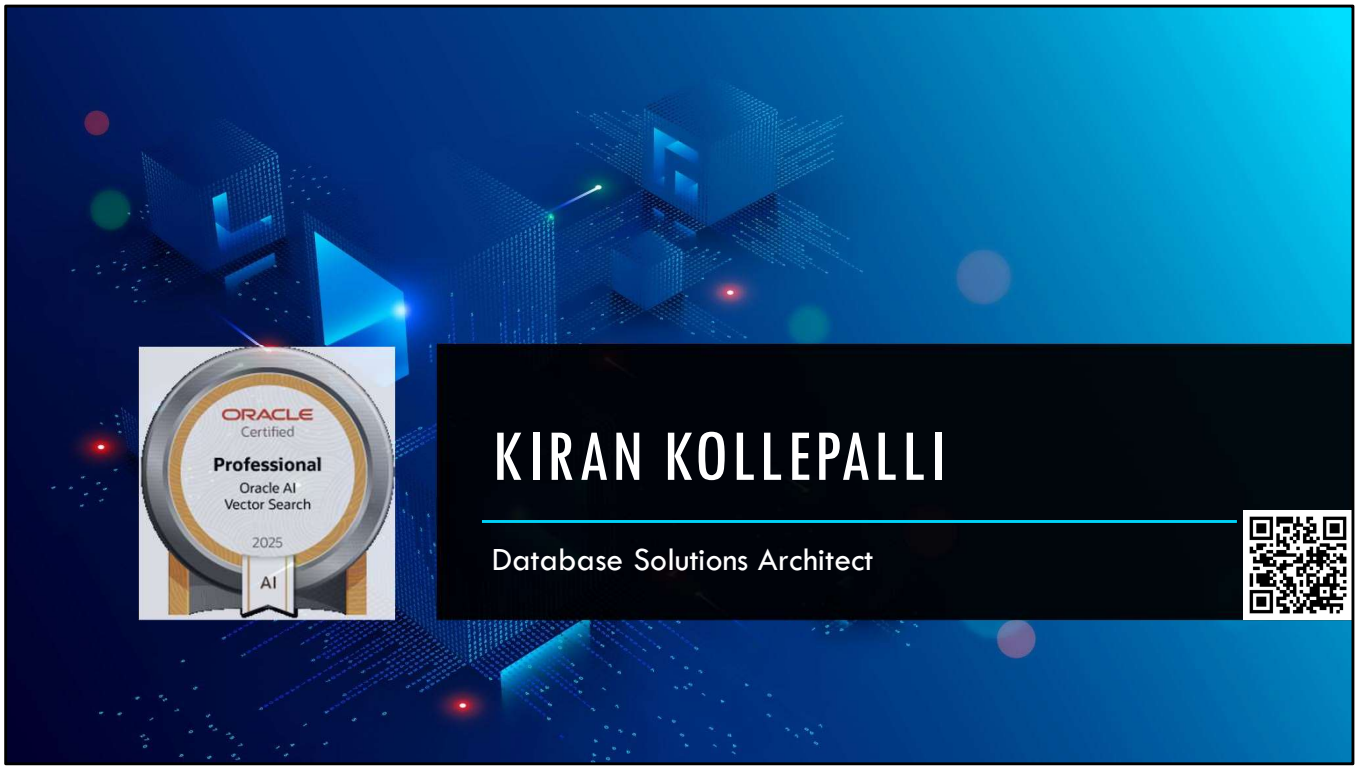

AI IN THE MODERN DATABASE
Kiran Kollepalli


ORACLE
Certified
Professional
Oracle AI
Vector Search
2025
AI
KIRAN KOLLEPALLI
Database Solutions Architect

# STORY LINE

Artificial Intelligence (AI)

