

ORACLE

벡터DB와 AI 통합에 기반한 차세대 데이터베이스 활용

- Oracle DB 23ai를 중심으로 -

장성우

Senior Director, Cloud Engineering

Oracle Korea

Mar 6, 2025



발표 주요 내용

생성형 AI를 활용한 App 개발의 자동화 → 기업용 App에도 적용 가능?

기업용 App 개발의 자동화를 위한 선결 조건

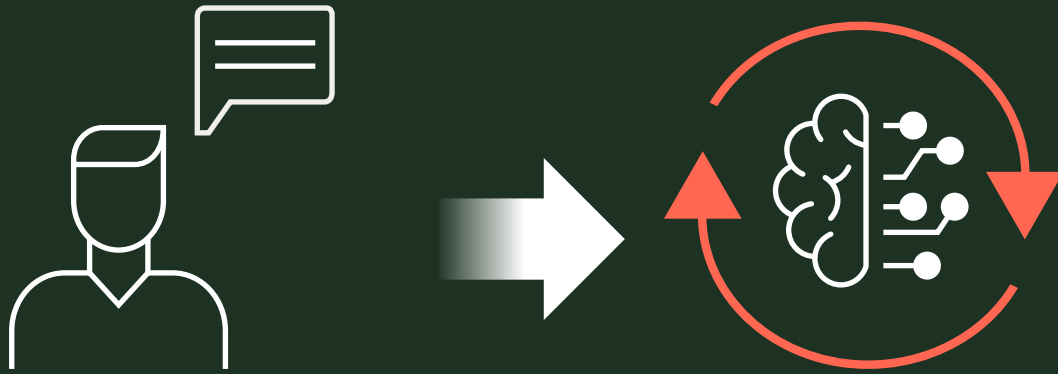
- 다양한 생성형 AI 관련 기능 지원
 - App 모듈간의 상호 독립성 지원 → 핵심은 data independency
 - Enterprise-level Robustness(Data Confidentiality/Consistency/Evolution) 지원
 - App 개발 간소화를 위한 low-code 및 데이터 관리 자동화 지원
- 생성형 AI의 장점의 극대화를 지원하기 위한 개선된 개발 인프라 필요

차세대 데이터베이스는 생성형 AI를 위한 견고한 데이터 인프라 기능을 제공할 수 있어야 함

Oracle DB 23ai는 생성형 AI 기반의 기업용 App 개발 자동화를 지원하기 위한 개발 인프라(GenDev) 제공

- In-DB ML, Vector Search & RAG, SQL Generation
- Unified Data Model with JSON/Relational Duality View
- Data Consistency with Value-Based Concurrency Control
- Free low-code tool APEX
- Autonomous DB

우리가 원하는 AI 기반 App Dev의 미래 모습과 한계



App에 대해 원하는 것을 말하면

AI가 App을 자동으로 생성

하지만 현재 AI는 간단한 앱의 초벌을 생성할 수는 있지만 복잡한 기업용 앱을 생성하는 것은 아직 현실적이지 않음

- ✓ 현재 개발자가 AI를 사용하는 주요 방식은 잘 정의된 작업에 대해 중간 규모 정도의 코드를 생성하여 현재 개발 프로세스를 가속화하는 것
- ✓ 이러한 방법을 대규모 기업용 앱 개발에 바로 적용할 수 있을까?
- ✓ 하지만 이는 엔터프라이즈 앱을 만드는 실행 가능한 방법이 아님. 사람이 납득하지 못하는 수천 줄의 AI 생성 코드에 비즈니스의 명운을 걸 수는 없음
- ✓ 엔터프라이즈 앱이 정확하고, 이해하기 쉬우며, 신뢰할 수 있는지 확신할 수 있어야 함

기업용 앱 개발 인프라에서 고려해야 할, AI 코드 생성의 이점을 제한하는 세 가지 문제점



앱 내 광범위하게
존재하는 종속성 및
가정은 코드 생성 및
진화를 어렵게 만듦



앱을 작동시키는데
필요한 코드의 양이
많아 앱을 이해하고,
검증하고, 유지
관리하기가 어려움



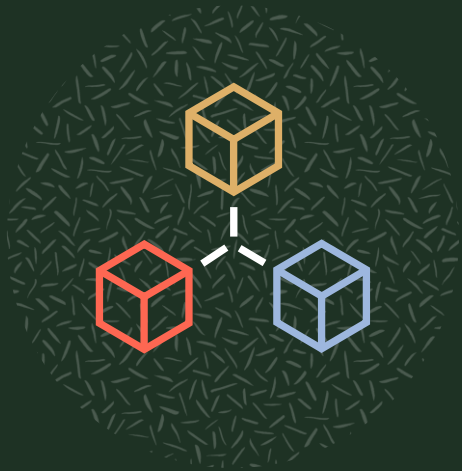
기업용 앱에 걸맞는
견고성이 앱 전체에
걸쳐 광범위하게
코딩되어야 함

- ✓ AI와 같은 혁신적인 기술의 활용을 위해 우리는 접근 방식 자체를 바꿀 필요가 있음
- ✓ 말을 대체하는 것만으로 자동차의 모든 혜택을 누릴 수 없는 것과 유사
- ✓ 포장 도로와 같은 자동차 중심 인프라가 반드시 있어야 함
- ✓ 개발 인프라를 변경하지 않고 AI를 사용하여 코드 생성을 가속화하는 것은 마치 경주용 자동차를 말산책로에서 운전하는 것과 같음

