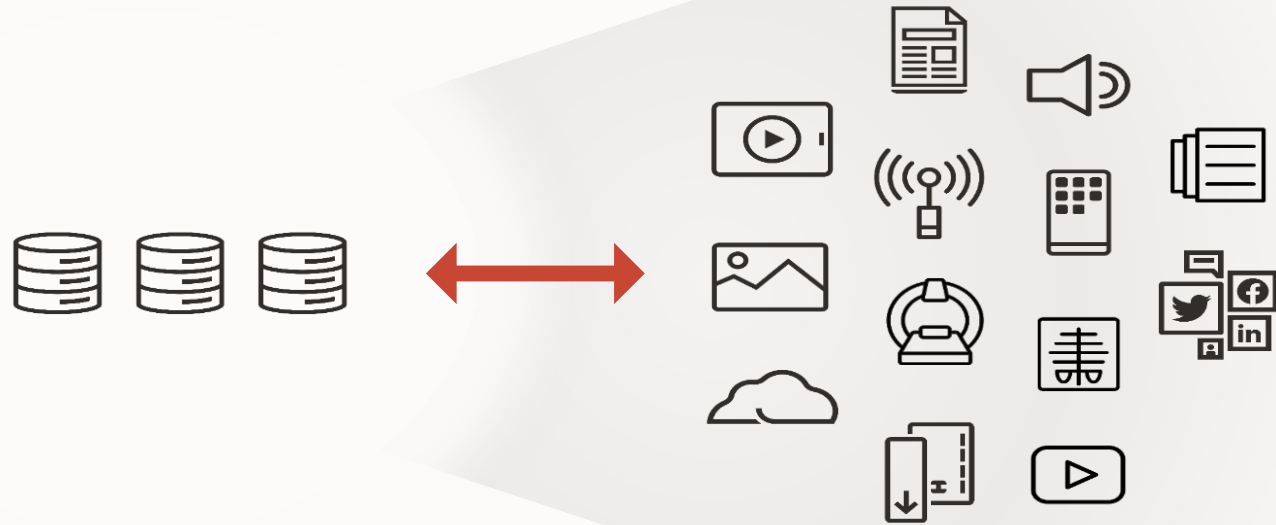# Extracting greater value from your data infrastructure

Ghassan Al Jamil
Business Development Director – Analytics
Oracle
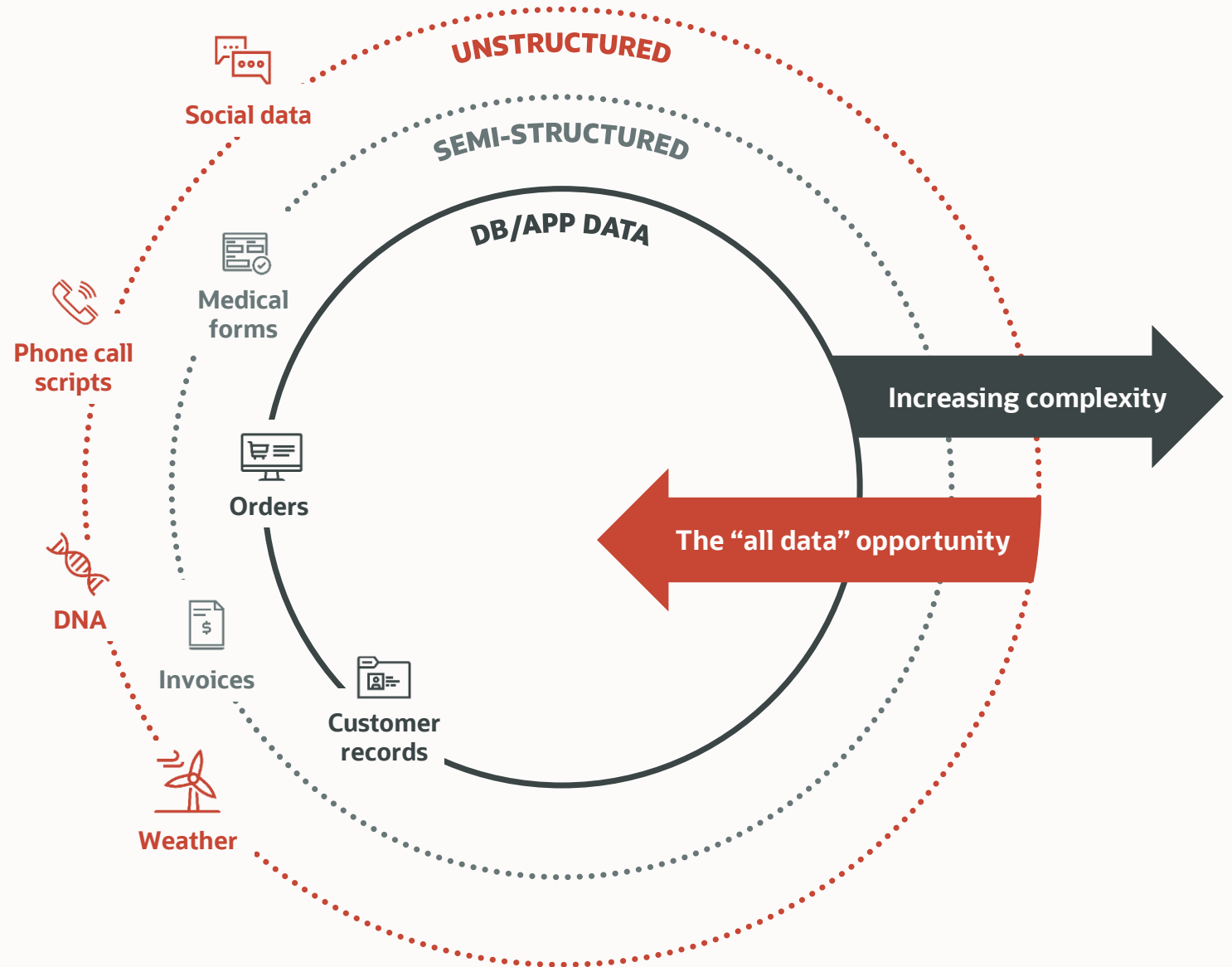
ORACLE

Using **all data** to innovate

Why is **analyzing** **all data** so hard?

