 초거대 언어모델(LLM)의 환각답변, 사내 정보유출 방지를 위한 RAG기반 LLMOps



문서 전처리 자동화	LLMs 선택사용
Flexible 임베딩	LLM 성능모니터링
파인튜닝	In-context Learning
가드레일	사용자 Feedback 반영



**ChatBot Builder** No-code 기반 챗봇 개발, 업무 Flow 적용



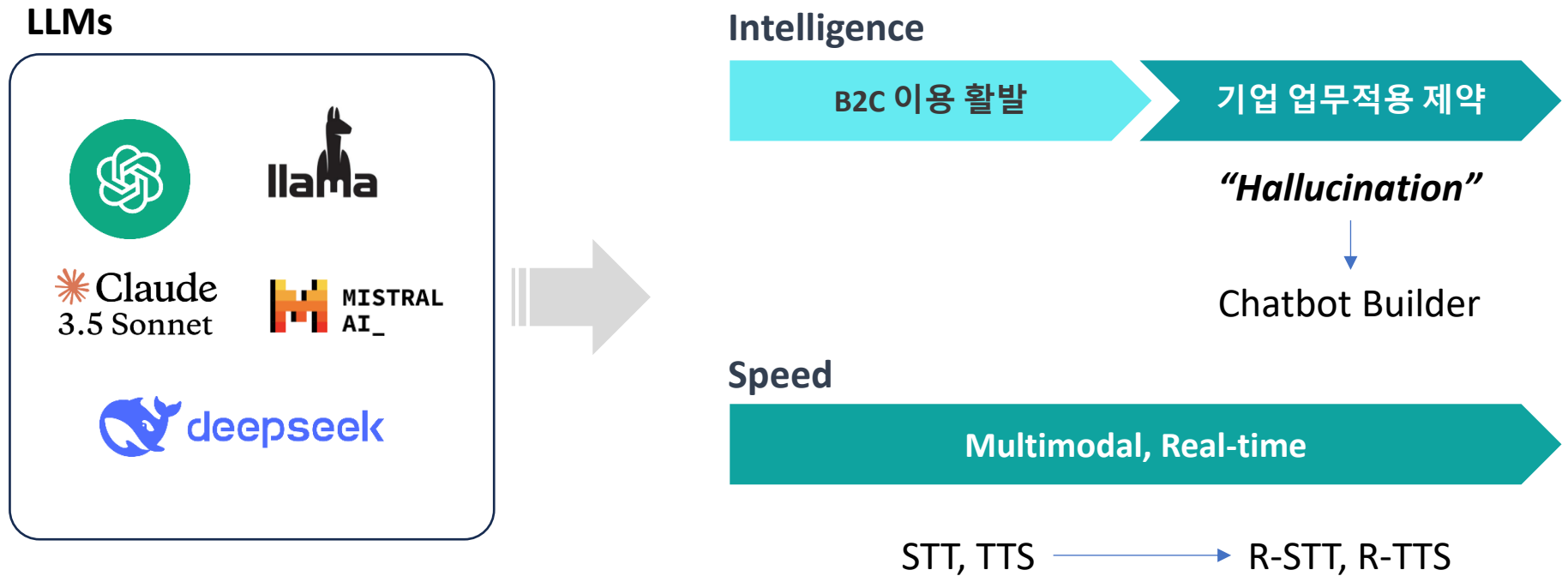

자체 Cloud 플랫폼 기술력 (OpenStack)

Software 개발 역량 (NIPASW품질 우수상 수상)

