Content=
Python nel Gruppo Sella
</*>

Federico Angaramo   Senior Data Scientist @Centrico

federico.angaramo@centrico.tech

Biella – 11 Giugno 2022

{centrico:

Agenda

- Who we are

- Dove viene utilizzato Python

- Use cases AI

{centrico:

Chapter=
Who we are
</*>

Who is Centrico?

Some numbers and facts:

- Open tecnology modular core banking without legacy dependencies
- >1,700 server in private cloud
- ~ 3 millions of accounting transactions per day
- > 1 billion API calls yearly
- ~ 200 (attended and unattended) RPA & AI flows

- Operational and IT backbone of Sella Group since 1968

- Starting from 2019, Open Banking outsourcer focused on bank and fintech core banking

- IT and BPO full outsourcing supported by 600 HC in Italy, 400 HC in Romania and 300 HC in India (900 IT and 400 BPO)

Italia Polo ICT e BPO Romania Polo ICT India
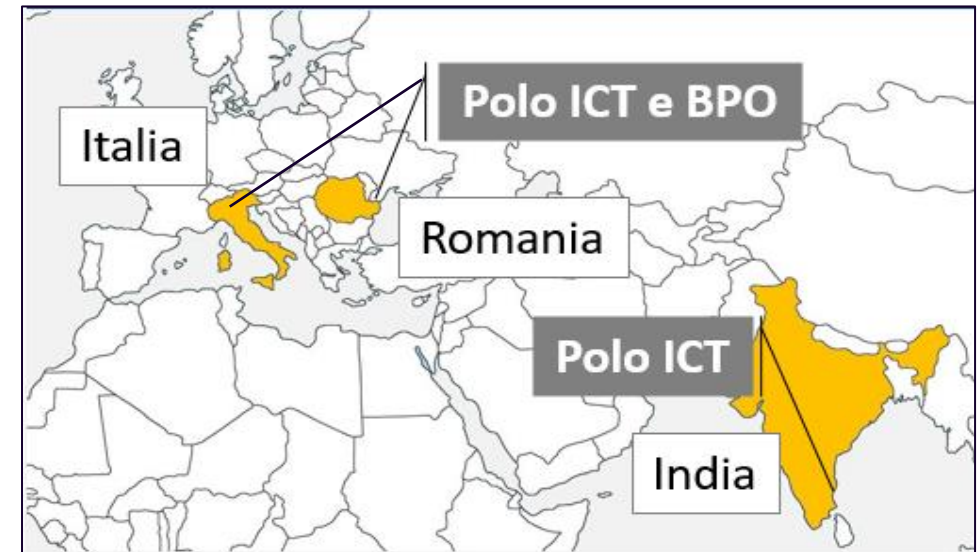
1,300 people in 3 countries

{ centrico :

We enable banks and fintechs to innovate thanks to our proven open banking system,

that is easy to integrate and customise,

and minimises time to market

Chapter=
Python nel Gruppo Sella
</*>

Dove viene utilizzato?

- Sistemi di build e deploy

- Schedulazioni (ETL / ecc)

- Automation

- Artificial Intelligence

{centrico:

Sistemi di build e deploy

Sistema di rilascio script SQL su database

- Prima versione nel 2008
- Python 2.6 -> Python 3
- Django
- Circa 28000 linee di codice
- Code coverage ~ 90%

Sistema di rilascio artefatti

- Prima versione nel 2011
- Python 2.7 -> Python 3
- Django
- Circa 18000 linee di codice
- Code coverage 98%

Sistema di build artefatti

- Prima versione nel 2014
- Python 2.7 -> Python 3
- Django
- Circa 16000 linee di codice
- Code coverage 95%

{centrico:

Schedulazioni

Principalmente per fare ETL

Possibile avere

- Ambienti Conda dedicati
- Utilizzare diverse versioni di Python

{centrico:

Ansible

Tool di RedHat, scritto in Python, utilizzato nel gruppo per fare Automation

Supporta nativamente:
- Sintassi Python per le regular expression
- Task per l'installazione di dipendenze e creazione di virtualenv
  - Modulo «pip»
- Plugin per estensione scritti in Python
  - https://docs.ansible.com/ansible/latest/dev_guide/developing_plugins.html

Conoscere Python ci facilità anche nel troubleshoting

{ centrico :

Artificial Intelligence

Il sistema centralizzato supporta la build e i deploy di due tipologie di applicazioni:

- *Web*

  - Esposizione di applicazioni web attraverso la specifica WSGI (Django, Flask, ecc..)