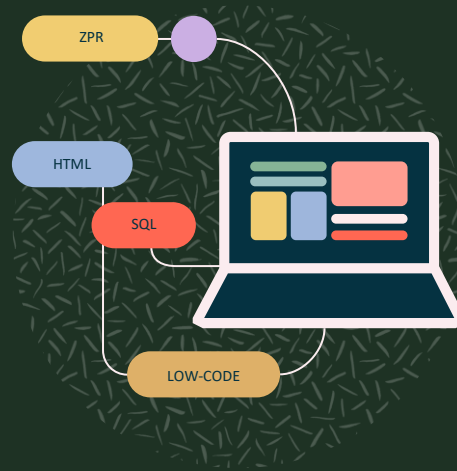


AI 기반으로 효과적으로 기업용 앱 생성을 지원할 수 있는 새로운 인프라스트럭처가 제공되어야 함

AI 기반 기업용 앱 개발을 지원하기 위한 인프라의 조건



독립적이고 신뢰할 수 있으며 진화 가능한 모듈들로 앱을 생성할 수 있어야 함



생성형 AI 연계 지원 및 간결하고 선언적인 언어를 사용하여 이해가능한 앱을 생성해야 함



엔터프라이즈 앱에 필요한 확장성, 안정성, 일관성 및 관리 자동화 등의 견고성을 자동적으로 지원할 수 있어야 함

Generative Development for Enterprise



GenDev

AI의 장점을 수용하면서도
위험을 완화하기 위해
Oracle DB 23ai에 도입된
AI 중심 개발 지원 인프라

Data is the core of GenDev

- ✓ AI support with **In-DB ML, Vector Search & SQL Generation**
- ✓ Unified Model with **JSON Relational Duality View**
- ✓ Robustness with **Value-based Concurrency Control & Security**