Agile AI Engineering

다양한 산업 도메인 전문성

# The Emergence and Impact of LLMs





# 할루시네이션(Hallucination) 현상

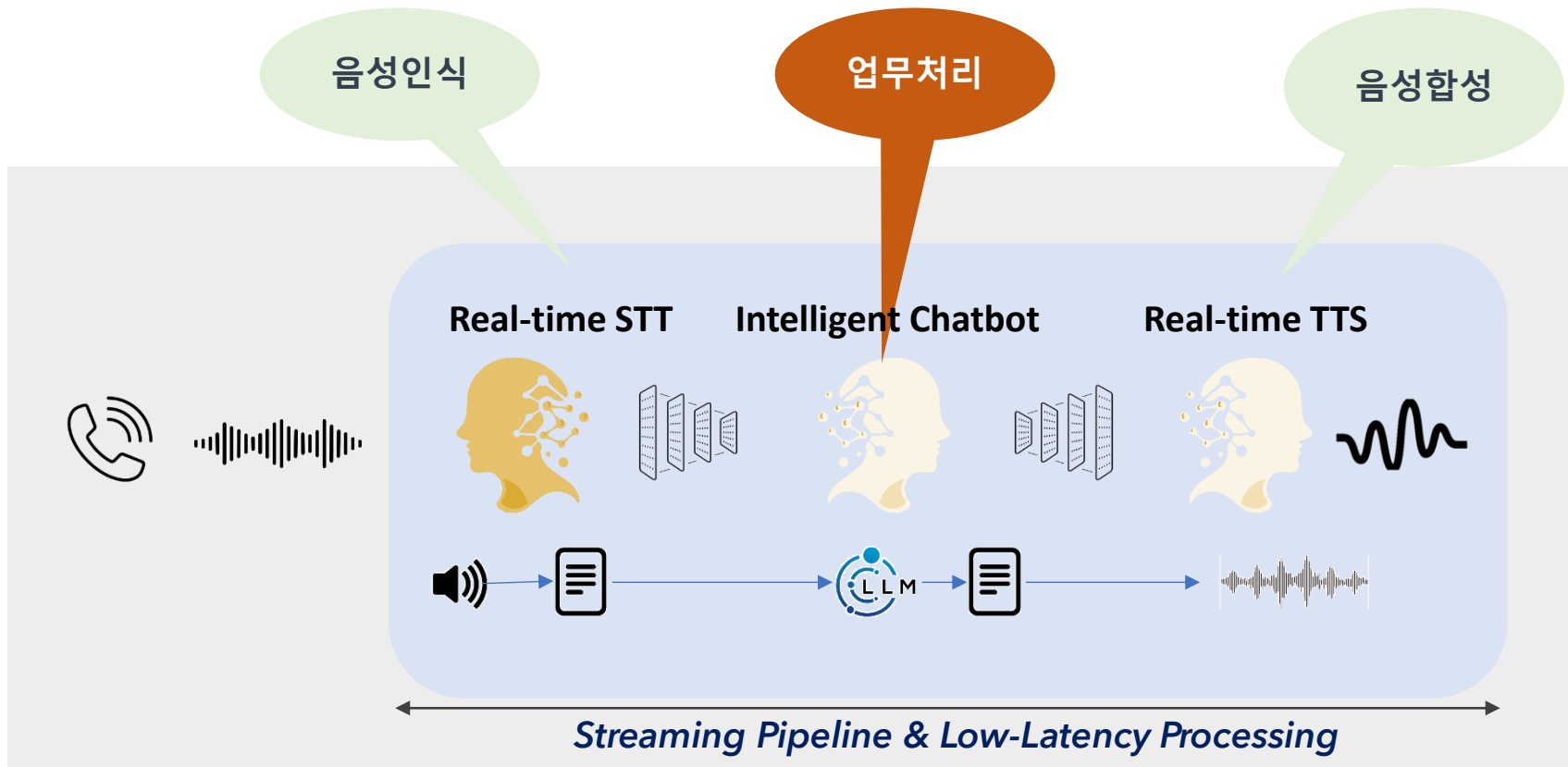
- 챗봇 확산에 가장 큰 걸림돌(Bottleneck)은 챗봇의 오답현상

**정의** 할루시네이션은 AI 챗봇(특히 LLM 기반 챗봇)이 사실과 다른 정보나 문맥적으로 부정확한 내용을 사실처럼 생성하는 현상을 의미함

특징	원인	
<ul style="list-style-type: none"> <li>① 사실처럼 보이지만 검증되지 않은 정보                             <ul style="list-style-type: none"> <li>- AI가 존재하지 않는 정보, 출처 불명 자료를 신뢰할 수 있는 사실처럼 제공</li> </ul> </li> <li>② 고객 질문과 연관되지만 부정확함                             <ul style="list-style-type: none"> <li>- 질문의 맥락을 고려하지만 실제 데이터베이스(DB) 또는 RAG 시스템과 불일치한 답변을 생성</li> </ul> </li> <li>③ 기존 문서 또는 데이터셋에 없는 정보 생성                             <ul style="list-style-type: none"> <li>- 챗봇이 참고할 수 있는 문서, DB, RAG 시스템 내에 없는 정보를 임의로 조합하여 새로운 내용을 생성</li> </ul> </li> <li>④ 상황에 따라 일관성 부족                             <ul style="list-style-type: none"> <li>- 같은 질문이라도 대화 맥락에 따라 다른 답변을 내놓을 수 있음</li> </ul> </li> </ul>	<ul style="list-style-type: none"> <li>• <b>LLM의 확률적 생성 방식</b> 대규모 언어 모델(LLM)은 텍스트를 예측하는 확률 모델이므로, 정답을 "기억"하는 것이 아니라 "패턴"을 학습하여 자동 생성  데이터셋에 없는 질문이 들어오면 기존 학습된 패턴을 조합해 답변 생성, 검증 과정 없음</li> <li>• <b>사전학습 데이터의 한계</b> 최신 데이터나 특정 기업내부 정보 등 도메인(DB 등)에 접근하지 못함</li> <li>• <b>고객의 모호하거나 불명확한 질문</b> 질문 불명확, 문맥부족하면, AI가 논리적 추론을 통해 틀린 정보를 보완하려는 경향</li> <li>• <b>훈련 데이터의 편향 및 오류</b></li> </ul>	<p><u>해결노력</u></p> <ul style="list-style-type: none"> <li>RAG 시스템</li> <li>프롬프트 엔지니어링</li> <li>사전 필터링 및 검증 시스템 도입</li> <li>Fact Checking 알고리즘 적용</li> </ul>

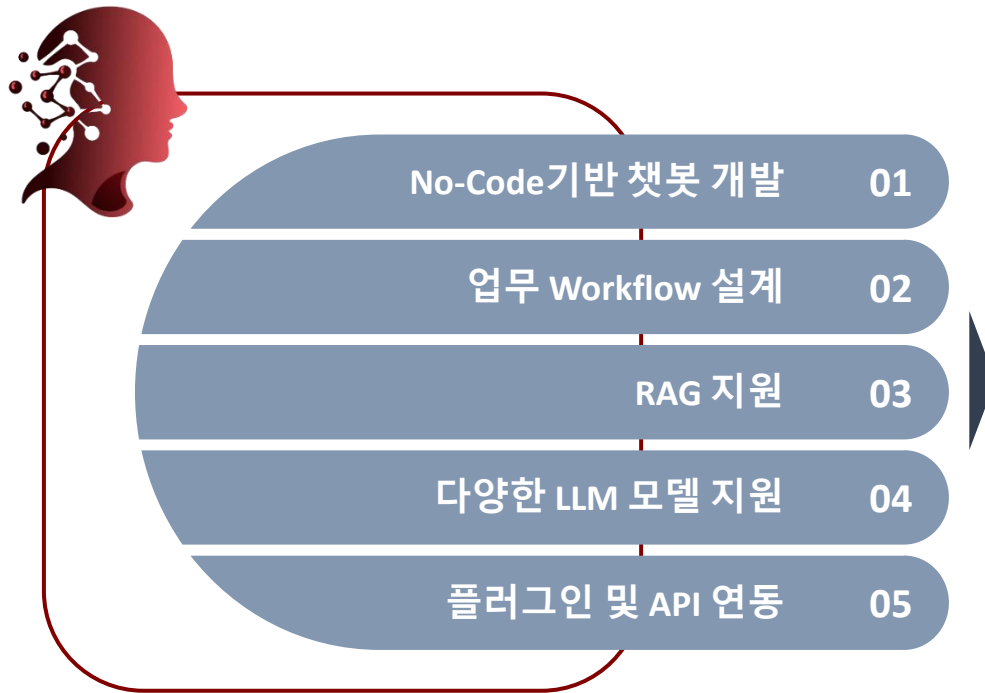
## SeiRen Suite Platform

- 음성인식+음성합성보다, 업무처리를 위한 LLM기반 챗봇의 성능이 핵심



## Seiren Chatbot builder

- 강력한 업무프로세스 설계(No-Code), 업무정보 RAG연동, GPT/로컬모델 연동, 업무용 DB



- **Langchain 등 Python 코딩없이 챗봇 개발**
  - 현업 사용자의 챗봇 관리 가능
- **업무 카테고리/단위업무 등 개별 로직 적용**
  - 업무단위별 처리/가이드 적용(할루시네이션 방지)
- **다양한 사내/업무지침 정보 적용**
  - 텍스트(PDF, MS-Word, HWP, TXT, Markdown, HTML, JSONL)
  - 구조화 문서(CSV, Excel)
- **범용 LLM 모델 지원**
  - OpenAI, Llama, CLaude, Deep seek
- **외부 연동 지원**
  - API기반 외부시스템 연결
  - 백엔드 서버 경유, DB 조회, API 형태로 제공
  - Oracle, MS-SQL, PostgreSQL, MySQL, MongoDB 등



## • Chatbot builder : Key Architectures

- 5가지 핵심 기능의 결합으로 비즈니스 챗봇 기능 완벽 구현, NLP+LLM(GPT) + RAG + Dialogue Management 하이브리드 모델

핵심 모듈	모듈별 세부 기능
NLU 모듈	<ul style="list-style-type: none"> <li>• 고객의 입력을 의도(Intent) 및 개체(Entity) 분류하여 처리</li> <li>• LLM 기반으로 가능 또는 NLP 모델(BERT)에 의한 처리 선택 가능</li> </ul>
RAG기반 지식검색	<ul style="list-style-type: none"> <li>• 문서(요금정책, 환불정책 현장점검 등)를 인덱싱하여 정확한 정보 검색</li> <li>• 검색된 정보를 바탕으로 LLM(GPT)가 환각 없이 응답하도록 유도</li> </ul>
대화관리	<ul style="list-style-type: none"> <li>• 대화 흐름(FSM, 스토리 기반 등) 관리</li> <li>• 멀티턴 대화 시 문맥 유지 및 필요한 정보 요청</li> </ul>
응답생성	<ul style="list-style-type: none"> <li>• RAG 검색 결과 기반 응답</li> <li>• LLM(GPT) 이용 답변 텍스트 생성 기능 이용</li> </ul>
비즈니스로직/ DB연동	<ul style="list-style-type: none"> <li>• 요금조회, VOC접수, 처리결과, 예약, 주문, 배달 처리 등 비즈니스 로직 구현</li> <li>• 데이터베이스(SQL, NoSQL) 연동 지원</li> </ul>