Oracle Vector Search

Oracle Select AI

SQLcl and MCP Server

Oracle Autonomous Database

Distributed AI with Golden Gate

AI Smart Scan with Exadata

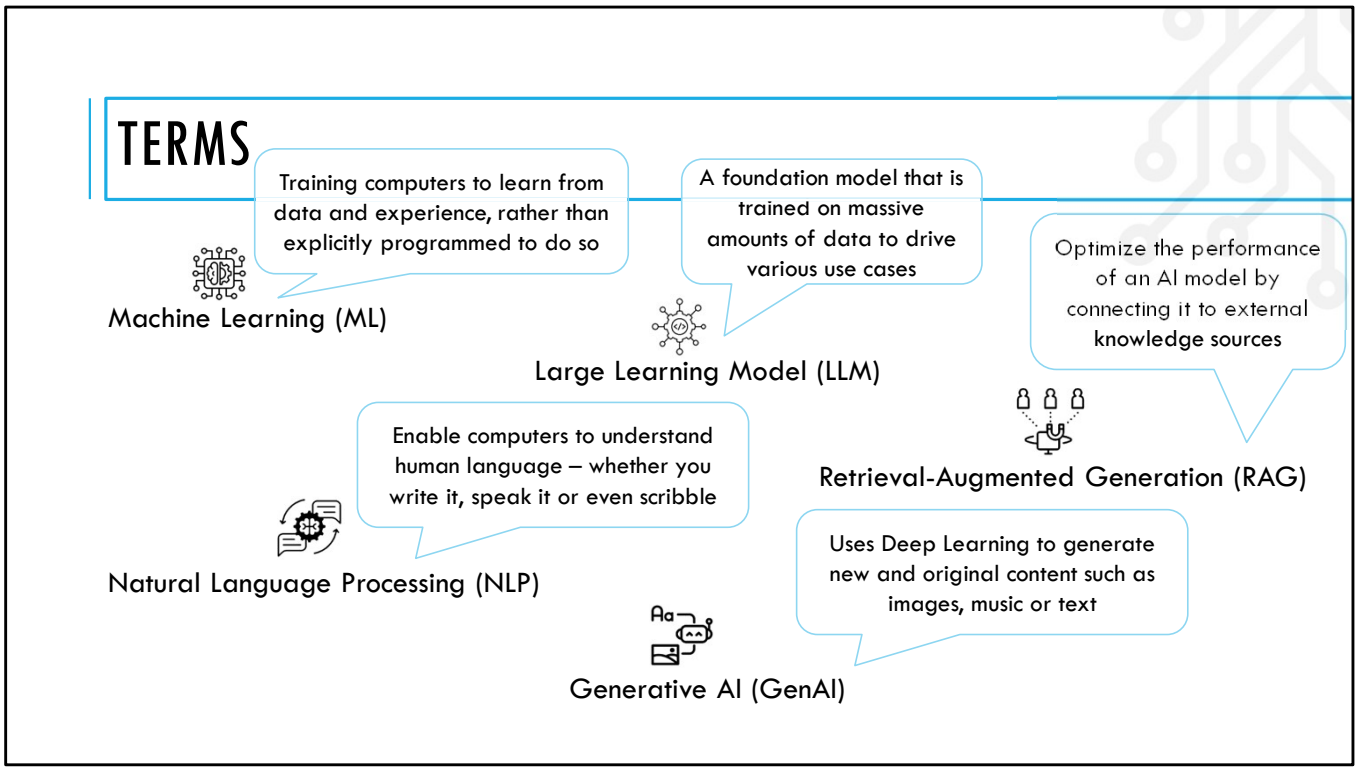
MySQL Heatwave

Oracle AI Data Platform

Best Practices

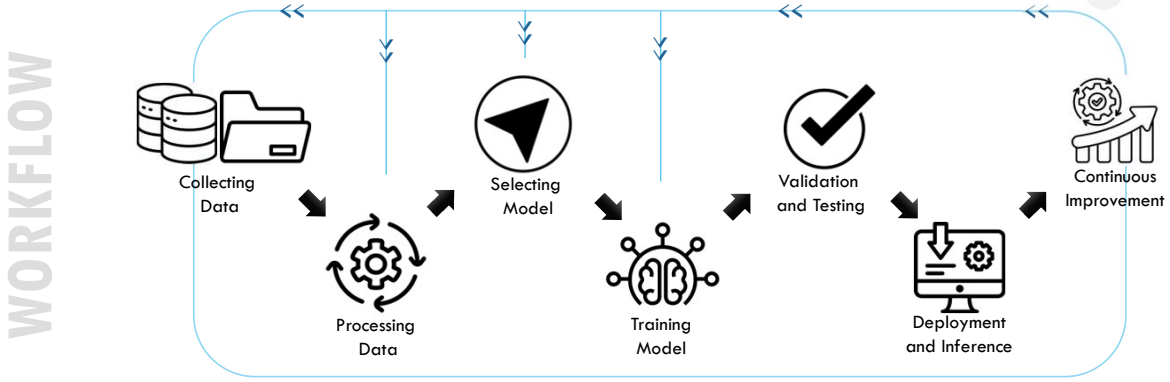
# ARTIFICIAL INTELLIGENCE (AI)

- **AI** is a technology that enables Computers and Machines to simulate:
  - Human Learning
  - Comprehension
  - Problem Solving
  - Decision Making
  - Creativity and Autonomy




- **RAG:** Facts and Context might be obsolete or old fashioned. We need to be current.
- **Augment** means to increase the size or value of something by adding something to it
- **GenAI** is a type of artificial intelligence that uses deep learning (Deep learning is a first bench student in a ML class) to generate new and original content such as images, music, or text
- **Deep learning** mimics neural networks of the human brain, it enables computers to autonomously uncover patterns and make informed decisions from vast amounts of unstructured data

# AI MODEL




An AI Model

**Collecting Data** – AI systems learn from large amounts of data such as text, images, numbers, streaming data etc

**Processing Data** – The data is cleaned, structured, and prepared for analysis. Correct the formatting and Datatype, Remove unrelated data and duplicates

**Selecting Model** – The choice of embedding model depends type of data, performance, cost etc – Regression, Classification, Clustering and so on

**Training Model** – Machine learning algorithms are trained on this data to find patterns and relationships

**Validation and Testing** – Evaluate how well the model performs by measuring its accuracy, precision, error percentage, etc.

**Deployment and Inference** – Once trained, the AI model can take new, unseen data and make predictions, recommendations, or decisions.

**Continuous Improvement** – Results are evaluated, and the system refines its model over time to become more accurate

# AUTOMATION vs AI

- Complete same task, same way, every time
- Strictly follow preset rules and commands
- Suited for tasks that don't change over time
- Error if result is different or unexpected

- Complete same task, dynamically respond, make decisions
- Make decision based on data, learning and information received
- Suited for complex tasks that require decision making
- Learns and evolves over experience


Automation is widely used to scale up production such as factories
AI is applicable where learning and dynamic decision making is important – chat bots

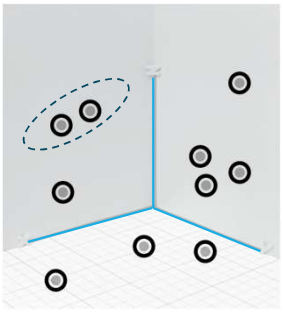
# ORACLE VECTOR SEARCH

**Vector**: AI models can convert complex, unstructured data like text, images, and audio into a series of numbers (a vector) that the computer can work with mathematically

The vectors are array of numbers that capture the complex relationships and meaning of the original data, understood by AI models

**[ 9.2341, 7.1234,... ]**

Data points with similar meanings are represented by vectors that are located close to each other in this vector space



# ORACLE VECTOR SEARCH

**Vector Data Type**: The new native vector data type stores vectors directly in tables with different dimension counts and formats allowing embedding model of your choice

ORACLE Database 23ai

You can either load vectors generated externally or load a model into the database to generate vectors from existing data

Hugging Face

This capability powers **Oracle AI Vector Search**, which lets you perform similarity searches on business data using simple SQL — allowing you to query based on meaning (semantics) rather than just keywords

Semantic search understands the context and provides personalized results, enhancing user experience

- Oracle database vector types: dense and sparse. Dense: nearly every dimension stores a non-zero value, Sparce: a large number of dimensions and few non-zero dimension values
- Formats: INT8, FLOAT32, FLOAT64, BINARY; Max Dimension is 65,535. Syntax: **VECTOR(<dimension> [, <format>])**

**Open Neural Network Exchange or ONNX** is an open standard format of machine learning models. To generate embeddings within the database, you can import and use vector embedding models in ONNX format. Oracle Database 23ai includes an ONNX runtime engine for running embedding models directly inside the database. Text, images, audio, video, or other data types are supported. Consider the following use cases:

- Using complex media input such as text or image for similarity search
- Perform text classification
- Perform reranking

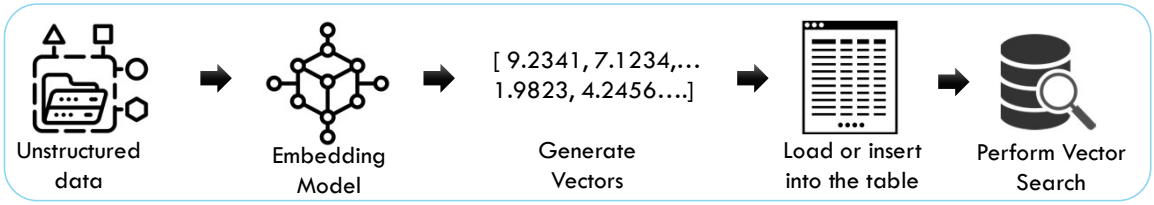
Oracle provides downloadable, augmented version of **Hugging Face's all-MiniLM-L12-v2 model in ONNX format**. This model, distributed by Hugging Face, can be loaded directly into the database via DBMS_VECTOR.LOAD_ONNX_MODEL or DBMS_DATA_MINING.IMPORT_ONNX_MODEL. The model available for download represents an augmented pipeline that includes tokenization and the required post-processing steps for generating vector embeddings

seamlessly with AI Vector Search in Oracle Database 23ai

# ORACLE VECTOR SEARCH

**Vector Data Type**: The new native vector data type stores vectors directly in tables with different dimension counts and formats allowing embedding model of your choice

ORACLE Database 23ai




Instead of searching through endless text documents, the vector databases make this process lightning fast by converting text into mathematical vectors based on word meaning and relationships

- Vectors are generated from various input data such as text, images, audio, video, or other data sources using deep learning models—so-called **embedding models**
- **dbms_vector** package helps to load an embedding model into the database

# ORACLE VECTOR SEARCH

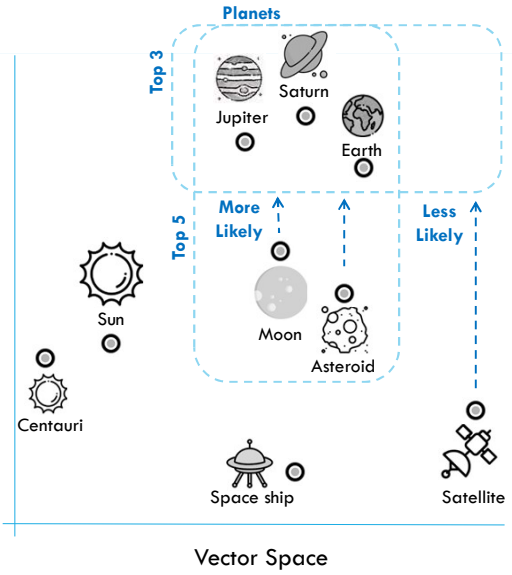
EXAMPLE

**User Query:**
"List of objects that are like planets"
on Data related to space objects

**SQL Version:**
```
SELECT object_name, object_property, object_category
FROM Objects_in_Space
WHERE upper(Object_category) like '%PLANET%';
```

**Vector Search using SQL:**
```
variable search_text varchar2(100);
exec :search_text := 'List of objects that are like
planets';

SELECT
 vector_distance(
 object_vector,
 (vector_embedding(
 all_minilm_l12_v2 using :search_text as data))) as
distance,
 object_name, object_property, object_category
FROM Objects_in_Space
ORDER BY 1
FETCH approximate first 3 rows only;
```




- **vector_distance** function is used to mathematically calculate distance between vectors (search two similar vectors)

- References:
  - https://www.killiansbytes.com/post/dive-into-ai-vector-search-in-oracle-database-23ai-alternate-title-a-negroni-and-the-oracle-databa
  - https://www.oracle.com/de/database/ai-vector-search/
  - https://livelabs.oracle.com/pls/apex/r/dbpm/livelabs/view-workshop?wid=4166
  - https://oracle-base.com/articles/23/ai-vector-search-23

# ORACLE VECTOR SEARCH

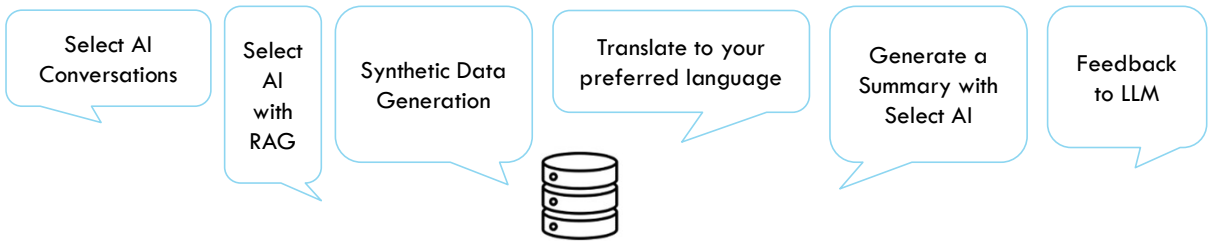
**Vector Search** feature has made its availability in many popular RDBMS




# ORACLE SELECT AI

**Converse** with your database using natural language via **Oracle Select AI** - powered by OCI GenAI

Use the AI keyword to perform actions - runsql, showsql, explainsql, narrate and chat

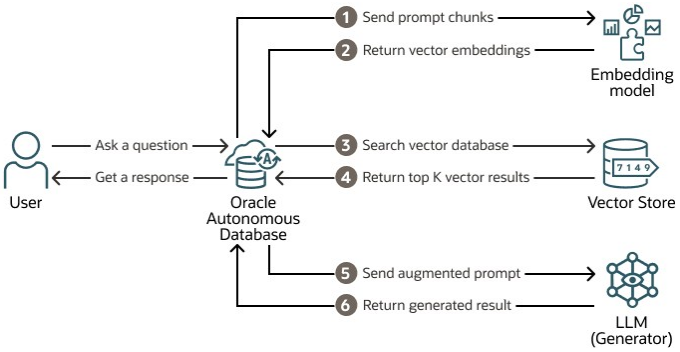



- Use with Oracle clients such as SQL Developer, OML Notebooks, and third-party tools
- **Select AI Conversations**: Interactive dialogue between the user and the system, where a sequence of user-provided natural language prompts are stored and managed to support long-term memory for LLM interactions
- **Select AI with Retrieval Augmented Generation (RAG),** supports your natural language prompt by retrieving content from your specified vector store using semantic similarity search. It provides up-to-date and more relevant natural language responses to your prompts
- **Synthetic Data Generation,** using random generators, algorithms, statistical models, and LLMs to simulate real data for developing and testing solutions effectively
- **Feedback**: Select AI enables you to provide feedback to help improve your selected LLM's ability to generate more accurate SQL queries.
- **Generate a Summary with Select AI**: Select AI enables you to generate a summary of your text, extract key insights from texts
- **Translate** your text into the language of your choice
- You must run DBMS_CLOUD_AI.SET_PROFILE in each new stateful database session (connection) before you use SELECT AI. If you are using a stateless connection, you must use the DBMS_CLOUD_AI.GENERATE function which enables you to specify the profile name in each invocation