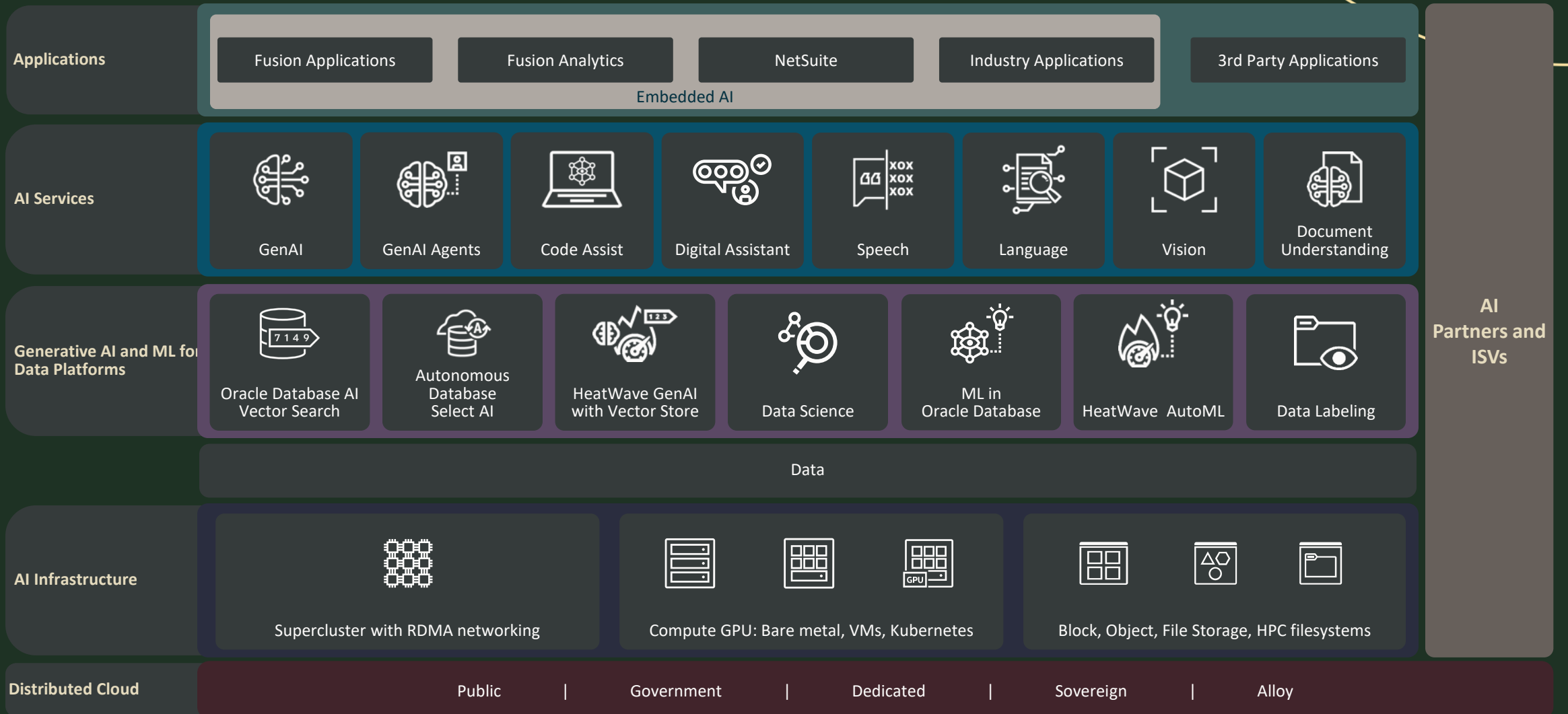
Low-code App Dev with APEX

- ✓ Free but strong with Blueprint features for enterprise AppDev

Autonomous Data Management

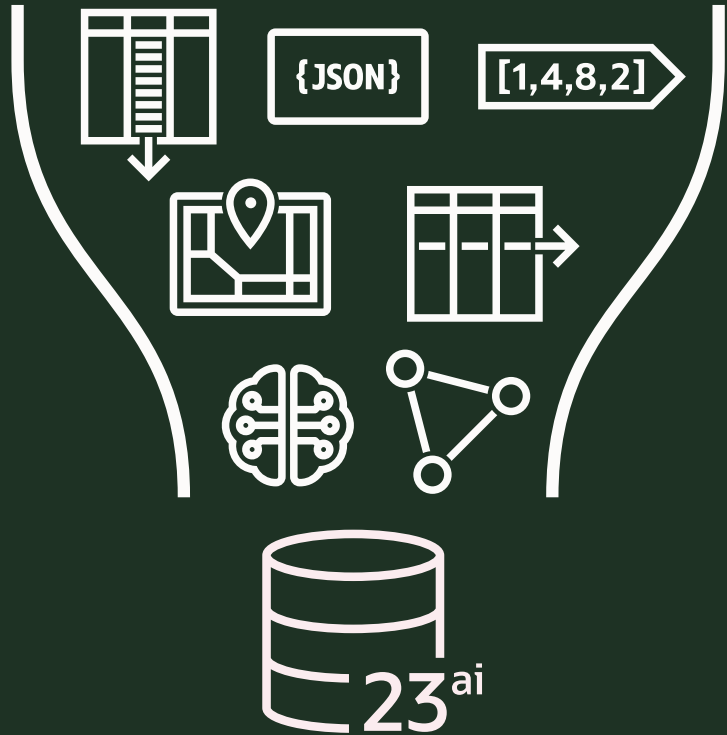
이러한 기능들을 기본 인프라로 제공함으로써 개발자들은 앱의 본질적인 기능에만 집중할 수 있도록 지원

The Oracle AI stack



Oracle Database

통합된 유희(Converged) 데이터베이스



- 모든 최신 데이터 유희, 워크로드 및 개발 스택에 적합
- 오라클 데이터베이스 23ai의 AI 벡터 검색 및 데이터베이스 내 AI/ML을 통한 데이터 분석 AI 지원
- 모든 데이터에 대한 최신 애플리케이션 및 분석 가속화

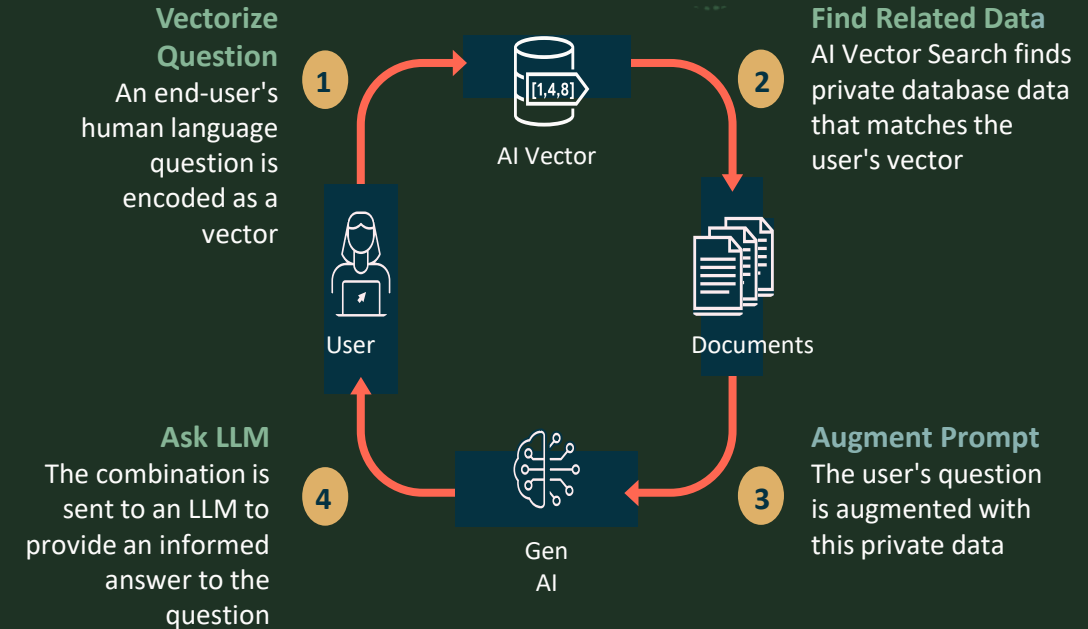


Database 23ai GenDev는 DB 내에서 다양한 AI 기능 직접 지원

- In-DB Machine Learning
- Vector Type and AI Vector Search
- Performant Vector Search with RAC, Parallel SQL and Security
- RAG(Retrieval Augmented Generation)
- SQL Generation with Select AI

