

ORACLE

The Future of Data and App Dev with Generative AI

오라클이 제시하는 생성형 AI와 데이터 관리 비전

Jung Ok Nah

VP / Head of Cloud Engineering, Oracle Korea

Jan 25, 2024



Oracle's Enterprise DNA

44+ years of helping
customers manage
the world's most
important data





생성형 AI의 시대

What is the Future of Data and App Dev with Generative AI?

생성형 AI의 중요성

IDC releases top Customer Experience (CX) predictions for 2024 and beyond.

 IDC FutureScape

60% of Asia's Top Firms to Pivot CX Budgets to GenAI, Phasing Out Old Software by 2027

By 2027, 60% of A2000 firms will reallocate 35% of CX spend on emerging products and skills as improved Generative AI renders entire software categories obsolete.

- By Lawrence Cheok, IDC, Asia Pacific

The public cloud services (PCS) market in South Korea is expected to grow continuously as enterprises expand their cloud workloads and the finance sector expands its cloud-based environment based on various references.

Generative AI and sustainability are expected to be major drivers in the market.

- By Sarah Park and Shouvik Nag, December 2023

생성형 AI와 기업의 과제



The Goal

Enterprises wish to leverage generative AI to process relevant data with natural language.

The Challenge

Pre-training, fine-tuning, or continuous training of LLMs on corporate knowledge bases is time consuming and expensive.

새로운 데이터 관리와 앱 개발의 미래



What is the future of Data and App Dev with Gen AI?

Data and App Dev
Innovations will enable radical
increases in productivity

How?

We will **generate** data solutions
instead of hand-coding them

Generate to Innovate!

Using 3 Transformational Strategies

오라클이 제시하는 미래 데이터 전략

Strategy 1 : Generate Use-centric Data

Strategy 2 : Generative AI

Strategy 3 : Generate Apps

오늘날 데이터를 다루는 모든 사람들은
데이터로 무엇을 하고 싶은지 뿐만이 아니라
어떻게 해야 데이터를 얻을 수 있는지도 시스템에 알려 줘야 합니다.

데이터 전문가



개발자



Low Code 개발자



사용자



이제 단순히 우리가 의도하는 결과를 서술(declare) 해 주기만 하면
데이터시스템은 그 결과를 생성(generate)하게 될 것입니다.

데이터 전문가



Declare
Data Intent

개발자



Declare
Data Usage Intent

Low Code 개발자



Declare
App Intent

사용자



Declare
User Intent

오라클 데이터 관리 및 App 개발 혁신 전략

오라클의 3가지 혁신 지원 전략 (HOW)

이제 단순히 우리가 의도하는 결과를 서술(declare) 해 주기만 하면
데이터시스템은 그 결과를 생성(generate)하게 될 것입니다.

데이터 전문가



Declare
Data Intent

개발자



Declare
Data Usage Intent

Low-Code 개발자



Declare
App Intent

사용자



Declare
User Intent

#1. Generate Use-centric data : 데이터 생성

데이터 전문가/ 개발자:

원하는 데이터 또는 데이터 사용 방법을 서술하면
DBMS가 데이터를 자동 관리

#2. Generative AI : 생성형 AI

사용자: 원하는 의도를 자연어로 서술하면 시스템이
자동 생성

#3. Generate Apps : 앱 생성

Low-code 개발자: 원하는 앱의 구조를 서술하면
APEX가 app frame 자동 생성

오라클 혁신 전략 #1: 데이터 생성(Generate Use-centric Data)

