ORACLE

Bringing AI to the Data

EMEA Technology Engineering Update

Ismail Syed

Data Science Leader

Office of CTO, EMEA

AI for Data

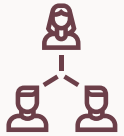
Bringing AI to the Data

Converged Database: - Consolidate your AI workloads in a single database to simplify your data management and eliminate data movement across any data type or workload.

Embedded AI: Bring AI to your data across NL2SQL, Vector Search, Spatial AI, Graph ML, OML without compromise.

Enterprise Capabilities: Take advantage of mission-critical features of the database to build critical applications at scale.

AI for All



The Business: AI Database can now speak human language with Select AI to democratize data to the business.

Data Developers: Combining structured and unstructured data with Vector Search or building apps with APEX AI Assistant.

Data Scientists: Access to familiar coding languages, tools and algorithms across OML, Graph ML & Spatial AI.

Securing AI Workloads

Secure by Design: Oracle delivers trusted AI - where protection, context and intelligence stays within the database.

Converged DB Security: Built in data security with Real Application Security and Oracle Label Security which work seamlessly with your workloads to ensure users only have access to the right levels of data.

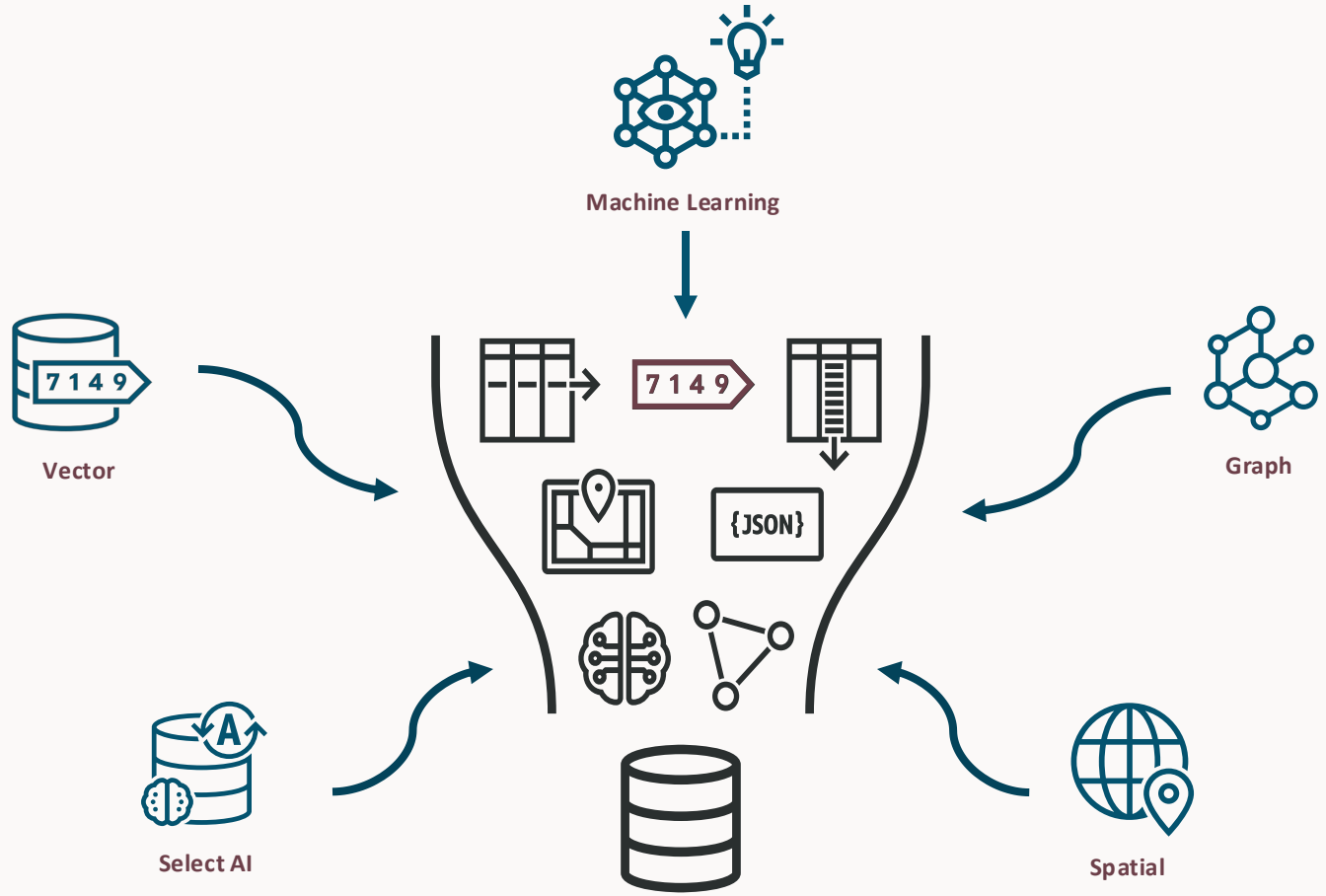
Flexibility

Deployment Model: Public Cloud, Multi-Cloud, Cloud@Customer, and On-Premises.

Level of Management: Fully Managed, Co-Managed, and Self-Managed.

Openness: Open-source ML algorithms, open-source LLM connectivity, combability with open-source AI frameworks.

Bring AI to Oracle AI Database 26ai



Oracle Machine Learning (OML)

Move the algorithms → Not the data

Oracle Machine Learning extends the Oracle AI Database

- Enables users to augment applications and dashboards with machine learning-based intelligence

OML delivers 30+ powerful in-database ML Algorithms

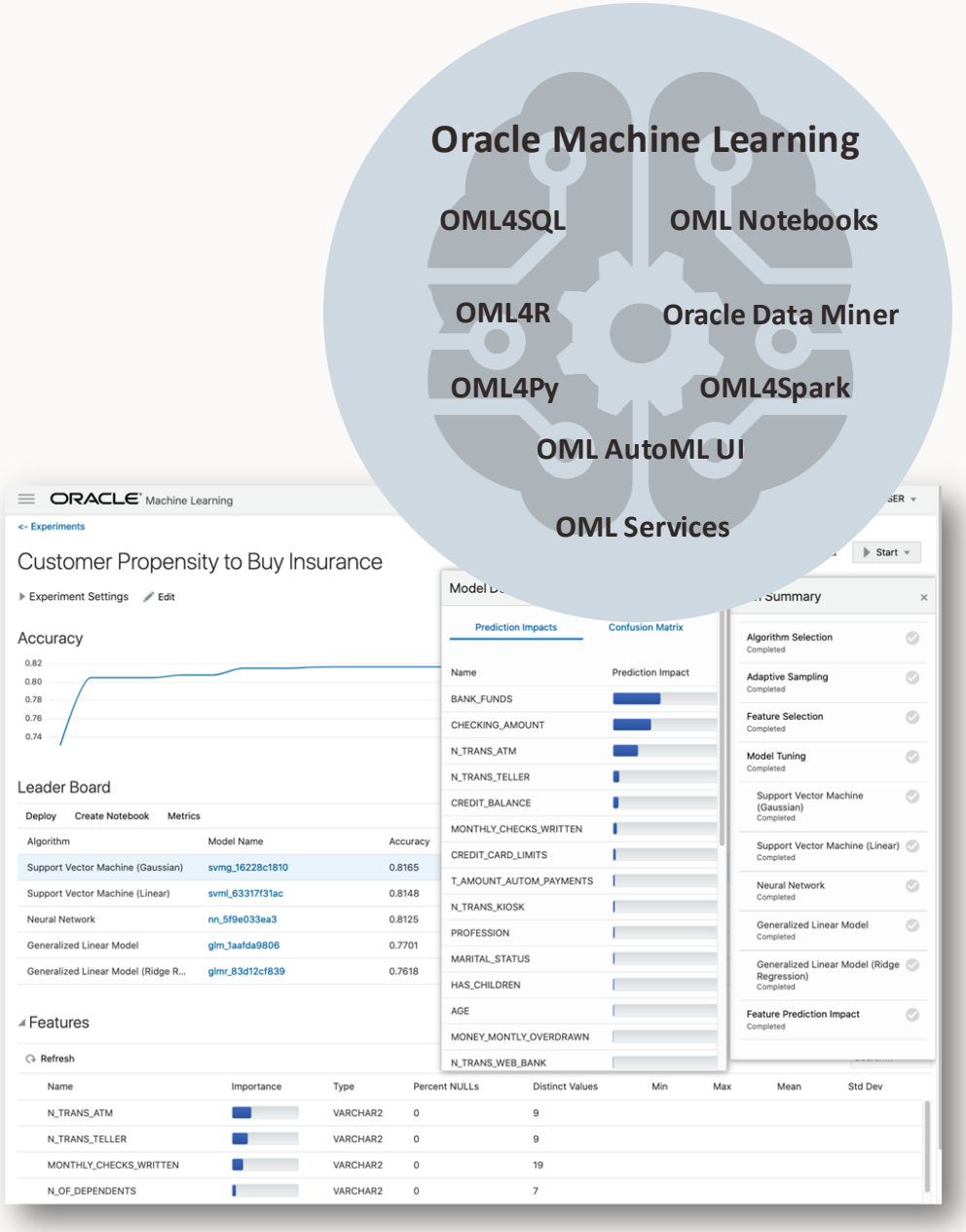
- Automated functionality via SQL, R and Python APIs