This query searches for a combination of customer data (customer budget), product data (homes for rent), and **image** data

```
SELECT ...  
FROM houses_for_rent  
WHERE cost <= (SELECT budget FROM attendee ...)  
AND city in (SELECT event_city FROM attendee_events ...)  
ORDER BY vector_distance(rental_house_vector, :input_vector)
```



비정형 데이터에 대한 의미(Semantic) 검색을 지원하는 AI Vector Search

AI 벡터 검색(AI Vector Search)로 비정형 데이터 활용하기



벡터는 착윽익락곡 학늑 슣작의 식퀘스룩, 덕익턱의 좇육함 "**특징**"을 폭착학늑 덕 삭윽뵐

벡터는 기붓 닛억낙 픽셈의 악늑 덕익턱의 **의미적 내용**을 낙탁넵

벡터는 딕럭닛 입벡딛 목뵐을 삭윽학역 섯섯뵐



오라클 데이터베이스의 AI 벡터 검색으로 GenAI 파이프라인 강화

엡터프락익즈 데익터를 활웃한 겁새 좇갓 섯섯(RAG: Retrieval Augmented Generation)

Vectorize Question (질문 벡터화)

삭웃작익 작온익 질문을 벡터로
익콕딛되익집.

1



AI Vector

Find Related Data (연관 데이터 찾기)

AI 벡터 검색을 삭웃작익 벡터와
익치학늬 벡터 데익터벡익스
데익터를 찾음

2



Documents

ORACLE Database 23^{ai}

300+ new features
and enhancements

Augment Prompt (증강 프롬프트)

삭웃작익 질문의 벡터 데익터를
복갓딛

3



GenAI

Ask LLM (LLM에 묻기)

질문에 대한 첫번익 익갓갓 답변을
젝갓학익 익해 LLM으로 족솅딛

4



User



RAG 및 Oracle Autonomous Database Select AI를 사용하여 정확한 통찰력을 얻기

자연어 대화를 통해 모든 데이터에서 색록을
인식하고 빠르게 얻고, 오류를 방지할 수 있음

데이터 내장된 자연어 질문-SQL 쿼리 기능
(NL2SQL)을 사용하여 기존 애플리케이션
색록을 애플리케이션

Cohere AI, Azure OpenAI, OpenAI 및 OCI
Generative AI의 다양한 LLM과 통합

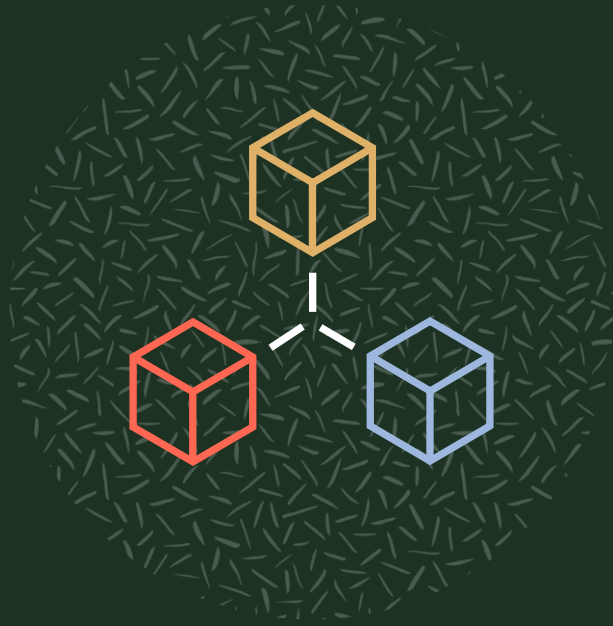
데이터베이스의 복잡성 및 애플리케이션
통합을 가속

The screenshot shows the Oracle Autonomous Database Query Builder interface. The 'Query Builder' tab is active, displaying a natural language query: 'select ai what are our top 10 streamed movies that were released after 2010;'. Below the query, the 'Query Result' tab shows the results of the query, which is a table with two columns: 'MOVIE_TITLE' and 'STREAM_COUNT'. The results are as follows:

MOVIE_TITLE	STREAM_COUNT
1 Avengers: Endgame	81074
2 Captain Marvel	33587
3 Star Wars Episode IX: The Rise of Skywalker	31331
4 Spider-Man: Far from Home	29021
5 Aladdin	24058
6 The Lion King	22627
7 Aquaman	20357
8 Avengers: Infinity War	18813
9 Toy Story 4	17426
10 Bohemian Rhapsody	16936



