



전사 AI 플랫폼

Athena
AURDA

: AI 챗봇 도입부터 전사 운영까지

Table of content

01 유라클 소개

02 AI Transformation

03 Our Approach

03 AI 솔루션

- 솔루션 개요 및 특징점
- 주요 기능
- 구축 사례

Uracle is : 안정성 및 성장성

유라클은 모바일 플랫폼을 비롯한 다양한 업종과 분야에서 시스템 구축 사업을 수행하고 역량을 축적해 온 25년차 전문기업입니다.

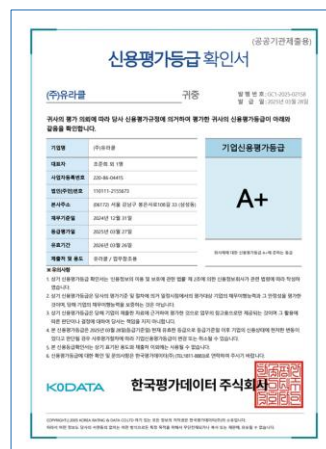


uracle

AI 플랫폼·LLM·자체 개발 솔루션 보유

회사명	(주)유라클	대표자	조준희, 권태일
주소	서울시 서초구 방배로 231 (방배동)		
사업분야	AI 솔루션 개발, 시스템 통합 구축·유지, 모바일 시스템 구축·개발		
설립년도 / 사업기간	2001.01 ~ 2026.03 (25년 2개월)		
보유인력현황	총 254명 (개발:218/지원인력:36)		

'25 신용평가등급
A+



최근 3년 평균 매출
350억+

2001 ~ 2009

2010 ~ 2019

2020 ~ 2025

- '08 벤처기업대상 중소기업청장표창 수상
- '07 (주)유라클 상호변경
- '07 정보통신부 핵심동력개발사업 차세대기술선정
- '01 '모바일로' 금융감독원 금융 신상품 대상
- '01 (주)아이엠넷피아 설립

- '19 SW산업발전유공자 대통령 표창 수상
- '19 대한민국 소프트웨어 기업경쟁력 대상
- '16 SW기업경쟁력 대상 IT솔루션 분야 최우수상
- '15 모바일 전자정부 행정자치부 장관 표창
- '11 스마트앱어워드2011 금융서비스 최우수상

- '25 생성형 AI 플랫폼 출시 (Athena AI 플랫폼)
- '25 AI 코드어시스턴트 제품 출시 (Athena Code)
- '24 고려대 HI AI 공동연구소 설립 및 LLM 개발
- '24 (주)유라클 KOSDAQ 상장 KOSDAQ
- '21 한국소프트웨어산업협회 회장 취임 (조준희 회장)

사업 분야

모바일·AI 플랫폼 등 SW, 시스템 구축,
운영 및 유지보수

주요 제품

AI Product Suite

Mobile Product Suite

주요 고객사

• 삼성



• 현대



• Major Financial Institutions



• Other Industries



Oracle is : 고객사 현황

1,000여개의 산업 별 전문 지식과 경험으로 유라클이 AI 도입 여정의 새로운 파트너가 되겠습니다.

공공

Public Sector



금융

Finance



제조

Manufacturing



서비스

Service



기타

Other Industries



Uracle is : AI 사업 수행 실적

2025년 1월부터 AI사업 시작 후 상반기 20건의 PoC를 거쳐, 상용 레퍼런스를 빠르게 확장하고 있습니다.

사업명	발주처	투입 S/W	주요내용
AI 개발플랫폼 구축	KODATA(한국평가데이터)	AI Platform	AI 개발플랫폼(LLMOps) 및 지식관리 시스템 구축
AI 서비스 플랫폼 구축	KCB(코리아크레딧뷰로)	AI Platform	No code 기반의 전사 AI 서비스 구축용 플랫폼 구축
협회 통합정보시스템 구축	대한의사협회	AI Platform	AI 플랫폼 기반 민원인·직원·광고심의 신청자용 AI 챗봇 구축
AI기반 FDS 검사시스템 구축	우리은행	AI Platform	과거 사례 시나리오 기반 FDS 검사시스템 구축
AIOps 플랫폼 및 고객 지원 챗봇 개발	인텔리안테크	AI Platform	생성형 AI 플랫폼 및 고객 지원용 RAG 챗봇 구축
AI 서비스 내재화 플랫폼 구축	현대건설	AI Platform	전사 생성형 AI 플랫폼 구축
新 모바일 워크플레이스 구축	GS건설	AI Platform	전 임직원용 AI 비서(기사 요약, 앱 검색) 구축
생성형 AI 시범과제 구축 용역	메리츠증권	AI Platform	실시간 문서 번역/요약 Agent 구축
AI기반 차세대 안전신고 및 위험징후 분석 서비스 기술 개발	행정안전부	LG AI 연구원 EXAONE LLM	신고 Agent 개발 및 차세대 안전신문고 시스템 연동
AI 코드 어시스턴트 도입	KT 스카이라이프	AI Code Assistant	코드어시스턴트를 통한 IE/JAVA기반 구형 웹서비스 전환
지식 검색용 RAG 챗봇 PoC	한국과학기술단체총연합회	AI Platform	정책 보고서 지식 검색 RAG 구축
홈페이지 고객용 RAG 챗봇 PoC	한국인공지능 SW산업협회	AI Platform	고객용 지식 검색 RAG 구축
경력관리 지식검색용 챗봇 PoC	한국건설기술인협회	AI Platform	경력 관리용 RAG 검색 테스트
학사행정 지식검색용 RAG 챗봇 PoC	서울과학기술대학교	AI Platform	학사행정 지식검색 RAG 테스트
지식 검색용 RAG 챗봇 PoC	인천소상공인종합지원센터	AI Platform	사내 직원용 RAG 검색 테스트
지식 검색용 RAG 챗봇 PoC	한국투자저축은행	AI Platform	내부직원용 지식 검색 RAG 구축
내부 지식 검색용 RAG 챗봇 PoC	NH저축은행	AI Platform	사내 문서 지식검색 RAG 구축
ERP 매뉴얼 지식검색 PoC	두산퓨얼셀	Athena GenAI Platform	ERP 매뉴얼 검색 용 RAG 구축
지식 검색용 RAG 챗봇 PoC	다이소	Athena GenAI Platform	직원용 매뉴얼 검색 RAG 테스트
AI를 활용한 적합 후보자 추천 PoC	커리어케어	AI Platform	맞춤형 채용 후보자 추천을 위한 Hybrid 검색 구축
품질검사 AI Assistant PoC	두산에너빌리티	AI Platform	검사원 선정 업무 자동화, 검사체크리스트 생성

AI Transformation : 3가지 구조적 변화

생성형 AI의 발전은 모델 기술을 넘어 소프트웨어와 서비스 구조 자체를 변화시키고 있습니다.

Transformation 01.



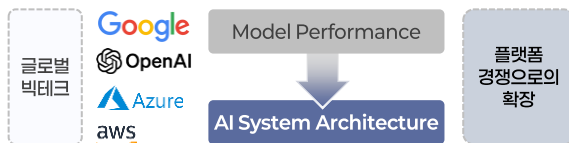
모델 경쟁의 종료

모델 성능의 상향 평준화로 인해, '얼마나 똑똑한가'보다 '어떻게 업무에 적용할 것인가'가 핵심 경쟁력이 되었습니다.

POINT 01 모델 성능 상향 평준화

- GPT-4, Claude, Gemini 등 주요 모델이 대부분 벤치마크에서 인간 수준 성능 도달
- MMLU benchmark: 대부분 모델 80~90% 이상

POINT 02 AI 경쟁 구조 변화



모델 자체보다 AI 시스템 설계가 중요

Transformation 02.



