VERSION 1

UK OUG | CONFERENCE DISCOVER 2024

3 Case Studies

Machine Learning in OCI

3rd December 2024

Žiga Vaupot

Žiga Vaupot

VERSION 1

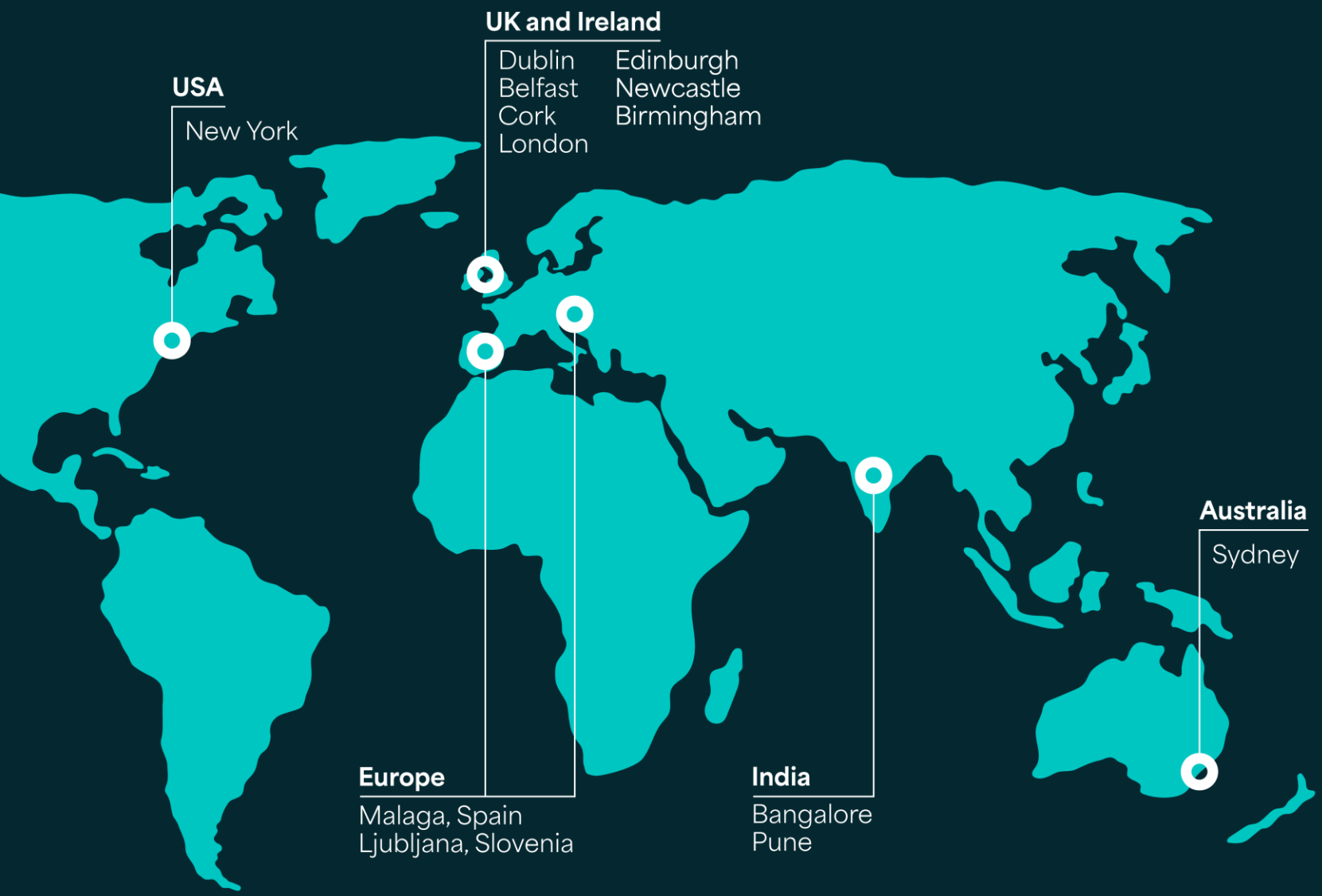
Country Manager, Slovenia
Principal Analytics Consultant

SIOUG
Board Member

Oracle ACE Pro

ziga.vaupot@version1.com
zigavaupot.blogspot.com
linkedin.com/in/zigavaupot

Version 1 at a glance



28 years of experience

13 global locations

6 strategic technology partners

Microsoft · aws · snowflake · ORACLE · outsystems · Red Hat

12 years as a great place to work

Best Workplaces™ Super Large Organisations Great Place To Work. UK 2023 · Best Workplaces™ Great Place To Work. IRELAND 2023 · Great Mid-size Workplaces™ Great Place To Work. INDIA 2023 · Great Place To Work. Certified · glassdoor BEST PLACES TO WORK 2023

Industry best practice

CYBER ESSENTIALS PLUS · ITIL® · ISO 20000 CERTIFICATION EUROPE · ISO 27001 CERTIFICATION EUROPE

Oracle ACE

The Oracle ACE Program

400+ technical experts helping peers globally

- The Oracle ACE Program recognizes and rewards community members for their technical and community contributions to the Oracle community
- 3 membership levels: Director, Pro, and Associate
- Nominate yourself or a colleague at ace.oracle.com/nominate
- Learn more at ace.oracle.com

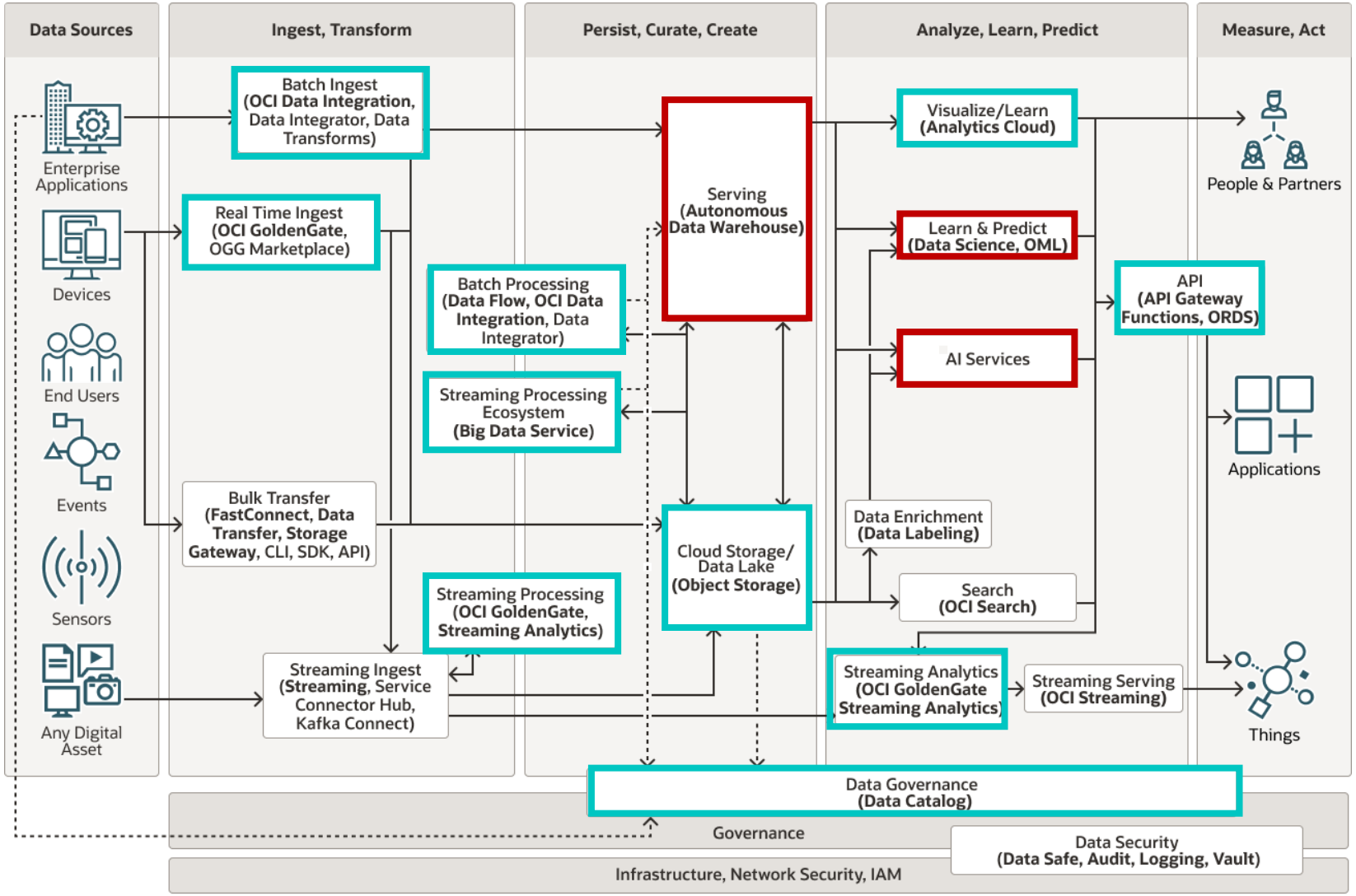
aceprogram_ww@oracle.com
Facebook.com/OracleACEs
@oracleace
Oracle ACE Program Group