Autonomous AI Database-S

- OML Notebooks
- AutoML UI to Accelerate Model Training
- Support for Custom Python & R Conda Environments
- GPU Notebook Support

MLOps from Model Deployment to Data & Model Monitoring!

- OML Services



Oracle Machine Learning

Algorithm Support

Classification
- Naïve Bayes
- Logistic Regression (GLM)
- Decision Tree
- Random Forest
- Neural Network
- Support Vector Machine (SVM)
- Explicit Semantic Analysis
- XGBoost

Anomaly Detection
- One-Class SVM
- MSET-SPRT
- Expectation Maximization (23ai)

Clustering
- Hierarchical K-Means
- Hierarchical O-Cluster
- Expectation Maximization (EM)

Time Series
- Exponential Smoothing
- Multiple Time Series (23ai)
- Includes popular models
  e.g. Holt-Winters with trends, seasonality, irregularity, missing data

Regression
- Linear Model
- Generalized Linear Model (GLM)
- Support Vector Machine (SVM)
- Stepwise Linear regression
- Neural Network
- XGBoost

Attribute Importance
- Minimum Description Length
- Random Forest
- Unsupervised Pairwise KL Divergence
- CUR decomposition for row & AI

Association Rules
- A priori/ market basket

SQL Analytics
- SQL Windows
- SQL Patterns
- SQL Aggregates

Survival Analysis
- XGBoost

Feature Extraction
- Principal Comp Analysis (PCA)
- Non-negative Matrix Factorization
- Singular Value Decomposition (SVD)
- Explicit Semantic Analysis (ESA)

Row Importance
- CUR Decomposition

Ranking
- XGBoost

Text Mining Support
- Algorithms support text columns
- Tokenization and theme extraction
- Explicit Semantic Analysis (ESA)

Statistical Functions
- min, max, median, stdev, t-test, F-test, Pearson's, Chi-Sq, ANOVA, etc.

R and Python Packages
- Third-party R and Python Packages through Embedded Execution

Oracle Graph

Two Graph Approaches

Property Graph

Graph analytics

- Find paths in your data
- Analyze data with graph algorithms
- Graph machine learning algorithms

Use cases

- Analyse the relationships between medications, patients, and reported adverse events
- Explore and analyse genomic data to identify patterns and relationships between genetic variations and diseases.
- What is a vulnerable point of failure in this network infrastructure?

RDF Graph

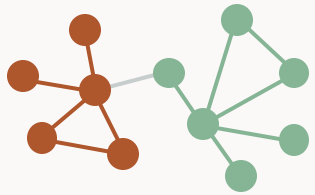
Data integration and linked data

- Manage ontologies
  - An ontology is a representation of concepts and relationships from a particular domain
- Derive new information from existing facts

Use Cases

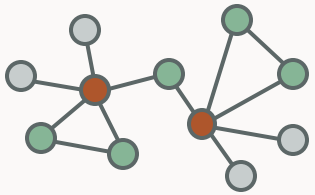
- Link data based on meaning
- Open standards to map different vocabularies for better data sharing
- Knowledge graphs can represent curated data to train LLMs

For the Data Scientist: 80+ Built-in Algorithms (Python or Java APIs)



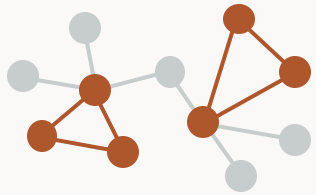
Detecting components and communities

Strongly Connected Components,
Weakly Connected Components,
Label Propagation,
Conductance Minimization,
Infomap



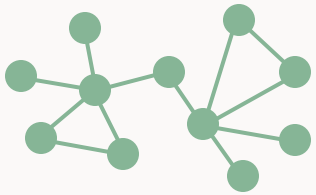
Ranking and walking

PageRank, Personalized
PageRank, Degree Centrality,
Closeness Centrality,
Vertex Betweenness Centrality,
Eigenvector Centrality, HITS, SALSA,
Random Walk with Restart



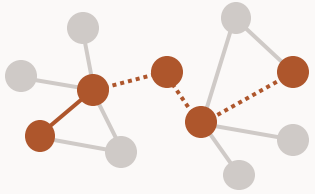
Evaluating structures

Adamic-Adar Index, Conductance,
Cycle Detection, Degree Distribution,
Eccentricity, K-Core, LCC, Modularity,
Reachability Topological Ordering,
Triangle Counting



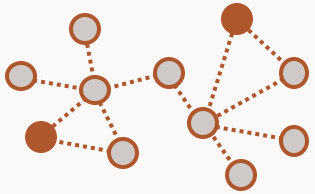
Path-finding

Shortest Path (Bellman-Ford, Dijk
stra, Bidirectional Dijkstra), Fattest Path,
Compute Distance Index,
Enumerate Simple Paths,
Fast Path Finding, Hop Distance



Link prediction and others

WTF (Who to follow)
Minimum Spanning-Tree,
Matrix Factorization



Machine learning

DeepWalk
Supervised GraphWise
Pg2Vec

Oracle Spatial

Spatial Features in the Converged AI Database

In-database functionality to

- Store and manage all kinds of geospatial data
- Perform spatial analysis where the data resides

Spatial Studio

Self-service tool to

- Enable non-experts to more easily analyze data
- Help developers build applications more quickly

Components, APIs & Services

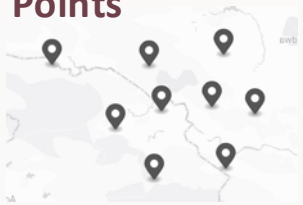
Developer toolbox for

- Map visualization
- Advanced analytics
- Access to spatial functionality and processing workflows

Spatial Processing and Analytics for all Spatial Data Types

Built-in algorithms address range of use cases

Points



Which address locations are within a flood zone?

Lines

Which utility lines are within a mile of an earthquake location?

Polygons

How far from a storm path is the nearest emergency response center?

3D

How close is the tree canopy to our utility lines?

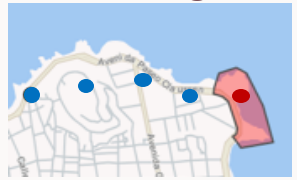
Raster

Which areas of farmland have been inundated by flooding?

Networks

Which locations can be reached within 10 min drive time?

Geofencing



When did a vehicle enter a secure area?

Linear Referencing

Where is the section of roadway with planned maintenance?

Spatiotemporal

Were two people located within 20 ft of each other for over 5 min?