- LLMs are subject to *hallucinations* and results are not always correct. For better results with natural language to SQL generation, use database views or tables with contextual column names or consider adding column comments explaining values stored in the columns

# ORACLE SELECT AI

**Converse** with your database using natural language via **Oracle Select AI** - powered by OCI GenAI



# ORACLE SELECT AI

EXAMPLE

```
SQL> select ai how many customers exist;

CUSTOMER_COUNT
--------------
         55500
```

```
SQL> select ai showsql how many customers exist;

RESPONSE
--------------------------------------------------------
SELECT COUNT(*) AS total_customers FROM SH.CUSTOMERS
```

```
SQL> select ai narrate how many customers exist;

RESPONSE
--------------------------------------------------------
There are a total of 55,500 customers in the database
```



Reference:

- https://docs.oracle.com/en-us/iaas/autonomous-database-serverless/doc/select-ai.html

# ORACLE SQLcl AND THE MCP SERVER

**The Model Context Protocol (MCP)** is an open standard, **USB-C port for AI**, defines a consistent way for applications to provide contextual information to AI models

Until MCP, the application layer was responsible for orchestrating context augmentation by using the model to enrich user questions with additional information

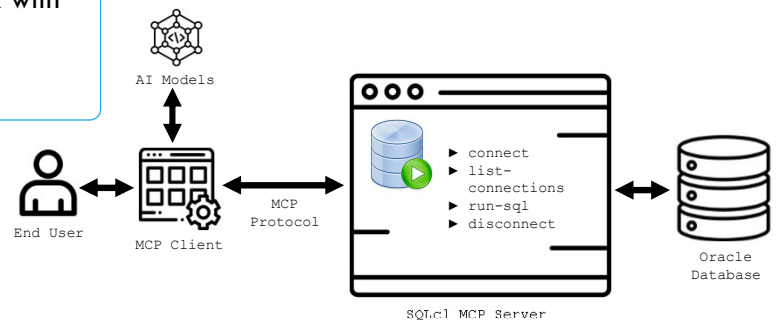
Oracle has evaluated and integrated MCP into our core developer tools, made available via our modern command line to the Oracle Database, **Oracle SQLcl**, manages the credentials and runs SQL and PL/SQL queries and scripts

Oracle SQLcl is shipped with the Oracle SQL Developer extension for VS Code, that can be run as an MCP Server and provides MCP tools that allow an AI assistant to securely connect to Oracle Database

- The absence of MCP protocol meant that every tool, every API, and so on, would have to tailor implementation for each platform or integration to offer the same functionality
- Oracle SQLcl, any AI client that supports MCP (e.g. Copilot, VS Code extensions, Claude, etc.) can connect to it. You can ask Copilot (or other LLMs) things like “Show me top 5 customers by revenue” and have it generate SQL, run it, and present results — all via the MCP Server.
- The requirements are simple: **1.** an Oracle Database to work with, **2.** Oracle SQLcl with one or more defined database connections **3.** your preferred IDE and LLM
- After establishing a connection, an AI client can use natural language to perform the full range of database operations. This includes executing SQL queries, invoking PL/SQL procedures, and running SQLcl-specific commands

# ORACLE SQLcl AND THE MCP SERVER

What do we need:
1. an Oracle Database to work with
2. Oracle **SQLcl**
3. your preferred IDE and LLM




Interact with DB via MCP

- **End User** → asks AI via MCP client → **AI Model** → sends command via MCP → **SQLcl MCP Server** → interacts with **Oracle Database** → returns results
- End User interacts naturally, AI models receives the user request, generate the query and goes MCP Server. The MCP Client uses MCP protocol to send commands to SQLcl MCP Server. The MCP Server exposes database operations, and translates AI agent requests into real SQL query
- This architecture lets AI models safely execute database tasks without direct database access, combining **AI reasoning** with **real data operations** securely

# ORACLE SQLcl AND THE MCP SERVER

EXAMPLE

**User Input:** "Show me the total sales by region for the last quarter."

**SQL generated by AI model:**

```
SELECT region, SUM(sales_amount) AS total_sales
FROM sales
WHERE sales_date >= ADD_MONTHS(TRUNC(SYSDATE, 'Q'), -1)
AND sales_date < TRUNC(SYSDATE, 'Q') GROUP BY region
ORDER BY total_sales DESC;
```

**MCP Command (json):**

```
{
  "command": "run-sql",
  "arguments": {
    "connection": "salesdb",
    "sql": "SELECT region, SUM(sales_amount) AS
total_sales FROM sales WHERE sales_date >=
ADD_MONTHS(TRUNC(SYSDATE, 'Q'), -1) AND
sales_date < TRUNC(SYSDATE, 'Q') GROUP BY region
ORDER BY total_sales DESC"
  }
}
```

**Output returned to MCP (json):**

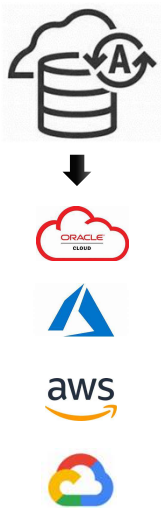
```
{
"status": "success",
"data":
   [
        { "region": "North", "total_sales": 1250000 },
        { "region": "West", "total_sales": 980000 },
        { "region": "East", "total_sales": 870000 },
        { "region": "South", "total_sales": 760000 }
   ]
}
```

Reference:

- https://blogs.oracle.com/database/post/introducing-mcp-server-for-oracle-database
- https://docs.oracle.com/en/database/oracle/sql-developer-command-line/25.2/sqcug/working-sqlcl.html
- https://docs.oracle.com/en/database/oracle/sql-developer-command-line/25.2/sqcug/using-oracle-sqlcl-mcp-server.html

# ORACLE AUTONOMOUS DATABASE (ADB)

**Autonomous Database** is an OCI service, supports all modern data types, workloads, and analytics



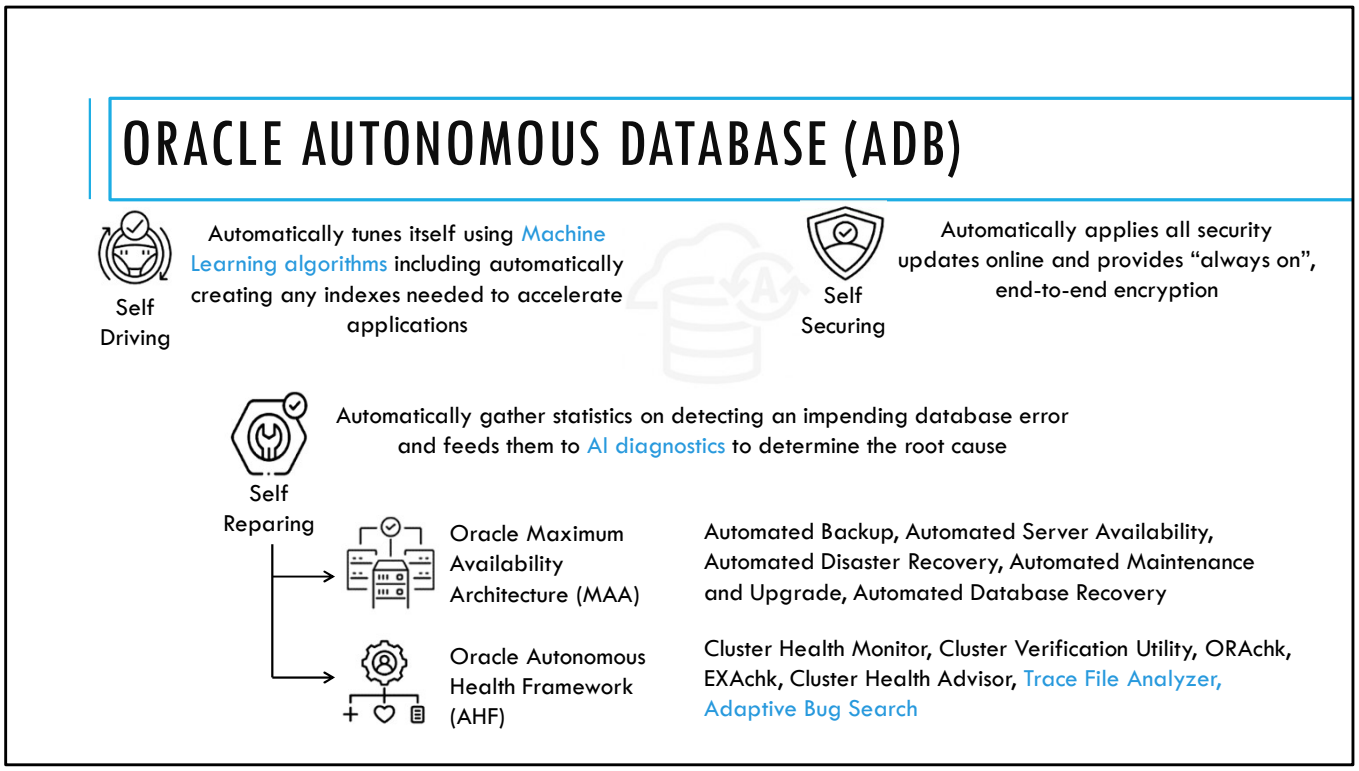
ADB includes

- Infrastructure Automation
- Database Automation
- Datacenter Operation Automation and Machine Learning

ADB features

- Native Vector Search for similarity search
- Built-in Oracle Machine Learning for building and deploying in-database models
- Converse in natural language with Select AI
- Generate Synthetic data using GenAI

- Easily access your data from spreadsheets with add-ins for Excel and Google Sheets
- Autonomous Database with Data Warehouse workloads gathers optimizer statistics automatically for tables loaded with direct path operations, and also for the for loads performed using the DBMS_CLOUD package
- Automatic Indexing automates the index management tasks in Autonomous Database. It monitors the application workload and creates and maintains indexes
- Automatic partitioning analyzes and automates partition creation for tables and indexes of a specified schema to improve performance and manageability in Autonomous Database, It is transparent to the application – enabled at schema level or table level




Oracle Maximum Availability Architecture (MAA) : Eliminate downtime, auto recovery from outage, eliminate data loss – zero data loss replication, offloading backups, offload reports

Oracle Autonomous Health Framework (AHF) : Goal is Availability and Performance

- Cluster Health Monitor – Monitoring metrics and diagnostic views
- Cluster Verification Utility - Baseline statistics, Configure automatic best practices
- ORAchk, EXAchk – Compliance and best practices, flag vulnerability, patches are deployed correctly
- Cluster Health Advisor - Discovers and predicts potential cluster and database problems
- Trace File Analyzer - Accelerates issue diagnosis, triage and resolution for the problems that involve Oracle Support Services. TFA makes sense of all the trace files and alert logs and takes appropriate action. It uses machine learning for anomaly detection and only collects the right logs to see which files are needed for resolution.
- Adaptive Bug Search - Discovers duplicate bugs, correlated issues and prioritizes them based on customer impact. It discovers bugs from over 400 Oracle products and uses machine learning to identify the bug – and help the developer to simplify or bypass the entire triage process

References

- https://www.oracle.com/autonomous-database/features/

- https://livelabs.oracle.com/pls/apex/dbpm/r/livelabs/livelabs-workshop-cards?p100_workshop_series=222
- https://blogs.oracle.com/machinelearning/post/announcing-select-ai-with-rag-on-adb
- https://www.oracle.com/us/products/database/autonomous-database-self-repairing-5116047.pdf
- https://www.oracle.com/oce/dc/assets/CONTC552D01FF25448BFA6CFA4E2D7758F6C/native/oracle-autonomous-database-strategy-wp.pdf?elqTrackId=4773d4684589422bbc56a54717d6b82c&elqaid=104046&elqat=2&