새로운 패러다임

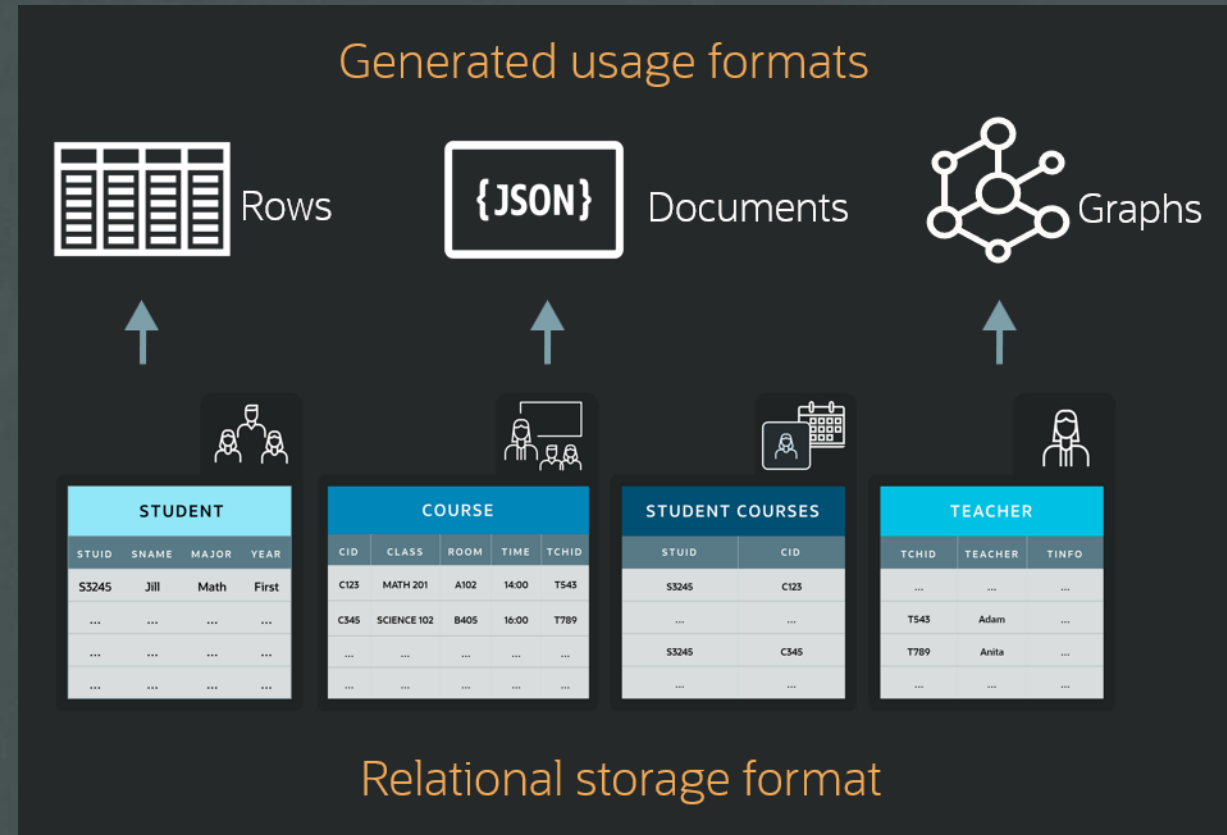
- ✓ 하나의 모델로 저장
- ✓ 사용자가 원하는 데이터 형태를 서술하면
데이터 변환과 검색/수정/저장은
자동으로 시스템이 지원

생산성을 높이는 오라클의 지원 방안

- ✓ 데이터 저장은 관계형 기반
- ✓ 데이터 저장/검색/수정은 사용자가 원하는
데이터 형태(JSON 혹은 Graph)로 선언
→ 어떠한 변환 작업도 필요 없으며 원하는 형태만
서술하면 시스템이 자동으로 변환 및 처리해 줌

최고의 생산성/효율성 지원 가능

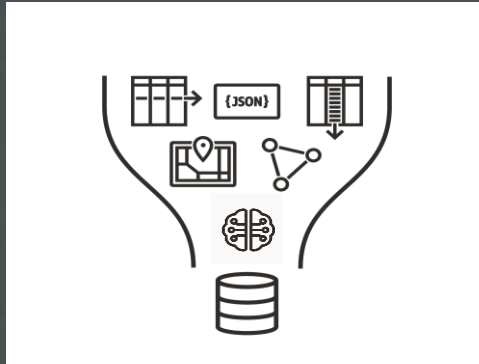
- ✓ 데이터 변환 작업과 통합 비용이 근본적으로 사라짐



오라클 혁신 전략 #1: 데이터 생성(Generate Use-centric Data)

오라클의 접근법: **All-in-One** Modern Apps을 개발하고 운영하기 위한 **Complete & Simple** 데이터 플랫폼을 제공

“Complete by Converged features, Simple by Autonomous Database”



Complete : Converged

- 모든 데이터 타입, 워크로드, 개발 스타일을 위한 **Complete** 지원
- 기존 장점(일관성, 확장성, 가용성, 보안) 유지/발전



Simple : Autonomous

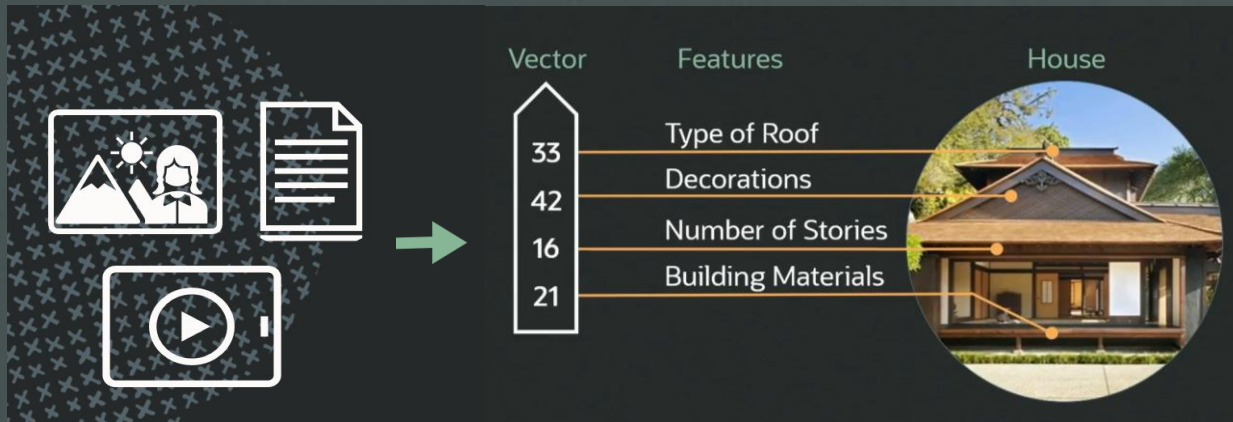
- 개발자, 운영자, 데이터 분석가를 위한 **Simplest** 클라우드 데이터베이스 서비스
- 데이터베이스 관리(운영, 튜닝, 보안강화)의 자동화로 본연의 업무에 집중