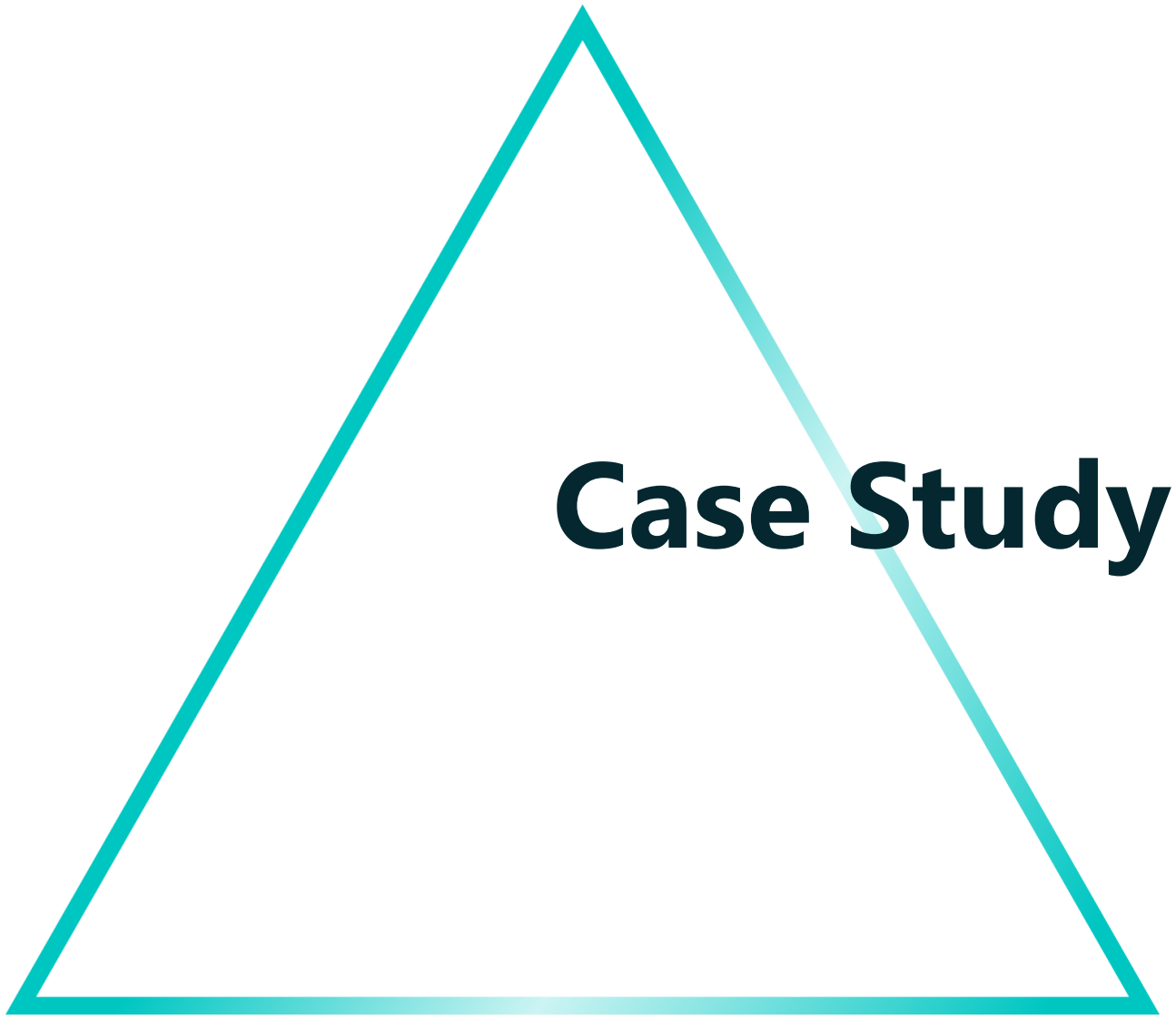
VERSION 1

UK OUG

Now, let's talk about the 3 Cases

ML/AI in OCI Data Lakehouse





Case Study #1

Customer #1

The largest manufacturer in the world of forgings for personal vehicle steering mechanisms. They are also known for their hand tools.

Business problem:

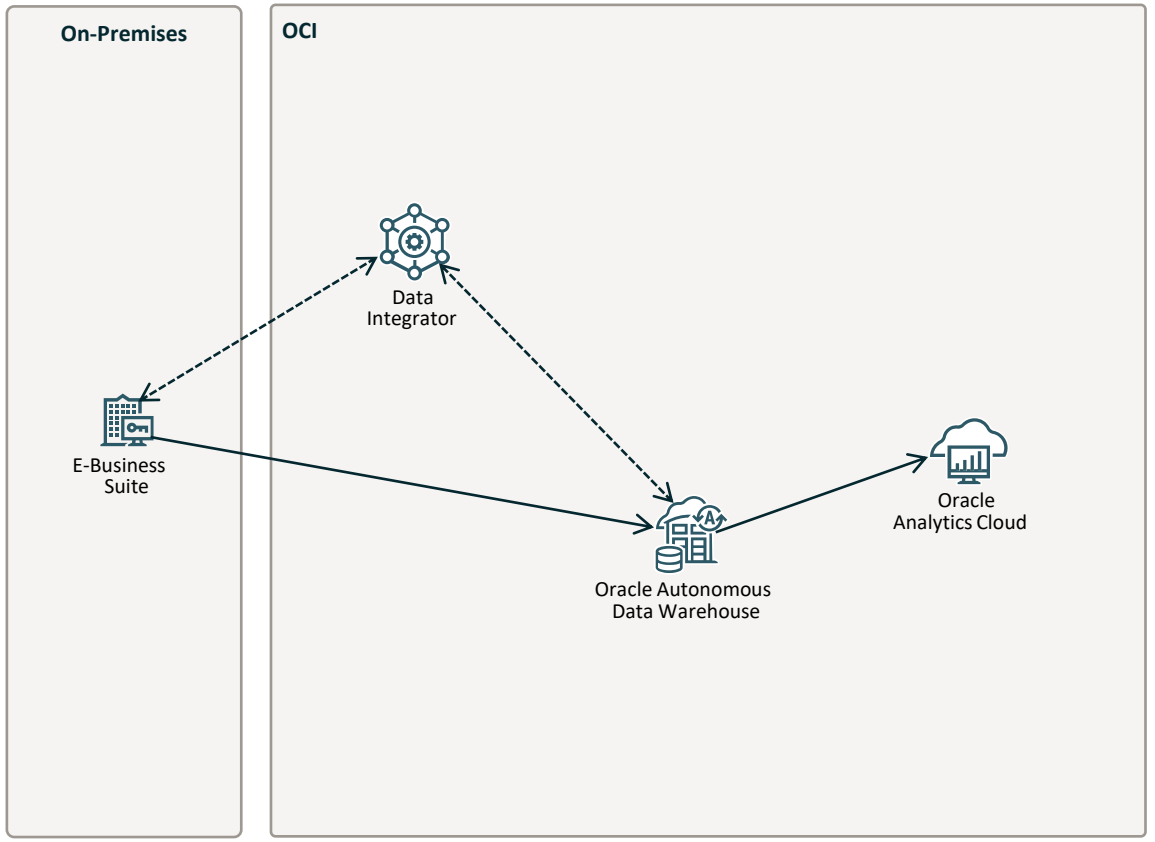
- Current sales planning and forecasting for its hand tools programme is still excel based or at best is using existing business analytics system.

Idea:

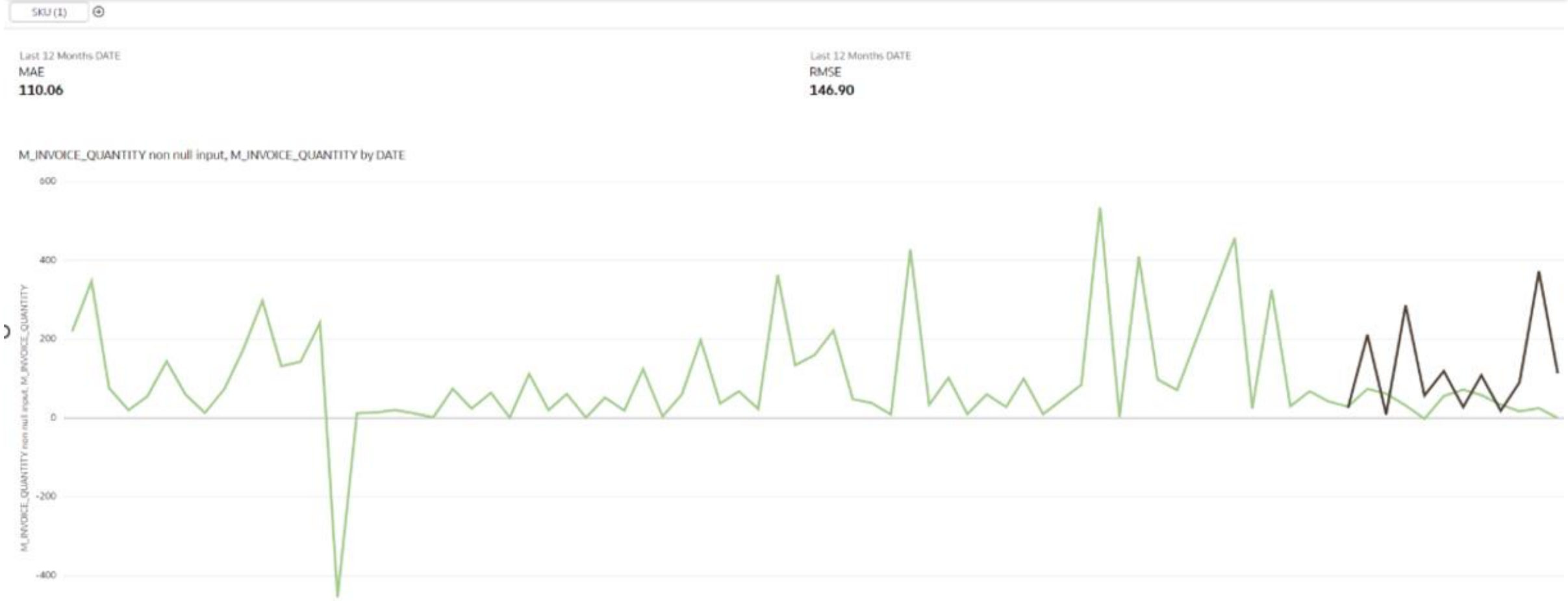
- Improve sales planning and forecasting with AI.