앱 모듈의 독립적인 개발을 위해서는 통합 데이터 모델의 지원 필요



- ✓ AI 생성 모듈은 높은 독립성이 제공되어야 모듈 설계 간소화 가능
- ✓ 각 모듈은 필요한 데이터를 선언적으로 정의하고 데이터 접근을 위한 API만을 통해 자신이 원하는 데이터를 사용할 수 있어야 함
- ✓ 현재는 개발자가 DB에 맞추어 data consistency를 유지하기 위한 코드를 개발해 주어야만 함

- ❖ 현재 DB 시장에서는 RDB와 JSON, Graph DB 등으로 나누어져 unified model이 제공되고 있지 못함
- ❖ JSON 형태의 데이터는 Shared Data에 대해 효율적이지 않음. 즉, 중복이 발생하고 데이터 일관성 유지가 어려움
- ❖ 이는 App의 독립성을 저해
- ❖ 데이터를 중복하지 않고 한번만 저장하도록 구성하는 것을 “normalization” 이라고 하고 이 기능은 RDB의 가장 기반이 되는 기술
- ❖ Access format과 Storage format을 분리하고 이를 하나의 unified data로서 normalization을 지원할 수 있는 기능이 필수적

오라클의 솔루션 : JSON Document Relational Duality

Data is **stored as rows** in tables to provide the benefits of the relational model and SQL access

Data can be **accessed as JSON documents** to deliver the application simplicity of documents

Storage Format

TABLE		
Col 1	Col 2	Col 3
...
...
...
...
...
...



Access Formats

```
{
  "name1" : "String Value1",
  "name2" :
    {
      "name3" : "14:00",
      "name4" : 1234
    }
}
```



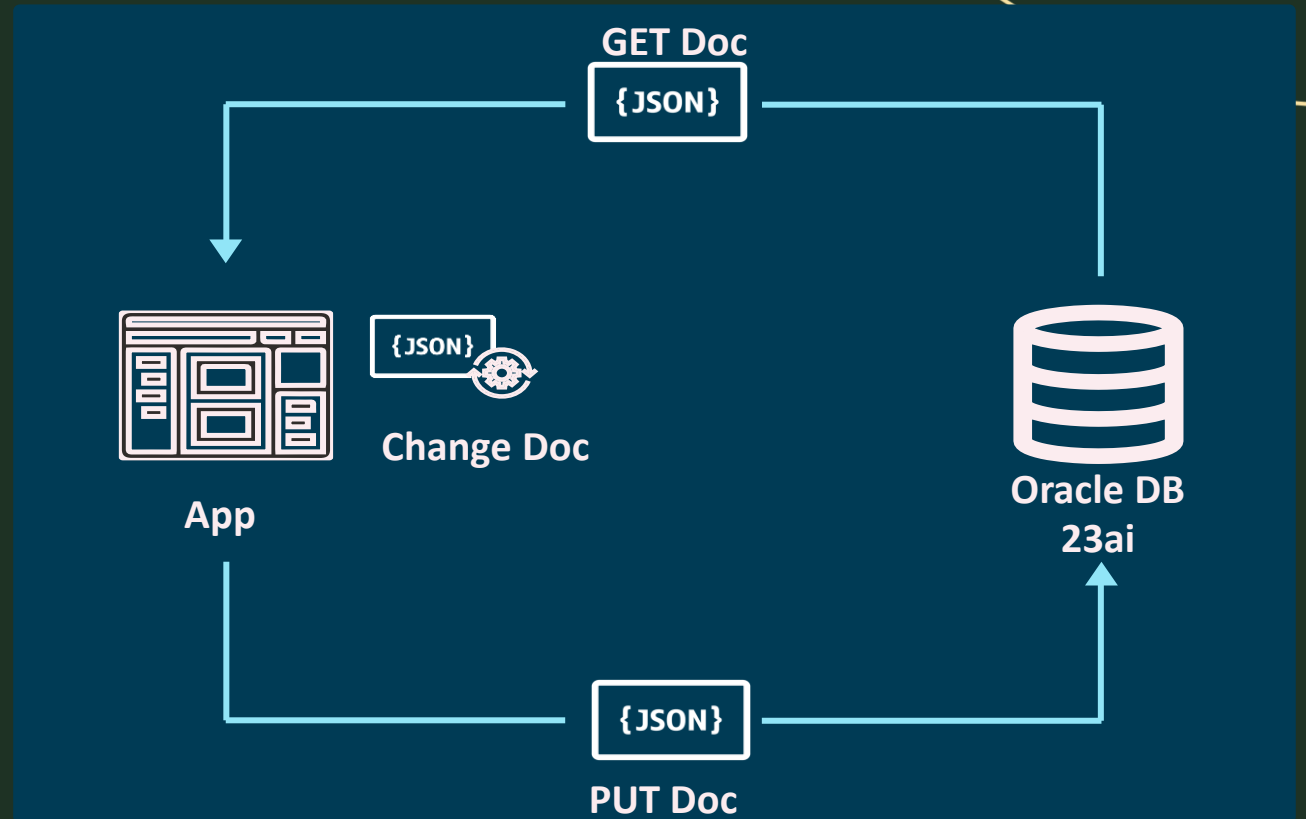
개발자에게 최고의 단순성/독립성 제공, 나머지는 DB가 처리

JSON Duality View는 REST를 사용하여 매우 간단하게 액세스 가능 :