오라클 혁신 전략 #2: 생성형 AI (Generative AI)

Vector Search: 오라클 DB 내에서 Vector Type 제공

- AI vector는 비정형 데이터 자체가 아닌, 그 비정형 데이터의 semantic content를 나타내는 특정 feature들을 담고 있는 데이터
- 비정형 데이터에 deep learning모델에 기반한 특정 transformation을 적용함으로써 얻어짐

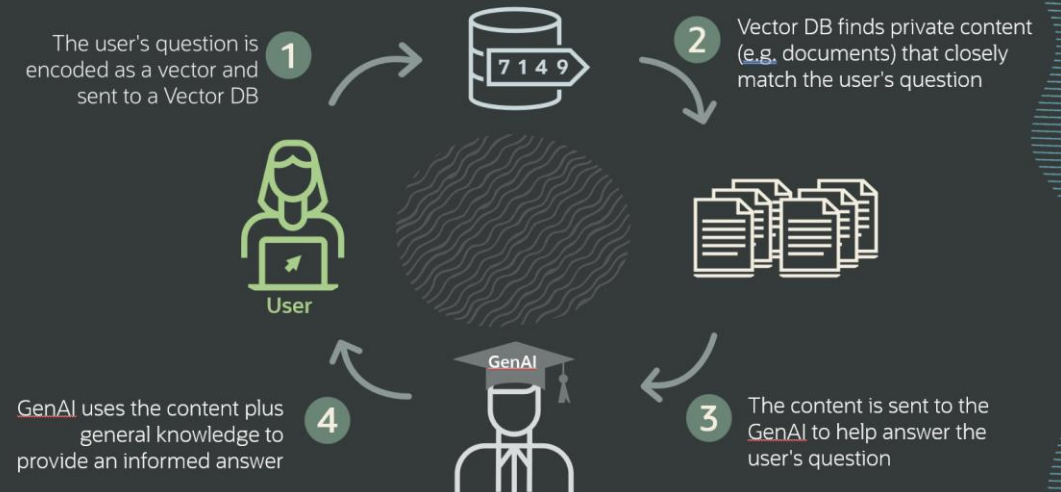
AI Vector Search 를 통하여 비정형 데이터 검색



Retrieval Augmented Generation (RAG)

- 기존 데이터/지식과 결합하여 정확한 대답을 만들기 위한 방법
- 벡터 데이터베이스는 질문에 대한 답변에 필요한 상세(Detailed)하고 비공개적인(Private) 콘텐츠를 검색하여 Generative AI를 강화합니다.

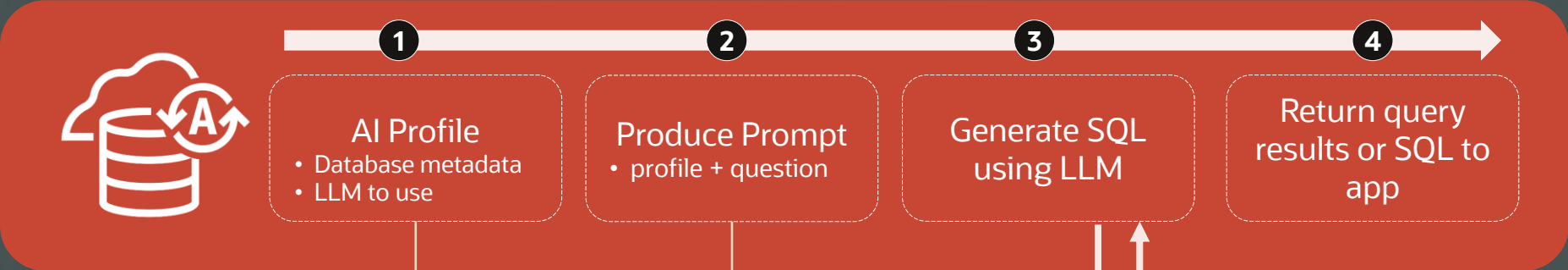
Retrieval Augmented Generation works like this



오라클 데이터베이스 Generate SQL using LLM

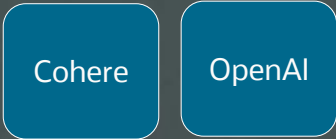


SELECT AI <ACTION>
Give me the average salary of
employees in each department;



dbms_cloud_ai.create_profile

dbms_cloud_ai.set_profile



REST

ACTION

runsql	return sql result set
narrate	return a conversational result
showsql	return the generated query
chat	general AI-chat conversation