Existing Technical Architecture

- 1st Oracle Autonomous Data Warehouse + Oracle Analytics Cloud implementation in Europe in 2018/2019
- Using (on-premises) Oracle E-Business Suite as a main data source Business Analytics System. The following business areas have been implemented:
  - Financials (GL, AP, AR, FA)
  - Procurement
  - Order Management
  - Inventory Management
  - Projects/Project Management
  - HR
  - Manufacturing (planned)

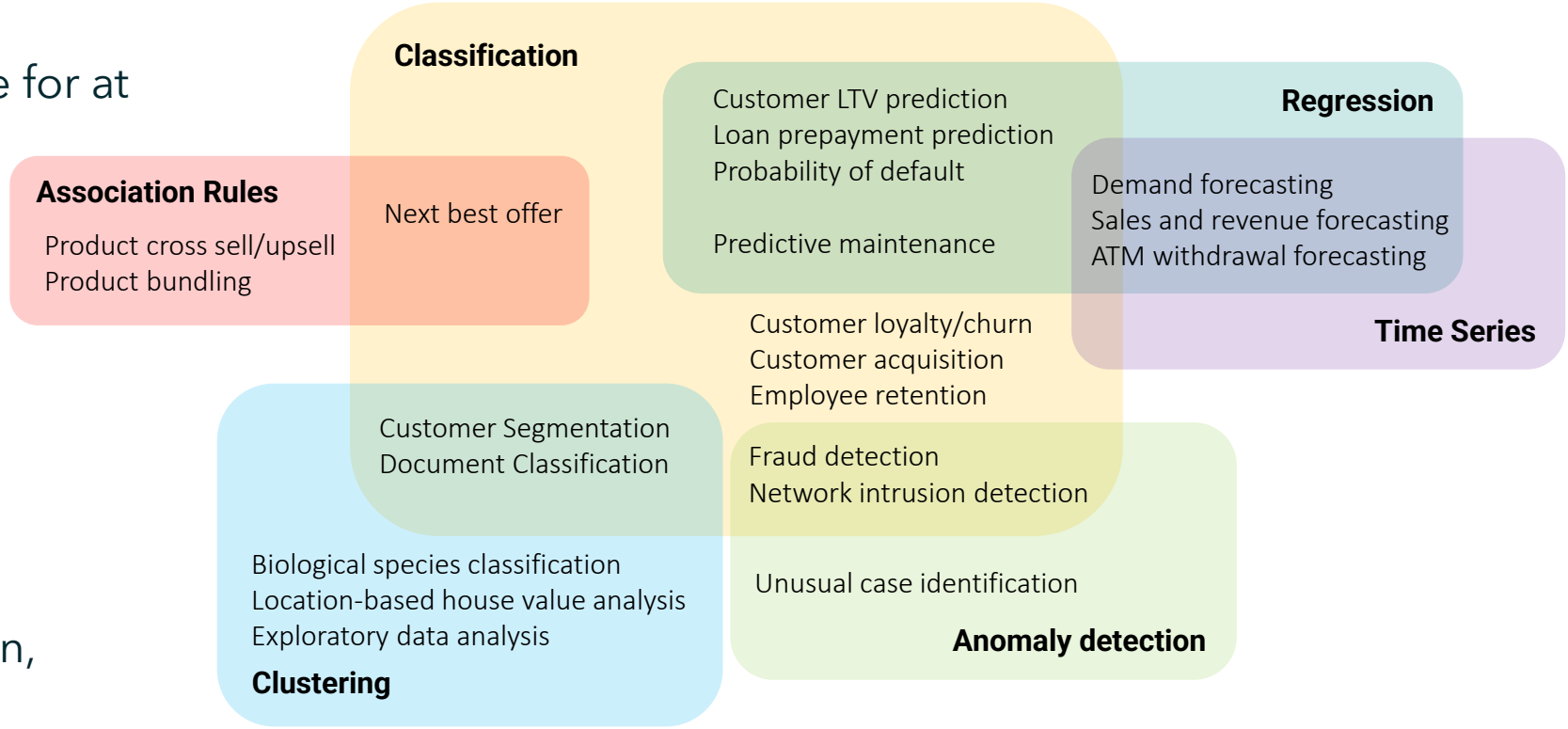


TASK: How can we use machine learning to improve sales planning in Hand Tools department?



Oracle Machine Learning (OML)

- OML is part of any Oracle database for at least 2 decades now.
- It's free option at no additional cost.
- Development options:
  - PL/SQL → OML4SQL
  - R → OML4R
  - Python → OML4Py
- Provides support for
  - supervised learning (regression, classification),
  - unsupervised learning (clusters, outliers),
  - plus, some other methods like time series, anomaly detection or market basket analysis.

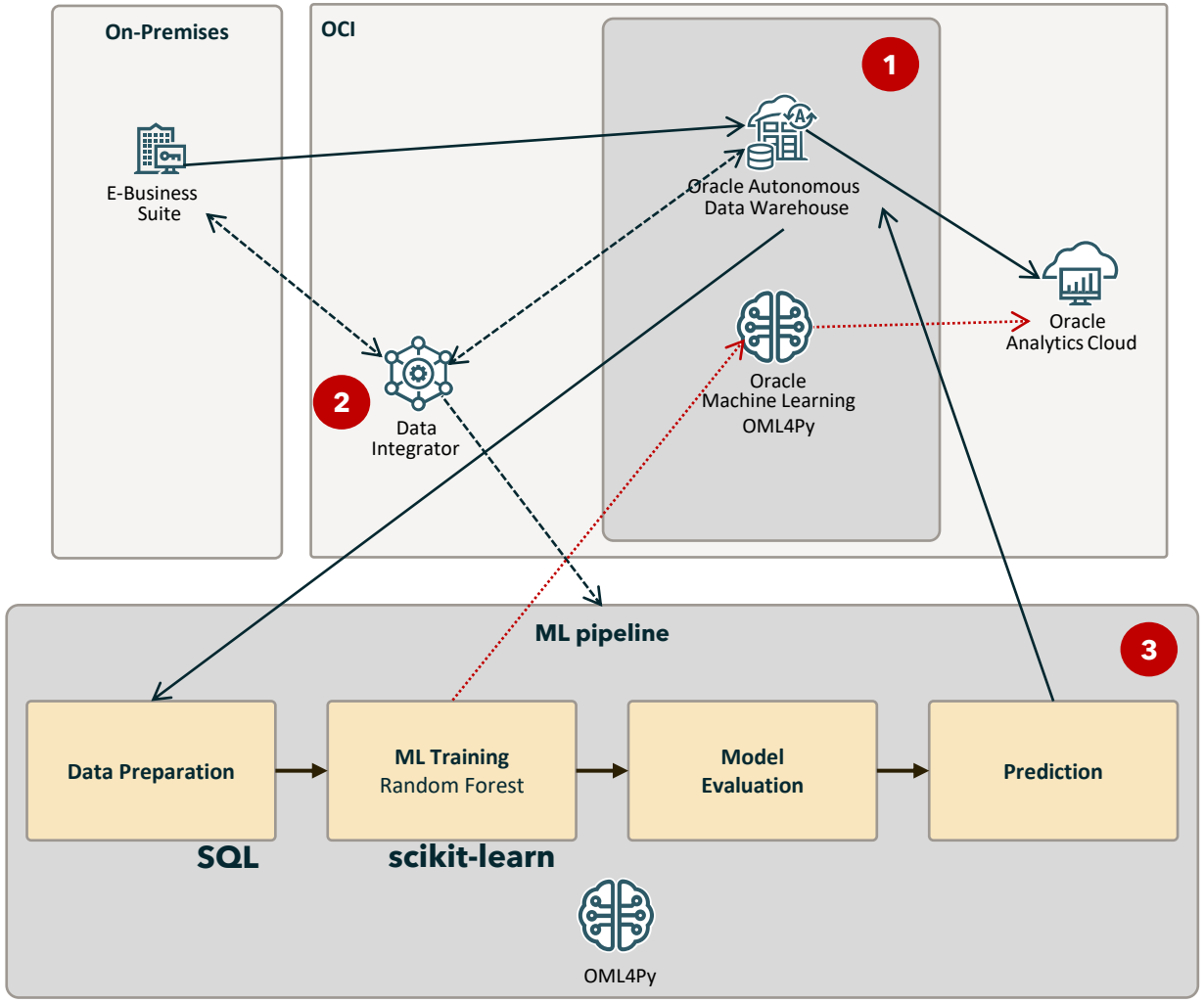


Key value proposition for the client

- all data stay in database
- no data movements is needed
- no additional tools (costs!) are required

Updated Technical Architecture*

*Architecture is actually unchanged, except it is using OML as additional service in ADW.