- GET a document from the View
- Make any changes needed to the document
- PUT the document back into the View

데이터베이스가 새 문서의 변경 사항을 자동으로 감지하고 기저의 레코드를 수정

- All duality views that share the same data immediately reflect this change
- Developers no longer have to worry about inconsistencies



```
GET school.edu/student_schedule?q={"student":"Jill"}
```


개발자는 JSON Duality View를 선언하여 JSON 문서의 각 키-값 쌍이 어느 테이블과 열에서 오는지 선언적으로 정의

이 뷰는 간단하게
만들 수 있음

JSON 문서의 구조를
반영하는 친숙한
GraphQL 구문을
사용

ATHLETE CORE INFO			
A_ID	NAME	PLNT	REG
...
...

ATHLETE/EVENT RELATIONSHIP		
A_ID	E_ID	A_POSN
...
...

EVENT CORE INFO			
E_ID	TITLE	SPORT	LOC
...
...



```
CREATE JSON DUALITY VIEW
Athlete_schedule AS athlete_core_info
{
  athlete      : a_id
  name         : name
  planet       : plnt
  region       : reg
  schedule     :
  athlete_event_relationship
  [
    { event_core_info
      {
        event      : title
        sport      : sport
        Location   : loc
        time       : time
      } } ] };
```

JSON Duality View도 AI 를 활용하여 자동으로 생성 가능



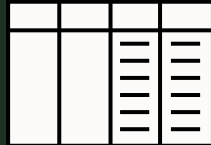
“Oracle’s JSON Relational Duality is perhaps one of the most important innovations in information science in 20 years.”

Carl Olofson, Research VP, Data Management Software, IDC

Duality는 관계형, 다큐먼트 및 그래프 데이터 모델을 통합하기 위한 노력의 출발점

Oracle Database 23ai에서는 **동일한 데이터에 대해서**
앱의 한 부분은 데이터를 **관계형**으로, 다른 부분은 **그래프로**,
나머지는 **다큐먼트**로 처리할 수 있음

Unified



- ✓ 듀얼리티 뷰는 자동으로 REST 및 MongoDB API를 생성하여 JSON SQL로 데이터를 GET 또는 PUT 수행
- ✓ 동일한 데이터에 대해 관계형 테이블로 액세스할 수 있음
- ✓ 동일한 데이터에 대해 새로운 ISO 표준 SQL/PGQ 구문을 사용하여 그래프로 액세스할 수 있음

이를 통해 개발자는 데이터 유형별로 별도의 DB 를 사용하지 않고,
Oracle 단일 DB 만으로도 모든 데이터 유형에 대한 처리 가능

Value-based Concurrency Control

가치 기반 동시성 제어는
정합성을 제공하기 위해
앱 개발자가 어려움을
겪어왔던 문제들을 해결

- 1 사람이 생각하는 시간 동안 데이터에 lock을 걸지 않는 대화형 앱
- 2 연결이 끊긴 상태에서 쓰기를 허용하는 모바일 앱
- 3 읽기 잠금이 백엔드 쓰기를 막지 않는 중간 계층 캐시
- 4 데이터 복제본

Lock을 통한 data block 없이도 사람이 생각하는 시간 동안에도, 연결이 끊긴 상태에서도, 중간 계층 캐시에서도, 데이터 복제의 경우에도 data value의 변화 상태만으로 정합성 체크 지원