# ORACLE AUTONOMOUS DATABASE (ADB)

**Autonomous Database – Data Studio** is a one-stop application of your analytics tool from multiple data sources

The Data Studio Tools enables you to load data from cloud and other diverse sources, analyzes it and gain insights from it, at no extra cost

**Load data**

Load data from CSV, Excel, Parquet, Local computer, Cloud Store

**Data Transform**

Transform Data for Analysis and other applications


**Table AI Assist**

Use GenAI to augment and fix data in your table

**Analysis**

View data using charts and reports

**Insights**

Search for data outliers and anomalies

**Catalog**

Find data and understand dependencies

- **Data Transform:** Transform without code, apply ML within data flow to score the data (analyze quality and relevance), automatically generate code, prepare for analysis
- **Table AI Assist feature:** Designed to enhance and prepare data tables using natural language prompts. It enables to add, remove, rename columns through AI interaction – add columns for day, month, quarter from a date column, or calculate distances between two co-ordinates and add as a new column
- **Analysis and Insights:** Detect anomalies using built-in AI and ML; Insights help to uncover hidden patterns and outliers
- **Catalog:** Organize and search data effectively by linking data catalog and cloud storage buckets

- **Ref:** https://docs.oracle.com/en/cloud/paas/autonomous-database/serverless/adbsb/adp-data-studio-overview-page.html

# DISTRIBUTED AI WITH GOLDEN GATE

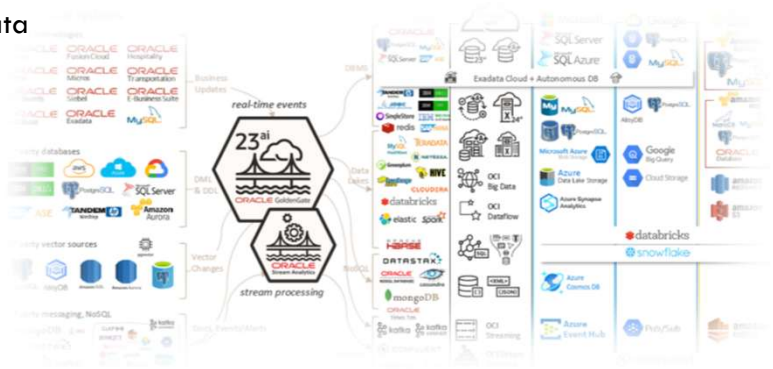
**Oracle GoldenGate** is a software product that replicates, filters, and transforms data between databases and other data sources

1000+ combinations of supported data platforms

Complete datatype coverage

Works with all Oracle DB versions, including Cloud/ADB

Data replication, Data transforms, Stream analytics



# DISTRIBUTED AI WITH GOLDEN GATE

**Oracle GoldenGate** is a software product that replicates, filters, and transforms data between databases and other data sources


1000+ combinations of supported data platforms

Complete datatype coverage

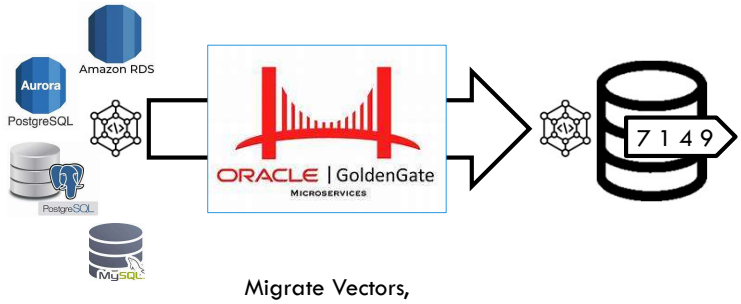
Works with all Oracle DB versions, including Cloud/ADB

Data replication, Data transforms, Stream analytics



# DISTRIBUTED AI WITH GOLDEN GATE

**Distributed AI:** Oracle GoldenGate 23ai can **replicate vector data** heterogeneously and homogeneously if the embedding algorithm is same in terms of type and dimension, across the databases, allowing continuous flow of vector data from one database to another, benefiting AI and machine learning applications




Migrate Vectors,
Replicate or Consolidate vector changes

Every moment your customers and infrastructure are changing; Digital events and transactions are continually flowing;
Golden gate captures these events in real-time to populate vector embedding data -> Benefit AI applications for training and fine tuning

# DISTRIBUTED AI WITH GOLDEN GATE

Oracle GoldenGate 23ai supports capture and delivery of the vector data type for Oracle Database 23ai, MySQL, PostgreSQL and most vendor derivatives to enable **real-time vector processing** for embedding flow


It keeps Retrieval Augmented Generation (RAG) and AI datasets up to date by seamlessly integrating multiple data sources

It enables replicate data and test the performance even if your embedding algorithm is same, but has different dimension

References:

- 20241018-oracle-goldengate-for-hroug-2024.pdf
- Live lab: https://livelabs.oracle.com/pls/apex/r/dbpm/livelabs/home
- https://www.youtube.com/watch?v=6arooYHyI4M
- https://blogs.oracle.com/dataintegration/post/how-to-replicate-postgresql-array-data-vectors-with-goldengate

# AI SMART SCAN WITH EXADATA

**Oracle Exadata** is an enterprise database platform that runs database workloads of any size, optimized for both Transaction Processing and Analytics


Exadata has evolved to support modern workloads like in-memory analytics, Artificial Intelligence(AI), and Machine Learning(ML), facilitating mixed-workload consolidation

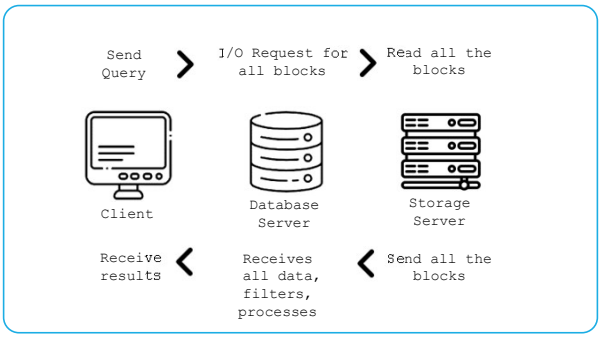
Exadata's **Smart Scan** technology sends only required rows to database node from storage instead of entire Oracle Block