1 **Machine Learning script** is prepared and stored in Oracle Autonomous Data Warehouse using Zeppelin Notebooks:

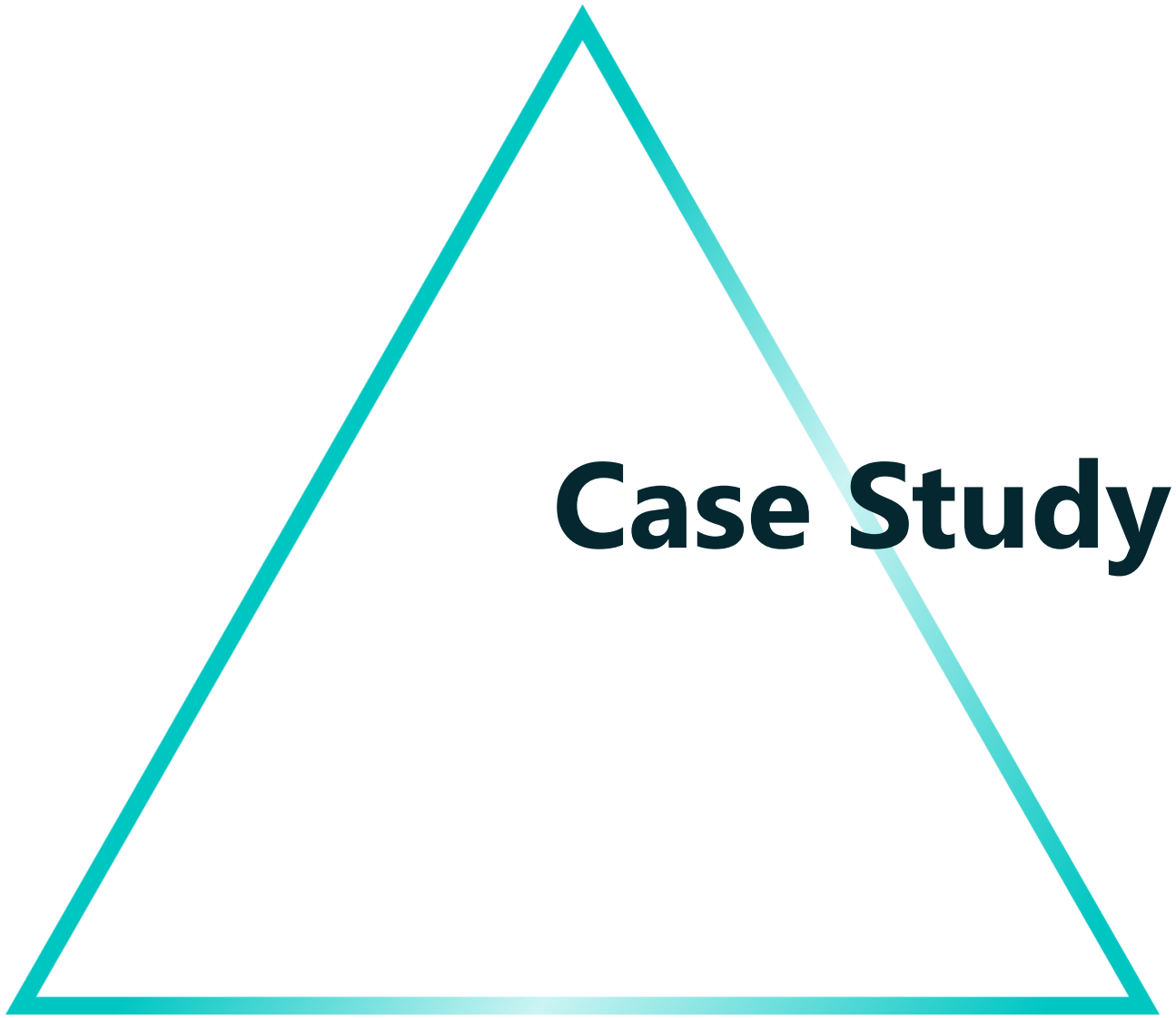
- SQL script is used to analyse and prepare data.
- Prediction model is trained using Random Forest algorithm in Python.
- Additional script is prepared to be executed via REST API call from ODI.

2 **ODI process**

- Prepares data in database.
- No data movement beyond database.
- Calls machine learning script using REST API call.
- Uses "prediction" results to populate data warehouse tables for predictive planning.

3 **ML Script Call**

- Runs python "REST API" script that calculates "predictions" for all SKUs (20000 SKUs).
- Parallel ML pipelines are executed to improve performance of model training.
- In case there is not enough data available, 12-month average is taken as predicted value
- For predictions model, default scikit-learn library (random forest) is used.
- Results are stored back into database tables.



Case Study #2

Customer #2

Top 5 European manufacturer of fuses, electrical switches and disconnectors.

Business problem:

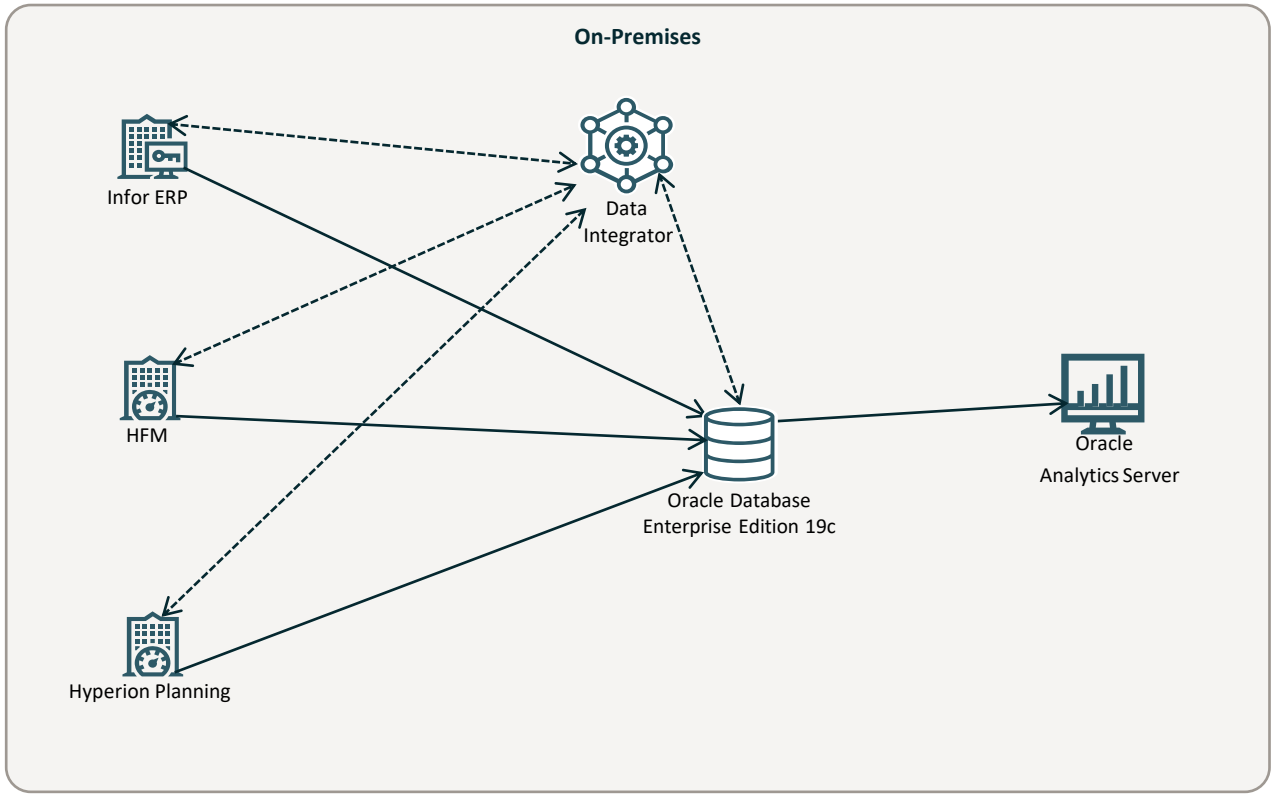
- Inventory levels are too high and tie up too much cash, which significantly affects cash flow.

Idea:

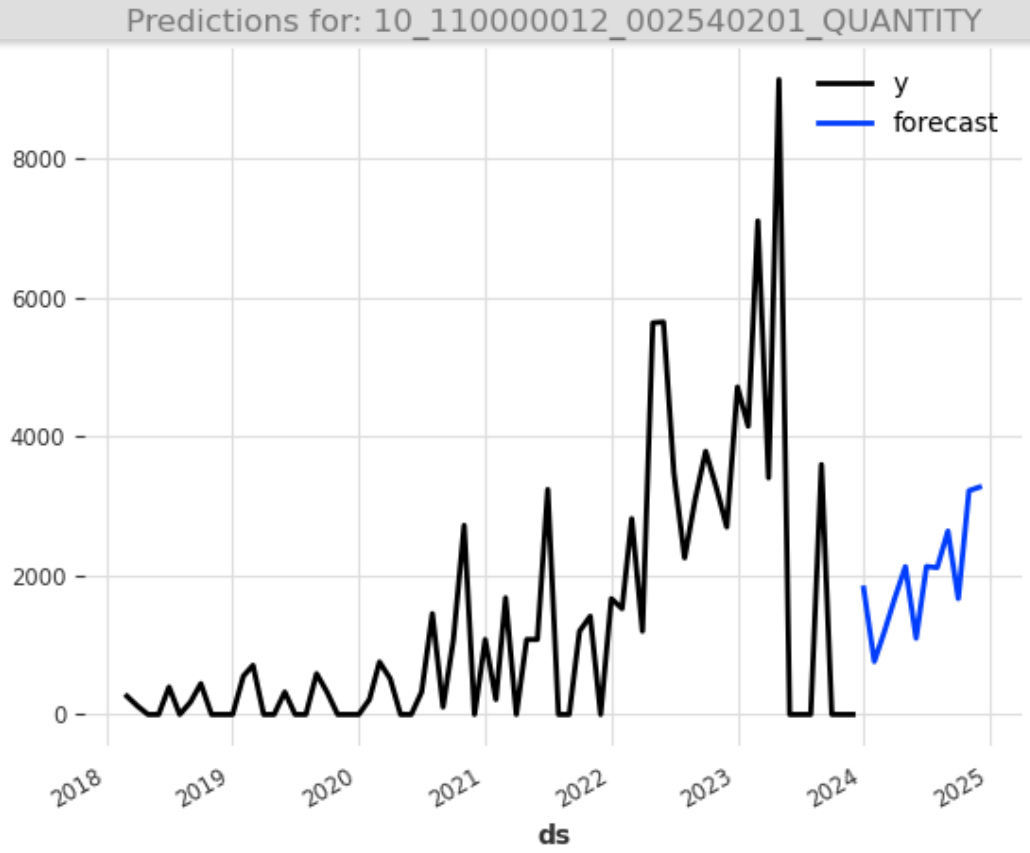
- Provide sales teams tools that would improve sales planning which would consequently affect inventory levels.

