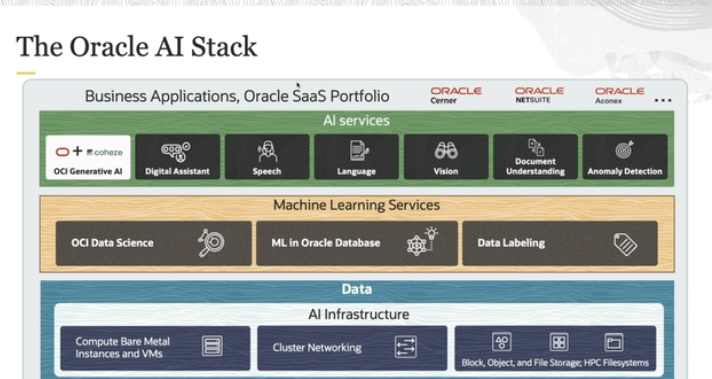
**ML services overview**

****

The image showcases the **Oracle AI Stack** architecture, which represents the different layers of services and infrastructure Oracle offers to help organizations leverage AI and machine learning technologies. Let's break it down based on the visual and the provided transcript:

**1. Business Applications (Top Layer):**

* At the top are **Business Applications** or Oracle SaaS (Software as a Service) solutions like:
  + **Oracle Cerner**: Healthcare solutions.
  + **Oracle NetSuite**: Cloud-based business management software.
  + **Oracle Aconex**: Construction management software.
* These applications consume AI services to optimize business processes, enhance decision-making, and improve user experiences. This could involve integrating AI into analytics systems, business processes, or applications.

**2. AI Services Layer:**

This layer includes pre-built **AI services** that businesses can directly use, which simplifies the process of incorporating AI into their solutions without needing to build models from scratch. The key services include:

* **OCI Generative AI (cohere)**: Tools for text generation using large language models (LLMs), useful for generating content or automating tasks.
* **Digital Assistant**: Oracle's AI-powered chatbots or virtual assistants for improving user interactions.
* **Speech**: Transcribes audio into text and provides other speech-to-text services (like the previous image).
* **Language**: Natural language processing (NLP) services for understanding, processing, and generating human language.
* **Vision**: Computer vision capabilities like object detection, image recognition, and image processing.
* **Document Understanding**: Extracts information from structured and unstructured documents, such as PDFs.
* **Anomaly Detection**: Identifies outliers or unusual patterns in data, useful for security, quality control, or financial applications.

**3. Machine Learning Services Layer:**

This layer provides tools for developing, training, and deploying machine learning models. It's aimed at data scientists and machine learning engineers:

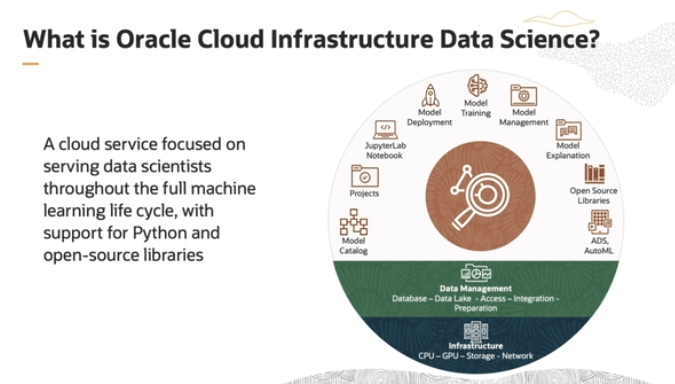
* **OCI Data Science**: A comprehensive environment for building, training, and deploying machine learning models, with capabilities like version control, collaboration, and deployment.
* **ML in Oracle Database**: Integrates machine learning algorithms directly into Oracle's databases, allowing users to train models using SQL without needing to move data.
* **Data Labeling**: Provides tools to label datasets, which is a critical step in supervised machine learning models. Labels help the model learn from training data.

**4. Data Layer (Foundation):**

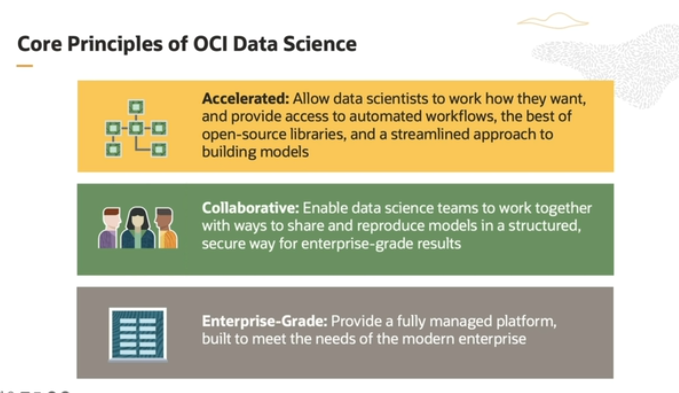
* **Data** forms the foundation for both AI and machine learning services. AI models and ML algorithms require vast amounts of data to learn patterns, make predictions, and automate tasks.
* Data services include:
  + **Compute Bare Metal Instances and VMs**: Offers powerful compute options for running AI and machine learning workloads.
  + **Cluster Networking**: High-performance networking required for scalable and distributed AI/ML tasks.
  + **Block, Object, and File Storage**: Different types of data storage solutions to accommodate structured and unstructured data, as well as **HPC Filesystems** for high-performance computing needs.

**Summary:**

* The **AI Stack** shows how Oracle structures its cloud AI offerings.
  + **AI Services** provide pre-built solutions for common AI tasks like speech, language processing, and vision.
  + **Machine Learning Services** enable more hands-on development of AI models using data science tools and integrated database capabilities.
  + At the base, the **Data Layer** provides the infrastructure needed to power AI and machine learning workloads.