## Seiren의 Anti-Hallucination(할루시네이션 방지)

- Seiren Chatbot은 정확한 답변을 유지하기 위해 안전장치(Safe-guard) 설정이 유연함

### 의도 분류 기반 질문 세분화

- 질문을 명확하게 구분하여 LLM이 모호한 추론을 덜 하게 됨
- 각 의도별 프롬프트를 따로 설정 가능 → 과적합 방지
  - 요금, 검침, 안전사고 등 주요 카테고리별로 개별 프롬프트 적용

### Rule기반 Fact-check 알고리즘

- 요금, 사용량, 검침 내용에서 일정값 이상 체크 → 범위 초과시 상담원 연결
- 발생확률이 맹 낮은 결과값(critical failure event) IF문에 의해 필터링
  - 상담원 연결로 안전한 고객상담 유지

### RAG 기반 검색 수행 후 답변

- 문서 기반 검색(RAG) 적용 + 프롬프트 강조(오답 방지)
- 검색 문서에 있는 정보만 활용하도록 유도하여 맥락 이탈 추론 감소
- 사용자 질문 → 의도 분류 → RAG 검색 (요금, 가격표 등) → 문서 검색 후 LLM 문장 생성 → Fact-checking 룰 적용 (숫자, 키워드 일치 확인)

# Chatbot builder 구성



**신규 프로젝트**

- 삼천리도시가스 챗봇 soohwa**  
챗봇
- 광양분기 김주원**  
챗봇
- 추론 김주원**  
에이전트

**신규 자료**

- 가스요금표**  
1 자료 4 k 단어 2 연결된 업  
가스요금표.xlsx 자료 관련 챗봇 및 에이전트 생성시 사용
- 태디노트-이경록-맹채인밋업2024...**  
2 자료 10 k 단어 2 연결된 업  
useful for when you want to answer queries about the 태디노트-이경록-맹채인밋업2024-Q2-발표자료-V05.pdf  
#태도, #o o
- 사규**  
1 자료 18 k 단어 4 연결된 업  
에이아이디뉴트리진의 사규입니다.  
#RAG태도
- AIthe사규(행정1k)**  
3 자료 6 k 단어 3 연결된 업  
useful for when you want to answer queries about the AN-007 자금관리 규정.pdf
- 01\_소장.xlsx...**  
1 자료 23 k 단어 2 연결된 업  
useful for when you want to answer queries about the 01\_소장.xlsx
- [2021-사업총괄본부-851] ...**  
1 자료 51 k 단어 2 연결된 업  
useful for when you want to answer queries about the [2021-사업총괄본부-851] 안전보건관리체계 자율점검표(필용).pdf

**기본API 사용자정의API 플로우 빌더**

- Wikipedia**  
위키피디아 검색을 수행하고 요약 및 킷 페이지 추출 API.  
소셜
- Github**  
온라인 소프트웨어 소스 코드 호스팅 서비스 API.  
도구
- YahooFinance**  
YahooFinance에서 최신 뉴스, 주식 시세, 관련 차트 연동 API 제공. Analytics, News, Ticker 3개 도구 제공.  
비즈니스, 금융
- DuckDuckGo**  
사용자 행동분석, 웹사이트 추적 방지 기능 제공 등 개인정보 보호 중심의 검색 엔진 API. AI Chat, News/Images/Video Search 등 6개 API제공  
검색
- ArXiv**  
다양한 연구 분야의 과학 논문과 기사가 보관된 방대한 저장소 접근 API.  
검색
- Stack Exchange**  
질문 및 많은 추천을 받은 순서대로 위에서부터 노출 영어로 된 내이버 지식인 스택 오버플로우 등 개발 관련 API. Questions과 Answers 두개 API ...  
검색, 도구

## LLM Chatbot 구축 사례

- 국내 다수의 기업에 챗봇 업무 적용 경험

대상 기업	Target	Project 상세
대교	KMS Chatbot	• 기업용 KMS를 활용하여 Azure 베이스의 내부용 GPT Chatbot 구축 진행
Wisely	FAQ Chatbot	• 고객 FAQ 대응을 위한 FAQ Chatbot 구축
Unicef	상담내용 요약	• GPT를 활용하여 상담내용 요약 및 정리 기능 구현
삼성전기	KMS Chatbot	• Small LLM(Vicuna 7b)와 GPT를 동시에 활용해 Dataiku 베이스의 내부 KMS용 Chatbot을 구축
유비벨록스	Data 전처리	• DB Table 구성을 위한 크롤링 데이터 전처리 사용
유비벨록스	검색 및 추천	• 상품 기능 조회 및 추천
카카오뱅크	KMS Chatbot	• 기업 KMS를 활용해 상담원 전용 GPT Chatbot 개발 진행
SK Shieldus	FAQ Chatbot	• FAQ를 활용한 Chatbot POC 진행
SK Shieldus	Text Analytics	• GPT를 활용한 TA 결과를 기존 PTL Model 기반의 TA 결과와 비교
Kolon FNC	Text Analytics	• GPT를 활용한 상담 TA/자동 QA/VOC 분석 진행
Kolon FNC	Chatbot	• FAQ/시나리오 대체/상품 검색 및 추천과 더불어 필요시 상담원 연결을 함께 지원하는 FNC Mall Chatbot 구축 진행 중
Metam	SLLM 기반 Hybrid Chatbot	• Small LLM 활용을 위한 Chatbot Framework 설계, Text Analytics 및 기능 모듈 개발



# Smart Realtime STT Engine

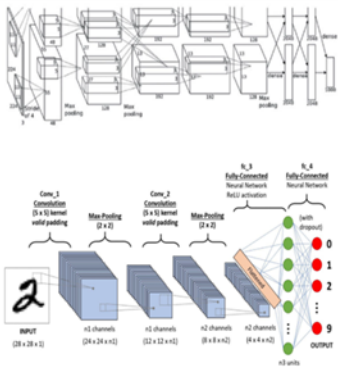
- 기존 Character 중심 Model이 아닌 CNN과 Transformer를 결합한 Conformer 구조, 사전 정렬되지 않은 음성과 문자열의 대응관계를 정렬하며 학습하는 CTC를 결합

## Conformer-CTC

### CNN

지역 정보(Local Feature)를 효과적으로 활용

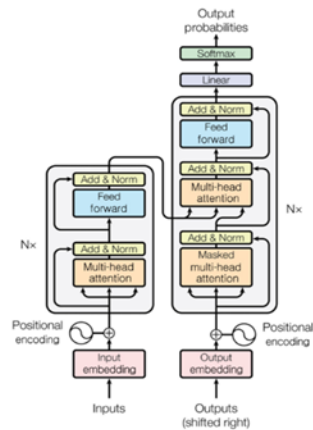
필터가 움직이면서 데이터의 지역적인 정보를 추출, 보존하는 형태로 학습이 이루어짐