UNSTRUCTURED

SEMI-STRUCTURED

DB/APP DATA

Social data

Phone call scripts

Medical forms

Orders

DNA

Invoices

Customer records

Weather

Increasing complexity

The "all data" opportunity

The Foundation
Modern Data Management Platform

Data Lakehouses on OCI

A unified, open analytics platform for all data
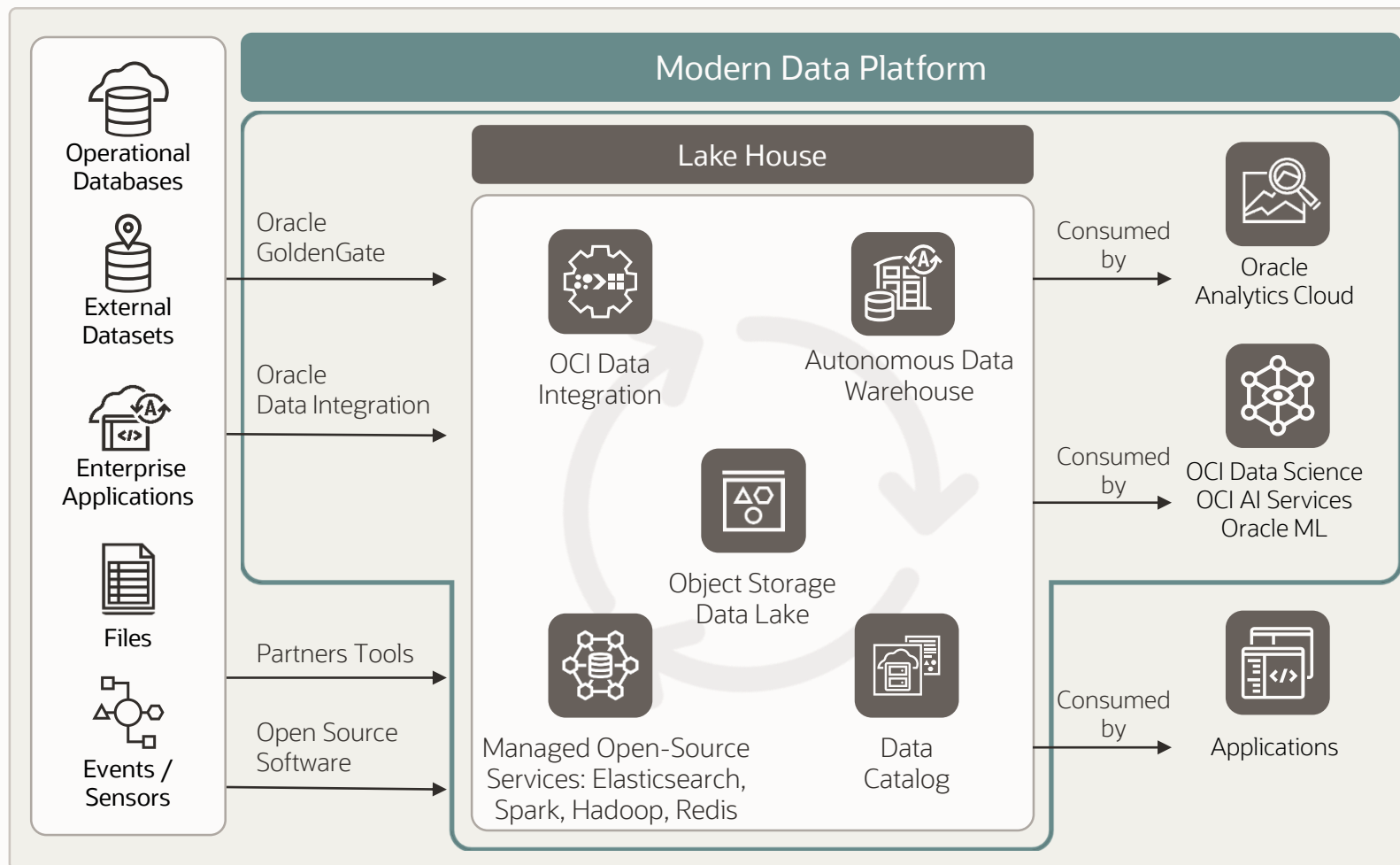
Comprehensive range of new services

OCI in action

The evolution of data management and analytics

Data warehouse

Data warehouse

Data Lake

Lakehouse

The Oracle Lake House - Five key elements

The core of the modern data platform

Autonomous Data Warehouse: Automated management with high-performance storage and analytics

Object Storage Data Lake: Low-cost storage for varied data

Managed Open-Source Services: Spark, Hadoop, Elasticsearch, Redis, etc. fully managed and integrated
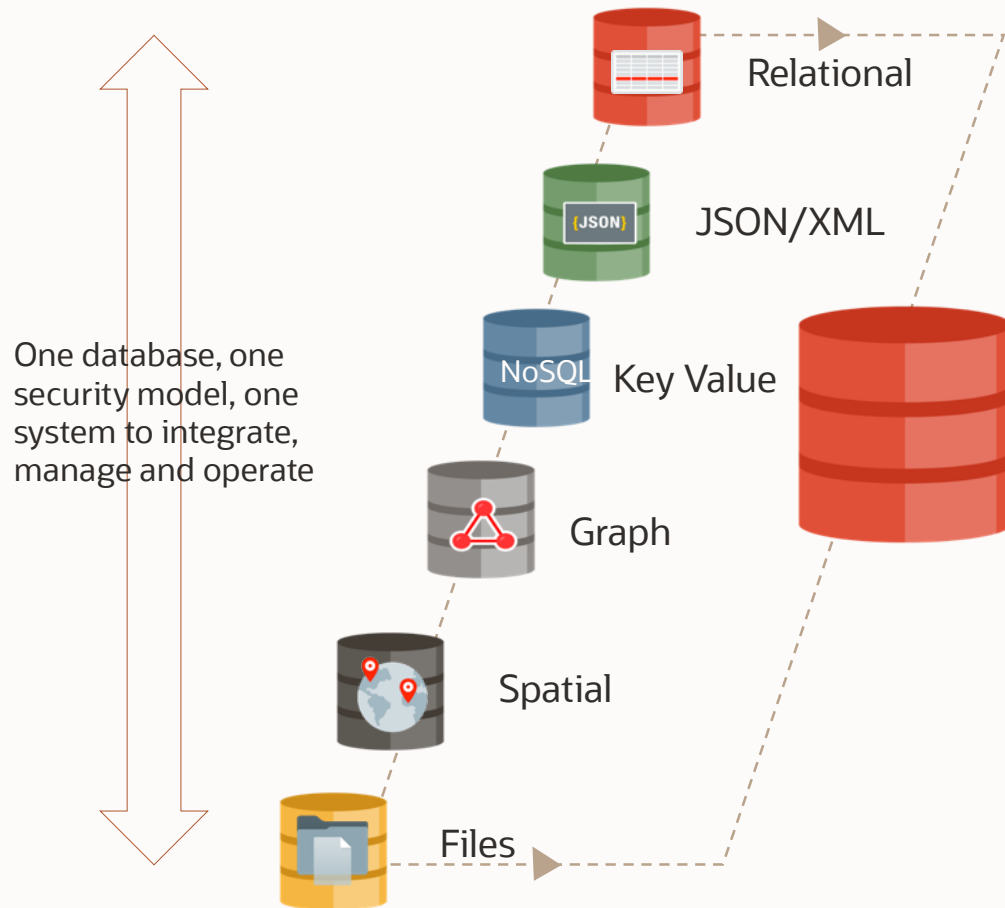
OCI Data Integration: Easily extract, transform, and load (ETL) data for analytics and data science. Design code-free data flows between data lakes and data warehouses

OCI Data Catalog: Maintains an inventory of assets used by both the data lake and data warehouse for data discovery