- *Batch*

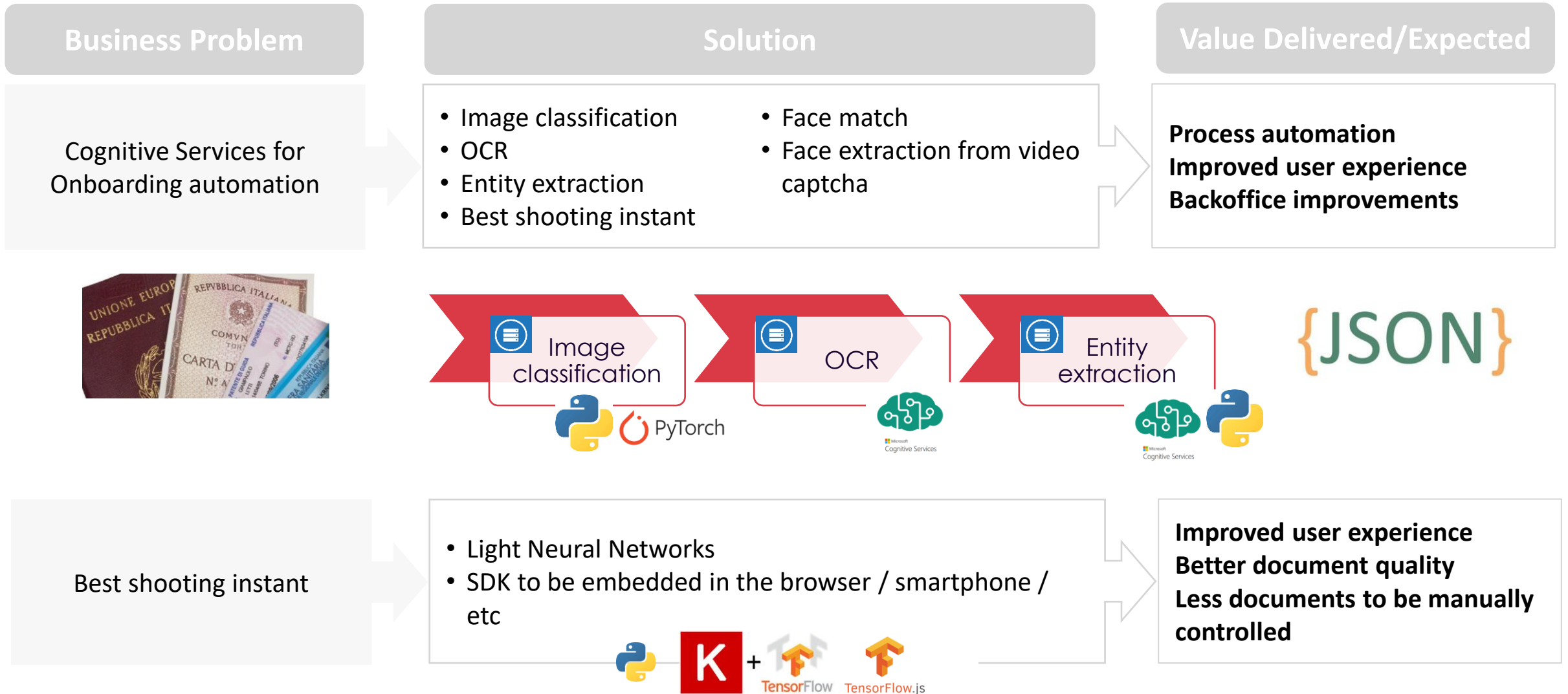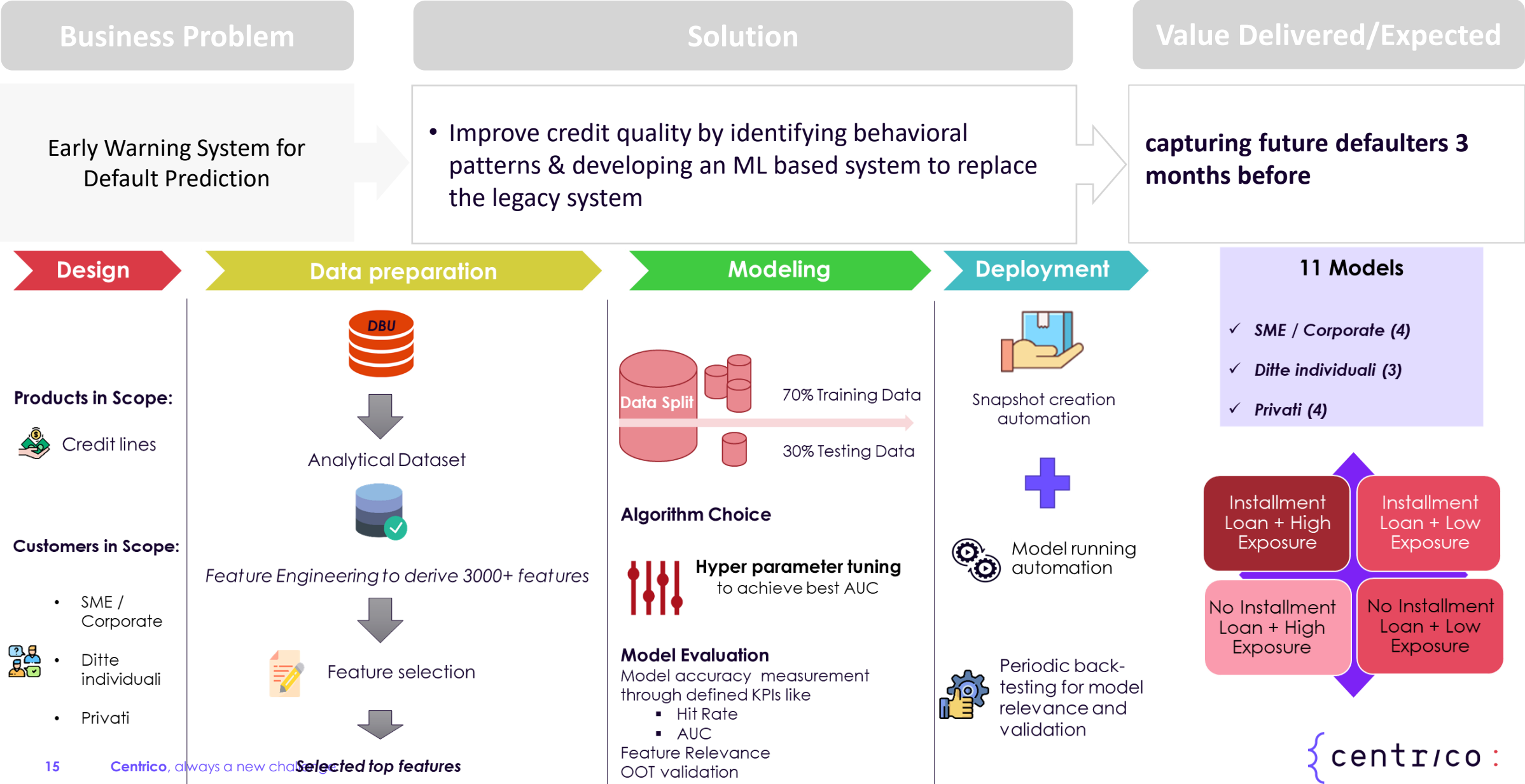 - *Script schedulati*