SW 구조 변화

사용자가 도구를 직접 조작하여 결과를 만드는 방식에서, AI가 복잡한 작업을 대신 수행하는 구조로 변화했습니다.

POINT 01 AI Agent 시장 성장

- AI agent 시장 2030년까지 약 45~50% CAGR 성장 전망 *Markets and Markets/Gartner*

POINT 02 기업 소프트웨어 구조 변화

- 2028년까지 기업 소프트웨어의 30% 이상이 AI agent 기반 기능 포함 *Gartner*

GitHub Copilot

OpenAI Operator

소프트웨어가 작업 수행 구조로 변화

Transformation 03.



SaaS → OaaS 등장

소프트웨어의 기능을 판매하는 것에서, 사용자가 원하는 최종 업무 결과를 직접 제공하는 서비스로 진화했습니다.

POINT 01 기업 AI 도입 목적 변화

- 생성형 AI 도입 기업 중 약 40% 이상이 업무 자동화 활용 *McKinsey AI Report*

보고서 작성

코드 생성

데이터 분석

고객 응답

POINT 02 Outcome 기반 소프트웨어 등장

- AI Agent 기반 소프트웨어는 기능 제공 중심 SaaS에서 결과 중심 서비스(OaaS)로 진화 *Gartner BCG/ITPro*

* OaaS : Outcome as Agentic Solution

기업은 기능이 아니라 결과를 기대

Our Approach : AI 시대의 설계 전략

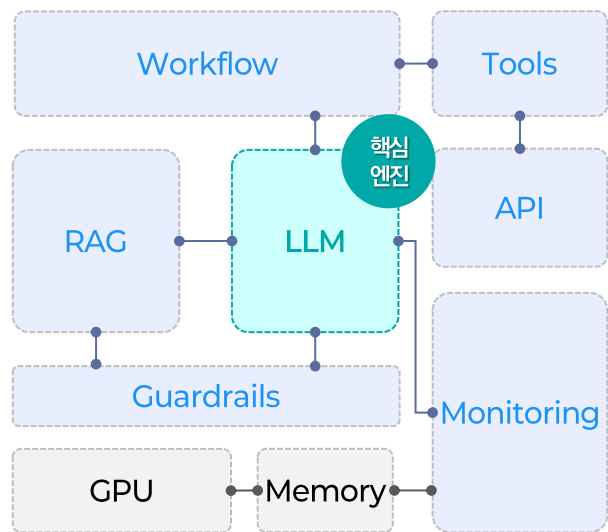
생성형 AI 의 변화에 대응하기 위해 AI 시스템 구조, 소프트웨어 방식, 서비스 모델까지 함께 설계하는 접근이 필요합니다.

Strategy 01.



복합 AI 시스템 구조

AI는 단일 모델이 아닌 **Workflow, Tool, Retrieval, Memory** 등 여러 요소가 결합된 구조로 설계가 필요합니다.

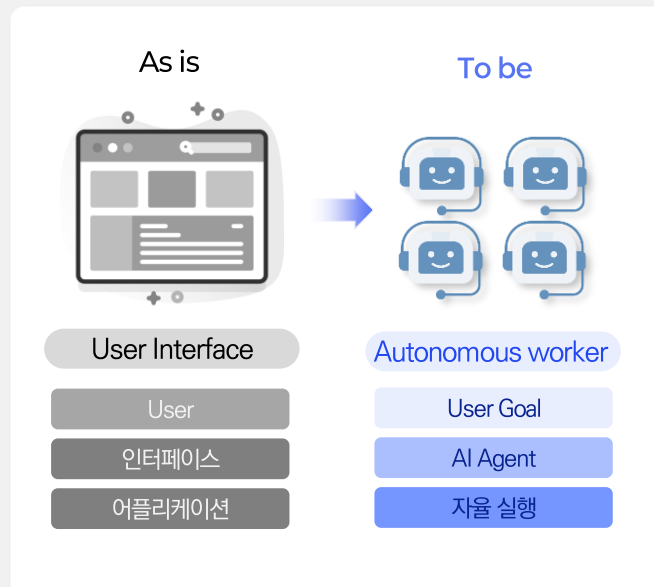


Strategy 02.



Agentic Software

AI 서비스는 단순 응답 생성이 아니라 계획, 실행, 도구 활용을 통해 업무를 수행하는 **Agent 기반 구조**로 설계가 필요합니다.



Strategy 03.



Outcome Automation

AI 서비스는 목표 기반 Agent Workflow를 통해 **업무 결과를 생성**하는 구조로 설계가 필요합니다.



이러한 구조를 구현하기 위해서 Agent 실행과 Workflow orchestration을 지원하는 AI 아키텍처가 필요합니다.

Our Solution : AI Product Suite (Athena/AURDA)

일부 부서의 실험이나 단기 PoC가 아닌, **그룹 전반의 업무에 AI를 점진적으로 확산시키는 통합 플랫폼을 제공합니다.**



아테나(Athena) AI Agent Builder platform

업무에 맞는 AI 기능을 단계적으로 확장

Expand: Extensions

다양한 업무 형태에 특화된 기능 모듈 결합

복합 질의(Agentic RAG)

Code Assistant

OCR

Core: Agent Builder

단순 답변을 넘어 업무 흐름을 수행하는 에이전트

워크플로우

멀티턴

Tool 연계

프롬프트

Base: RAG Pipeline

기업 내부 지식 자산화 및 검색 품질 확보

문서 파싱/임베딩

벡터 검색

Re-ranking



오르다(AURDA) Ops platform

확장된 AI 서비스의 안정적 운영을 위한 인프라/거버넌스

Govern: Governance

데이터 및 기능 접근 통제

IAM / SSO

RBAC

Audit

Core: Operation

확장된 AI 서비스의 안정적 운영

API Gateway

CI/CD (GitOps)

Auto-Scaling

Base: Infrastructure

AI 실행을 위한 기본 인프라 레이어

GPU 가상화(MIG)

스토리지 관리

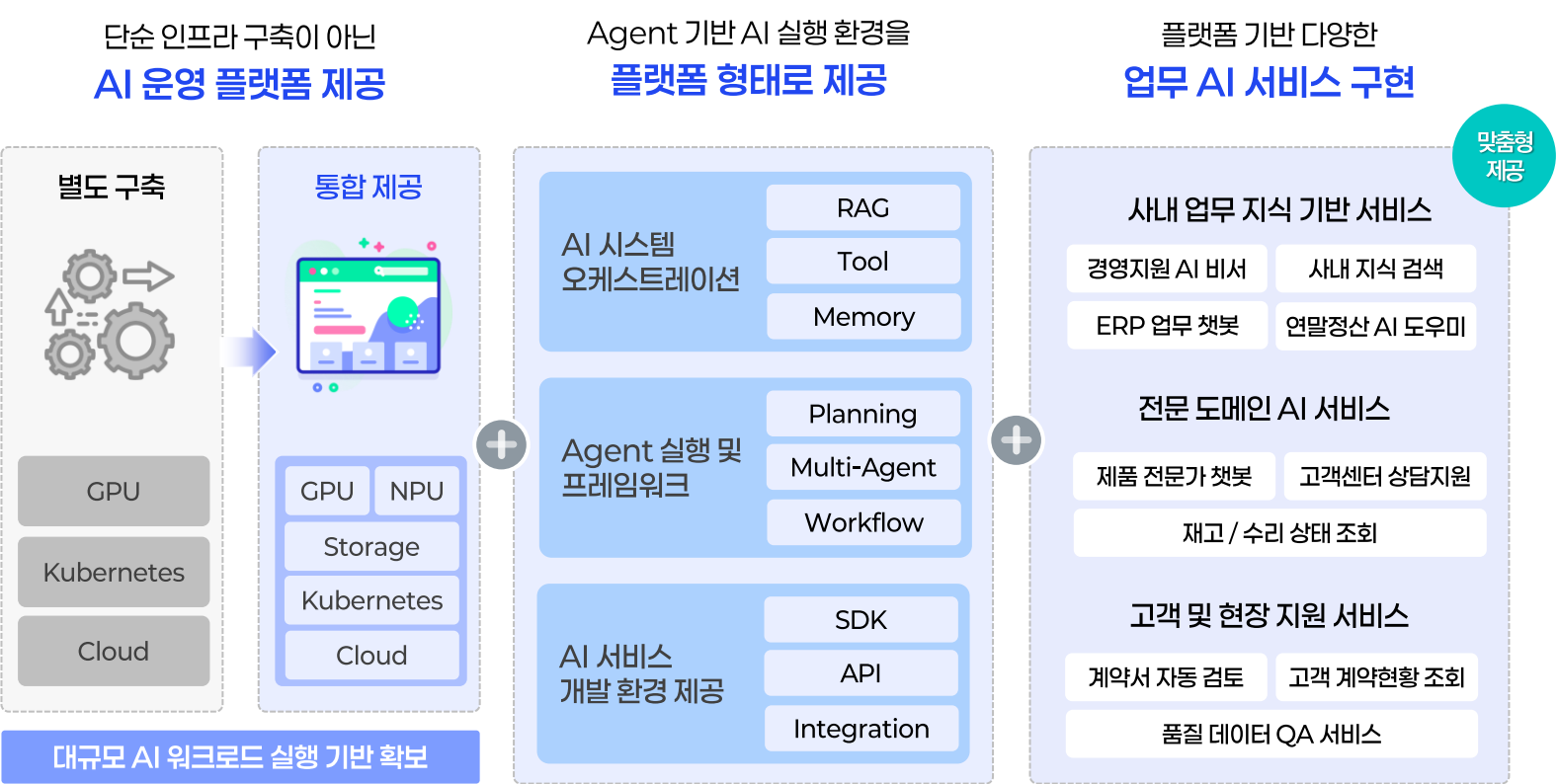
네트워크

Key Capabilities : 플랫폼 통합 아키텍처

AI 인프라부터 Agent 플랫폼, 그리고 실제 서비스 구축까지 하나의 플랫폼 구조로 통합 제공합니다.

통합성 AI 기술력

AI 인프라부터 서비스까지의 통합 플랫폼 구조



플랫폼 도입 효과

AI 인프라부터 Agent 플랫폼, 서비스까지 통합된 기업 AI 운영 체계 확보

1

AI 서비스 구축 속도 향상

- 인프라 모델 연동 반복 최소화
- 서비스 로직 개발에 집중

2

안정적인 운영 체계

- 인프라와 Agent 플랫폼 통합 관리
- 배포·모니터링·운영 통제 일원화

3

지속 가능한 AI 확산

- 신규 AI 서비스 추가에 유연한 구조
- 단일 서비스가 아닌 다수 서비스로 확장

Key Capabilities : **고도화된 RAG 처리**

단계별 복합 추론을 수행하는 Agentic RAG 기술을 적용하여 고도화된 RAG 처리를 지원합니다.

정확성

지능형 검색

차세대 RAG 아키텍처 지원

기술 흐름

“ **검색** 기반 RAG 구조 ”

Standard RAG 구조

- 단일 검색 기반
응답 구조

사용자의 질문을 그대로 검색에
사용하여 관련 문서를 검색하는
기본 구조
- 일방향
(One-Way)
실행 구조

검색 결과가 부정확하거나 내용이
부족하더라도 LLM이 바로 답변을
생성하는 구조
- 단일 검색
중심 응답

단순 정보 검색만 가능하며, 여러
정보를 결합하거나 비교/분석하는
복합 질의 해결 한계

65세 남성이 1억 담보로 '종신형'과 '10년 확정형'을
선택할 때, 1년 동안 받는 수령액은 수차상 어느 쪽이 더 큰가요?

“ **Agent** 기반 RAG 구조 ”

Agentic RAG의 확장성

- 질문 의도 분석
및 쿼리 최적화

LLM이 질문의 의도를 파악하고
쿼리·로직을 재작성하여
검색 정확도 극대화
- 순환형(Loop)
피드백 구조

검색 결과가 불충분하면 추가 검색
및 정보 보완을 수행하여
정답률 향상
- 다단계 추론 및
복합 질의 해결

Agent 기반으로 서로 다른 문서의
정보를 비교·분석·계산하여
심층적 해답 도출

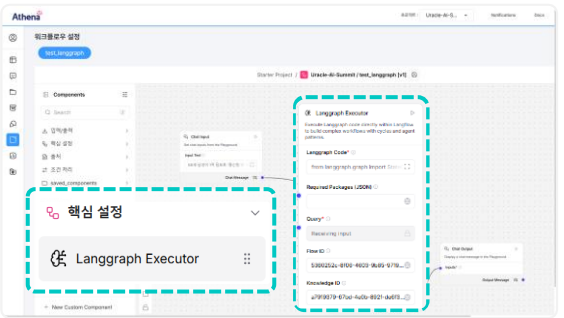
확장형
기술

Agent Builder

- Workflow
- Tool Registry
- MCP
Repository
- Multi-turn
- 대화 이력
- 대화형
인터페이스

+

 **LangGraph**



Agentic RAG 구현 복잡도를 플랫폼으로 단순화

Key Capabilities : Agent 개발 생산성 확보

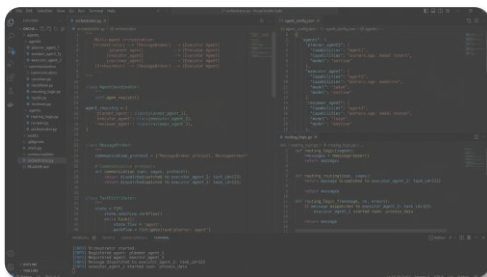
Agent 개발을 위한 코드 작성과 협업 과정을 AI Code Assistant 기반으로 지원합니다.

생산성

개발 자동화

AS-IS

IDE 기반 개발



〈IDE 개발 예시〉

RAG 파이프라인 코드 기반 구현

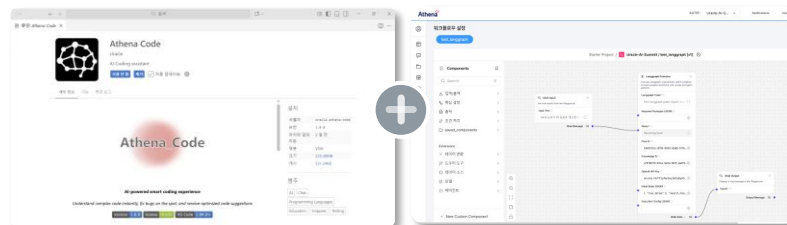
Tool / Memory / Retrieval 연결 구현

Multi-agent orchestration 직접 구성

Agent 구현 복잡도 증가

TO-BE

IDE + Agent 플랫폼 기반 개발



〈IDE 내 Athena Code 설치 예시〉

〈플랫폼 내 Workflow 예시〉

VSCode / Eclipse IDE 지원

API 기반 대규모 문서 처리 지원

코드 자동 생성 및 개선 지원

오케스트레이션 톨 로직 관리

프로젝트 구조 기반 코드 분석 지원

AI 실행 인프라와 개발 코드 분리

개발 복잡도를 줄이고 협업 개발 환경으로 생산성 향상

핵심 효과

개발 패러다임의 변화

IDE 기반 Agent 협업 개발 환경 확보

- VSCode / Eclipse IDE 기반 개발 지원
- Code Assistant 기반 코드 생성 및 리뷰
- 팀 단위 Agent 협업 개발 환경 제공

RAG 및 Agent 구현 복잡도 감소

- PDF 파싱 · 청킹 · 임베딩 등 전처리 로직 플랫폼 처리
- Tool / Memory / Retrieval 연결 구조 단순화
- Agent orchestration 코드 부담 감소

AI 서비스 개발 및 확장 속도 향상

- Agent Workflow 기반 서비스 구현
- 대규모 문서 처리 및 AI 실행 인프라 플랫폼 제공
- 서비스 개발 → 운영까지 일관된 개발 구조 확보

Key Capabilities : Agent 통합 운영 관리

생성형 AI의 발전은 모델 기술을 넘어 소프트웨어와 서비스 구조 자체를 변화시키고 있습니다.