So what is Oracle Cloud Infrastructure Data Science? OCI Data Science is the cloud service focused on serving the data scientist throughout the full machine learning life cycle with support for Python and open source. As you can see by icon in this graphic, the service has many features such as model catalog, projects, JupyterLab Notebook, model deployment, model training, management, model explanation, open-source libraries, and AutoML.



**1. Accelerated:**

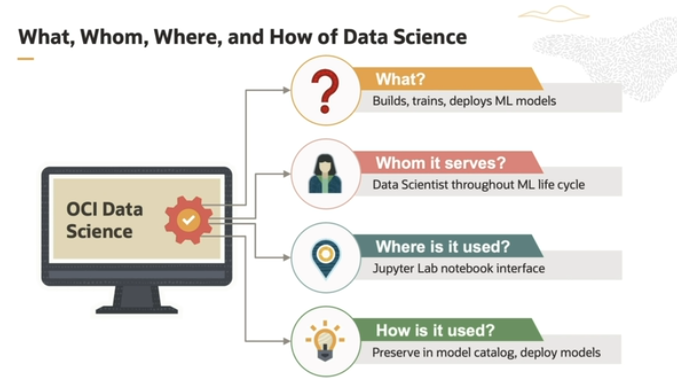
* This principle is about making it faster and easier for individual **data scientists** to do their work.
* **OCI Data Science** gives them tools like open-source libraries (which are pre-made blocks of code they can use) and easy access to powerful computing resources, without having to worry about managing servers or hardware.
* It also includes Oracle's special tools that make certain tasks easier and faster for them.

**2. Collaborative:**

* This principle focuses on **teamwork**.
* Data science teams can **share their work**, like models and code, so they don’t have to redo the same tasks.
* Everything is organized in a way that makes it easy to track, reproduce results, and ensure models are safe and reliable, helping teams work more efficiently together.

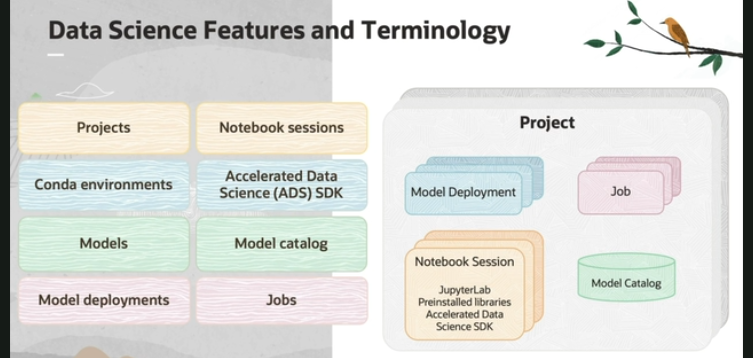
**3. Enterprise-Grade:**

* This means it’s designed to meet the high standards of **big businesses**.
* OCI Data Science comes with **security features** to protect data and make sure only the right people have access.
* Oracle manages everything in the background (like storage, server updates, and maintenance), so users don’t have to worry about technical details—they can just focus on using data to solve business problems.



So having seen the core principles, let's drill down a bit more into the specifics of OCI Data Science. So again, what is OCI Data Science? It's a cloud service to rapidly build, train, deploy, and manage machine learning models. Whom does it serve? It serves data scientists and data science teams throughout the full machine learning life cycle with support for Python and open source.

Where it is being used? Users work in a familiar JupyterLab Notebook interface, where they write Python code. And how it is used? So users preserve their models in the model catalog and deploy their models to a managed infrastructure.



**1. Projects:**

* **Projects** are like folders where data science teams can organize their work.
* Inside a project, they store things like code, models, and documentation.
* These are **collaborative workspaces** where everyone can access the shared resources and work together.
* There’s no limit to how many projects you can create.

**2. Notebook Sessions:**

* **Notebook sessions** are where data scientists actually do their work—writing code and running experiments.
* It’s an environment called **JupyterLab** with pre-installed tools that help with building and training machine learning models.
* Data scientists can interact with the code in real time, test it, and train models.
* These sessions run on managed servers, and users can choose whether they need **CPU or GPU** power and how much storage, without worrying about setting up servers manually.

**3. Conda Environment:**

* **Conda** is a tool that manages environments and packages for Python programs.
* In OCI Data Science, it helps you easily install, update, and manage the necessary software libraries (like tools for machine learning) in your notebook sessions.
* It also lets you switch between different setups (called environments) without hassle.

**4. Accelerated Data Science (ADS) SDK:**

* The **ADS SDK** is a Python library created by Oracle to help data scientists speed up their work.
* It simplifies many common steps, such as:
  + Connecting to data.
  + Exploring and visualizing data.
  + Training models using **AutoML** (which automates model training).
  + Evaluating how well the model performs.
  + Explaining the models’ results.
* ADS also makes it easy to interact with other Oracle services, like object storage.

**5. Models:**

* A **model** is a mathematical representation of your data that helps make predictions or understand patterns.
* You create models inside notebook sessions and store them within projects.

**6. Model Catalog:**

* The **Model Catalog** is a place where you can store and manage models.
* It’s like a library where you keep all the details about your models, such as the scripts, metadata, and versions.
* You can share these models with your team, and anyone can load a stored model back into their own notebook session.

**7. Model Deployments:**

* **Model Deployments** allow you to take a model from the catalog and deploy it as an **HTTP endpoint**.
* This means you can turn your model into a service that can make predictions in real-time through an API.
* Think of it as making the model accessible via the web, so other applications can request predictions whenever needed.

**8. Jobs:**

* **Jobs** are predefined tasks that run on OCI's infrastructure.
* You can set these jobs to **automate tasks** like retraining a model or running machine learning workflows regularly, without manual intervention.