### Transformer

콘텐츠 기반 광역정보(Global Interaction)를 잘 포착함

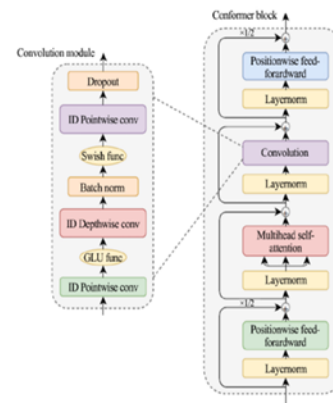
Self-attention에 기반의 트랜스포머 구조는 long distance interaction을 포착하는 능력과 high training efficiency로 시퀀스 모델링에 주로 사용됨



### Conformer

음성인식을 위한 Convolution-Augmented Transformer를 제안

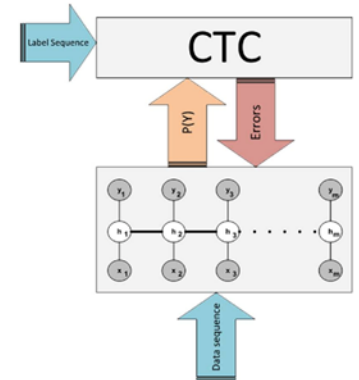
CNN과 self-attention을 유기적(organically)으로 결합하는 방법



### CTC

타겟 단어/음소 시퀀스 간에 명시적인 Alignment 정보 없이도 학습 가능

Unsegmented sequence data 학습을 위해 사용



## Smart Realtime STT Engine 특징점

- 경쟁사와 비교했을 때, AN Smart STT Engine은 다음의 네 가지 특징점이 있음

장점	특징
모델 학습시간	<ul style="list-style-type: none"> <li>타사 대비 더 많은 양의 데이터가 사전 학습되어 있음 (Saltlux: 약 10,000 시간, 당사: 14,000 시간 이상/25,000 시간 모델 학습 중)</li> <li>방언, 숫자, 단위, 외래어 등을 포함한 다양한 한국어 데이터를 사전에 학습</li> </ul>
모델 경량화	<ul style="list-style-type: none"> <li>모델의 경량화를 유지하면서 최적의 정확도 보장 (Customized 결과 약 95%의 Performance, 3 min Call data 기준 평균 처리 시간 3~4초) (Demo 측정 기준 경쟁사 STT보다 3.6배 빠른 STT 결과 도출)</li> </ul>
모델 차별성	<ul style="list-style-type: none"> <li>한국어에 최적화(한국어 어휘의 대부분을 포함하는 약 3000개의 Token 엄선 Model)</li> <li>Fine-Tuning을 통해 원하는 Domain/Data에 최적화된 Customized STT Engine 사용 가능</li> <li>STT Output 정확도를 위한 추가적인 Language Model 개발 및 결합</li> </ul>
추가제공 기능	<ul style="list-style-type: none"> <li>STT Labeling Tool 제공을 통한 Training Data Labeling Process 자동화, 지원 및 가속화</li> <li>STT에 필요한 화자 분리 Model 제공</li> </ul>

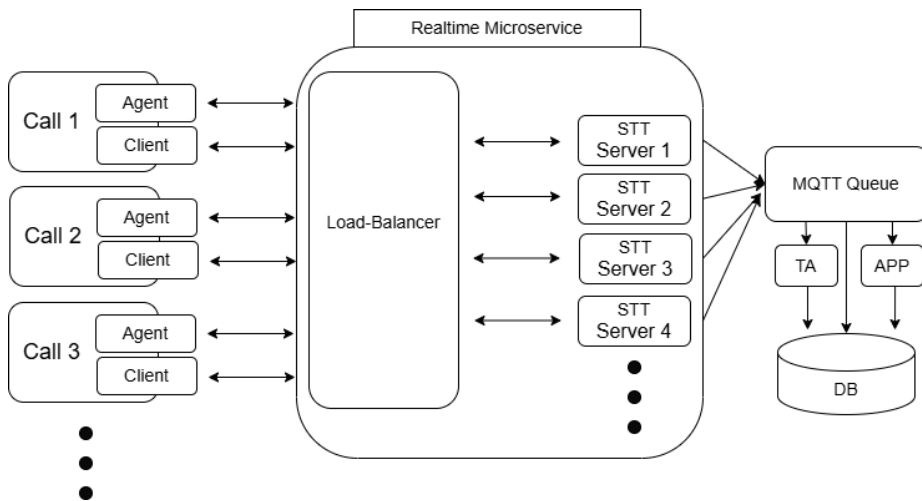
# STT Engine의 아키텍처

- 실시간 서비스 및 배치 처리 서비스 구현

## 실시간 서비스(Realtime Service)

- 콜, 채팅 데이터가 발생할때마다 실행되는 실시간 서비스
- IPCC 이벤트 발생 → 실시간 STT 서버: STT 결과 생성후 DB 저장
- Backend : Request → Realtime Microservice 서버에서 프로세스 처리
- UI 화면에 보여져야 할 TA(Text Analysis) 항목들을 수행후, 그외 항목들은 TA후 결과를 DB 저장