안정성 자원 최적화



Key Features : 멀티 모델 관리

Private·Public LLM을 단일 카탈로그에서 등록·제어하고 전사 AI 자원의 거버넌스 체계를 실현합니다.

Private · Public LLM 관리

내부 vLLM 서버 모델과 외부 API 모델을 단일 카탈로그에 등록하여 중앙 집중식 제어

코드 수정 없는 무중단 모델 전환

AI Gateway를 통해 이기종 모델을 단일 API로 통합하여 관리 콘솔에서 연결 모델만 교체하여 반영

부서·프로젝트 단위 AI 자원 격리

RBAC 및 네임스페이스 기반으로 모델 사용 권한과 파라미터 상한선을 부서별 적용 가능

모델 품질·운영 현황 가시성 확보

모델 별 응답 속도·토큰 사용량·성공률을 실시간 추적 및 이상 징후 조기 감지 및 대응

모델 관리

모델 찾기

모델 목록 | 10개씩 보기

모델별칭	설명	모델타입	공급자	상태	기본 모델 여부	사용 여부	등록자
gpt-oss-120b		llm	openai-compatible				
gemini-2.5-flash		llm	google				
gpt-4o		llm	openai				
gpt-oss-20b			openai-compatible				
kulim-embed			athena				
ko-reranker			athena				

용도별 모델 분리 관리

모델 통합 등록

LLM

text-embedding

Rerank

모델 별 정책 설정

기본값 제어

파라미터
기본값 설정

사용여부

모델 수정

LLM

공급자

OpenAI

모델명

gpt-4o

엔드포인트 URL

엔드포인트 URL을 입력해주세요.

API Key

API Key를 입력해주세요.

기본값 설정

사용여부

사용

Key Features : RAG 파이프라인 빌더

복합 문서 파싱부터 Agentic RAG 기반 검색까지, 검색 파이프라인 전 과정을 플랫폼에서 직접 설계합니다.

비정형 문서의 구조화된 지식 변환

PDF·HWP·MS Office 복합 문서 파싱 및 Parent-Child 청킹 전략 설정

하이브리드 검색 기반 정밀 답변

키워드 + 벡터 검색 결합 및 Reranker로 검색 정확도 극대화

벡터 DB 통합 인프라 즉시 제공

Milvus 기반 벡터 DB 플랫폼 내 통합, 별도 인프라 구축 불필요

Agentic RAG로 복합 질의 처리

질문 의도 분석 → 반복 검색 → 복합 추론으로 단순 RAG 한계 극복

파서/청킹 전략 설정

문서에서 텍스트를 추출하고 분석하는 방식을 설정합니다. 기본 파서를 통해 문서 내용을 읽고, 청킹 전략으로 검색에 적합하게 분할합니다.

업스테이지 도큐먼트파스

아테나 파서

사이냅소프트 도큐멘타라이저

업스테이지 도큐먼트파스

상용·오픈소스 파서 지원

상용 파서 연동 지원

OCR 미적용

테스트 파일 선택

청킹 결과 실시간 검증

적용 전 품질 확인

미리보기 파일

문서명	파일 경로
sample_200pages.pdf	0317upstage/sample_200pages.pdf

임베딩 설정

선택한 임베딩 모델을 통해 텍스트를 의미 기반의 벡터로 변환하고, 벡터 저장소에 저장하여 질문과 가장 관련 있는 정보를 빠르게 검색합니다.

임베딩 모델

kullm-embed

벡터 저장소

MILVUS

검색 설정

유사도

하이브리드

전체 텍스트 검색

키워드 검색과 의미 검색을 결합한 방식으로 문서를 찾습니다.

유사도 검색

질문의 의미와 유사한 내용을 담고 있는 텍스트 조각을 검색합니다.

하이브리드

전체 텍스트 검색의 정확성과 유사도 검색을 동시에 활용하여 검색합니다.

검색설정·Reranker

검색 정확도 및 품질 향상

최종 검색 결과 수 Rerank 후 LLM에 최종 제공될 문서 개수를 설정합니다.

1

5

5

Key Features : 프롬프트 관리

프롬프트 설계·버전 관리·비교 검증·전사 자산화까지 AI 서비스 품질의 핵심을 체계적으로 관리합니다.

변수 기반 프롬프트 템플릿 설계

질문·검색 결과 등 실행 시점에 바뀌는 값을 변수로 정의해 하나의 템플릿으로 다양한 상황에 재사용

환경별 버전 관리 및 즉시 롤백

프롬프트 수정 시 신규 버전 생성, 개발·검증 등을 환경별로 적용하여 품질 저하 시 이전 버전 즉시 복구

멀티 모델·파라미터 비교 검증

동일 프롬프트를 복수 모델과 파라미터 조합으로 실행하고 검증하여 배포 전 품질 확인

프롬프트 템플릿 기반 전사 자산화

Few-shot·요약·번역 등 자주 사용되는 프롬프트 패턴을 공유하여 전사 표준 프롬프트 자산으로 관리

프롬프트 편집

< 템플릿D

모델 선택

gpt-oss-120b

파라미터 설정

온도 1

0.0

최대 응답 길이 1

2000

최소 1

고급 옵션 사용

Top-K 1

0

프롬프트 생성

System

당신은 도움이 되는 비서입니다.
아래 주어진 '관련법령'과 '내부규정'을 바탕으로 질문에 가능한 한 최선을 다해 답변하세요.
단, 사전 지식을 활용하지 않고 주어진 '관련법령'과 '내부규정'만을 이용해 답변하세요.

즉시 실행 테스트

설계 중인 프롬프트 결과를 실시간으로 검증

테스트 실행

멀티 모델 응답 비교 검증

동일 프롬프트의 모델별·파라미터별 결과 즉시 확인

< 프롬프트 비교

템플릿D 버전 1

당신은 도움이 되는 비서입니다. 아래 주어진 '관련법령'과 '내부규정'을 바탕으로 질문에 가능한 한 최선을 다해 답변하세요. 단, 사전 지식을 활용하지 않고 주어진 '관련법령'과 '내부규정'만을 이용해 답변하세요. ----- 관련법령 : {knowledgeA} 내부규정 : {knowledgeB}

system

question : {input}

human

변수 값 입력

변수 이름

변수 값

knowledgeA

변수값을 입력해주세요.

knowledgeB

변수값을 입력해주세요.

템플릿B 버전 1

당신은 "Context"의 내용을 참고하여 Question에 맞는 카테고리를 알려주는 비서입니다. 단, Context의 category만 대답하세요. 예시 1 질문 : 유라클의 대표 제품은 무엇일까? 답변 : 회사소개 예시 2 질문 : 논술 전형에서 합격률 결정하는 기준을 정리해줘 답변 : 입학전형 --- Context : {context} ---

system

Question : {input}

human

변수 값 입력

변수 이름

변수 값

context

변수값을 입력해주세요.

input

변수값을 입력해주세요.