OCI Generative AI Service

고성능 사전 학습 모델

최소한의 노력으로 고객의 비즈니스 요구 사항을 충족할 수 있도록 Meta 및 Cohere의 고품질 모델을 제공

Fine-tuning 지원

고객의 데이터로 모델을 Fine-Tuning하여 특정 문제를 해결하거나 특정 도메인을 최적화

OCI 완전 관리형 호스팅 서비스

모든 처리 및 데이터 저장은 OCI 내부에서 이루어지며, 지역간 클라우드간 통신 없음

LLM 데이터 스토리지 및 LLM 프로세싱, 모델 버전, 업그레이드 관리는 OCI의 완전 관리형 서비스 형태로 제공

비공개적이고 안전하게 고객 데이터 보호

고객이 제공한 학습 및 추론 데이터는 안전하게 보호되며 다른 고객이 접근 불가. Cohere 또는 메타와 데이터가 공유되지 않음.

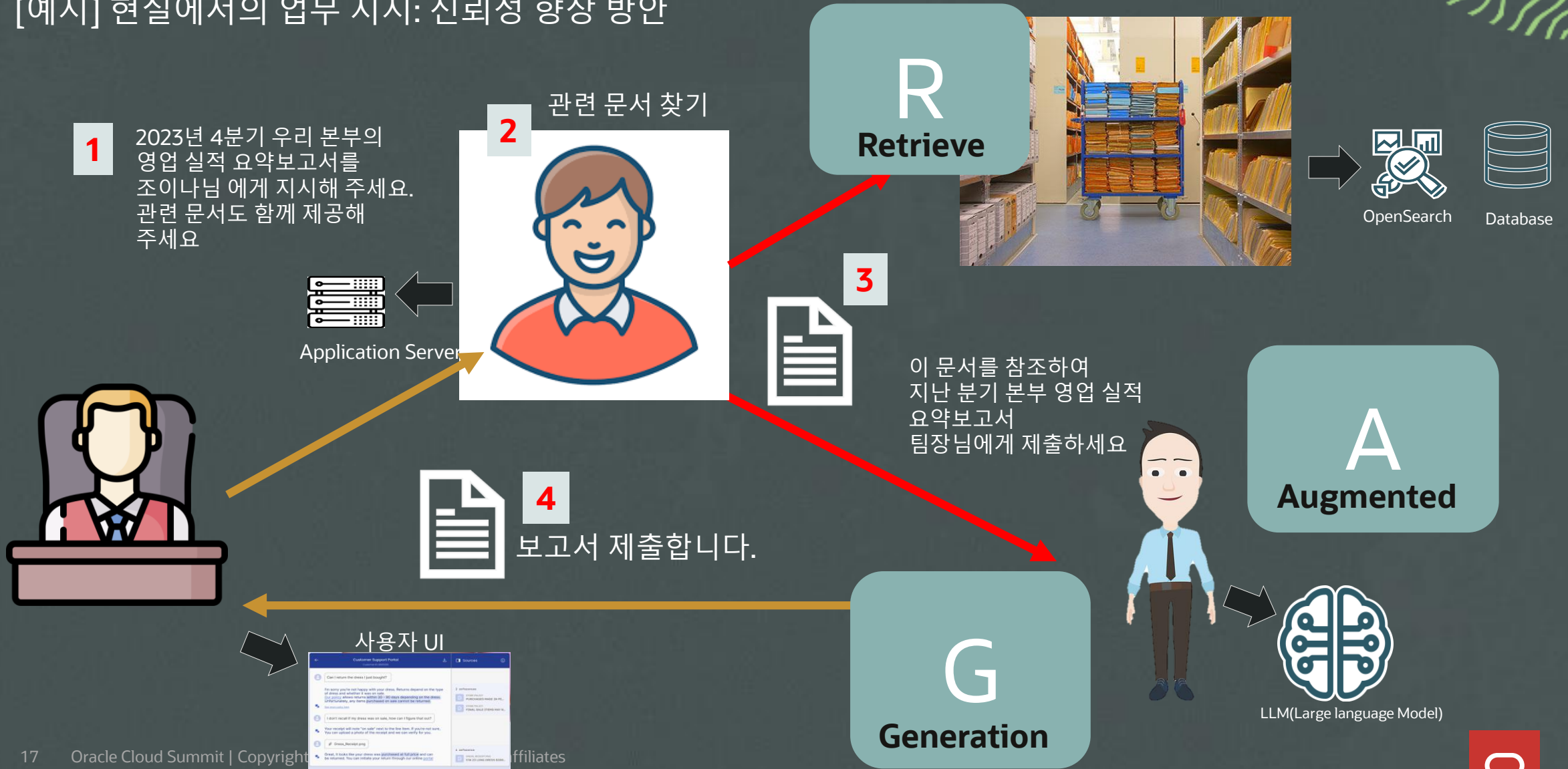
Global announcement

GA Today!