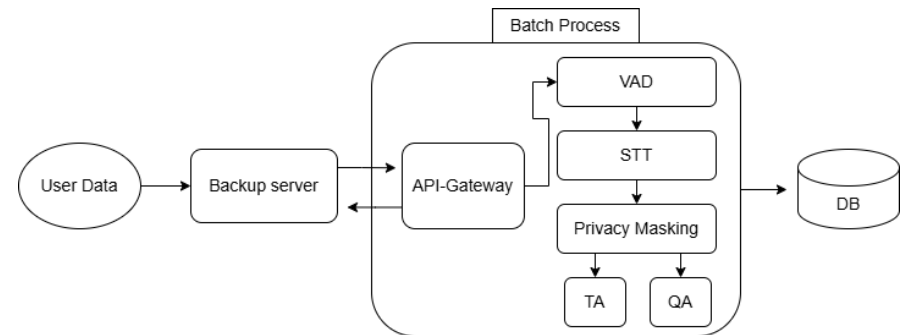
<Realtime STT workflow>



## 배치 서비스(Batch Service)

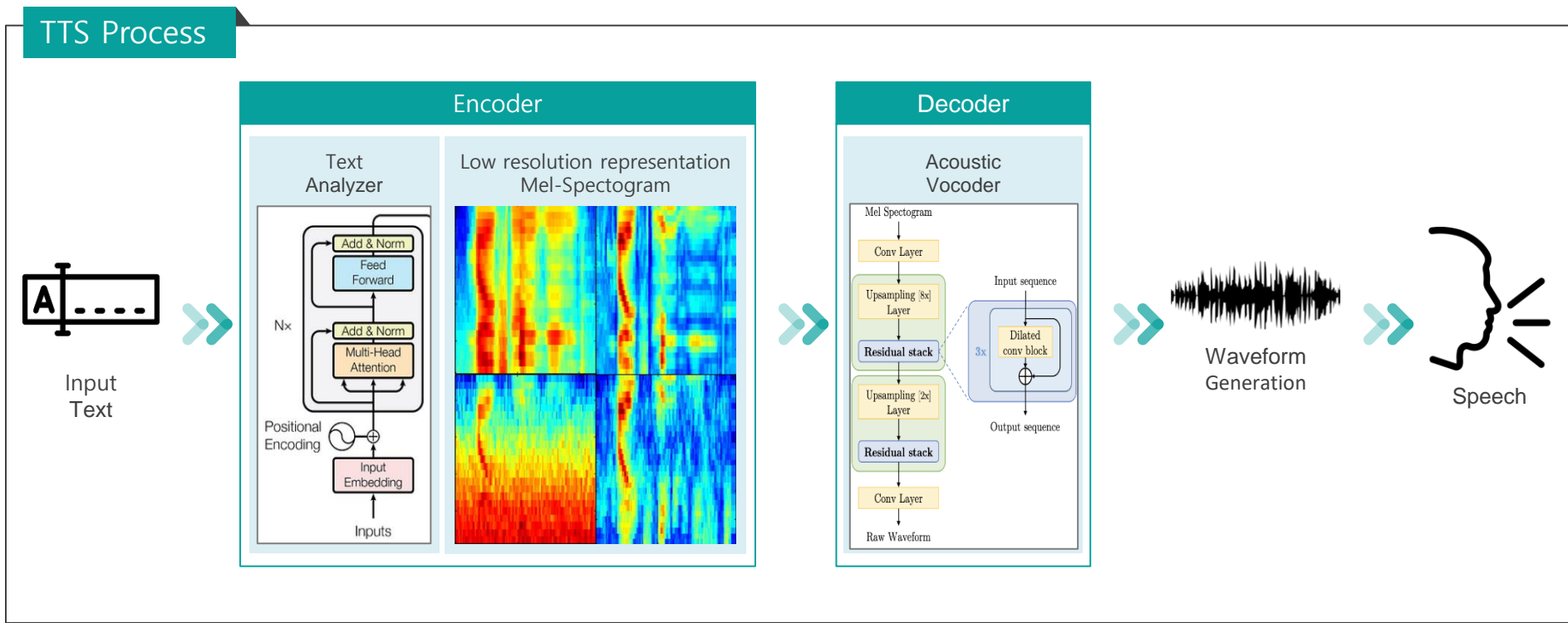
- 배치로 처리되는 기능을 포함하는 서비스
- 30분 단위로 VAD, STT, 마스킹 모듈이 함께 실행되고 결과가 DB에 저장
- Daily Batch로는 게시판 데이터에 대한 TA와 타겟 콜 데이터에 대한 QA 처리

<Batch STT workflow>



# TTS(Text To Speech)

- 음성 합성(Speech Synthesis), 입력한 Text를 AI Model로 음성을 변환하는 기능을 TTS(Text To Speech)라고 하며, 아래와 같은 절차를 통해 텍스트를 음성으로 변환하는 작업을 진행



## TTS Main Process

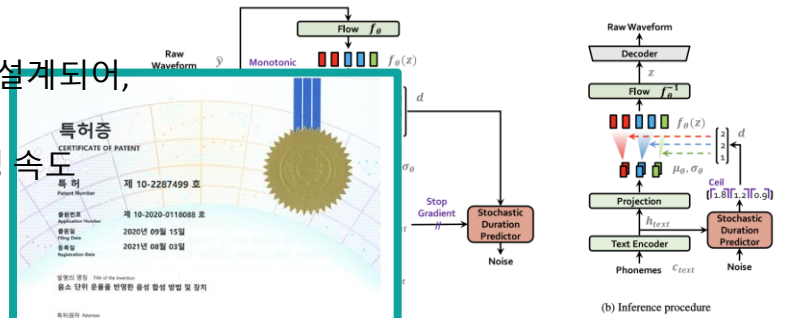
1. 입력한 Text에 대해 Encoding 단계에서 Text Analyzer를 활용해 전처리 및 해당 Text에 맞는 Melspectrogram을 추론
2. Encoding 단계에서 인식, 추론한 언어와 Melspectrogram을 Decoder에 입력하여 합성 음성(Speech) 형태로 출력

# TTS 알고리즘 설명

- 당사 TTS엔진은 MeloTTS 모델과 GPTSoVITS 모델 기반의 우수한 알고리즘으로 구현됨

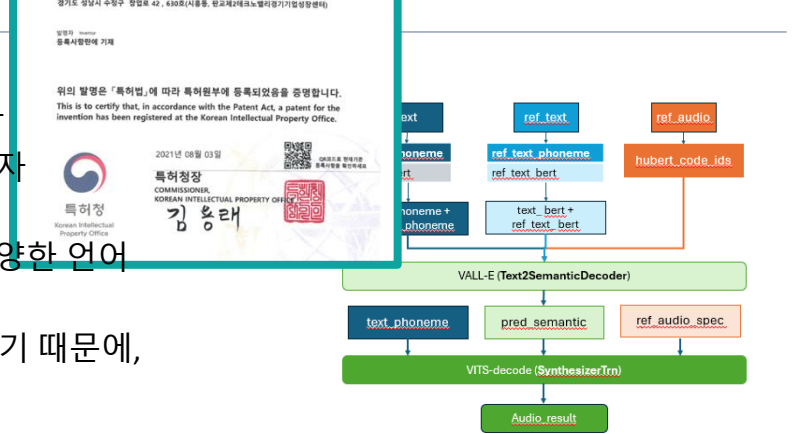
## MeloTTS 기반 모델

- Sequence-to-Sequence(Seq2Seq) 모델 + 신경망 보코더(Neural Vocoder) 조합 모델
- 자연스러움 강화: VITS 아키텍처[1]를 기반으로 설계되어, 더욱 유창하고 자연스러운 음성 생성 가능
- 효율성 최적화: 구조화된 처리 방식 덕분에 합성 속도 빠르면서 품질 유지
- 음소(Phoneme) 기반 처리: 언어 규칙이 명확한 한국어에서 발음 정확성이 우수함



## GPTSoVITS 기반 모델

- 대규모 생성형 모델 적용, 적응성과 표현력 강화
- 표현력 & 적응성: 소량의 데이터로도 다양한 화자 스타일에 적응 가능
- End-to-End 학습: 음소 변환 의존도 최소화 → 다양한 언어 구조에 유연하게 대응
- 실시간 음성합성: 고품질 오디오를 직접 생성하기 때문에, 대화형 AI 등 실시간 응대에 적합함



# AICC: 컨택센터 + AI기술

- 인공지능(AI)을 활용해 컨택센터의 업무 효율성과 고객 경험을 개선하는 시스템

시스템간 상호 연결성: STT(음성 텍스트화) → 감정 분석 → 상담원 실시간 어시스트 → 상담 후처리 자동화



## STT, TA 적용 사례

- 당사 STT, TTS 및 Text Analytics 적용 기업 사례



STI(STT+TA) Process

- KB신용정보 Call Data Fine-Tuning 및 STT 자동 전사
- TA(Keyword Extraction, 추출형 요약, 고객감정 긍정/부정 분류, Category 분류) 모델 개발, STT + TA 자동 파이프라인 구축