Key Differentiator - Oracle converged database capabilities are broad

Synergies across features makes the whole better than the sum of the parts

Relational

JSON/XML

NoSQL   Key Value

Graph

Spatial

Files

One database, one security model, one system to integrate, manage and operate

- **Integrated low code application development (APEX)**

- **Machine Learning**

- **Node.js, Python, many others**

- **In-memory**

- **Spatial, Graph support**

- **NoSQL (JSON, key-value, wide column, XML)**

- **Blockchain tables**

- **Multitenant**

- **REST APIs**

- **Deployment choice (on-prem, cloud, hybrid)**

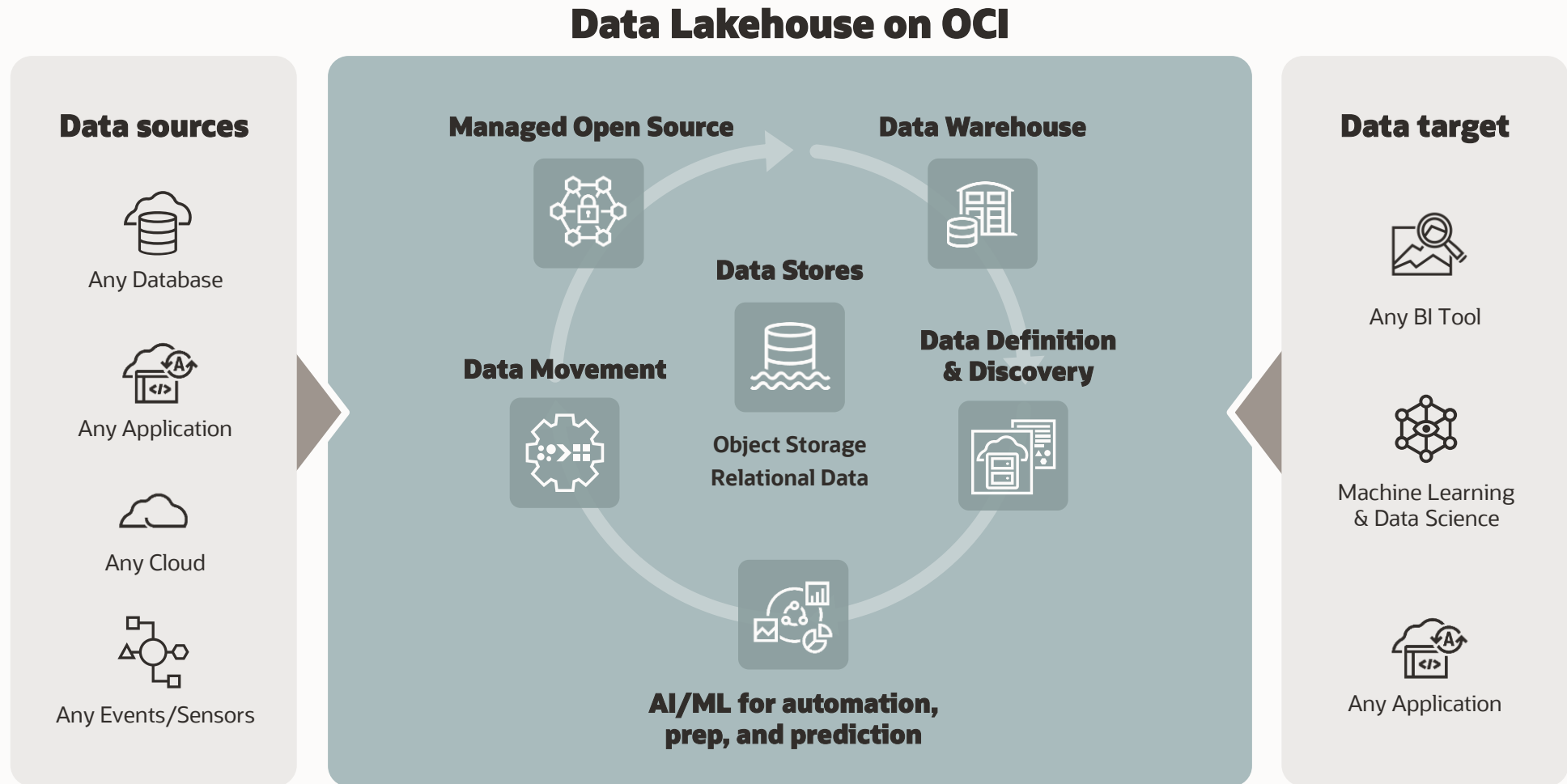# Key Differentiator: Easy deployment of multiple data models
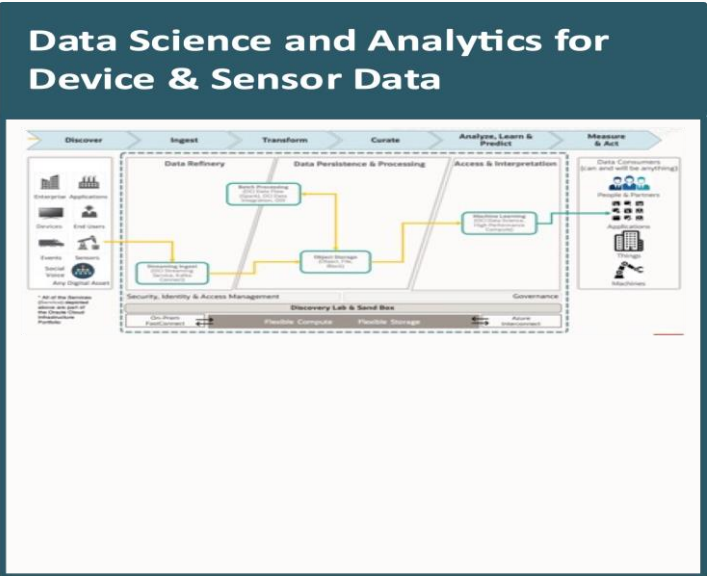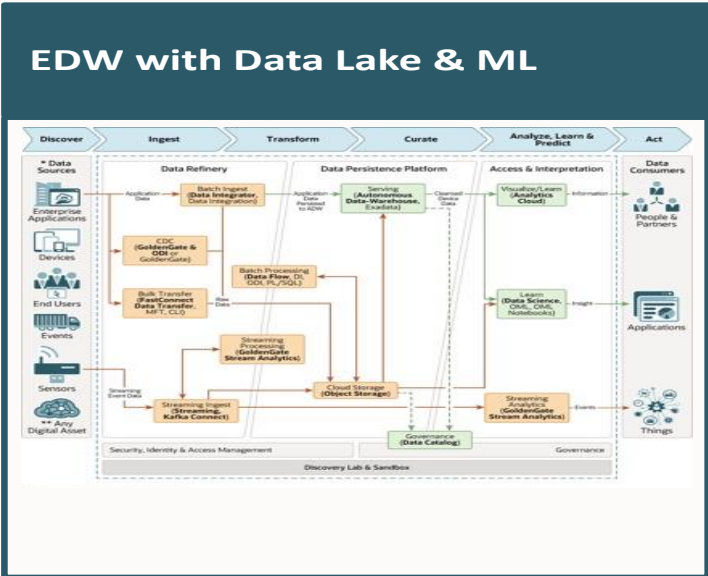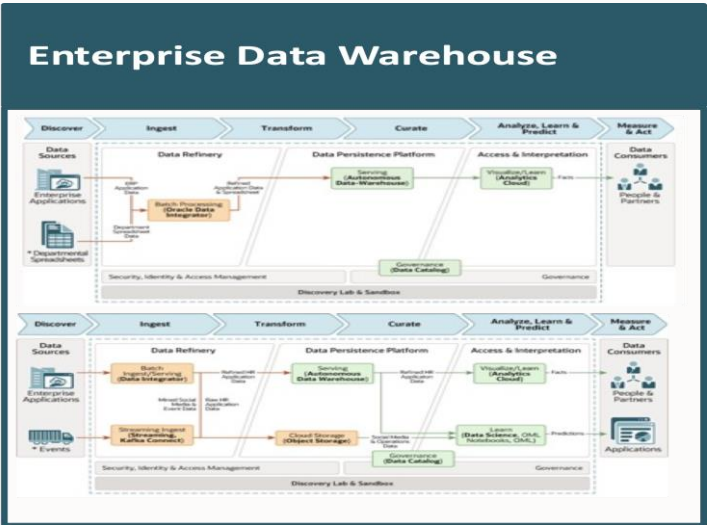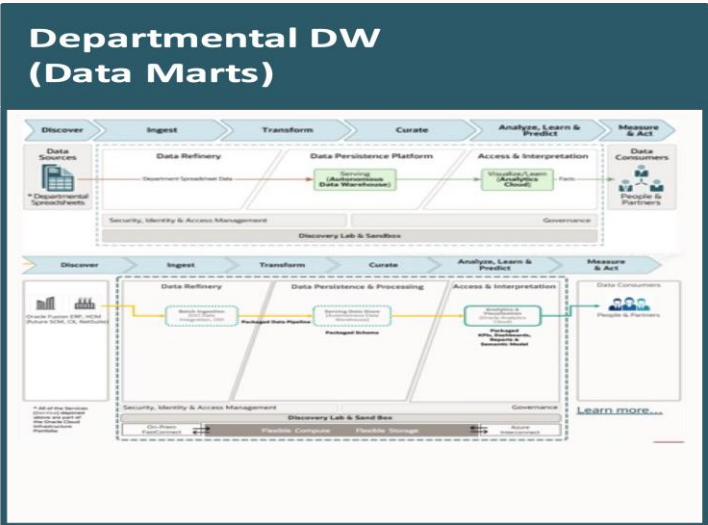
Data Lakehouses on OCI