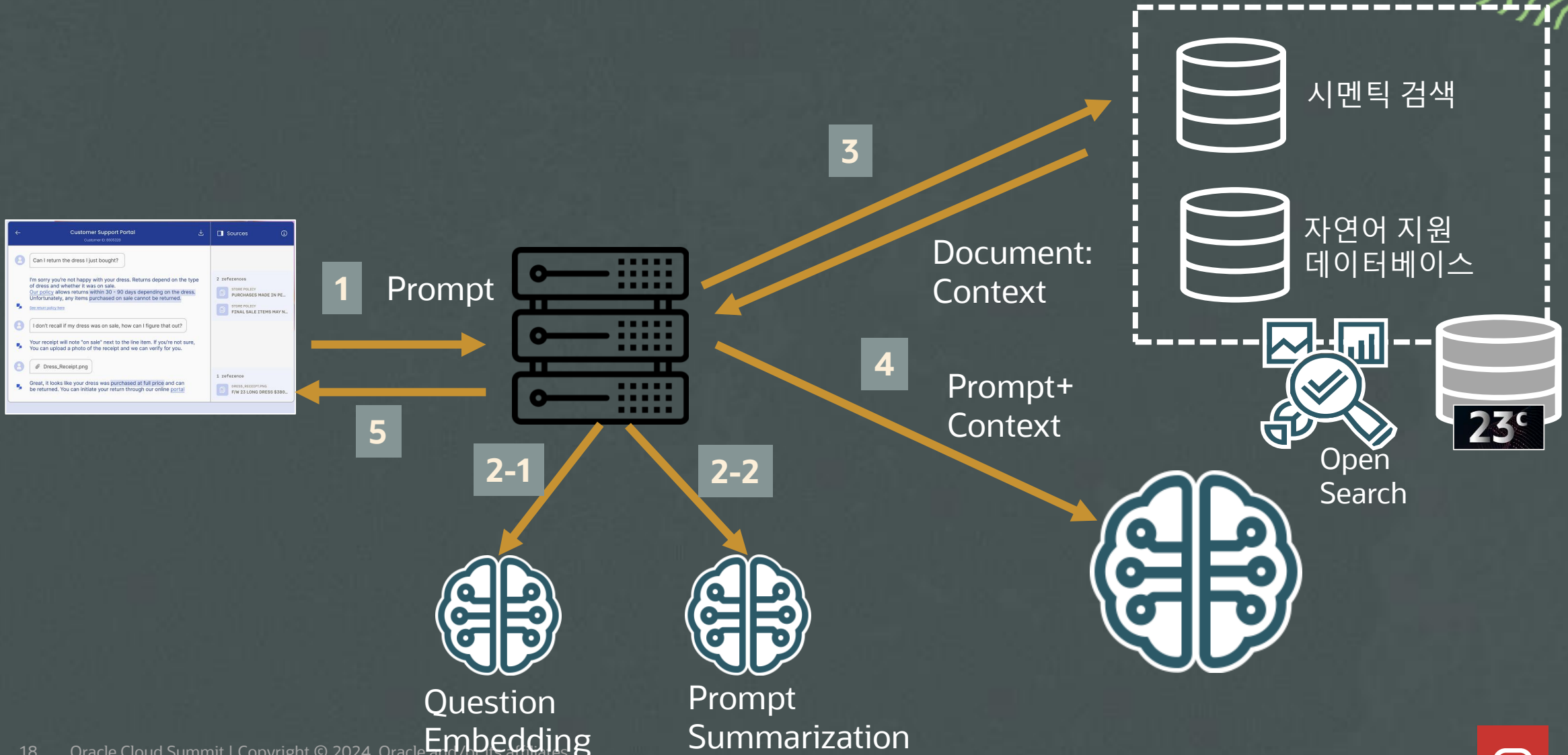
The screenshot shows the Oracle Cloud Generative AI overview page. The top navigation bar includes the Oracle Cloud logo, a search bar, and the region 'US West (Phoenix)'. The main content area is titled 'Generative AI overview' and features a dark banner with the text: 'Power your apps with large language models and generative AI. OCI Generative AI is a fully managed service that provides a set of state-of-the-art, customizable LLMs that cover a wide range of use cases for text generation. Use the playground to try out the models out-of-the-box or create and host your own fine-tuned custom models based on your own data on dedicated AI clusters.' Below this, there are three metrics: 'Dedicated AI clusters' (7), 'Custom models' (3), and 'Active endpoints' (12). A 'Get started' section includes a 'Go to playground' button and a 'Playground' card. The 'Playground' card describes it as a visual interface for exploring hosted pretrained and custom models. Below the 'Get started' section are three cards: 'Dedicated AI clusters' (Spin up dedicated hardware units for fine-tuning custom models and hosting them), 'Custom models' (Create custom models by fine-tuning the base models with your own dataset), and 'Endpoints' (Create and manage endpoints to host your custom models). On the right side, there is a 'Resources' section with links to 'All documentation', 'Rest API reference', 'Workshops', 'Tutorials', and 'Pricing'.