STT 고도화  
TTS 개발  
음성기반  
감정분석(SER)  
상담 Assistant Tool 개발  
VOC 고도화

- 기존 Smart STT(Pretrained 14000 hours) 고도화 작업(25000 hours)
- 생성형 기반 TTS 개발
- 음성 기반의 고객/상담사 응대 감정 분류 모델 개발
- SLLM 기반의 맞춤법 검사기 개발
- Elastic Search 기반의 KMS 검색 기능 개발
- 금지어/부정어 인식 및 관리 기능 개발
- SLLM/PTL 기반의 상담 요약 모델 개발, 자동 QA 모델 고도화 개발



TA Process

- SK Argos System 연동
- TA(상담 자동요약 3종, Category 분류) 모델 개발 및 LLM 결과 비교



STI(STT+TA) Process  
Chatbot

- Kolon FNC Call Data Fine-Tuning 및 STT 자동 전사(실시간 STT)
- LLM 적용을 위한 고객 개인정보 Masking(NER) 모델 개발
- TA(화자분리 요약, 키워드 추출, 고객감정 긍정/부정 분류, 브랜드 추출, 자동 QA) 모델 개발, STT/Chat/Board Data+TA 자동 파이프라인



STI(STT+TA) Process

- Unicef Call Data Fine-Tuning 및 STT 자동 전사
- TA(상담 자동 요약, 상담 긍/부정분류), STT+TA 파이프라인 구축



STI(STT+TA) Process  
FAQ Chatbot

- Wisely Call Data Fine-Tuning 및 STT 자동 전사
- TA(Keyword Extraction, Category 분류) 모델 개발, STT/Chat Data + TA 자동 파이프라인 구축
- 고객 FAQ 대응을 위한 GPT기반의 Chatbot 구축

## 고객 분석 및 상담업무 효율화 분석(1/2)

- 고객센터 운영에 필요한 다양한 분석 및 NLP+LLM 특화 모델링

분석항목	기술분류	내용 및 사용 알고리즘
고객 의도 분석	NLP/LLM	<ul style="list-style-type: none"> <li>• <b>Text 데이터를 활용하여 고객의 의도를 분석하는 분류 모델 개발</b> - KoBERT, KoELECTRA, GPT 기반 Zero-shot 분류, CNN, LSTM</li> </ul>
고객 감정 분석	NLP/LLM	<ul style="list-style-type: none"> <li>• <b>고객 불만·만족·중립 감정 자동 분류 및 화남 등 세부 감정 추출</b> - KoBERT, KoELECTRA, LSTM+Attention, RoBERTa, 감정사전 기반 분석</li> </ul>
키워드/엔티티 추출	NLP/LLM	<ul style="list-style-type: none"> <li>• <b>주요 키워드 및 상품/브랜드와 같은 도메인 특화 개체(Entity) 추출</b> - TF-IDF, 정규표현식(Regex, 속도·정확성), Bi-LSTM + CRF, Transformer 기반(BERT 등), NER(Named Entity Recognition)</li> </ul>
상담 유형 분석	NLP/LLM	<ul style="list-style-type: none"> <li>• <b>상담 내용을 토픽(주제)별로 자동 분류하여 유형화</b> - LDA(Latent Dirichlet Allocation), BERTopic, Top2Vec, GPT 기반 주제 분류</li> </ul>
상담 요약	NLP/LLM	<ul style="list-style-type: none"> <li>• <b>고객사의 요구에 맞는 상담 내용 핵심 요약 모델 개발</b> - TextRank(추출 요약), BART, Pegasus, GPT(생성 요약), KoBART(한국어)</li> </ul>
개인정보 마스킹	NLP	<ul style="list-style-type: none"> <li>• <b>고객 개인정보(주민번호, 계좌번호 등)를 자동 탐지 및 비식별화</b> - 정규표현식(Regex), 패턴 매칭, KoBERT + 마스킹 처리 모델 등</li> </ul>



## 고객 분석 및 상담업무 효율화 분석(2/2)

- 고객센터 운영에 필요한 다양한 분석 및 NLP+LLM 특화 모델링

분석항목	기술분류	내용 및 사용 알고리즘
자동 QA	NLP/LLM	<ul style="list-style-type: none"> <li>• 고객 질문에 자동으로 답변 생성 및 적합한 문서 검색</li> <li>- Retrieval-Augmented Generation(RAG), Dense Passage Retrieval(DPR), GPT, KoGPT, BM25(검색기반)</li> </ul>
금지어/부정어 인식	NLP	<ul style="list-style-type: none"> <li>• 상담 중 욕설, 비속어, 부정적 표현 감지 및 고객 감정 상태 파악</li> <li>- 욕설 사전, 감정 사전 기반, KoBERT, RoBERTa, 규칙 기반 패턴 매칭</li> </ul>
맞춤법 검사	NLP	<ul style="list-style-type: none"> <li>• 상담 텍스트 실시간 맞춤법 및 띄어쓰기 교정</li> <li>- PyKoSpacing, Hanspell, KoBERT 기반 문장 교정 모델, 네이버 맞춤법 API</li> </ul>
상담 KMS 구축	NLP/LLM	<ul style="list-style-type: none"> <li>• 상담 지식 DB 구축 및 자연어 검색, 유사사례 자동 추천 시스템</li> <li>- Dense Retrieval, SBERT, FAISS, KoBERT 기반 유사도 검색, GPT 기반 FAQ 생성</li> </ul>
상담 품질 자동 평가 (QM)	NLP	<ul style="list-style-type: none"> <li>• 상담 내용에서 법적고지, 친절표현 여부 점검</li> <li>- 패턴 매칭, KoBERT, 감정 분석 모델</li> </ul>
이탈고객 예측	NLP/ML	<ul style="list-style-type: none"> <li>• 상담 내용+이력 기반 해지·이탈 가능성 예측</li> <li>- 의도 분석 + 고객행동 예측 모델(GBM, XGBoost)</li> </ul>

## SeiRen RAG 시스템

- LLM모델, 프롬프트, 자원관리, 검색 증강, 보안 기능 등을 갖춘 LLMOps운영 방법 도구

### 다양한 모델

로컬/원격  
모델 지원

- Mistral, Llama3, 허깅페이스, OpenAI API 등 다양한 모델을 지원
- Ollama를 통한 로컬 모델 지원
- 모델간 성능 비교 테스트

