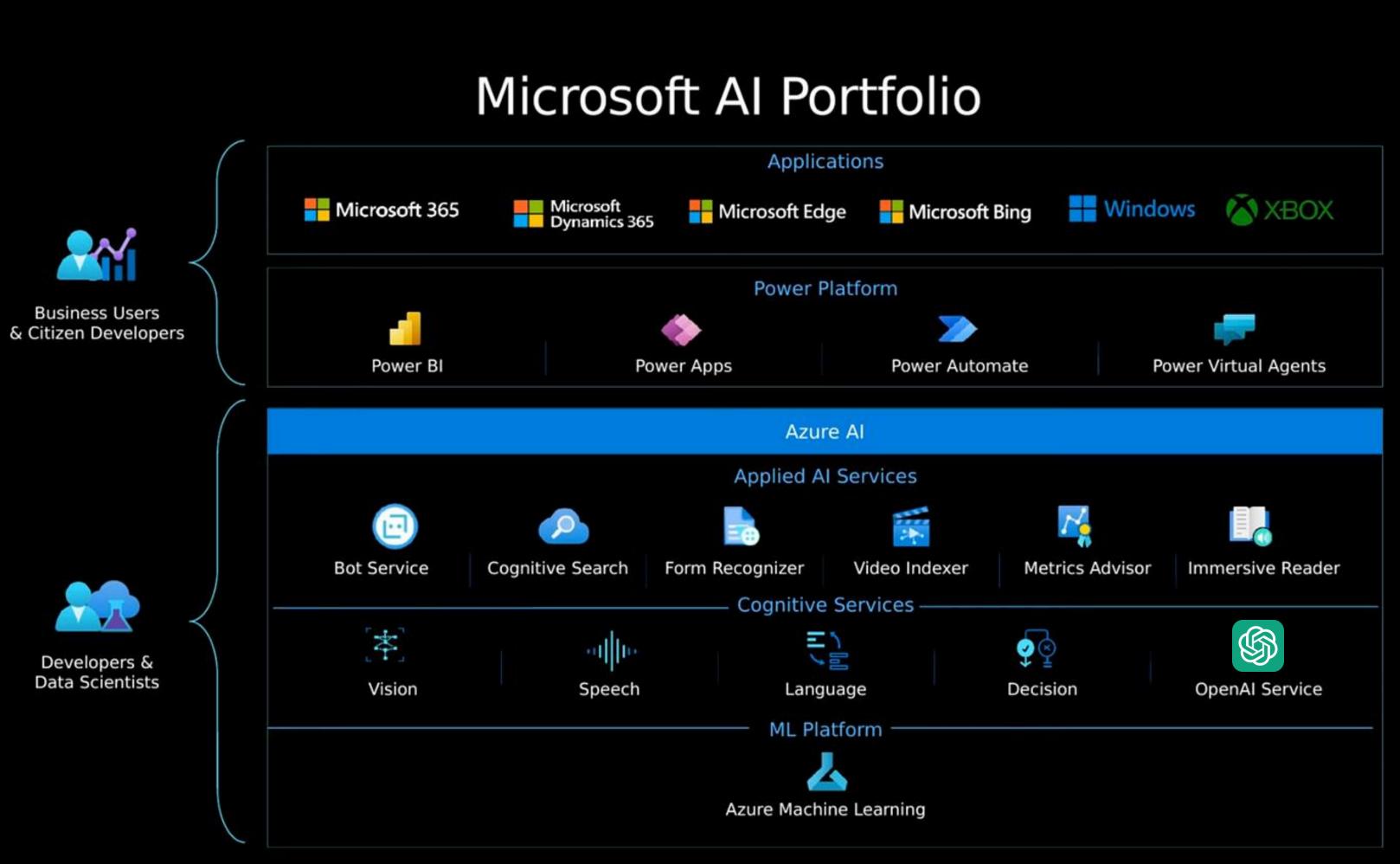


improdve

## Azure (Open) AI showdown

Andrea Saltarello Adjunct Professor @ Politecnico di Milano Founder @ Improove





### **AZURE MACHINE LEARNING**



### **Azure Machine Learning in a nutshell**

A cloud-based machine learning service capable of executing models based on a variety of frameworks.

Models are managed via **Azure ML Studio** 

- Build new models with a designer
- Create models out of templates
- Import models based on known libraries (ML.NET, TensorFlow, PyTorch, ONNX, ...)
- Create models via AutomatedML

The service is free, but training and predictions are billed on a *per compute* basis

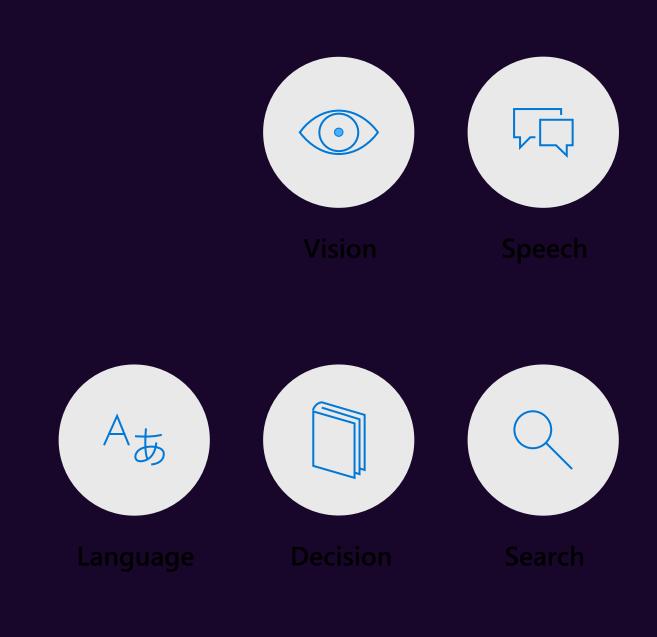
### **COGNITIVE SERVICES**



### **Azure Al Services**

a group of services, each supporting different, generalized prediction capabilities

- Simple REST APIs or SDKs
  - -.NET
  - -Go
  - -Java
  - -Node.js
  - Python
- Price-based Service Level Agreement



### **APPLIED AI SERVICES**



### **Azure Document Intelligence**

Azure Form Recognizer is a cognitive service capable to identify and extract key-value pairs and table data from form documents stored as PDFs and images.

Unsupervised learning allows the model to understand the layout and field data without manual data labelling or intensive coding.

You can also do supervised learning with manually labelled data. Models trained with labelled data can perform better and can work with more complicated documents.

Two pricing tiers:

- Free: 500 pages/month ullet
- SO:
  - Custom, Invoice: \$50 per 1000 pages •
  - Pre-built: Layout, Receipt, business Card: \$10 per 1000 pages  $\bullet$

https://www.improove.tech/tv/episode/3365/Azure-AI-Document-Intelligence

### **Training requirements**

**Document requirements:** 

- Format must be JPG, PNG, PDF (text or scanned), or TIFF. Text-embedded PDFs are best because ulletthere's no possibility of error in character extraction and location.
- If your PDFs are password-locked, you must remove the lock before submitting them.  $\bullet$
- PDF and TIFF documents must be 200 pages or less, and the total size of the training data set ightarrowmust be 500 pages or less.
- For images, dimensions must be between 600 x 100 pixels and 4200 x 4200 pixels.  $\bullet$
- If scanned from paper documents, forms should be high-quality scans. ullet
- Text must use the Latin alphabet (English characters). •

### **AZURE OPENAI**



### **OpenAl** @ AZURE

Starting from January 2023, **OpenAI** models are available on **Azure** due to a 2019 deal by which Microsoft got exclusive rights to the models.

Gated access, must declare the exact use case to be allowed in.

To date, Azure is the only choice for GDPR-compliant OpenAI adoption.

API compatible with OpenAI, SDKs are available for .NET and Python

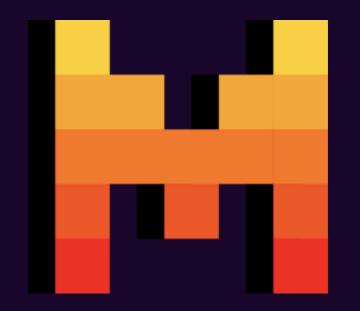
### MODELS.MORE();

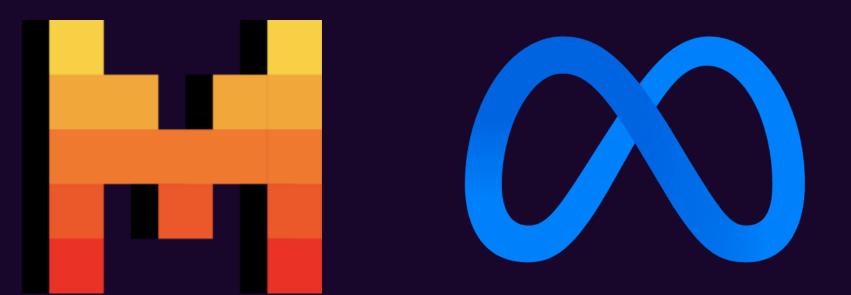


### **Azure Al Studio**

More third-party models, such as **Meta's and Mistral's**, are available via the **Azure AI Studio** application.

Models are deployable to VMs, with some of them supporting PaaS/serverless compute as well.





### Grazie!





# 17/06/2024 >>AICCONF

### **4a EDIZIONE**