{ centrico :

Chapter=
Use Cases AI
</*>

Use Case Marketing

Business Problem	Solution	Value Delivered/Expected
Segmentation for Marketing Strategies	• Identify clusters defined by their financial behavior and preferences	**Customized segment-wise campaign journeys**
Churn Modelling	• Develop a churn prediction model that helps in early & accurate detection of potential churners, and also identify churn drivers	**Model helped to capture potential churners up to 6 months in advance**
Propensity Models	• Models to identify customers who are likely to buy a product • Scores generated for every customer based on recent data • Design a system to identify top factors impacting propensity to purchase a product	**Improved campaign redemption Provided incremented revenue**
Model Monitoring	• Monitor model performances • PowerBI Dashboard to monitor all the models	**Identify when model is drifting RCA helper**

Centrico, always a new challenge

{ centrico :

Use Case Onboarding

Business Problem	Solution	Value Delivered/Expected

Cognitive Services for Onboarding automation

- Image classification
- OCR
- Entity extraction
- Best shooting instant
- Face match
- Face extraction from video captcha

Process automation
Improved user experience
Backoffice improvements

Image classification — PyTorch

OCR — Microsoft Cognitive Services

Entity extraction — Microsoft Cognitive Services

{JSON}

Best shooting instant

- Light Neural Networks
- SDK to be embedded in the browser / smartphone / etc

K + TensorFlow TensorFlow.js

Improved user experience
Better document quality
Less documents to be manually controlled

centrico :

Use Case Risk / Crediti / Antifrode

Business Problem	Solution	Value Delivered/Expected

Business Problem

Early Warning System for Default Prediction

Solution

- Improve credit quality by identifying behavioral patterns & developing an ML based system to replace the legacy system

Value Delivered/Expected

capturing future defaulters 3 months before

Design	Data preparation	Modeling	Deployment	11 Models

11 Models

- ✓ *SME / Corporate (4)*
- ✓ *Ditte individuali (3)*
- ✓ *Privati (4)*

Products in Scope:

Credit lines

Customers in Scope:

- SME / Corporate
- Ditte individuali
- Privati

Data preparation

DBU

↓

Analytical Dataset

↓

Feature Engineering to derive 3000+ features

↓

Feature selection

↓

Selected top features

Modeling

Data Split — 70% Training Data / 30% Testing Data

Algorithm Choice

Hyper parameter tuning to achieve best AUC

Model Evaluation
Model accuracy measurement through defined KPIs like
- Hit Rate
- AUC
Feature Relevance
OOT validation

Deployment

Snapshot creation automation

+

Model running automation

Periodic back-testing for model relevance and validation

Installment Loan + High Exposure	Installment Loan + Low Exposure
No Installment Loan + High Exposure	No Installment Loan + Low Exposure

centrico

Chapter=
Perchè Python nel 2022?
</*>

Perché Python nel 2022

- Grande community
- Linguaggio flessibile, utilizzabile per molti scopi diversi
- In ambito AI è quasi uno standard di fatto
  - Migliaia di librerie
  - Molti ruoli
    - Data Scientist
    - Data Engineer
    - SW Developer

{centrico :

{centrico: