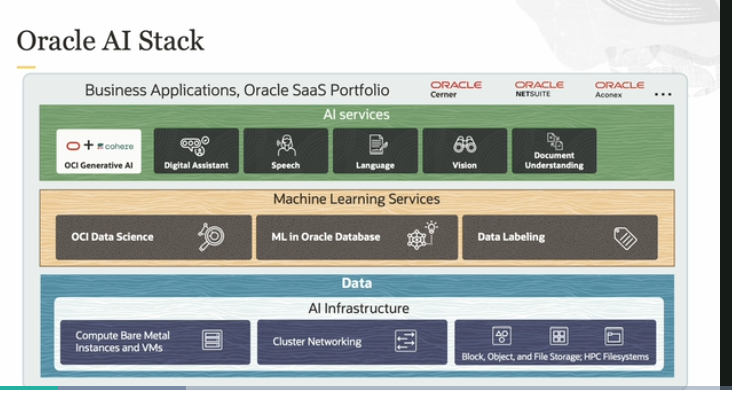
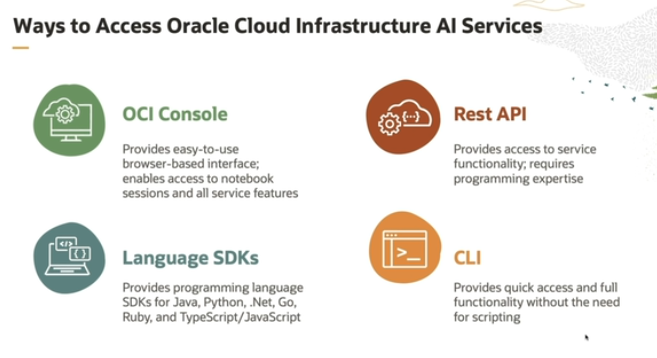


Oracle has been focusing on bringing to the enterprise at every layer of our stack. It all begins with data and infrastructure layers. OCI AI services consume  data. And AI services, in turn, are consumed by applications. This approach involves extensive investment from infrastructure to SaaS applications. Generative AI and massive scale models are the more recent steps.



Oracle AI is the portfolio of cloud services for helping organizations use the data they may have for the business specific uses. Business applications consume AI and ML services. The foundation of AI services and ML services is data. AI services contain pre-built models for specific uses. Some of the AI services are pre-trained, and some can be additionally trained by the customer with their own data. AI services can be consumed by calling the API for the service, passing in the data to be processed, and the service returns a result.

There is no infrastructure to be managed for using AI services



 OCI AI Services provide multiple methods for access. The most common method is the OCI Console. The OCI Console provides an easy-to-use, browser-based interface that enables access to notebook sessions and all the features of all the data science as well as AI services.

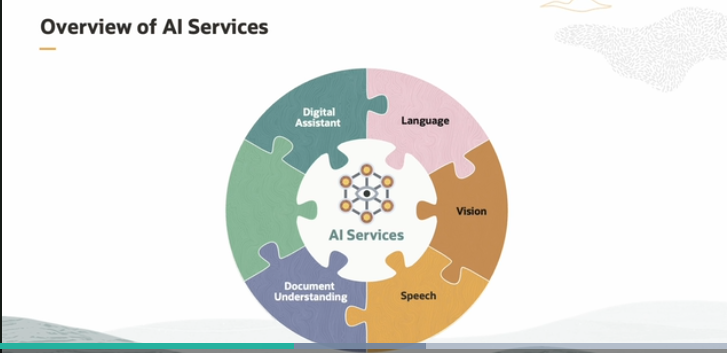
REST API, the REST API provides access to service functionality but requires programming expertise, and API reference is provided in the product documentation. Language SDKs, OCI also provides programming language SDKs for Java, Python, TypeScript, JavaScript, .NET, Go and Ruby. The command line interface provides both quick access and full functionality without the need for scripting.

Wha are SDk ?

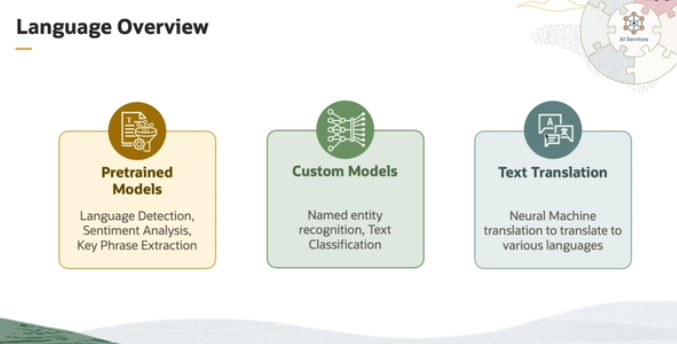
An SDK (Software Development Kit) is a collection of tools, libraries, documentation, and code examples that developers use to build software for specific platforms or frameworks. SDKs typically provide everything needed to create applications in a particular environment, including:

1. **Libraries**: Pre-written code that developers can use to add functionality to their apps.
2. **Documentation**: Guides and references explaining how to use the SDK and its components.
3. **Tools**: Utilities for testing, debugging, and compiling software.
4. **Code Samples**: Example projects or snippets that show how to implement features using the SDK.
5. **APIs**: Interfaces for interacting with other software components, systems, or platforms.

For example, if you want to develop an app for Android, you would use the Android SDK. If you're building on blockchain, you might use a specific blockchain SDK for that platform.



OCI AI Services is a collection of services with pre-built machine learning models that make it easier for developers to build a variety of business applications. The models can also be custom-trained for more accurate business results. The different services provided are-- Digital Assistant, Language, Vision, Speech, Document Understanding. Let us explore each one of these services.