Key Features : 에이전트 빌더

No Code/Low Code 기반으로 에이전트 로직을 시각적으로 설계하고 MCP 기반 외부 시스템 연동까지 단일 환경에서 구현합니다.

LC/NC 기반 에이전트 설계

Drag & Drop 인터페이스로
LLM·메모리·도구 등 에이전트 구성
요소를 연결해 서비스 로직 완성

사전 정의 컴포넌트 제공

ReAct·Zero-shot 에이전트,
프롬프트 템플릿, 파서 등 검증된
노드를 라이브러리에서 즉시 활용

설계-테스트-배포 단일 환경 제공

워크플로우 실행 결과를 즉시
확인하고 단계별 Trace, 클릭으로
REST API 엔드포인트 생성

MCP 기반 외부 시스템 표준 연동

MCP 프로토콜로
DB·API·파일시스템 등 이기종
외부/내부 시스템 즉시 연결

워크플로우 설정

Agent_RAG

Components

Search

입력/출력

핵심 설정

출처

조건 처리

saved_components

Extensions

데이터 변환

도우미 도구

데이터 소스

모델

에이전트

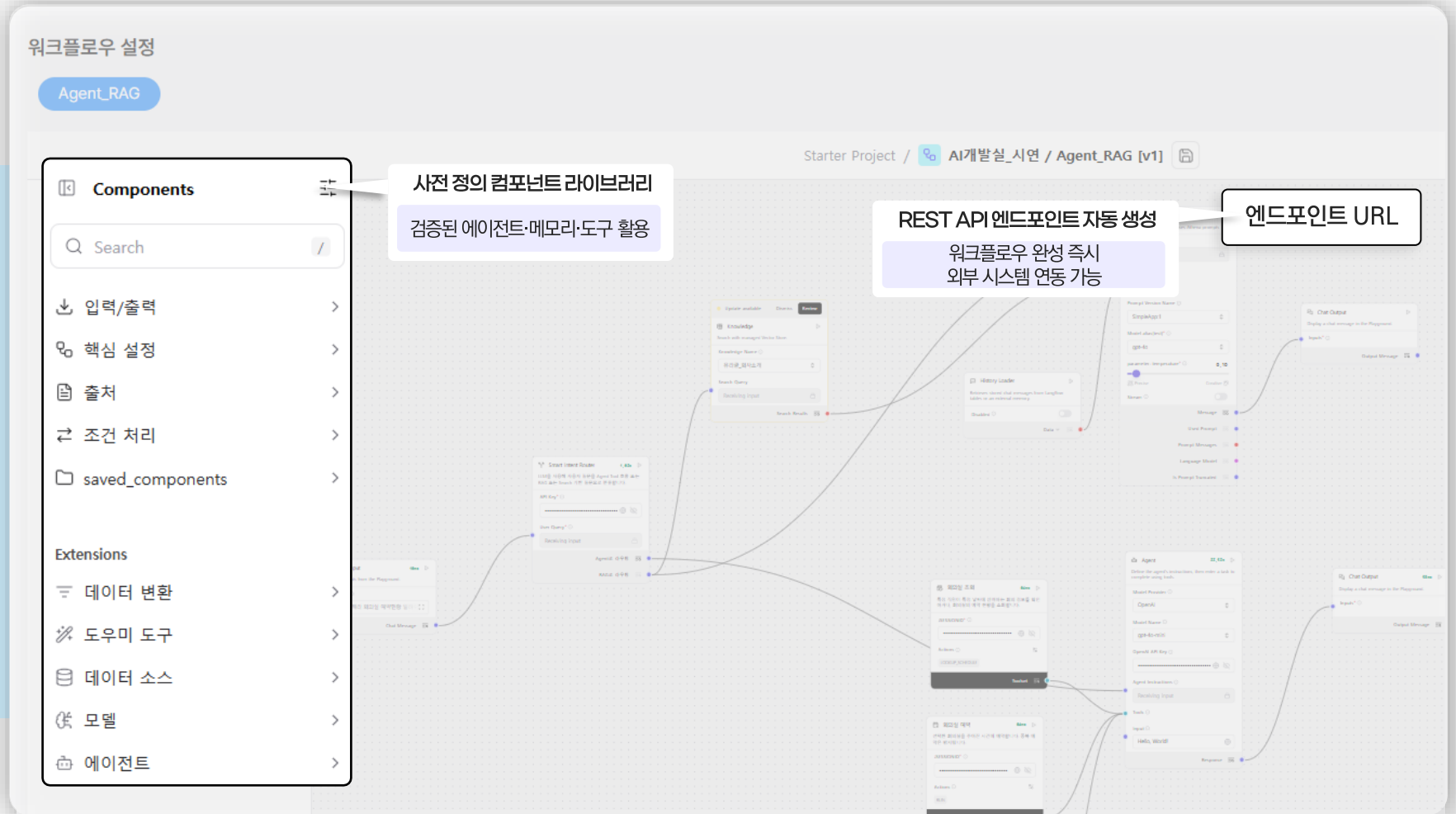
사전 정의 컴포넌트 라이브러리

검증된 에이전트-메모리-도구 활용

REST API 엔드포인트 자동 생성

워크플로우 완성 즉시
외부 시스템 연동 가능

엔드포인트 URL



Key Features : AI 서비스 배포 관리

워크플로우 버전 관리부터 무중단 배포·롤백, 프롬프트 자산 관리까지 AI 서비스 운영 전 과정을 통합 지원합니다.

워크플로우 버전 관리 및 이력 추적

에이전트 워크플로우를 Git 기반으로 버전별 저장 및 개발·운영 환경을 분리하여 변경 이력 통합 관리

운영 배포 전 자동 검증 체계

Promote 시점에 가드레일 작동 여부·배포 요건을 자동 테스트하여 배포 후 정상 응답 최종 확인

무중단 배포 및 즉시 롤백

신규 버전 배포 중에도 이전 버전이 동시 운영되어 장애 발생 시 즉시 이전 버전 복구

유연한 배포 환경 대응

로컬·Docker·Kubernetes 등 실행 환경에 따라 배포 워크플로우를 커스터마이징

CI/CD 이력

testapp

상태

전체

운영 배포 시작 일자

YYYY.MM.DD HH:MM

운영 환경 배포 대상 설정

앱·버전·서버 정보 입력만으로
운영 배포 즉시 실행

CI/CD 목록 | 10개씩 보기

앱 이름	앱 버전
testapp3	4
testapp3	4
testapp3	4
testapp3	4

앱 운영 배포

1 운영 서버 연결

2 운영

운영 서버 연결 운영 서버 연결을 위해 로그인 정보를 입력해주세요.

앱 이름

testapp3

버전

4

운영 Target *

http://172.20.0.10:14009/

Username *

admin@uracle.co.kr

Target URL 및
Username 저장

☒ * 저장 후 운영 배포를 완료하셨을 경우, 다음부터는 URL과 Username이 자동 입력됩니다.

Password *

Password를 입력하세요.

Key Features : AI 인프라 운영 관리

GPU 자원 최적화부터 네트워크 보안, 통합 관측성까지 AI 서비스 구동에 필요한 인프라 전 영역을 안정적으로 운영합니다.