Open & flexible: analyze any database, any application, from anywhere

Data Lakehouse on OCI

Data sources

Any Database

Any Application

Any Cloud

Any Events/Sensors

Managed Open Source

Data Warehouse

Data Stores

Data Movement

Data Definition & Discovery

Object Storage
Relational Data

AI/ML for automation, prep, and prediction

Data target

Any BI Tool

Machine Learning
& Data Science

Any Application

OCI in action

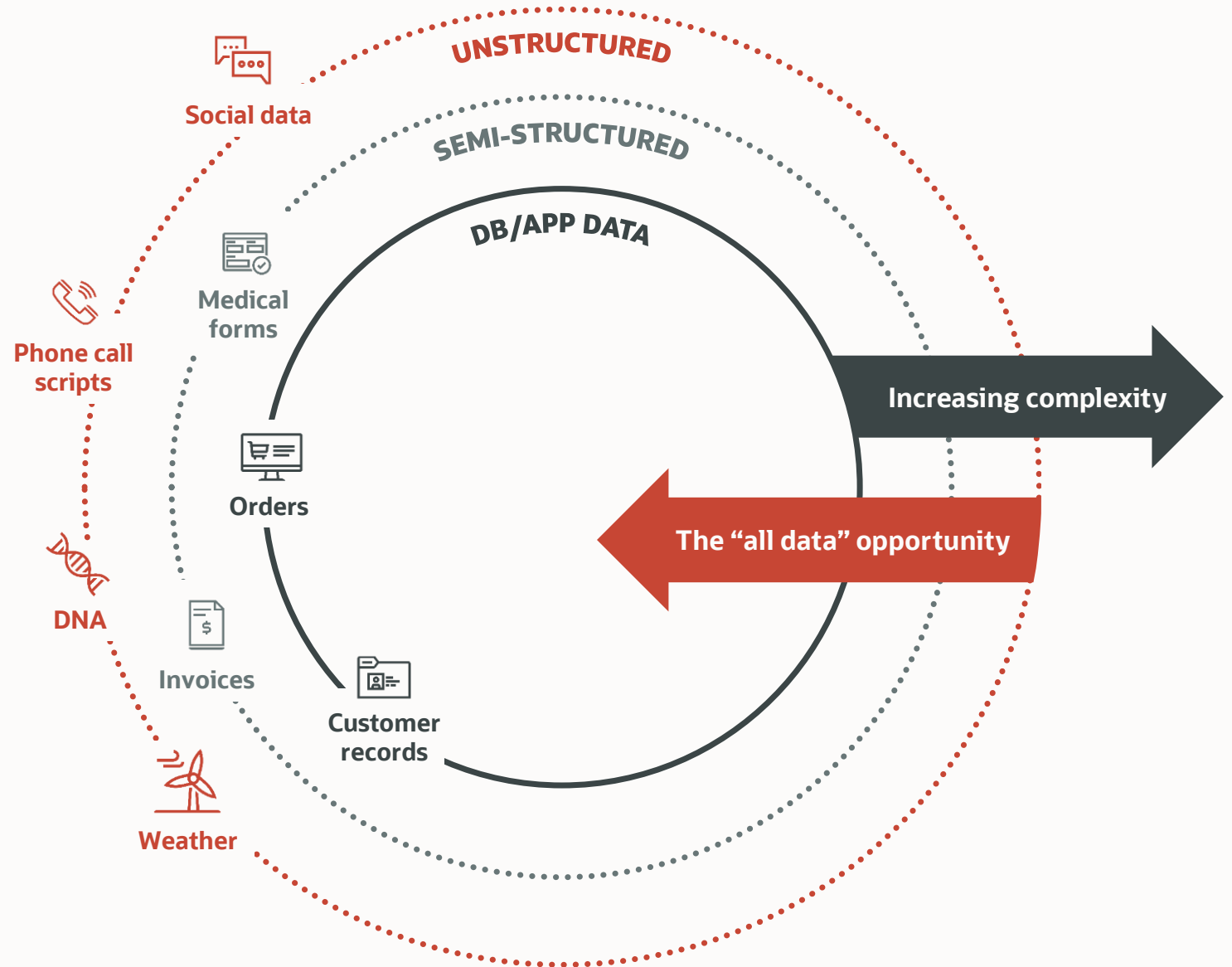Build new and modernize existing data lakes