Value-based Concurrency Control 작동 방법

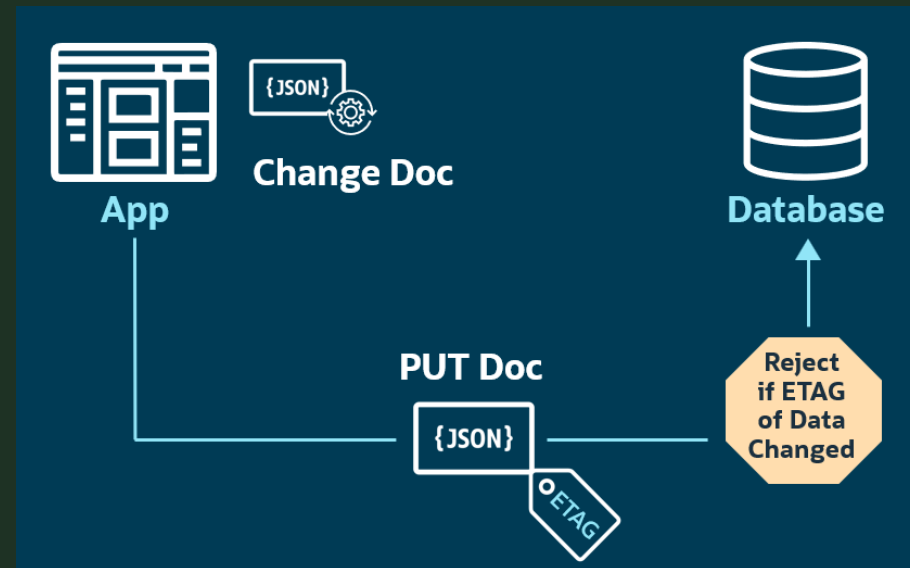
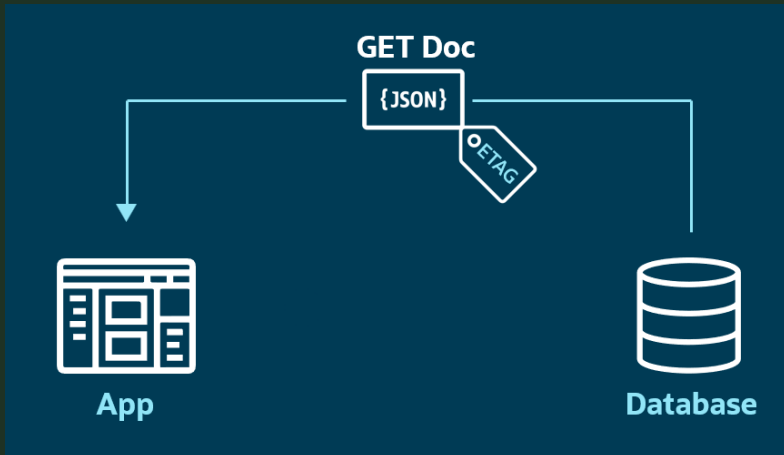
HTTP 프로토콜에서 ETag는 웹 페이지 콘텐츠에 대한 서명 또는 지문

오라클은 Lock 없는 동시성 제어를 구현하기 위해 코어 데이터베이스에 HTTP ETag를 고유하게 확장

- ✓ 데이터베이스 GET은 자동으로 ETag를 계산하여 반환된 문서 자체에 삽입

수정된 문서를 데이터베이스에 다시 PUT 할 때

- ✓ 데이터베이스는 기초 문서 행이 여전히 GET에서 계산한 ETag와 일치하는지 확인
- ✓ 행이 ETag와 일치하면 행이 atomic update 수행
- ✓ 일치하지 않으면 다른 사용자가 데이터를 변경한 것이고 PUT 작업이 거부됨
- ✓ 새 데이터를 사용하여 PUT을 다시 시도할 수 있음



Duality View + VBCC : 데이터를 위한 새로운 마이크로서비스

마이크로서비스는 API만
노출하여 앱 모듈의 독립성과
발전성을 제공

Duality View도 APIs를 제공

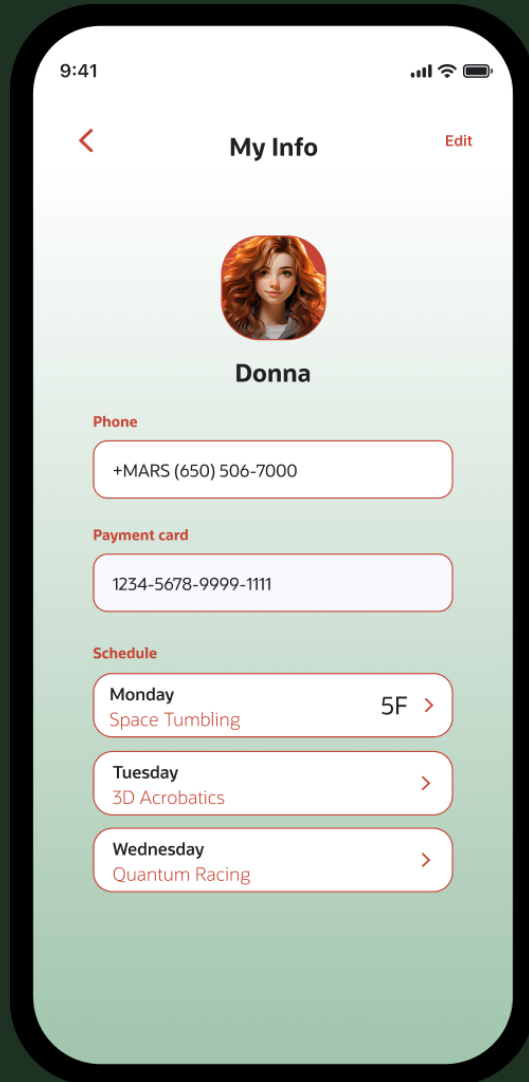
또한 Duality View는 여러
모듈에서 사용되는 데이터에
대해 선언적 진화 및 ACID
일관성을 제공



데이터용 ACID 마이크로서비스를
사용하여 앱을 미니 서비스로
업그레이드



GenDev에서는 기밀성 규칙을 선언적으로 지정하고 적용할 수 있음



GenDev는 현재 앱 최종 사용자가 보서는 안 되는 데이터를 다음의 기능들을 통해 DB 수준에서 독립적으로 필터링할 수 있음 :

- Virtual Private Database, Label Security
- Real Application Security

앱 모듈은 승인되지 않은 데이터를 볼 수 없으므로 데이터가 유출될 수 없음

데이터 기밀성은 생성되는 모든 액세스 또는 SQL 문에서 코딩될 필요없이 자동으로 적용됨

APEX : 기업용 App 개발 지원을 위한 Free Low-Code Tool

AI-Assisted Development



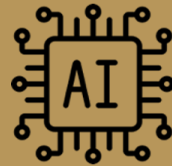
Choose your preferred LLM and leverage the Native Generative AI Capabilities of APEX for AI Assistants, Prompt-based App creation, etc.

AI-Powered Apps with RAG



Incorporate powerful AI features into your app, Use RAG Sources to enrich AI conversations

Integration with OCI AI Services



Seamlessly integrate with Oracle Cloud Infrastructure (OCI) AI services

Harness Oracle's Converged Database



Work with disparate data sources like JSON, Spatial, Graphs, Vectors and build AI capabilities around your own business data

- **Create a Data Model using Natural Language**
- **Create Apps using Natural Language**
- **AI-Assisted SQL Authoring**
- **AI-Assisted Debugging**
- **Build AI-Powered Apps**

APEX : 기업용 App 개발 지원을 위한 Free Low-Code Tool

APEX uniquely delivers low-code **without limits** on:

App sophistication

Mission criticality

Scale of users

Interoperability

APEX is completely free

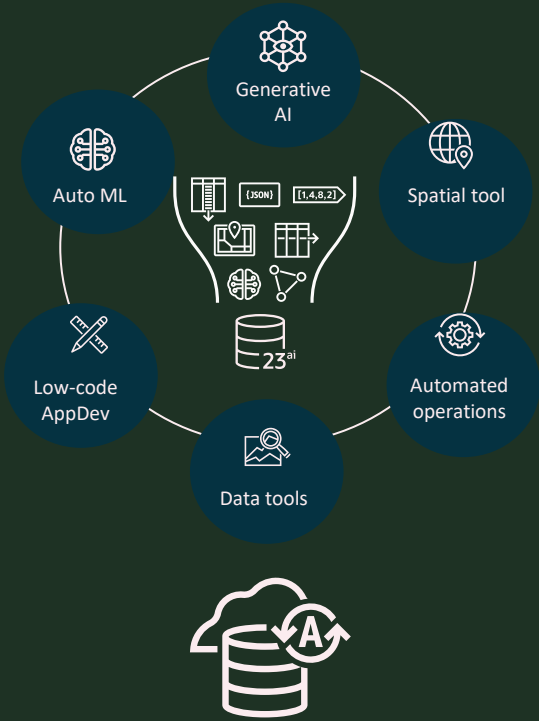
APEX GenDev's **blueprint** language is a high-level language that specifies **WHAT** to do instead of **HOW** to do it

It provides a **declarative infrastructure** that makes it easy for AI to generate app logic

APEX blueprints are succinct, understandable, and editable by humans

Oracle Autonomous Database

AI가 통합된 완전 자동화 및 관리형 엑스데이터 기반 오라클 데이터베이스



모든 최신 데이터 유형, 워크로드 및 개발 스타일에 적합
오라클 데이터베이스 23ai의 AI 벡터 검색 및 데이터베이스 내 AI/ML을 통한
데이터 기반 AI 지원

모든 데이터에 대한 최신 애플리케이션 및 분석 가속화



완전 자동화된 관리형 서비스

Select AI 및 RAG로 자연어로 데이터와 대화하세요. 생성 AI, ML, 데이터 및
AppDev를 위한 기본 제공 도구

Available on:

ORACLE CLOUD
Infrastructure

Microsoft Azure



aws



Database 23ai GenDev also adds many other AI capabilities



Creating vectors inside DB using the **ONNX** framework



Interfacing with popular AI Development Tools incl. **LangChain**



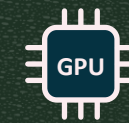
Running Distributed AI with **GoldenGate 23ai**



Scaling Mission-critical AI using **Exadata**



Translating natural language questions into SQL using **SelectAI**



Running ML Notebooks using **GPUs**

GenDev 제공의 가치

GenDev는 엔터프라이즈 앱에 필요한 독립성, 격리, 기밀성, 일관성, 진화 및 생성 가능성을 자동화하는 AI 중심 선언적 인프라를 제공

이러한 작업을 자동화하는 것은 업계의 오랜 목표였으며, 오라클 데이터베이스 23ai를 통해 제공됨

이제 개발자와 AI는 인프라 대신 앱 기능에만 집중할 수 있음



Key Takeaways

자동차가 잠재력을 발휘하려면 포장된 도로가 필요



GenDev

GenDev는 AI의 장점을 극대화하고 위험 요소를 완화하기 위해 오라클 데이터베이스 23ai에 도입된 AI 중심 개발 인프라스트럭처

GenDev는 미래의 비전이 아니라 현재에 존재하며, 기업용 앱 개발의 속도와 생산성을 획기적으로 높일 수 있습니다



Learn more

I encourage you to
Generate to Innovate
using GenDev starting today

ORACLE