GPU 자원 효율 극대화

MIG 기술로 단일 고성능 GPU를
최대 7개 독립 인스턴스로 분할하여
자원 낭비 최소화

부서·프로젝트 단위 자원 격리

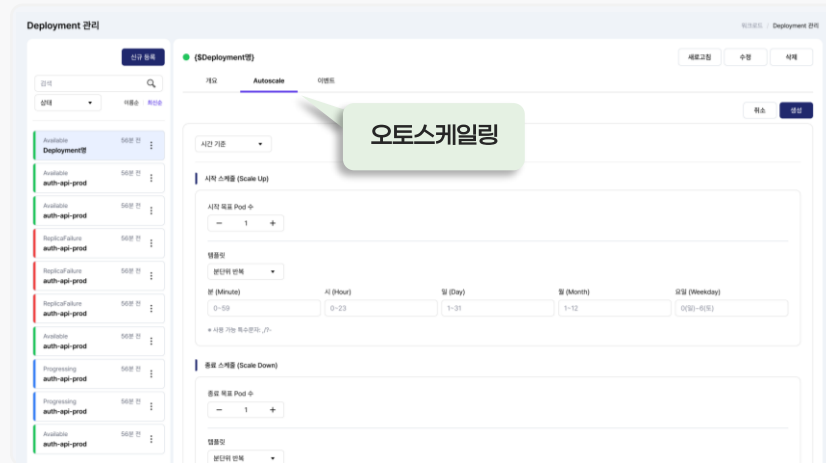
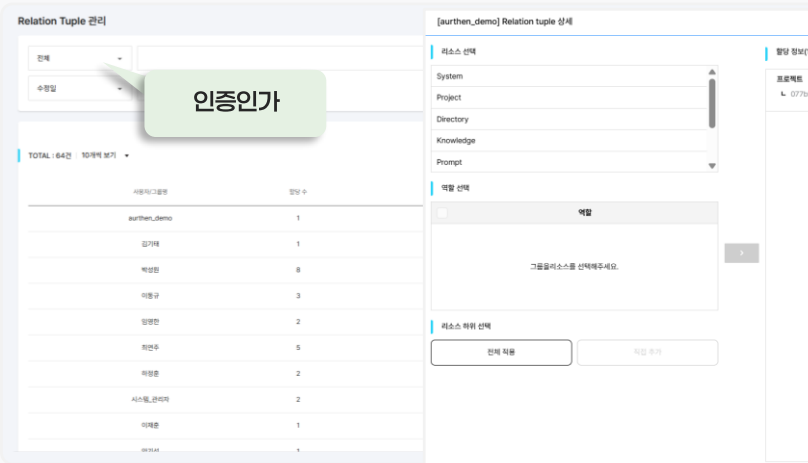
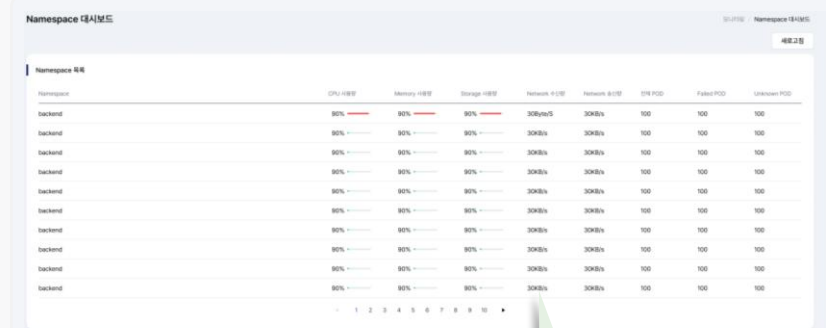
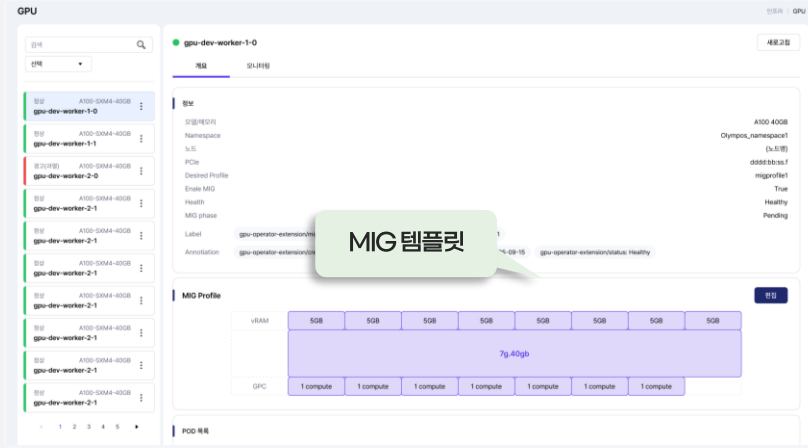
네임스페이스별
CPU·메모리·스토리지 상한선을
적용하여 데이터 간섭·유출 원천 차단

멀티테넌트 인증·인가 체계

Ory Stack 기반 JWT 토큰
·SSO 통합 인증으로 중앙 집중식
접근 제어 및 독립적 인가 정책 적용

고가용성 및 장애 자동 복구

서비스 부하에 따른 오토스케일링과
컨테이너 장애 자동 복구로 AI 서비스
중단 없는 운영 체계 확보



Key Features: 통합 모니터링

AI 서비스 운영 현황을 종합적으로 파악하기 위한 통합 모니터링 체계를 제공합니다.

Full-Stack 통합 관제

하드웨어·애플리케이션 전 계층
메트릭 실시간 수집 및
유효 자원 식별

로그 수집 및 감사 추적

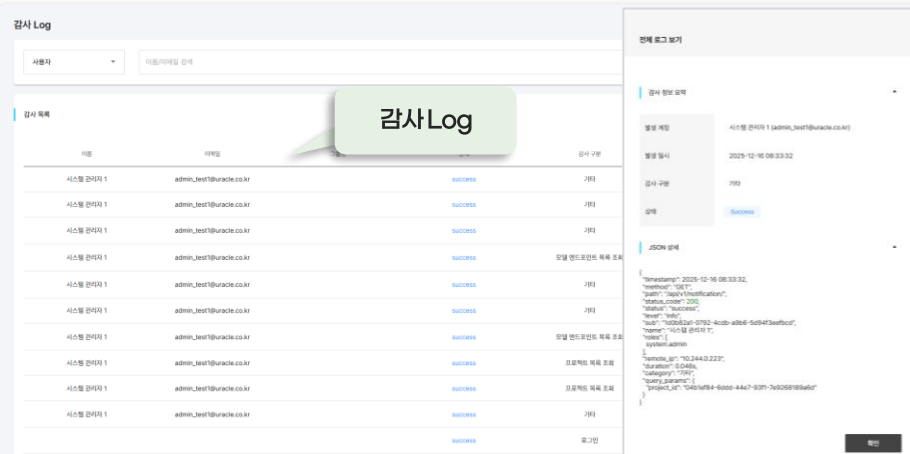
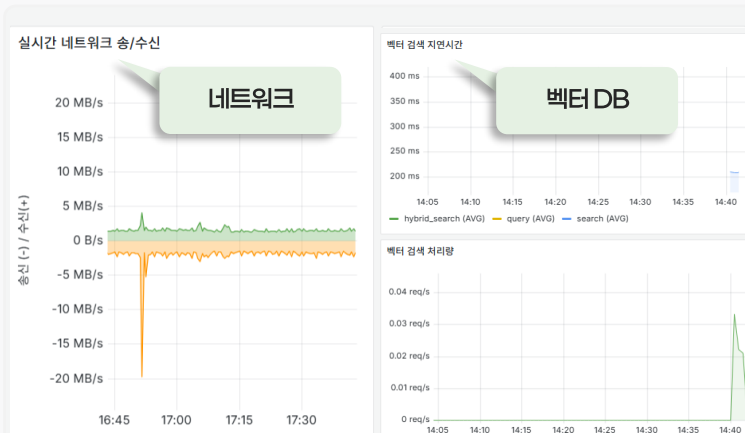
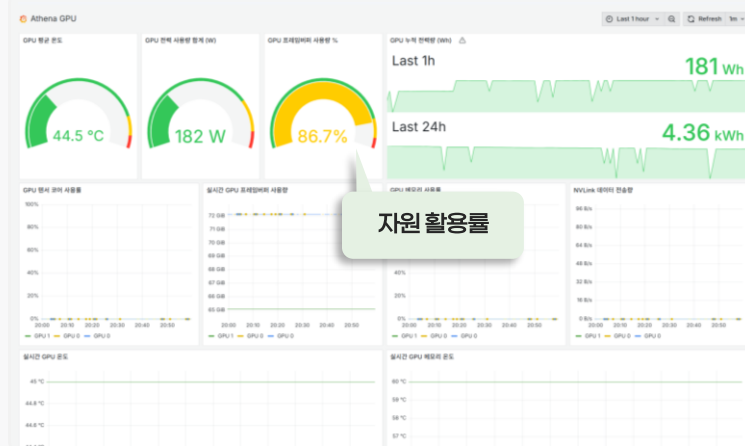
로그인·API 호출·데이터 조회 등
사용자 전 활동을 감사 로그로
기록·추적