RAG (Retrieval Augmented Generation)

[예시] 현실에서의 업무 지시: 신뢰성 향상 방안



RAG(Retrieval Augmented Generation) 시스템 아키텍처



오라클 혁신 전략 #3: 앱 생성(Generate Apps)

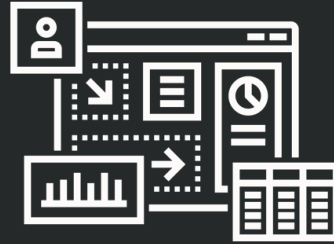
시각적 도구와 메타데이터를 통해
앱의 의도를 선언하면 자동적으로 앱이 생성됨

APEX는 workflow, mission critical SaaS,
mobile 등의 개발에도 사용될 수 있음

사용자가 선언한 데이터 구조를 활용하여
앱 생성 프레임 지원
→ Enterprise level의 low code 지원

현재 **Cerner Application** 개발에 **APEX** 사용 중

Low-code platforms generate apps by
declaring the intent of the app using visual tools and metadata



Low-code is far more productive than coding

Today, low-code is great for generating simple apps

The challenge for low-code is generating
enterprise-grade apps

Oracle APEX is the leader in addressing this
challenge

APEX is extending low-code to handle more app use cases



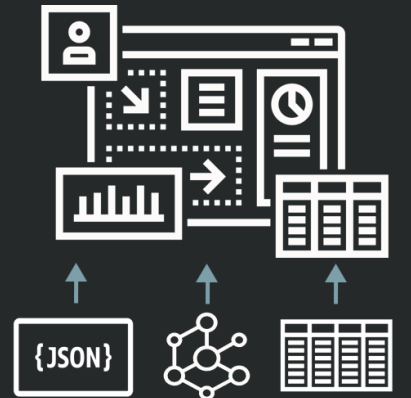
In the near future,
the large majority of
enterprise apps will be built
using low-code

Low-code workflow

Low-code for mission critical SaaS

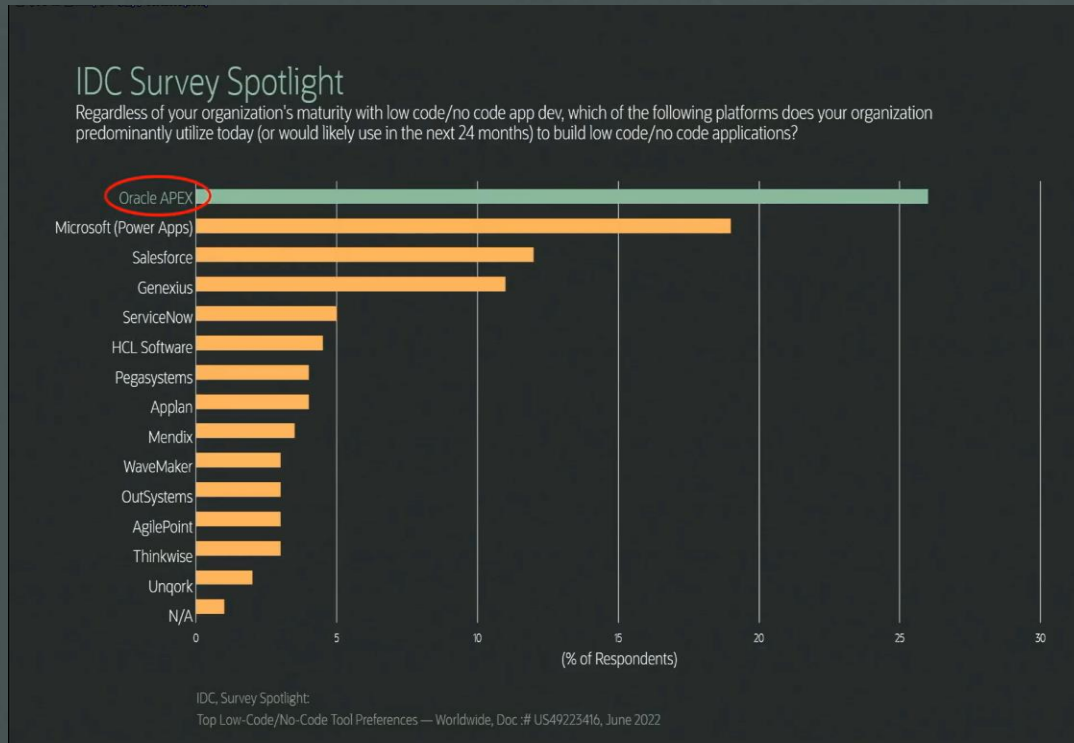
Low-code for mobile

Low-code interoperability using
GraphQL, and JavaScript



오라클 혁신 전략 #3: 앱 생성(Generate Apps)

Oracle APEX : 데이터베이스 중심의 웹 애플리케이션 개발 플랫폼으로서 low code 지원
APEX는 현재 가장 많이 활용되는 low code 솔루션 (IDC Survey, 2022)



- ✓ 데스크탑과 모바일 앱 개발
- ✓ 레고 방식으로 원하는 위젯을 선택하면 해당 코드를 자동으로 생성



- ✓ 데이터베이스 데이터의 시각화 및 관리
- ✓ SQL과 DB 기능을 최대한으로 지원

오라클의 AI 전략



Oracle AI Strategies: OCI Native AI Services & GenAI

엔터프라이즈 기업의 생성형 AI 요구사항

생성형 AI를 통한 비즈니스 혁신



Legacy Data with Gen AI

Oracle은 클라우드 애플리케이션 및 데이터베이스 포트폴리오 전반에 생성형 AI 통합



Smart and Secure Gen AI

OCI Gen AI Service는 사용하기 편리하고 안전한 SOTA LLM 모델 제공, Fine-Tuning 및 전용 모델 지원



Infrastructure for Gen AI

OCI는 생성형 AI의 모델 학습 및 운영을 위한 초고속 AI 인프라 제공

오라클 생성형 AI 의 특징점

시작부터 “엔터프라이즈 고객 요구”에 충족할 수 있도록 설계되었습니다.



Designed for enterprise

최고의 데이터 관리 플랫폼, AI 인프라 및 비즈니스 앱을 갖춘 우수한 모델에 Access하십시오



Tailored to your data

OCI Generative AI 서비스를 통해 자신의 데이터를 활용해 모델을 개선할 수 있습니다.



Security and privacy

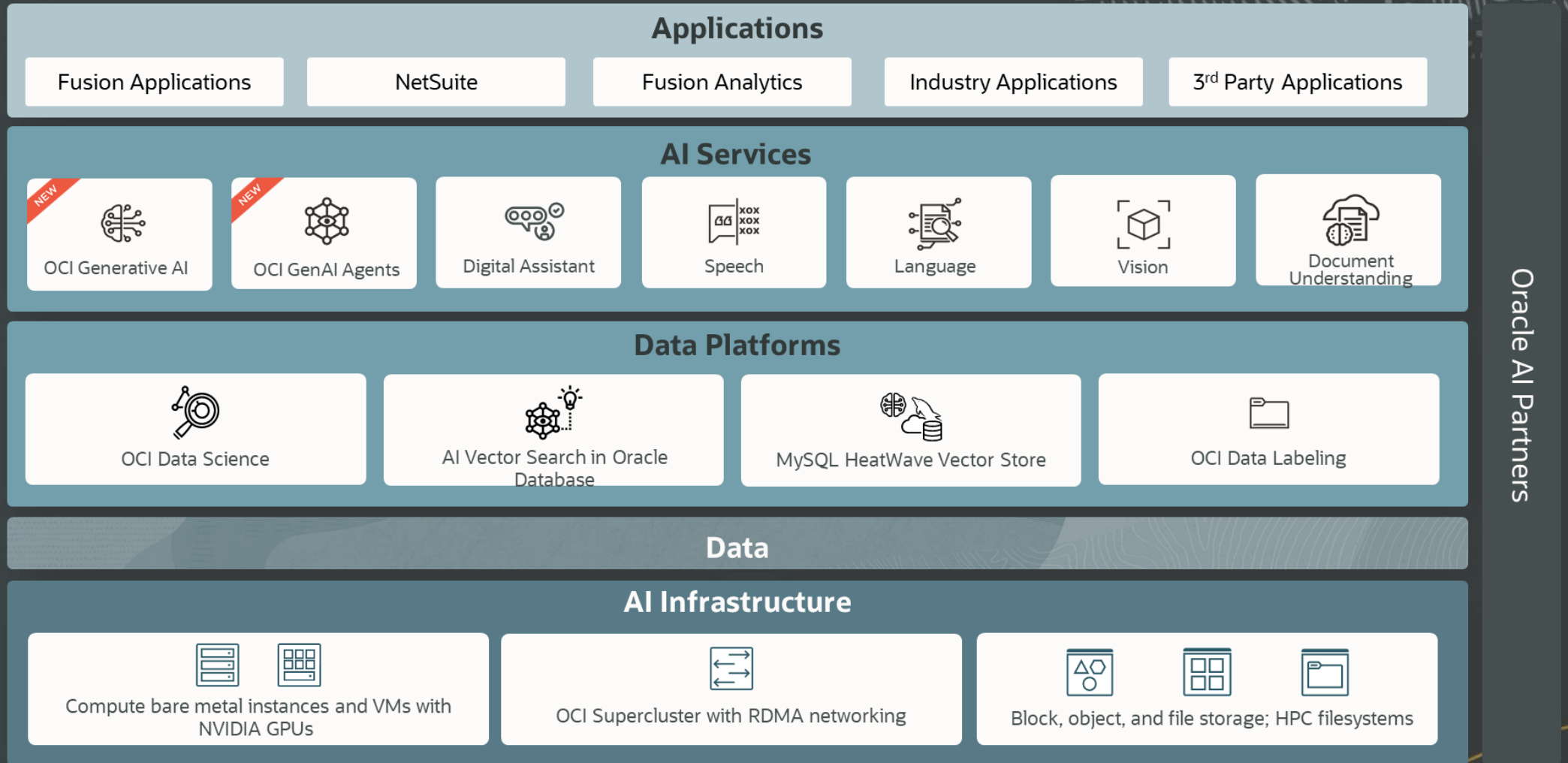
귀하의 데이터는 LLM 제공업체와 공유되거나 다른 고객에게 공개되지 않습니다.