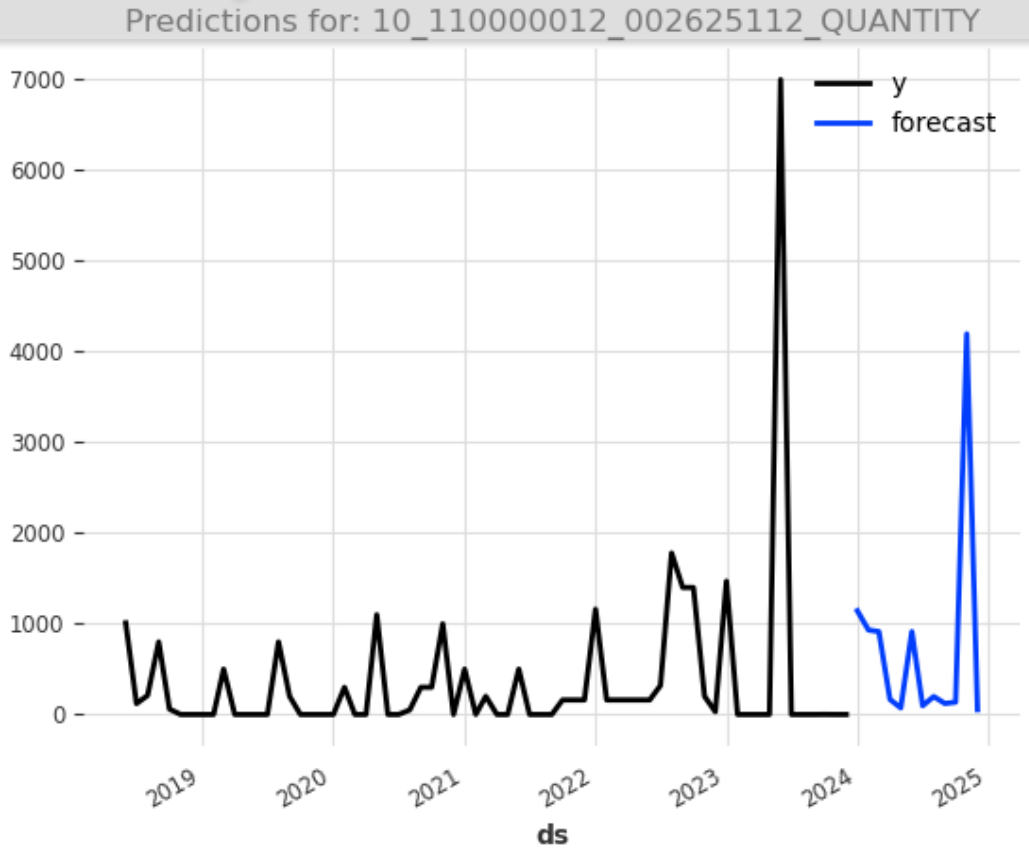
Existing Technical Architecture

- Business intelligence solution (started with OBIEE 10g, OAS today) since 2009.
- Analytics system comprises of
  - Oracle Analytics Server is used for reporting.
    - 30+ subject areas live and are constantly updated/added
  - Data Warehouse with data from key data sources:
    - INFOR ERP
    - Hyperion Financial Management (HFM)
    - Hyperion Planning
- Current sales planning process:
  - Everything is done manually in Hyperion Planning.
  - Data, that sales planners are using as help for planning, resides in company's Analytics System (Oracle Analytics Server)



TASK: Create additional Scenario dimension member *Prediction* that is a result of a machine learning solution that predicts future sales for 12 months ahead by all SKUs by all markets



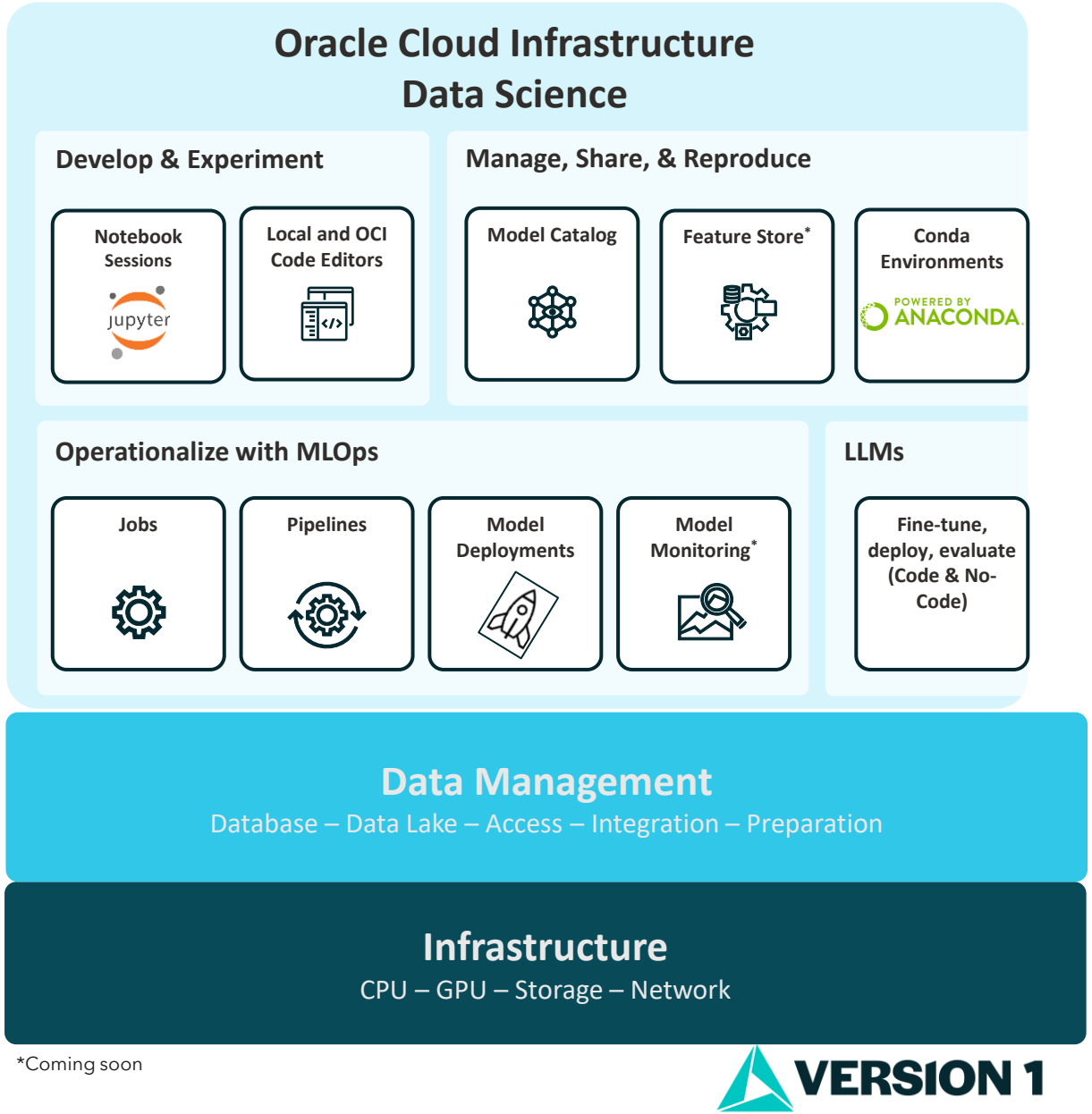


OCI Data Science

- Accelerate and automate the entire end-to-end data science lifecycle
- Use favorite open-source Python tools and frameworks
- Fine-tune and deploy Large Language Models (LLMs) without writing code
- Enterprise-grade MLOps with flexible interfaces and unlimited scale
- Collaborate with teammates on shareable and reproducible data science assets
- Run large-scale workloads with access to GPUs and distributed data processing and model training
- Pay only for on demand infrastructure with no additional markeup or overhead

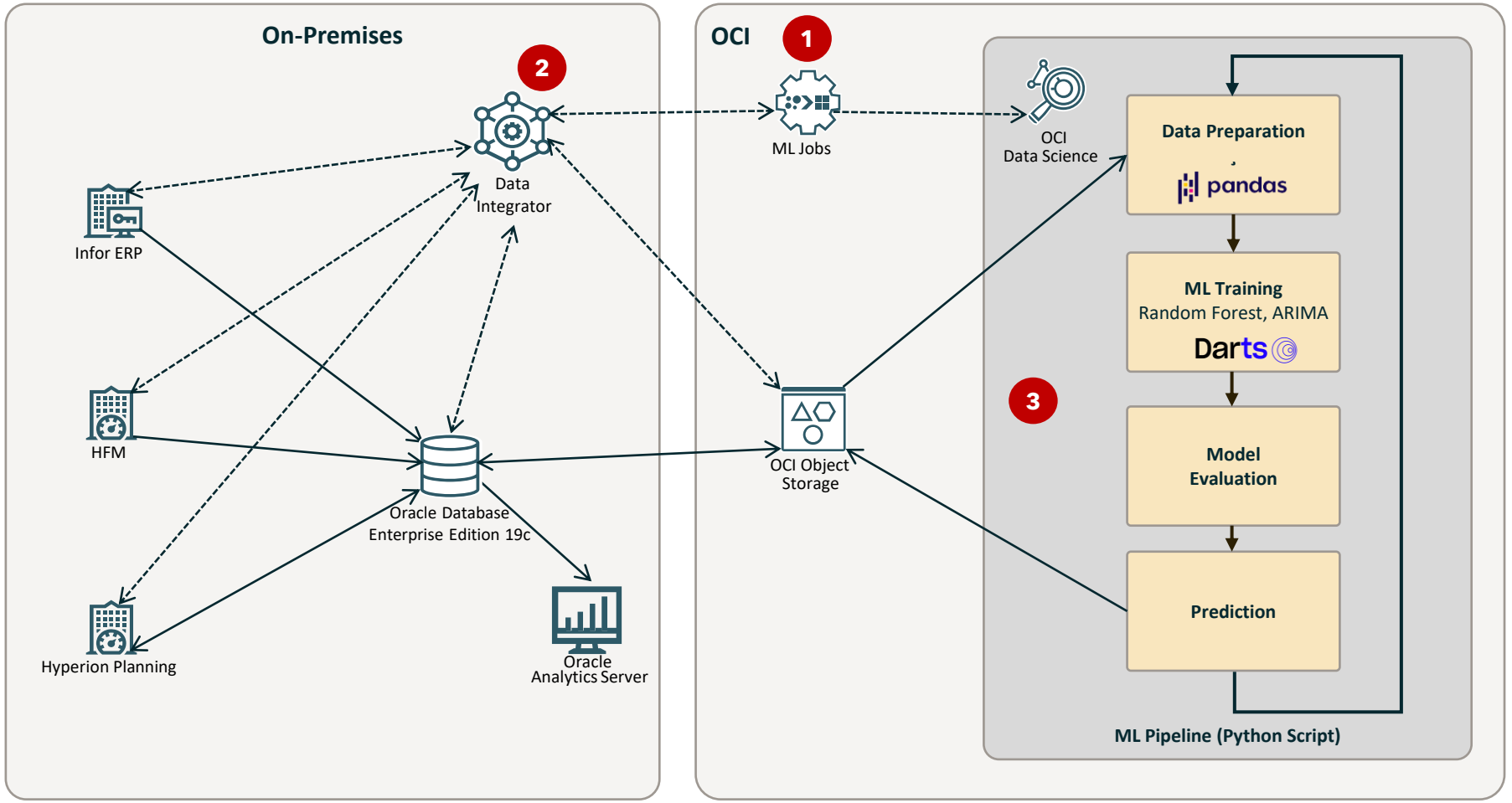
Key value proposition for the client

- Low cost of deployment
- Flexible and comprehensive data science platform for the future.



*Coming soon

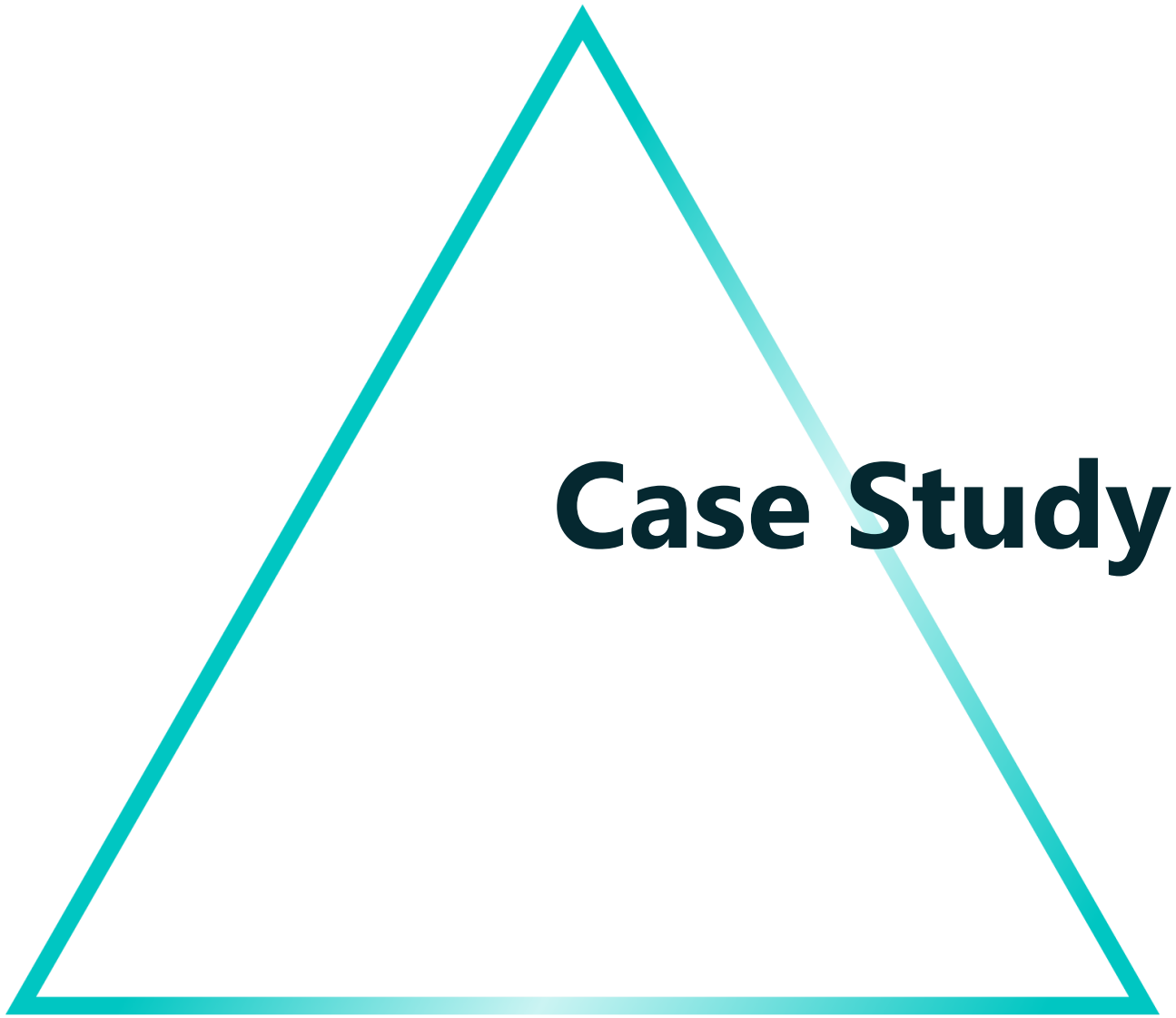
Updated Technical Architecture and ML process



1. **Python script** is prepared using OCI Data Science Jupyter Notebook and "attached" in ML Job as job artifact.

2. **ODI process**
   - Prepares data in database.
   - Uploads data to Object Storage.
   - Calls ML Job to run "predictions"
   - Retrieves data from object storage after "predictions" are prepared.
   - Populates tables in data warehouse, staging area and Hyperion Planning in the end.

3. **ML Job Run**
   - Runs python script that calculates "predictions" for all SKUs across all markets
   - Script reads data from Object Storage
   - For prediction, we are using Darts library which is dedicated time series predictions library.
   - At the end of the process, "predictions" are stored back to Object Storage.



Case Study #3

Customer #3

One of the leading producers of stainless steel and special steels in Europe.

Business problem:

- As part of company's digital transformation, the company wants to measure key manufacturing data in real-time to be able to react to anomalies on time.

Idea:

- Implement solution that would capture MES and sensor data, analyze it in real time, and take action in case of predicted anomalies.

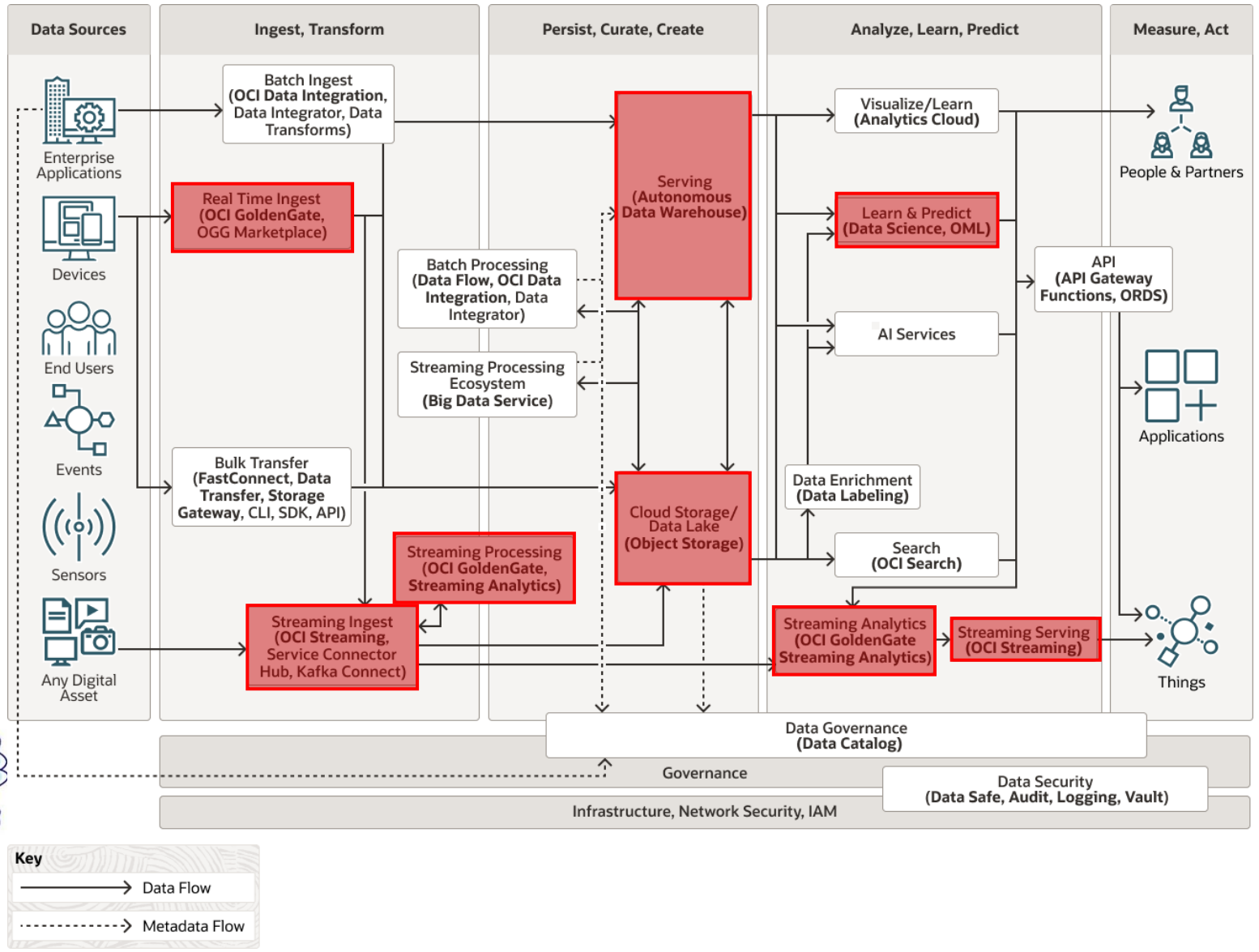
After initial assessment …

Three are main **business requirements** and **focus areas**:

- Replicate MES data to the cloud in real time.
- Connect to “brokers” and capture real-time data from sensors and meters in manufacturing process.
- Develop ML/AI solutions that would help them to optimise use of their resources, identify anomalies or make predictions (i.e. water consumption or electricity usage)

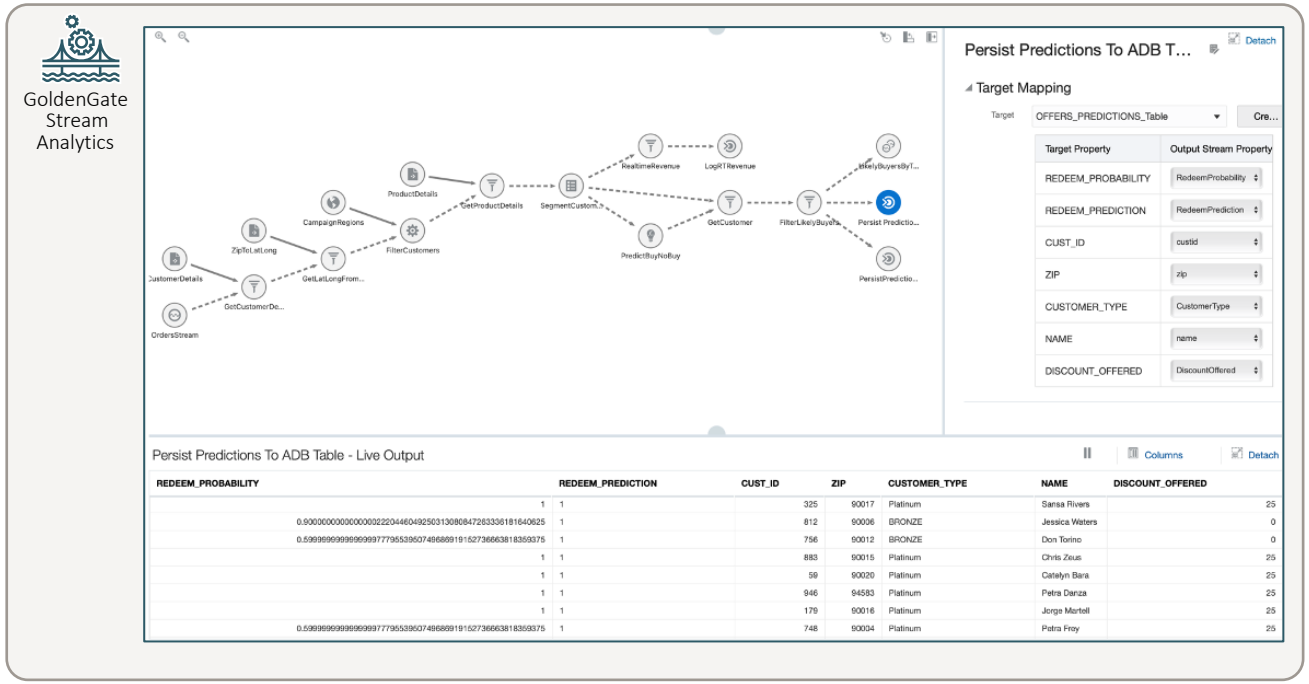
Proposed solution

- Use of Golden Gate for MES data replication.
- OCI Streaming for stream data capture and processing that data with GoldenGate Stream Analytics
- OCI Data Science for development and deployment of ML/AI models



Oracle GoldenGate Stream Analytics

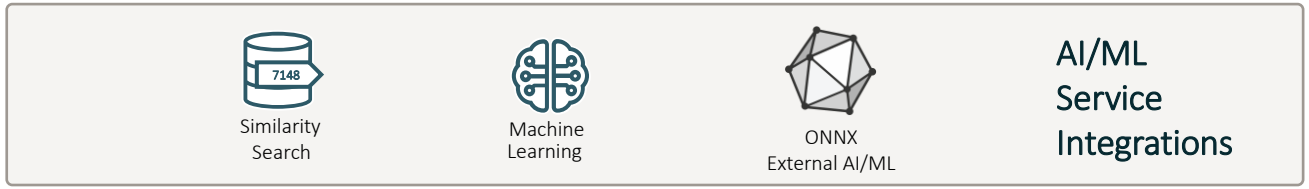
Events



Insights

Alerts

Data products



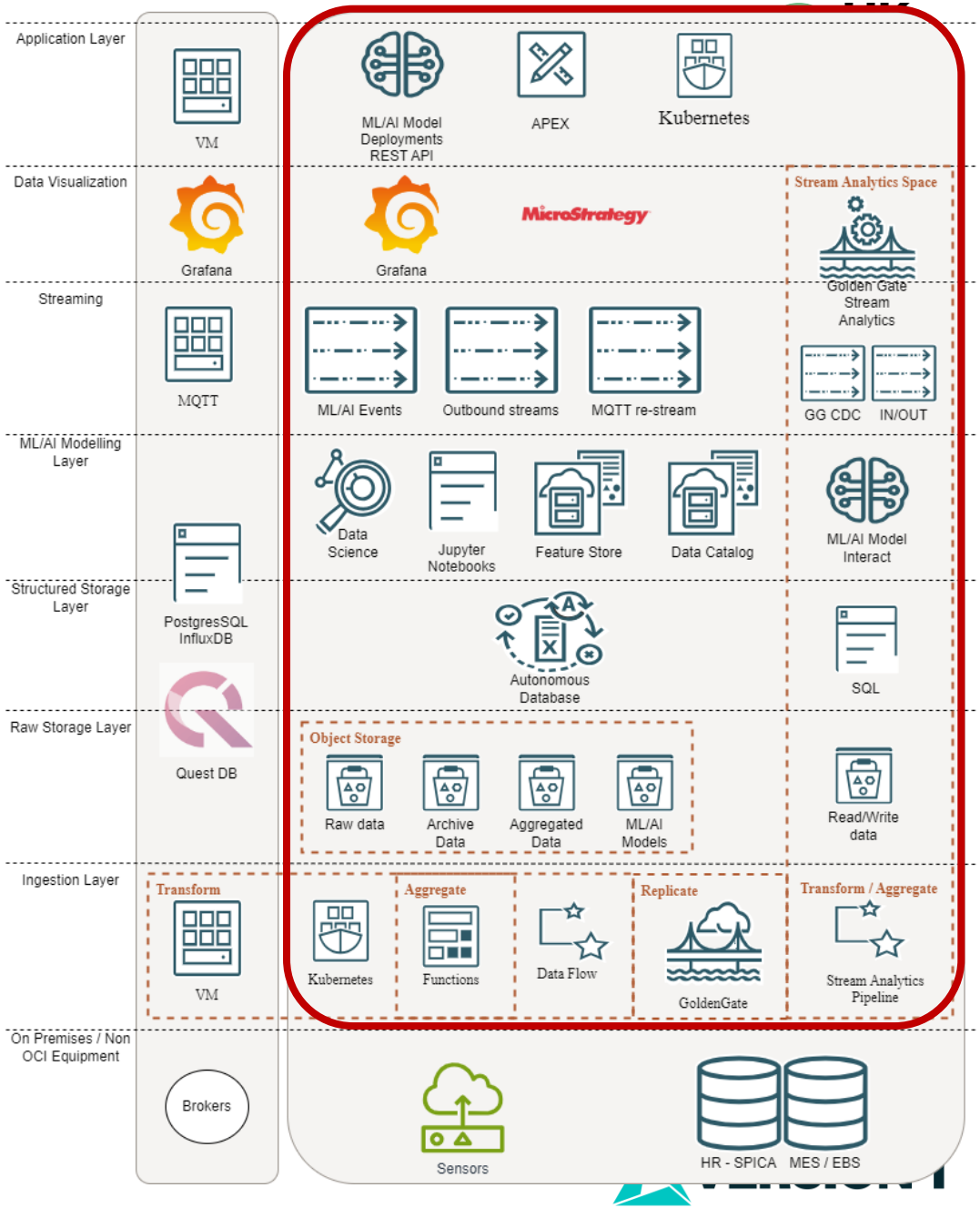
Ingest Events	Select Processing Patterns	Build Event Pipelines	Serve Data Downstream
100+ supported sources	Pre-build patterns to improve developer efficiency and time-to-value	Leverage geo-fencing, machine learning and other reference data within the stream	Data can be delivered to Kafka, databases, or can be staged for external ETL jobs.

… and after Solution Design

- It's not just some OCI products, but we are talking about full OCI Data Lakehouse implementation that includes several services such as:
  - Oracle Autonomous Database
  - OCI Compute
  - OCI Object Storage
  - OCI Streaming
  - OCI Data Flow
  - OCI GoldenGate (replications)
  - OCI GoldenGate Stream Analytics
  - OCI Data Science
  - OCI Functions
  - Kubernetes
  - (unfortunately, no Oracle Analytics Cloud)
- Services are organised into logical layers which are planned to support MES and sensor data real-time streaming and ML/AI analysis.
- The 1st phase of the project is about data management platform rather than data analysis, which is planned for the 2nd phase.

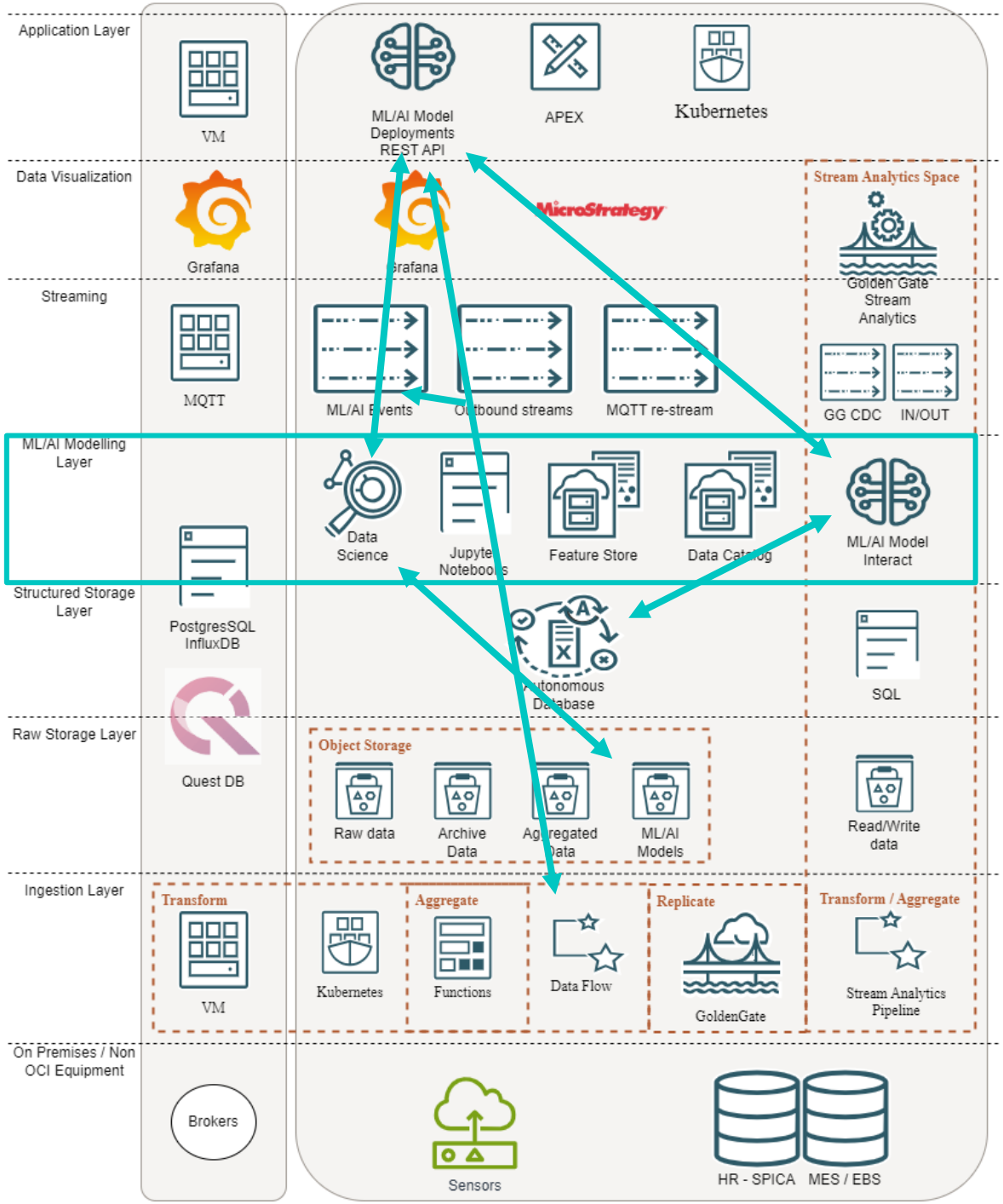
Key value proposition for the client

- They already use EBS in OCI, this project is natural extension of their OCI usage.
- Enterprise-level services in one place.



ML/AI Layer

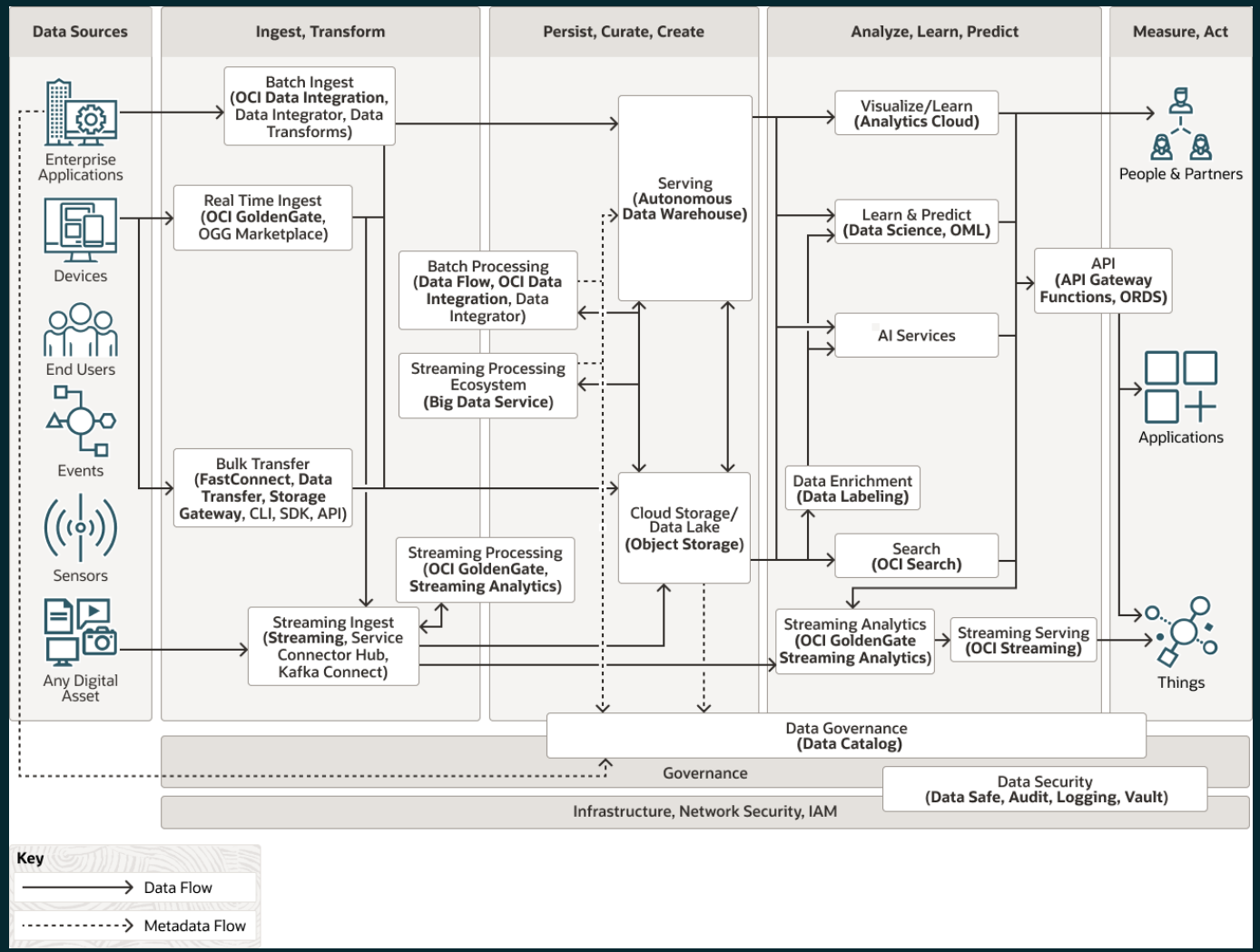
- Advanced tools for machine learning
  - Model lifecycle management
  - Jupyter Notebooks for interactive development
  - Support for popular libraries:
    - TensorFlow
    - Scikit-Learn
    - PyTorch
- Oracle OCI Cloud SDK for connectivity with other services in Oracle OCI Cloud:
  - Data Catalog
  - Feature Store
  - Object Storage, Functions, Data Flow, etc.
- Easy collaboration for multiple analysts' teams.
- Serverless Docker and Kubernetes technology for running models and notebooks.
- REST API for easy calls and responses of ML/AI models, suitable for integration into event-based streaming systems.



Demo: GGSA & ML

Conclusion

- Use machine learning in Oracle Cloud Infrastructure for **its powerful, scalable, and secure environment** that simplifies model development and deployment.
- Use Oracle Machine Learning (OML) for integrated, scalable, and automated machine learning **directly within Oracle databases and cloud environment**.
- Use OCI Data Science for its seamless, high-performance platform that **simplifies and accelerates the entire machine learning lifecycle**.
- Use machine learning in Oracle GoldenGate Stream Analytics to make **real-time predictions and automate actions based on streaming data**.



VERSION 1

UK OUG