### 에이전트툴

실시간 데이터  
사용

- 실시간 데이터 조회
  - 에이전트, 툴
- LLM 함수 호출
- ReAct를 기반 에이전트 지원

### 워크플로우

노코드 기반  
구축

- 워크플로우를 통한 노코딩
- 대화 변수 지원, 다양한 로직 구성
- 다양한 모델, 툴, 지원

### RAG 파이프라인

다양한 문서  
지원

- 최적의 인덱싱, 실시간 테스트
  - 벡터, 풀텍스트, 하이브리드
- 모델 맞춤형 프롬프트 생성
- 실시간 모델간 성능 비교

### 통합 운영

지식생성  
배포까지

- 프로토타입에서 운영까지 지원
- 강력한 디버깅 기능
- 지식생성에서 배포까지 제공
- 입/출력에 따른 보안 기능

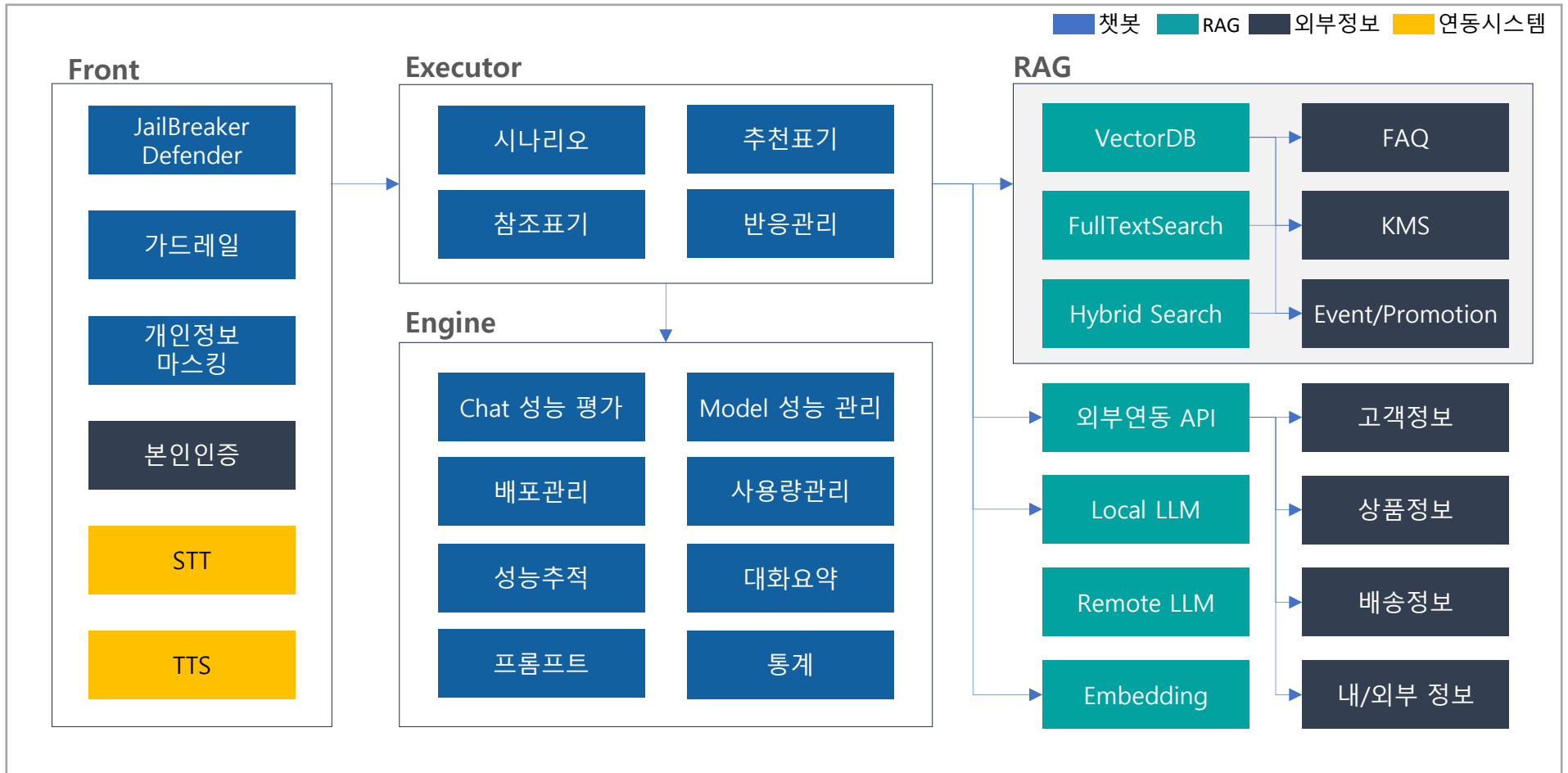
### LLMOps

평가 모니터링

- 대화에 대한 평가
- 모든 대화 모니터링
- 시간경과에 따른 성능 모니터링을 통한 개선기회 도출

# RAG & Chatbot 통합 아키텍처

- RAG 기반의 정보 외에도 고객정보 상품정보 등 다양한 내/외부 데이터를 연동 가능하며, 운영시 성능 및 개인정보 관리에 중점을 둠



# 강력한 RAG 최적화 제공

## 1. 다양한 RAG 기술 적용

- 다양한 임베딩 모델 지원
- 벡터검색, 텍스트 검색, 하이브리드 검색 지원
- 오프라인 인덱싱 지원

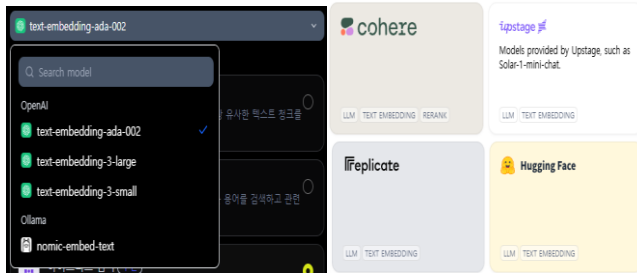
## 2. Database Vector Store, 관리용이

- 문서별 페이지별 버전 관리로 최신성 등 관리용이
- 문서별 권한 관리 용이
- 지식 분할 등 확장성 지원

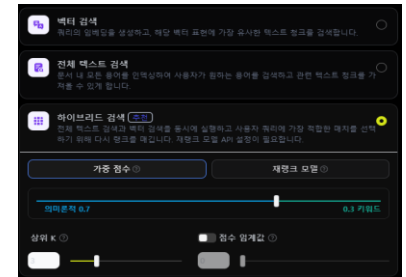
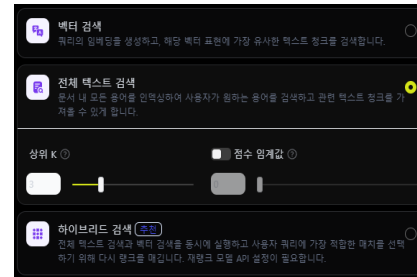
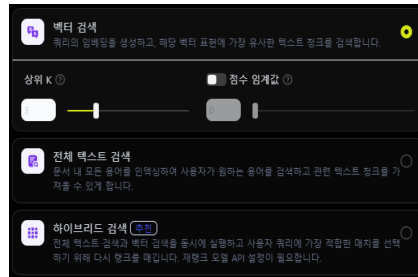
## 3. 각종 편의 기능 제공

- 인덱싱된 문서 결과 비교 전환율 제공
- 문서 최적 chunk\_size 함수 제공
- 임베딩 모델에 따른 평가 기능 제공

### 다양한 임베딩 모델 지원 및 추가 가능



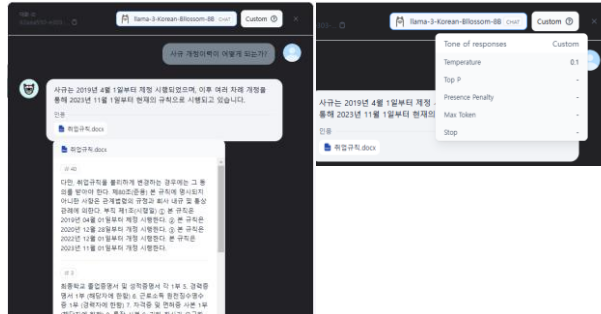
### 다양한 인덱싱 방식 지원



### 청킹 결과 제공



### 검색 결과 제공 (로그에서도 확인)



### 엑셀파일 임베딩

년월	지역	매출액	담당자
01월	이서울	950	이서울
02월	이경기	800	김경원
03월	이인천	100	박민철
04월	이경남	700	고경남
05월	이경북	600	소경남
06월	이서울	950	이서울
07월	이경기	800	김경원
08월	이인천	100	박민철
09월	이경남	700	고경남
10월	이경북	600	소경남
11월	이서울	950	이서울
12월	이경기	800	김경원