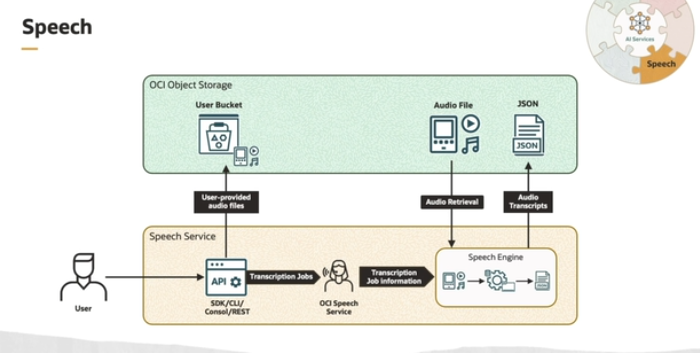
OCR Language allows you to perform sophisticated text analysis at scale. Using the pre-trained and custom models, you can process unstructured text to extract insights without data science expertise. Language services are provided by pre-trained models, custom models, and text translation. Pre-trained models include language detection, sentiment analysis, key phrase extraction, text classification, named entity recognition, and personal identifiable information detection.

Custom models can be trained for named entity recognition and text classification with domain-specific data sets. In text translation, natural machine translation is used to translate text across numerous languages.



Using OCI Vision, you can upload images to detect and classify objects in them. Pre-trained models and custom models are supported.

In image analysis, pre-trained models perform object detection, image classification, and optical character recognition. In image analysis, custom models can perform custom object detection by detecting the location of custom objects in an image and providing a bounding box. The custom image classification builds a model to identify objects and scene-based features in an image.



The OCI Speech service is used to convert media files to readable texts that's stored in JSON and SRT format. Speech enables you to easily convert media files containing human speech into highly exact text transcriptions.

**1. OCI Object Storage:**

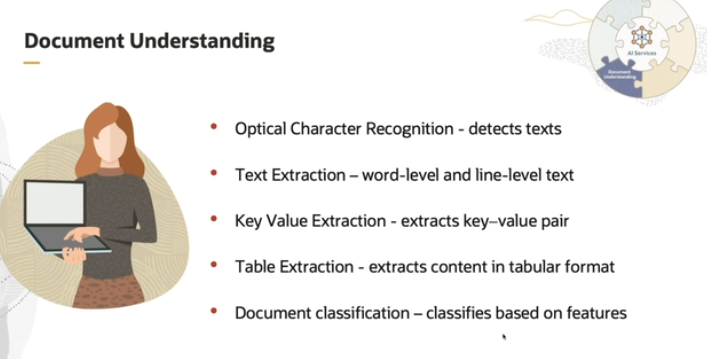
* **User Bucket**: The user uploads audio files to a storage bucket in OCI. These files can be in various formats like .mp3 or .wav.
* The **audio files** are stored in OCI Object Storage, which is a scalable and durable storage solution.
* The user-provided audio files are sent from the bucket for processing.

**2. Speech Service:**

* The **Speech Service** can be accessed through various methods:
  + **SDK**: A software development kit for developers to programmatically interact with the service.
  + **CLI**: Command Line Interface for interacting with the service via terminal or command line.
  + **Console**: The Oracle Cloud Console interface for managing the service.
  + **REST API**: You can use RESTful APIs to trigger speech recognition jobs.
* A **Transcription Job** is created when the user initiates a request to transcribe an audio file. The job sends information about the transcription request to the service.

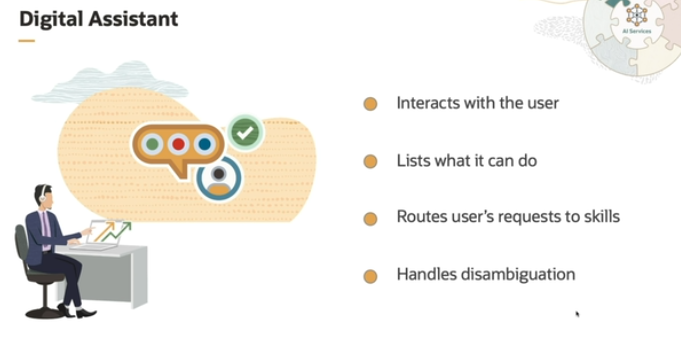
**3. Speech Engine:**

* The **Speech Engine** retrieves the audio file from OCI Object Storage.
* It processes the file and runs the **speech-to-text transcription**.
* The transcription result is returned in **JSON format** with the audio transcripts, which can then be accessed by the user.



Using OCI Document Understanding, you can upload documents to detect and classify text and objects in them. You can process individual files or batches of documents. In OCR, document understanding can detect and recognize text in a document.

In Text Extraction, document understanding provides the word-level and line-level text and the bounding box coordinates of where the text is found. In Key Value Extraction, document understanding extracts a predefined list of key-value pairs of information from receipts, invoices, passports, and driver IDs. In table Extraction, document understanding extracts content in tabular format, maintaining the row and column relationship of cells. In Document classification, the document understanding classifies documents into different types.



Oracle Digital Assistant is a platform that allows you to create and deploy digital assistants, which are AI-driven interfaces that help users accomplish a variety of tasks with natural language conversations. When a user engages with the Digital Assistant, the Digital Assistant evaluates the user input and routes the conversation to and from the appropriate skills.

Digital Assistant greets the user upon access, upon user request, lists what it can do and provide entry points into the given skills. It routes explicit user requests to the appropriate skills, and it also handles interruptions to flows and disambiguation. It also handles requests to exit the bot.