Predictable performance and pricing

신뢰할 수 있는 성능과 가격을 제공하는 전용 단일 테넌트 AI 클러스터에 모델을 배포하십시오.

오라클 클라우드 Ecosystem: AI Stack



기업 고객을 위한 오라클 생성형 AI/Data Platform



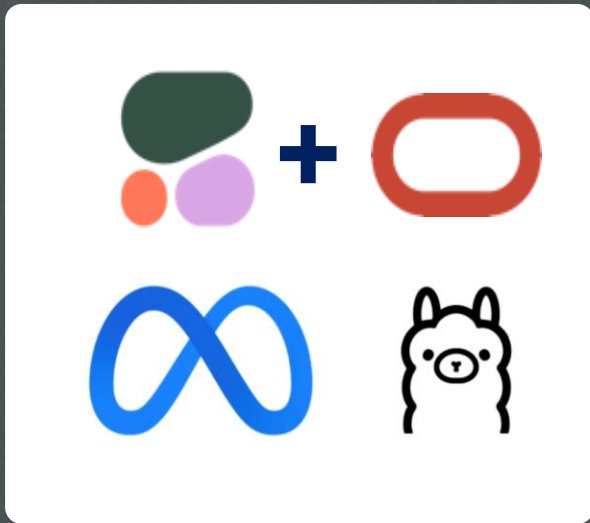
Model



DATA



INFRASTRUCTURE



Oracle
Database 23C



OpenSearch



MySQL
Heatwave



ADW



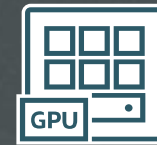
ATP



EXACI



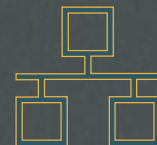
EXACC



H100/A100
/A10 GPU



Cloud Native
OKE



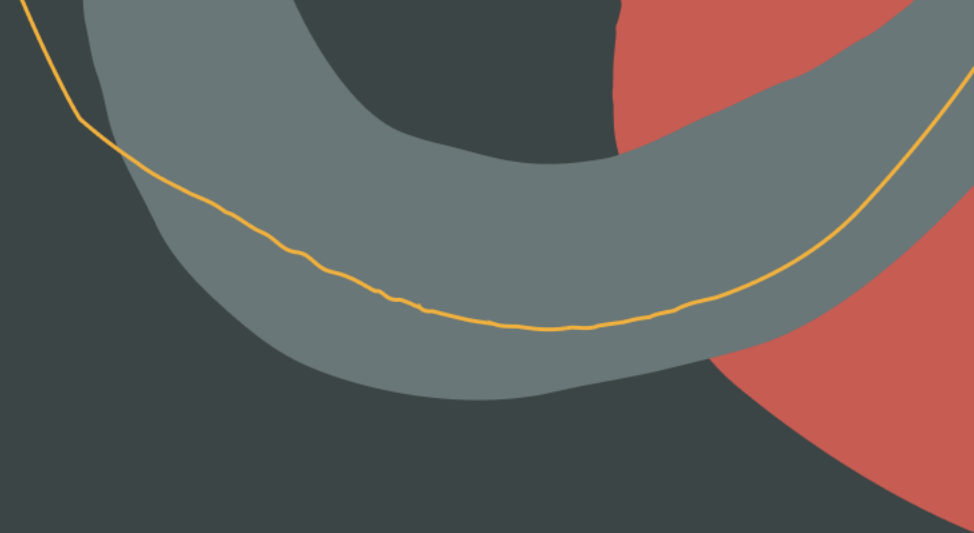
Super Cluster



Chatbot

맺음말:

- 데이터는 다양한 포맷으로 생성, 저장, 처리, 관리되고 있고, 이러한 다양한 데이터를 어떻게 기업의 지식(Knowledge)으로 만드는지가 기업 경쟁력의 핵심입니다. 복잡하고 다양한 데이터 관리 문제를 해결하기 위해서 오라클의 **Converged Database**를 활용하시기 바랍니다.
- 오라클이 제공하는 **Vector Database, Generate SQL** 등 **최신 AI 기반 기술**을 통해 이제 데이터 전문가, 개발자, 사용자는 **원하는 데이터와 앱 생성을 쉽게** 할 수 있습니다.
- 오라클은 **End-to-End Enterprise AI/Data Platform**을 통해서 **안전한 Gen AI** 서비스 개발 및 운영 플랫폼을 **고객이 원하는 모든 형태의 분산 클라우드(Distributed cloud)**로 제공합니다.



Thank you

Jung Ok Nah

VP, Head of Cloud Engineering
Oracle Korea

ORACLE