Every data warehouse becomes a lakehouse

Build intelligent SaaS analytics on a lakehouse

# Innovate with Business-First AI

It all starts with **data**

UNSTRUCTURED

SEMI-STRUCTURED

DB/APP DATA

Social data

Phone call scripts

Medical forms

Orders

DNA

Invoices

Weather

Customer records

Increasing complexity

The "all data" opportunity

What customers want for their production AI
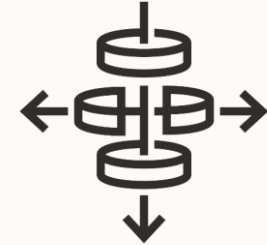
AI that works for your enterprise scenarios

- Pretrained, customizable models for your industry
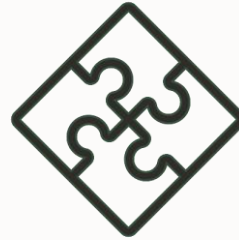
Collaborate with a single, consistent experience

- Single discovery and publishing experience for models, features, datasets and labels
- Consistent APIs

Open and extensible platform

- Use favorite open-source tools and frameworks
- Run your AI models anywhere
- Portability- no lock-in

Oracle AI

Unified AI/ML platform spanning cloud services, apps and data assets

Applications

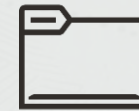|					
OCI Digital Assistant	OCI Language	OCI Speech	OCI Vision	OCI Anomaly Detection	OCI Forecasting

AI Services

OCI Data Science	Oracle Database Machine Learning	OCI Data Labeling

Machine Learning Services

Data

Introducing OCI AI Services

Digital Assistant

Language

Speech

Vision

Anomaly Detection

Forecasting

Machine Learning Services

Data

Language
Understanding text

Common use cases

- Analyze customer feedback – product reviews, support tickets, call center interactions, social media posts

- Streamline and automate document-centric processes – contracts, RFPs, financial documents

Challenges

- Language understanding is context dependent, which is different at every organization

- Interpreting different languages with high fidelity is challenging

- Large models are expensive to train (BERT, GPT-3)

NEW

Introducing OCI Language Service

Pretrained models trained on industry data to perform language analysis with no data science expertise needed

Detect the language of your text

Identify key phrases and entities in text

Classify text content into more than 600 categories and subcategories to support data analysis

Aspect-level sentiment analysis

In 2021, Oracle released the OCI Language service. It uses pre-trained machine learning models to make text analytics easy I love this AI service – it is great!

Language Detection	English (0.991)		
Named Entity Recognition	2021 [**DATE**] (0.9914)	Oracle [**ORG**] (0.9942)	
Sentiment Analysis	Oracle **(Positive)**	AI Service **(Positive)**	OCI Language service **(Positive)**
Key Phrase Extraction	OCI Language service (0.9710)	Pre-trained machine learning models (0.9710)	
	text analytics (0.9710)	AI service (0.9710)	Oracle (0.9763)
Text Classification	science and technology / computer science (0.9851)		

Speech
Transcribing audio into text

Common use cases

- Call center analytics
- Media search and discovery
- Generating close captions for your media files
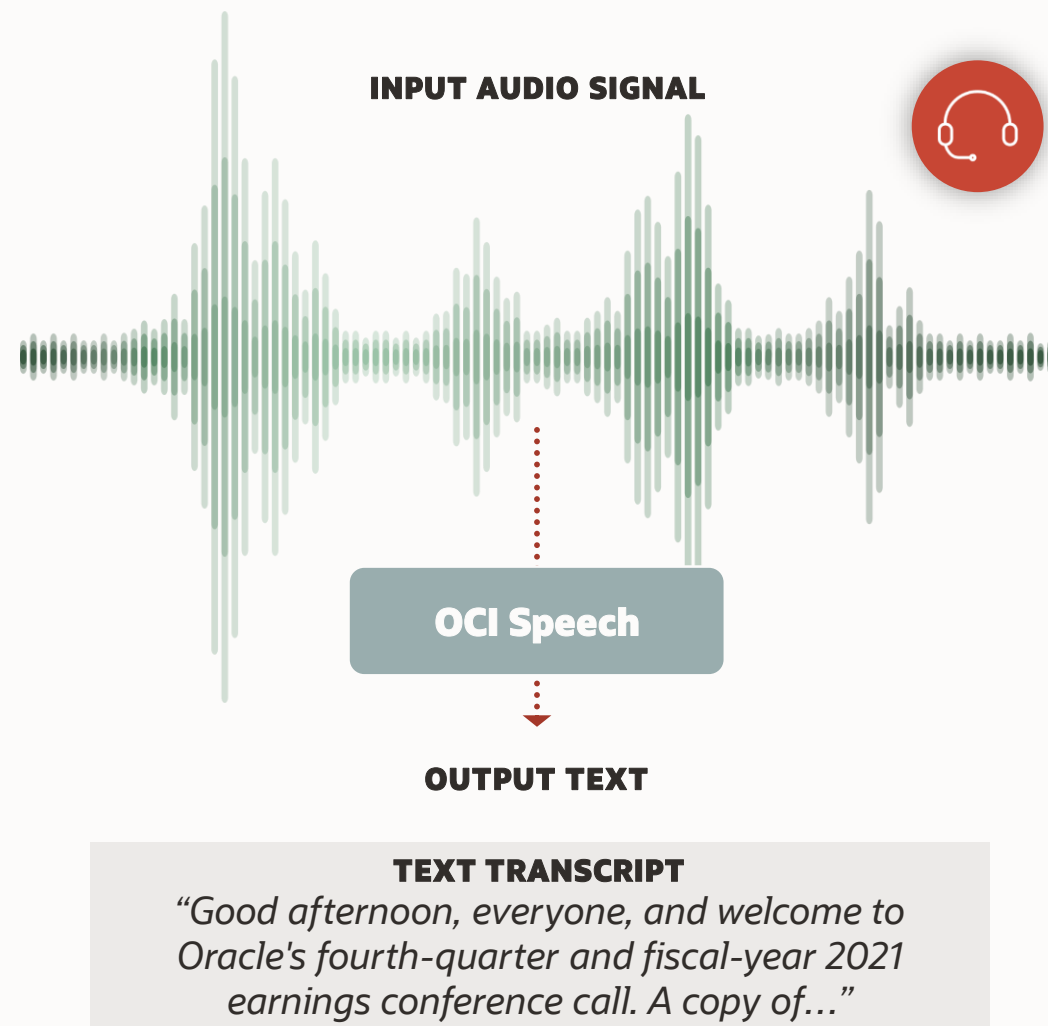- Powering conversational experiences

Challenges

- Requires data science expertise and curating large amounts of data to account for various languages, accents, and background noises
- Moving large amounts of data across cloud or applications introduces latency
- Expensive to do at scale

NEW

Introducing OCI Speech Service

- Automatically transcribe your audio and video files into text using advanced deep learning techniques.

- No data science expertise required

- Process data directly in object storage, no data movement needed

- Generate timestamped grammatically accurate transcriptions

INPUT AUDIO SIGNAL

OCI Speech

OUTPUT TEXT

TEXT TRANSCRIPT
"Good afternoon, everyone, and welcome to Oracle's fourth-quarter and fiscal-year 2021 earnings conference call. A copy of…"

Computer Vision
Recognizing content in images

Common use cases

- Digital asset management - annotate your documents, images and videos to make them discoverable and searchable

- Identify visual anomalies in your data

- Business process automation - recognize and extract text from your scanned documents or images (invoices, receipts, IDs, tax forms, etc.)

- Detect and classify objects in images

Challenges

- General-purpose models don't work for specific business scenarios

- Training custom models require data science expertise and significant investment in data labeling

- Understanding unclear text is difficult - handwritten, tilted, shaded, rotated, etc.

NEW

Introducing OCI Vision Service

Provides pretrained, customizable computer vision models and fully managed model infrastructure

Automatically classify documents into different types based on visual appearance, language, and extracted keywords.

Hybrid models that combine visual and text algorithms deliver more accurate results

Complete integration with OCI Data Labeling simplifies data labeling and provides more accurate models
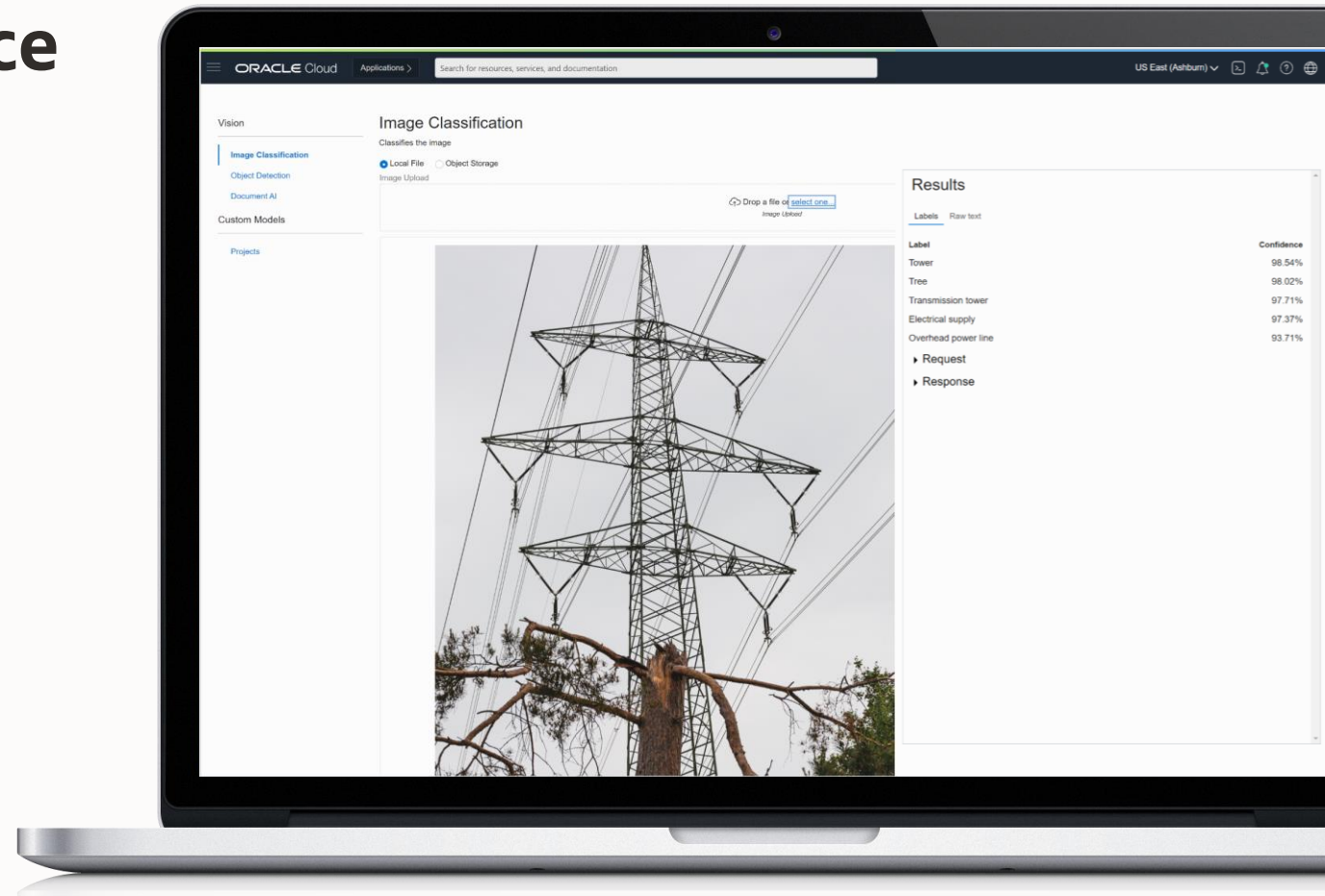
OCI Vision Provides

Image Analysis

- Image Classification
- Object Detection

Auto design	99.26%
Car	99.26%
Bumper	99.23%
Wheel	98.95%

Document AI

- Text Recognition
- Document Classification
- Language Detection

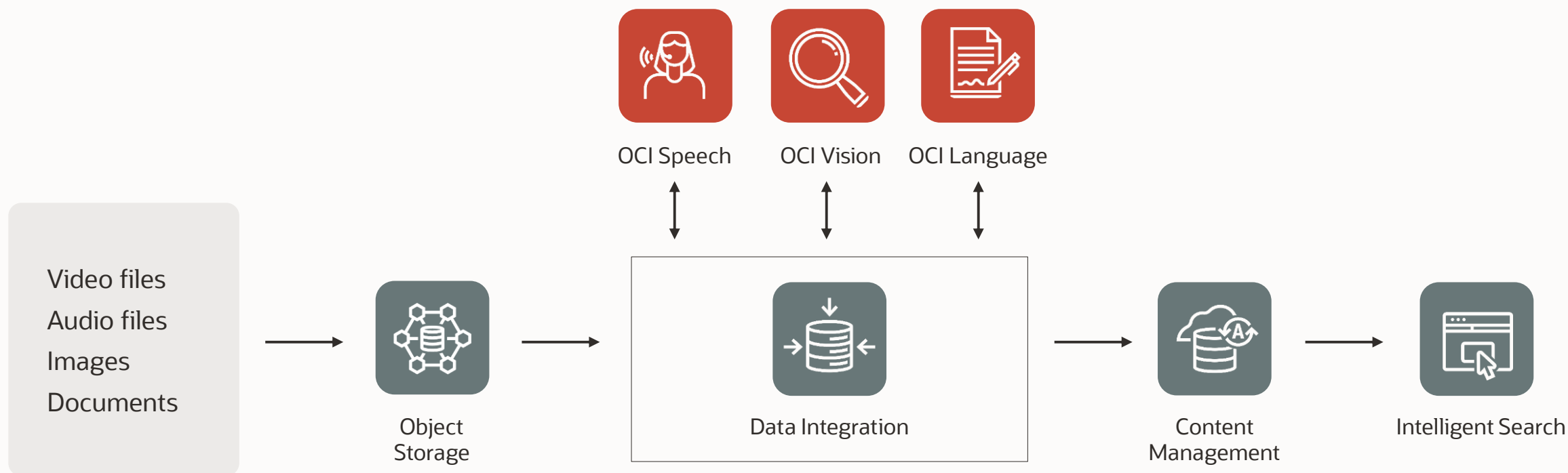- Table Extraction
- Key Value Extraction

Digital Asset Management / Media Indexing

OCI Speech   OCI Vision   OCI Language

Video files
Audio files
Images
Documents

Object Storage

Data Integration

Content Management

Intelligent Search

Anomaly Detection
Finding anomalies in your time-series data

Common use cases

- Manage and monitor assets

- Detect fraudulent behavior

- Monitor resource production and consumption in real time

- Identify irregular behavior in service operations

Challenges

- Building and scaling accurate anomaly detection models requires skilled data scientists

- Time-series data is rarely clean

- Basic anomaly detection techniques are prone to false alarms and errors

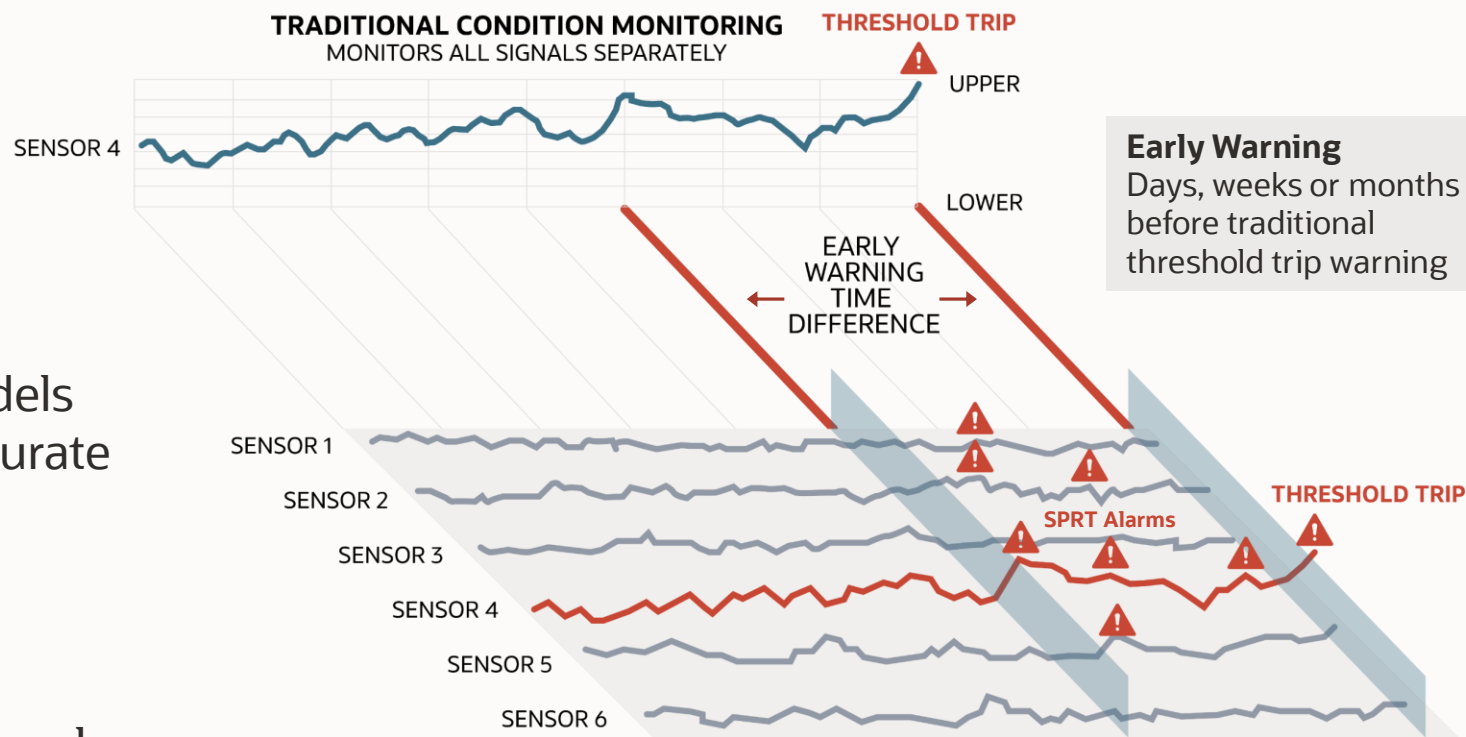- Machine learning models lack accuracy when faced with multiple sensors

NEW

Introducing OCI Anomaly Detection Service

- Builds multiple anomaly detection models and automatically selects the most accurate to flag critical incidents earlier

- Automatically identifies and fixes data quality issues

- Based on industry-leading, proven anomaly detection techniques (MSET-2)

- Detects anomalies that span across multiple sensors to increase accuracy

TRADITIONAL CONDITION MONITORING
MONITORS ALL SIGNALS SEPARATELY

THRESHOLD TRIP

SENSOR 4

UPPER

LOWER

EARLY WARNING TIME DIFFERENCE

Early Warning
Days, weeks or months before traditional threshold trip warning

SENSOR 1
SENSOR 2
SENSOR 3
SENSOR 4
SENSOR 5
SENSOR 6

SPRT Alarms

THRESHOLD TRIP

Forecasting
Predict your critical business metrics

Common use cases

- Predict business performance metrics – revenue, sales, profit…

- Forecast resource requirements – equipment, call center agents, data center assets…

Challenges

- Good business forecasting is complex, and needs a variety of algorithms that account for business variables

- Business performance varies over time and location (and context)

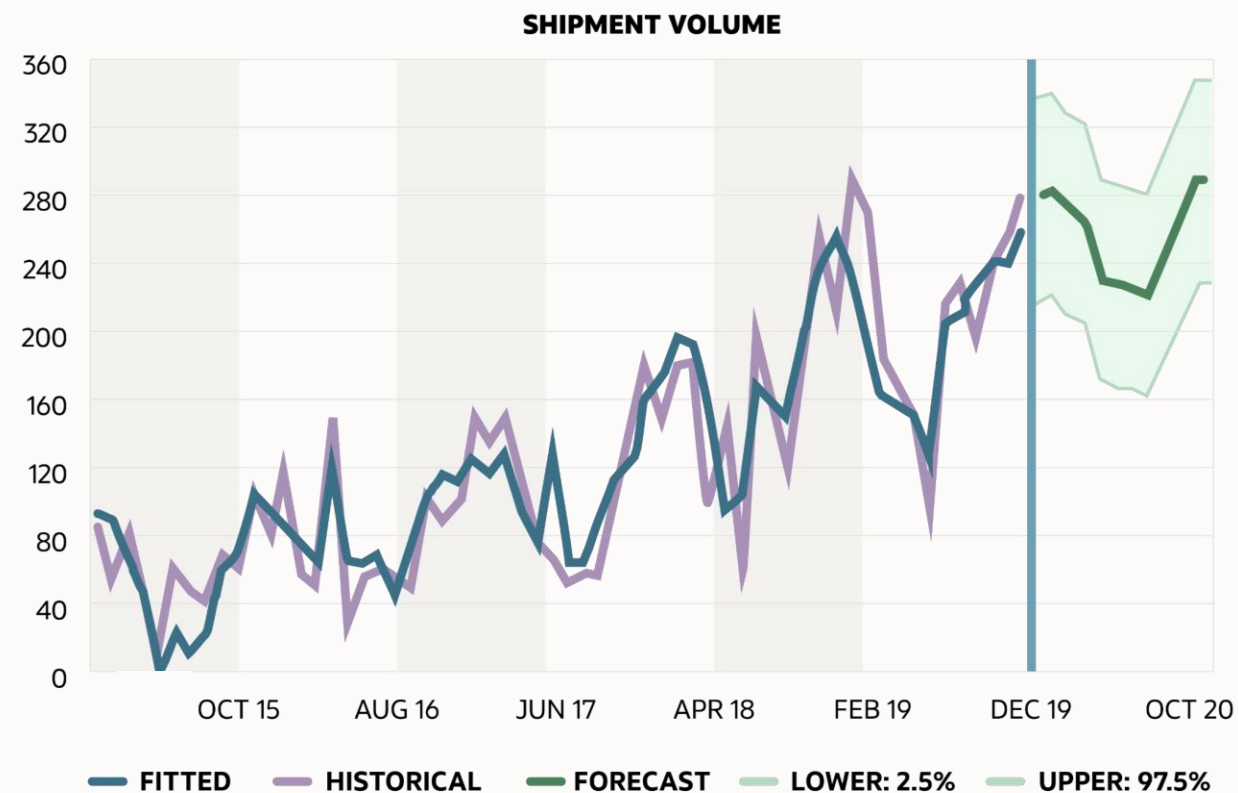- Explainability is critical to aid decision makers