실시간 네트워크 트래픽 분석

네트워크 In/Out 트래픽 흐름을
실시간 모니터링하여 이상 징후
즉시 감지

Vector DB · AI 서비스 성능 가시화

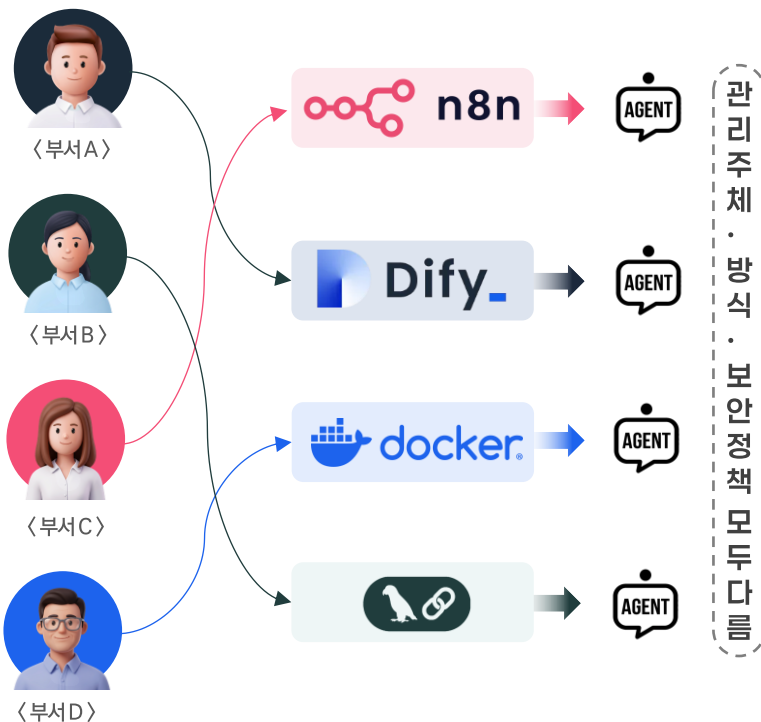
Vector DB 검색 지연시간
·처리량·인덱싱 상태 모니터링으로
RAG 서비스 품질 보장



Our Roadmap : Universal AI Runtime

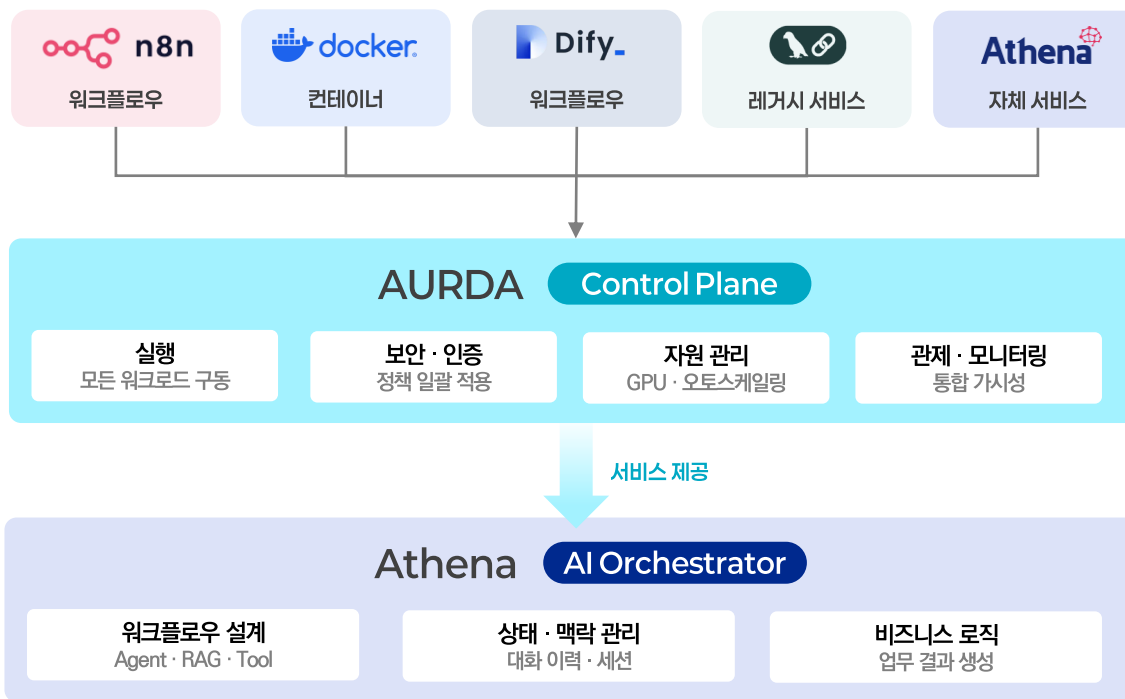
N8N, Docker, 레거시 Agent 등 이기종 AI 서비스를 Athena·AURDA 단일 환경에서 실행·제어합니다.

이기종 AI 환경의 사일로화



서비스별 독립 운영으로 보안 정책·자원 관리·모니터링 기준이 상이하여
전사 AI 거버넌스 수립 불가

Control Plane 기반 통합 운영



Aurda Control Plane이 모든 워크로드의 실행·보안·자원을 단일 기준으로 처리하고
Athena가 비즈니스 로직을 오케스트레이션

Use Cases : 한국평가데이터 - AI 개발플랫폼 구축 (LLMOps + KMS)

표준화된 AI 개발 환경과 사내 지식관리 시스템을 통합 구축하여 전사 AI 서비스 개발·운영 체계를 확립할 예정입니다.



Use Cases : 코리아크레딧뷰로 – 전사 AI 서비스 플랫폼 구축

약 1TB 규모의 사내 문서를 AI 지식으로 전환하고, OCR 기반 비정형 문서 처리부터 전사 RAG 서비스까지 통합 AI 플랫폼을 구축했습니다.