Exadata's **AI Smart Scan** is an optimization that accelerate AI Vector Search queries by offloading compute-intensive vector processing operations from the database servers to the intelligent storage servers

# AI SMART SCAN WITH EXADATA

AI Smart Scan, sends an **Intelligent Database (iDB) message** containing the query criteria, the vector distance calculation logic and Top-K filter, to the storage servers

In **conventional architectures**, full table scans move all data blocks across the network to the database server for filtering (WHERE clauses) and projection (SELECT list), consuming network bandwidth and database server CPU



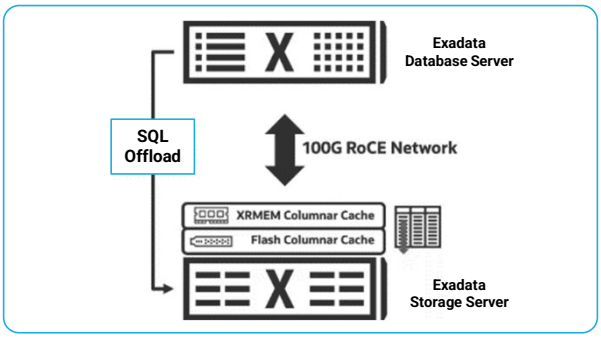

Conventional Architecture

Client/Application issues a SQL query - Database Server analyzes the query, determines it's a full segment scan that can be offloaded

# AI SMART SCAN WITH EXADATA

AI Smart Scan, sends an **Intelligent Database (iDB) message** containing the query criteria, the vector distance calculation logic and Top-K filter, to the storage servers

With **SQL Offloading**, most of the data (the raw vectors that do not match the criteria) never leaves the storage tier, eliminating a massive network bottleneck and reducing the CPU load on the database servers




SQL Offloading

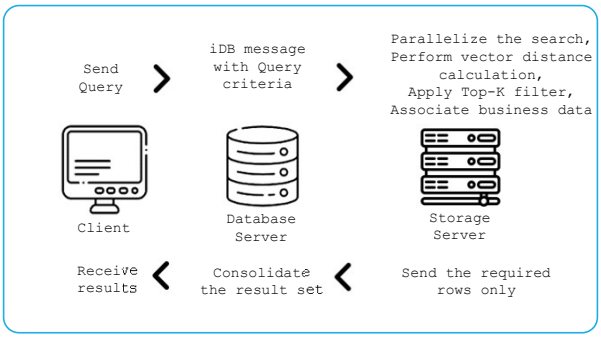
Storage Server Processing (The Core of AI Smart Scan): The Exadata software (cellsrv) performs the CPU-intensive vector distance calculations (e.g., Cosine Similarity) directly within the storage server memory/flash

# AI SMART SCAN WITH EXADATA

AI Smart Scan, sends an **Intelligent Database (iDB) message** containing the query criteria, the vector distance calculation logic and Top-K filter, to the storage servers

Exadata's **AI Smart Scan** optimizes this by sending SQL filters and column projections to the storage servers, returning only relevant rows and columns to the database server

AI Smart Scan is central to optimizing AI Vector Search query performance, delivering low-latency responses




AI Smart Scan Architecture
Exadata System Software 24ai (24.1)
Database release update 23.7

Acceleration mechanisms include:

- Extends Smart Scan
- Leverages Exadata hardware
- Faster AI queries
- Transparent execution: the process is automatic and transparent, requiring no code changes
- Compute Offload: Intensive vector distance calculations and Top-K filtering are executed on storage servers, not the database server.
- Parallel Processing: Exadata's scale-out architecture allows these offloaded operations to run in parallel across multiple storage servers.
- Data Reduction: Only filtered results (top K vectors) are sent back to the database server, minimizing network traffic.
- Minimized Network Traffic
- Hardware Optimization: AI Smart Scan leverages Exadata's ultra-fast storage tiers (XRMEM, Smart Flash Cache) and low-latency RDMA network

References:

- https://blogs.oracle.com/exadata/post/exadata-ai-smart-scan-deep-dive#:~:text=the%20video%20below.-,Overview,ExaDB%2DXS%2C%20or%20Autonomous.
- https://blogs.oracle.com/exadata/post/exadata24ai

# ORACLE AI DATA PLATFORM

The **Oracle AI Data Platform** is an OCI service, provides the platform and the framework to create data analytics pipelines


Enterprise-grade scalable, and secure infrastructure for mission-critical workloads

It helps you quickly build, orchestrate, and manage and automate end-to-end data workflows across your organization

Integration with external catalogs, including Autonomous Database (ADB) and Object Storage

It can Analyze your data, generate insights, and create reports

- Seamless connectivity to Oracle and third-party applications
- It empowers enterprises to accelerate AI development and derive insights – making data AI ready
- End-to-end AI lifecycle tools – workspace, notebook, workflow automation, fully integrated DIE, model training and deployment

# ORACLE AI DATA PLATFORM

The **Oracle AI Data Platform** is an OCI service, provides the platform and the framework to create data analytics pipelines




Create AI Data Platform

Review IAM Policies and Roles

Set Up Compute

Ingest and Organize Data

Explore Notebooks

Steps for Oracle AI Data Platform

# ORACLE AI DATA PLATFORM

Useful for large-scale data pipelines, transforming raw data into usable formats, manage data workflows, and ensuring data quality


USE CASE

Discover, analyze, and generate insights from data to detect anomalies and fraud in real time

Feed corporate documents, PDFs, and emails into a knowledge base to power AI assistants using LLMs

Reference: https://docs.oracle.com/en/cloud/paas/ai-data-platform/index.html

# MYSQL HEATWAVE

**MySQL HeatWave**, a managed service on OCI, is a In-Memory query accelerator, supports Transactional Processing and Real-time analytics without ETL


MySQL HeatWave has native vector processing, supports vector datatype and In-Memory hybrid-columnar storage format for vector columns

Data stored in Object storage can be queried via MySQL Autopilot which can infer the schema by sampling and loading the data into memory of HeatWave cluster