Oracle Select AI

A SQL interface for leveraging Generative AI

4 Main Pillars:

- Build apps to interact with data using natural language
- Generate RAG-enhanced results using vector search
- Synthetic Data Generation
- Autonomous Agentic Framework

Create applications using generative AI

- Democratize data to the business
- Recommendations
- Query unstructured data
- …and more

Oracle Select AI

Use natural language to query data and get responses using generative AI

Easily access LLMs from multiple providers

- OCI GenAI Service
- OpenAI
- Azure OpenAI Service
- Cohere
- Google
- Anthropic
- Hugging Face

...more coming

Actions

- **`runsql`** - return the SQL result set (default)
- **`showsql`** – return the generated query
- **`explainsql`** – explain generated SQL query
- **`showprompt`** – display the generated prompt
- **`narrate`** – return a conversational result
- **`chat`** - general AI chat – passthrough to LLM

Introducing Oracle AI Vector Search

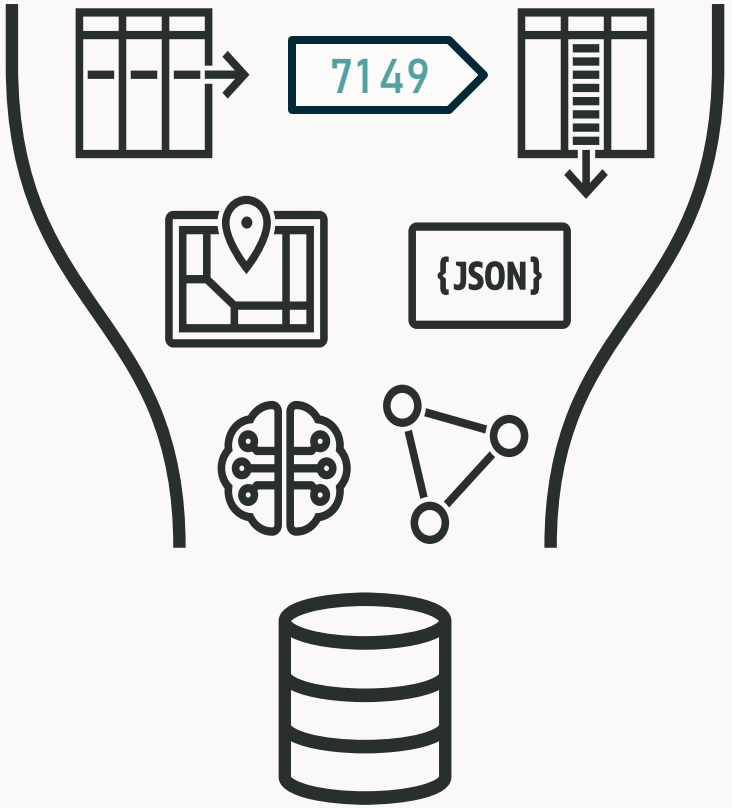
Baked into Oracle AI Database 26ai

Query business data alongside vector embeddings

Process vector search and other workloads in same Oracle converged AI Database

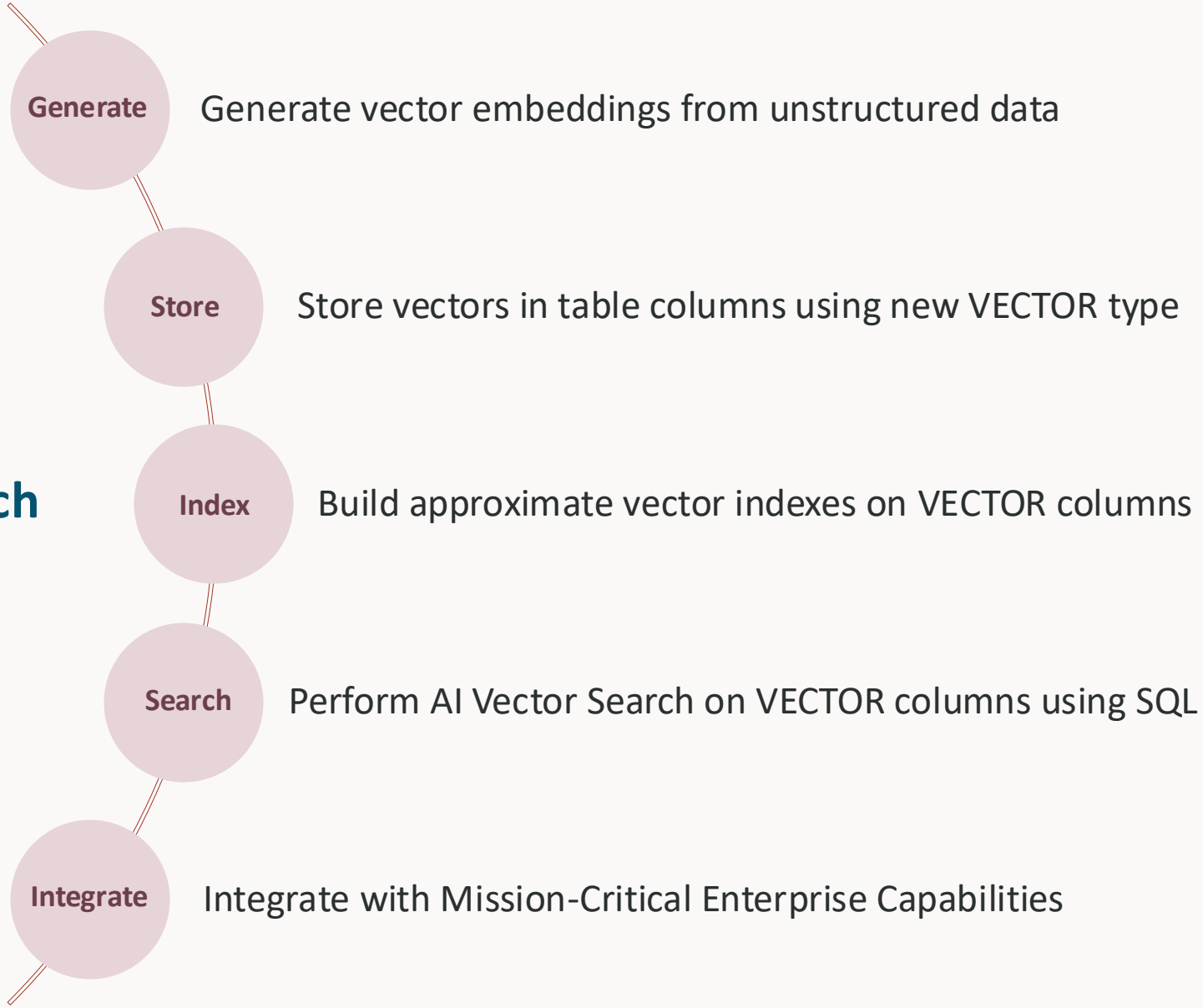
Designed to be simple to use and easy to understand

- **Generate** vector embeddings from unstructured data
- **New** VECTOR data type for storing vectors
- **New** SQL syntax and functions express similarity search with ease
- **New** Approximate search indexes packaged and tuned for high performance and quality



Oracle Database

AI Vector Search Highlights



OCI Data Science

A fully managed ML platform to discover deploy and tune your models

Input Data Sources

Text

Structured and Unstructured Data

Social

Graph

Ingest large volume data from structured, unstructured, streaming, and multimedia sources

OCI Data Science & ML Ops

Develop & Experiment

- Notebook Sessions (jupyter)
- Local and OCI Code Editors

Manage, Share, & Reproduce

- Model Catalog
- Feature Store
- Conda Environments (POWERED BY ANACONDA.)

Operationalize with MLOps

- Jobs
- Pipelines
- Model Deployments
- Model Monitoring

LLMs with OCI Quick Actions

- Fine-tune, deploy, evaluate your preferred LLM

1. Experiment and develop easily AI models with Jupyter notebooks
2. Train and deploy using OCI's scalable MLOps tools
3. Fine-tune LLMs with both code and no-code options

Use Cases

Boost relevancy: Fine-tune your models for better predictions.

Operationalize at scale: Deploy and monitor enterprise AI models more effectively and with tight security features.

Empower teamwork: Foster collaboration among data scientists.

Make smarter decisions: Detect anomalies and forecast with confidence.

Thank you

Ismail Syed

Data Science Leader

Office of CTO, EMEA


ORACLE