# 001	60 문자
"년월":"2024-01-01 00:00:00";지역:"서울";매출액:"950";담당자:"이서울"	
# 002	60 문자
"년월":"2024-02-01 00:00:00";지역:"경기";매출액:"800";담당자:"김경원"	
# 003	60 문자
"년월":"2024-03-01 00:00:00";지역:"인천";매출액:"100";담당자:"박민철"	
# 004	60 문자
"년월":"2024-04-01 00:00:00";지역:"경남";매출액:"700";담당자:"고경남"	
# 005	60 문자

# Smart Vision

- 이미지 프로세싱 및 다양한 분석(Image Analytics) 기술 플랫폼
  - 객체 및 패턴인식 (Object Detection & Pattern Identification) 기술
  - 유사도기반 이미지 검색 기능(Image similarity & Search technology)
  - 생성 알고리즘 기반 타겟 이미지 생성기술(Image Generation; GAN)
  - 다양한 산업내 OCR 기술(OCR in images)
  - Background Removing 기술



## □ 적용사례

Video Analytics	OCR in Images	Image Similarity	차량정보 자동인식
<p><b>9 나인포즈 앱 출시</b></p> <p><b>Video Analytics Tech.</b></p> <ul style="list-style-type: none"> <li>• Video Pre-processing</li> <li>• Pose Detection</li> <li>• Swing Analysis</li> </ul> <p>1 내 스윙동작 자동 분석 &amp; 점수</p> <p>타이거우즈의 스윙과 자동비교분석</p>	<p><b>OCR in Images</b></p>	<p><b>Image Similarity</b></p> <p>유사 이미지 탐색 단계</p> <ol style="list-style-type: none"> <li>1 이미지 전처리 (원본 이미지, Grayscale 변환, 채널 제거)</li> <li>2 이미지 특징 추출</li> <li>3 유사 이미지 Reranking (1.1 이미지 유사도 비교, 1.2 이미지 유사성 측정 지표)</li> </ol>	<p><b>차량정보 자동인식</b></p> <p>시각만 사진 판독을 통한 차량 정보 자동화 처리</p> <p>약 1,000여종의 자동차 models &amp; 연간 30만 건의 중고차 대용 업무용 AI 개발</p> <ul style="list-style-type: none"> <li>• 승용차: 약 900종(국산 250, 수입차 650)</li> <li>• 승합차: 약 100종 수준</li> <li>• 화물차: 약 70종 수준</li> <li>• 특수차: 약 50종 수준</li> </ul> <p>색이묘트</p> <p>YOLO, MobileNets (CNN), Tesseract (OCR), Color Histogram, XGBoost, IsolationForest</p> <p>인식값: 차량, 색상, 번호판, 파손, Fake</p> <p>대용량인 처리기, 추가 업무 확장, 차량가치 평가, 대용량용 연계, 차량등록</p>

# Smart Vision

- 영상기반 품질관리

## AI/DT Visual Fault Inspection

- 현재 모든 반도체 설비 내외부에는 CCTV가 설치되어 있음  
(문제 발생 시 추적하는 기능으로 사용)
- 실시간 Monitoring하는 기능의 Motion 감지 카메라는 특정 설비에 적용 중
- 특정 설비에 적용하는 Motion 감지 기능은 Image 비교하여 정상여부를 판단하고 Alarm을 띄우는 역할을 하고 있음

특징

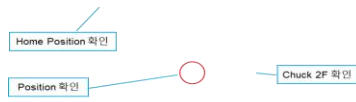
Recipe에 따른 정상 유무를 판단



양영준

Process		④DIW	⑤IPA	⑥IPA+N2
Time(동영상sec)		60~79s	85~90s	95~140s
Dispenser1	Chemical	Home	X	X
Dispenser2	DIW	Action	Home	X
Dispenser3	N2	X	X	Action
Dispenser4	IPA	X	Action	Action
Fixed nozzle	DIW	X	X	X
Jet spray	DIW+N2	X	X	X
Chuck position		2F	2F	2F
Chuck pin		X	X	X
Chamber		X	X	X
Door		X	X	X

김원석

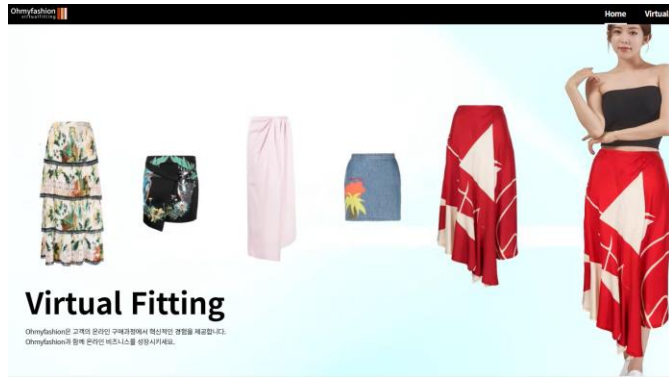


# Smart Vision : Virtual Fitting

- 생성형 AI기반 가상착의(Virtual Fitting) 솔루션
  - Body Segmentation, Posture detection, Deformation & Generative AI 알고리즘을 적용함
  - 표준모델(Base Model) 또는 아바타 방식이 아닌 개인용 맞춤, 본인 사진 업로드 가상피팅
  - 사진 업로드 후, 약 2~3초 후 피팅 이미지 생성,클라우드 기반 Computing



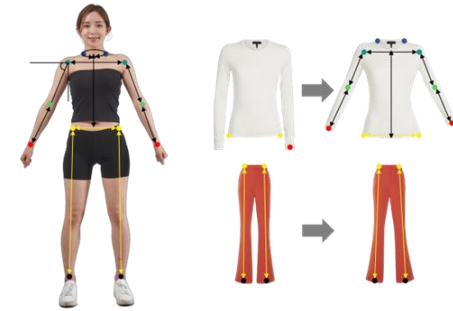
## Ohmyfashion.ai



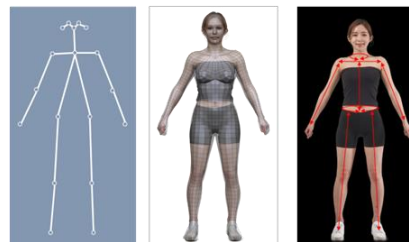
### 1 Body Parts Segmentation



### 3 Clothing deformation by posture and size



### 2 Posture detection & Parts Size



### 4 Generative Fitting Output











# InterRoid

- Voice Call Automation with SeiRen

**(주)인터로이드**

Home | <https://www.interroid.com>

eMail | [marketing@interroid.com](mailto:marketing@interroid.com)

Tel | 02-563-2244

Add | 서울시 강남구 논현로503, 17층