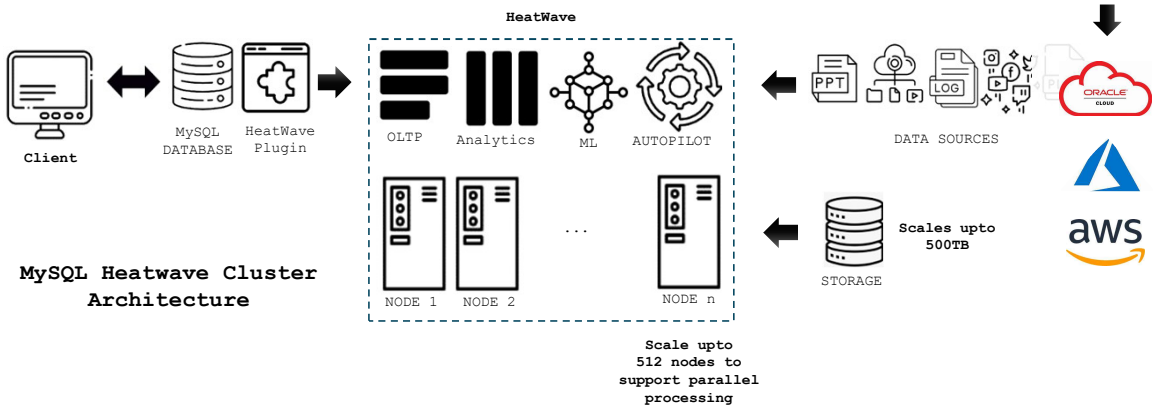
HeatWave AutoML, offers simple SQL stored routine for training a model and generating predictions

HeatWave GenAI, lets you perform natural-language searches using either in-database or external LLM without AI expertise

- HeatWave Lakehouse (similar to Oracle External Table): `Create table… Engine=Lakehouse…` ; MySQL Autopilot can sample the data and infer the schema of the data and return back 'create table…' ddl for the data lying in object store with length and precision ; it supports incremental data load into Lakehouse, when the data is stored with limited retention (e.g., keep last 30 days of data)
- Heatwave AutoML: Offers a simple SQL interface for training and utilizing predictive machine learning models, making it accessible to both novice and experienced ML practitioners. Training a model and generating predictions only requires a single CALL or a SELECT statement, which can be easily integrated into your applications

# MYSQL HEATWAVE

**MySQL HeatWave**, can analyze data from relational MySQL database, analyze data from object store (csv), analyze exported data from other databases



# MYSQL HEATWAVE

**MySQL HeatWave**, It can help quickly summarize product reviews. It can also help translate and analyze sentiment on demand


USE CASE

--- New review ---
Review: The T-Shirt is fantastic as it is not only comfortable, but it is also made of sustainable material. The organic cotton feels like normal cotton while being environment-friendly. However, the washing instruction is difficult to follow

--- Sentiment of the new review ---
**POSITIVE**

# CROSS PLATFORM PERSPECTIVE

The future of data platforms clearly points toward **AI-assisted** optimization, automation, and semantic search

| Feature | Oracle | MySQL | PostgreSQL | SQL Server |
|---|---|---|---|---|
| Native vector store | ☑ | ☑ | ☑ | ☑ |
| Vector Search | ☑ | ☑ | ☑ | ☑ |
| In-database LLM | ☑ | ☑ | ☑ | ☑ |
| Natural Language Query | ☑ | ☑ | ☑ | ☑ |
| Database Tuning | ☑ | | | |
| Database Diagnosis | ☑ | ☑ | | |
| In-Memory Column Store | ☑ | ☑ | | ☑ |
| AI Platform | ☑ | ☑ | | ☑ |

While Oracle has tightly integrated AI and vector capabilities in 23ai, other databases like SQL Server, MySQL HeatWave, and PostgreSQL are evolving rapidly through extensions and cloud-native features. The future of data platforms clearly points towards AI-assisted optimization, automation, and semantic search

# CROSS PLATFORM PERSPECTIVE

The future of data platforms clearly points toward **AI-assisted** optimization, automation, and semantic search

MySQL AI, an option of MySQL Enterprise Edition is Optimized for execution on CPUs, provides built-in **GenAI** capabilities, Automated **ML**, In-built LLMs, and a Vector Store

PostgreSQL has new open-source extensions along with **pgvector** to make a better AI database - **pgai** and **pgvectorscale**. It brings AI workflows to the database, supports AI chat, also enhancing vector search for high-performance

SQL Server 2025, in public preview, has the integration of AI directly into the database engine, enabling **built-in vector search** capabilities, generate embeddings in-line, provide a natural conversational experience

References:


- https://blogs.oracle.com/mysql/post/announcing-mysql-ai
- https://dev.mysql.com/doc/mysql-ai/9.4/en/myai-introduction.html
- https://aka.ms/sqlcopilot
- https://learn.microsoft.com/en-us/ssms/copilot/copilot-in-ssms-install
- https://learn.microsoft.com/en-us/sql/sql-server/what-s-new-in-sql-server-2025?view=sql-server-ver17
- https://www.tigerdata.com/blog/making-postgresql-a-better-ai-database
- https://docs.tigerdata.com/ai/latest/sql-interface-for-pgvector-and-timescale-vector/

# BEST PRACTICES

Using AI tools need **properly designed schemas**;
If your schema isn't human-readable, data can't be machine-understandable

Use masking or tokenization before embedding or inference especially with PII and financial data. **Tag and classify** columns that can be safely exposed to AI ("public data")

Grant the least required **privileges** and take extreme caution while providing access to Production database; instead use a sanitized, read-only replica

Regularly **audit** queries executed to detect anomalies. Collect **feedback** to detect model drift or poor accuracy. Implement human validation checkpoints for AI results that impact customers.

# ROAD AHEAD

As **AI matures**, the Databases will become intelligent — capable of predicting failures, optimizing themselves, and even conversing through natural language

DBAs and Developers will manage not just tables and queries, but also AI models, vector data, embeddings, and semantic search

As databases expose data to AI systems, ensuring data privacy, masking, and regulatory compliance becomes a top priority

# CONCLUSION

AI doesn't replace Developers or DBAs,
It makes us more effective, productive and unlocks more opportunities

By providing access to historical, operational, financial, and documentation data libraries, you let GenAI serve as a force multiplier in many ways for many people throughout your enterprise

The list of potential use cases is growing by the day, because frankly, if you can imagine it, you can probably do it

# GET STARTED

What's most important is that you **get started**.
AI technology is moving fast, and you don't want to fall behind your competition

LiveLabs


NL.OUG
NL ORACLE USER GROUP
Please fill in your evaluations


# AI IN MODERN DATABASE

Kiran Kollepalli