NEW

Introducing OCI Forecasting Service

Forecast any time series metric – product demand, revenue, # of service requests

Builds multiple models and automatically select the most accurate one for your business

Deliver forecasts with explainability, which brings transparency to predicted results

SHIPMENT VOLUME

Legend: — FITTED — HISTORICAL — FORECAST — LOWER: 2.5% — UPPER: 97.5%

X-axis: OCT 15, AUG 16, JUN 17, APR 18, FEB 19, DEC 19, OCT 20

Y-axis: 0, 40, 80, 120, 160, 200, 240, 280, 320, 360

Oracle AI

Unified AI/ML platform spanning cloud services, apps and data assets

Data Labeling

Label your data for model training

Common use cases

- Label images and classify them to create training data for computer vision models

- Label text for identifying entities, topics, sentiment for natural language processing models

Challenges

- Most data is unstructured, but accurate machine learning requires labeled data sets for training

- Difficulty in labeling a wide variety of data types

- Inconsistent integration across different AI services and ML tools

- Lack of consistent labeling experience for multiple types of model training

NEW

Introducing
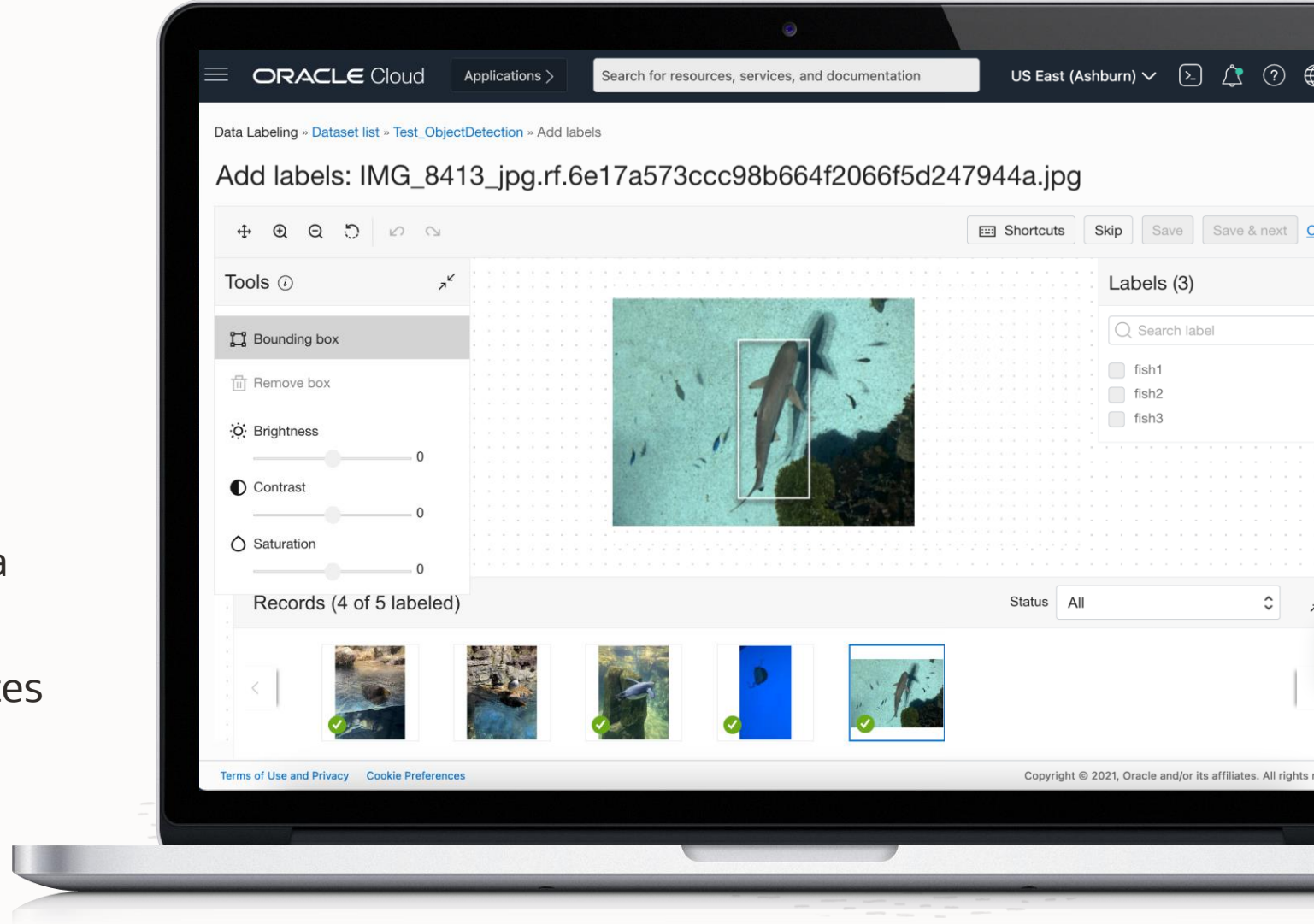# OCI Data Labeling

Build enriched, labeled datasets to
more accurately train AI and machine
learning models

Label images, text, and documents

Simplifies data labeling by providing a
consistent UI experience

Label data faster with custom templates
and multiple annotation formats

Provides easy exportability

Open and extensible

Use your favorite open-source tools and framework with OCI Data Science

Use your AI anywhere using ONNX

NEW: Accelerated Data Science SDK available as an open-source library

Let's do this
Try Oracle AI today

» **oracle.com/ai**

Try a Live Lab

Sign up for Preview services

Register for an upcoming webinar