배경

AX(AI Transformation)를 통한
서비스 경쟁력 강화 및 업무 효율성 향상
필요성 증가

고객 주요 과제

AI 추진단 4개 운영 중, 전사 직원이
활용할 수 있는 플랫폼 환경 필요

약 1TB 규모 사내 문서의
체계적 AI 활용 체계 미흡

다양한 비정형 문서 처리를 위한
고정밀 OCR 시스템 필요

적용 플랫폼

아테나

오르다

업스테이지 파서

구현 서비스

AI 플랫폼 구축

- 오픈소스 LLM·RAG·데이터 파이프라인
기반 전사 통합 AI 플랫폼
- 망 분리 환경 대응 아키텍처 구성

고정밀 OCR 시스템

- Upstage 상용 OCR 연동
- 스캔본·사진·표·PDF 등 다양한 비정형
문서 인식 및 민감정보 마스킹 처리

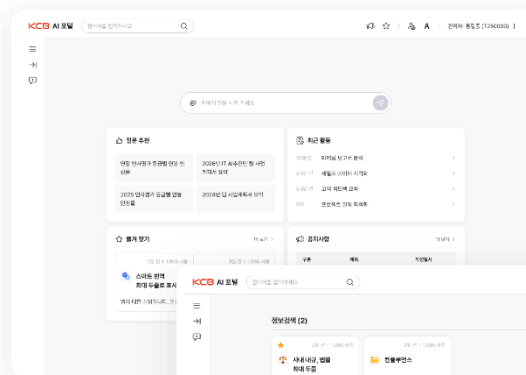
지식 검색· 정보제공 서비스

- 약 1TB 문서 카테고리별 벡터화
- RAG 기반 자연어 질의응답·요약·번역,
근거 문서 함께 제공

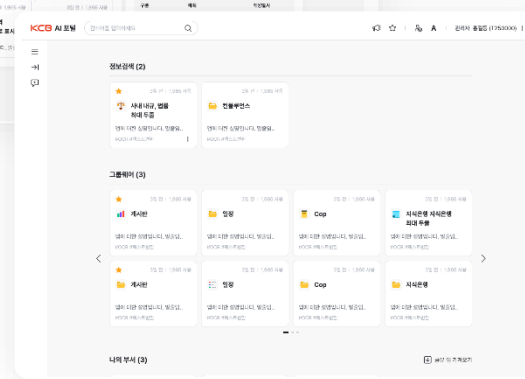
관리자· 운영 환경

- 부서·개인별 AI Agent 생성·공유
- 사용 통계·성능 모니터링
- RBAC 기반 권한 관리 및 DRM 연계

〈사용자 AI포털 메인 UI〉



〈개인별 템플릿 UI〉



약 1TB 규모 사내 문서
AI 지식화로 전사 검색
정확도 95% 이상 확보



Upstage OCR
연동으로 비정형 문서
인식·처리 자동화 실현



전사 표준 AI 플랫폼 기반
부서별 AI 서비스 자율 확장
환경 구축

Use Cases : 현대건설 - AI 서비스 내재화 플랫폼 구축

RAG 기반 사내 지식 활용부터 개인화 AI 비서까지, 전사 표준 AI 서비스 내재화 플랫폼을 구축 완료했습니다.

배경

AI 활용 업무 개선 트렌드 가속화에 따른
전사 AI 활용 환경 제공 필요성 증가

고객 주요 과제

사내 지식의 체계적 AI 활용 체계 부재

보안 환경 내 최신 AI 기술 도입 요구

현업 니즈 맞춤형 서비스의 지속 발굴

적용 플랫폼

아테나

오르다

구현 서비스

일반 챗봇 · 사내 지식 검색

- 범용 업무 챗봇 및 구매계약서 검토
/품질규정 검색 등 특화 지식 기반 서비스

나만의 비서

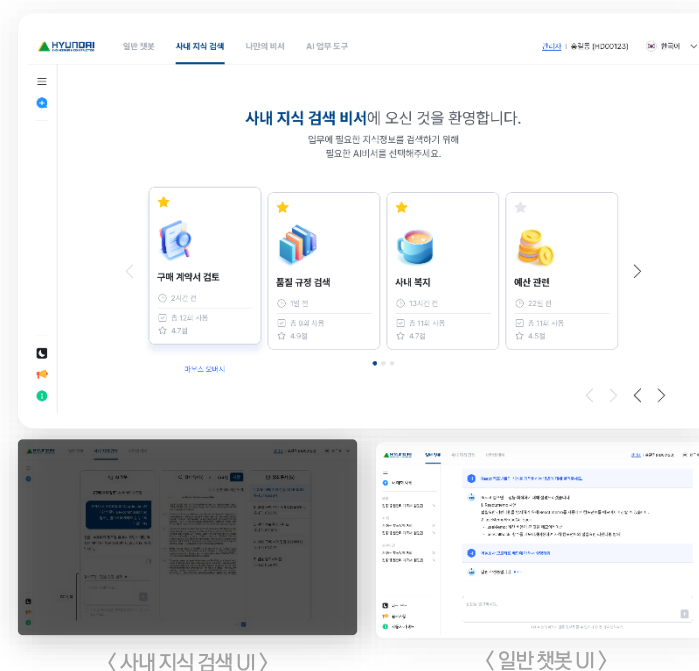
- 사용자별 개인화 AI 비서 생성
- LLM 모델 선택 및 지식 문서 직접 등록

AI 업무 도구

- 문서번역(DeepL), 회의록 자동 생성
- 문서작성 도구 등 업무 자동화 도구 제공

관리자 시스템

- 사용자·부서별 이용통계
- 금칙어 설정 등 응답 품질 및 운영 관리



임직원 업무 생산성 향상
및 사내 지식 활용 가치
극대화



SSO·OTP 기반
안전한 AI 서비스 제공



PC·모바일 통합 서비스로
업무 환경 전반 AI 내재화

Use Cases : 인텔리안테크 – AIOps 플랫폼 및 고객 지원 챗봇 개발

방대한 제품 데이터를 기반으로 24시간 고객 문의에 대응하는 지능형 챗봇과 전사 AI 서비스 확장을 위한 AIOps 플랫폼을 구축했습니다.

배경

제품 사양·모델 비교·호환성 등 다양한
고객 문의에 대응하는 지능형 챗봇 및
전사 AI 서비스 확장을 위한 전략적
AIOps 플랫폼 필요

고객 주요 과제

고객 문의 대응 자동화 및
24시간 지원 체계 부재

제품 사양·매뉴얼 등
방대한 내부 데이터의 AI 활용 체계 미흡

향후 다양한 AI 서비스 확장을 위한
표준 플랫폼 부재

적용 플랫폼

아테나

오르다

구현 서비스

AIOps 플랫폼 구축

- Kubernetes 기반 전사 AI 운영 플랫폼 구축
- 클라우드·온프레미스 모두 대응

고객 지원 챗봇

- 제품 사양·모델 비교·재고·수리 현황 등 내부 데이터 기반 답변 제공
- 24시간 한국어 지원

Agent Builder 및 CI/CD

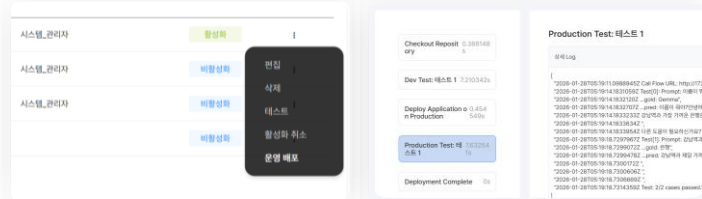
- 노코드 기반 AI 에이전트 설계·배포 환경 제공
- 모델 변경 시 자동 테스트 및 재배포

AI 거버넌스 및 보안

- 정책 기반 데이터 접근 제어·자원 할당·비용 통제
- 개인정보 마스킹 및 RBAC 기반 권한 관리



< 홈페이지 내 고객 챗봇 UI 예시 >



< CI/CD UI >



제품 전문 챗봇으로 고객
문의 24시간 자동 대응 및
지원 응대 시간 단축



기간제 시스템 연동으로
실시간 정보 기반 능동적
고객 응대 실현



전사 AIOps 플랫폼 기반
신규 AI 서비스 신속 확장
체계 확보

감사합니다.

H. uracle.co.kr
E. sales@uracle.co.kr
T. 02-3479-4400

uracle^U