ADALTAS

8. End-to-end ML platforms

Reducing the complexity

Until now we saw...

ADALTAS

- Getting a model to production takes many steps
- Many different stakeholders will participate
- Many different tools will be used
- Even after the deployment the pipelines will evolve

  => The process is complex and it takes lots of time and maintenance

Solution

ADALTAS

- A tool that treats multiple steps from development to deployment
  => end-to-end ML platform
- More and more platforms is appearing
- Several are open-source, mostly they are managed
- They mostly started as in-house solutions (Google, Uber, Netflix, AirBnB...)
- They focus on different parts of the process (no unified definition)
- Can be framework-specific (TensorFlow Extended) or framework agnostic (Liminal) (both open-source)
- Cloud providers offer many services, including MLOps toolstack

Platform comparison

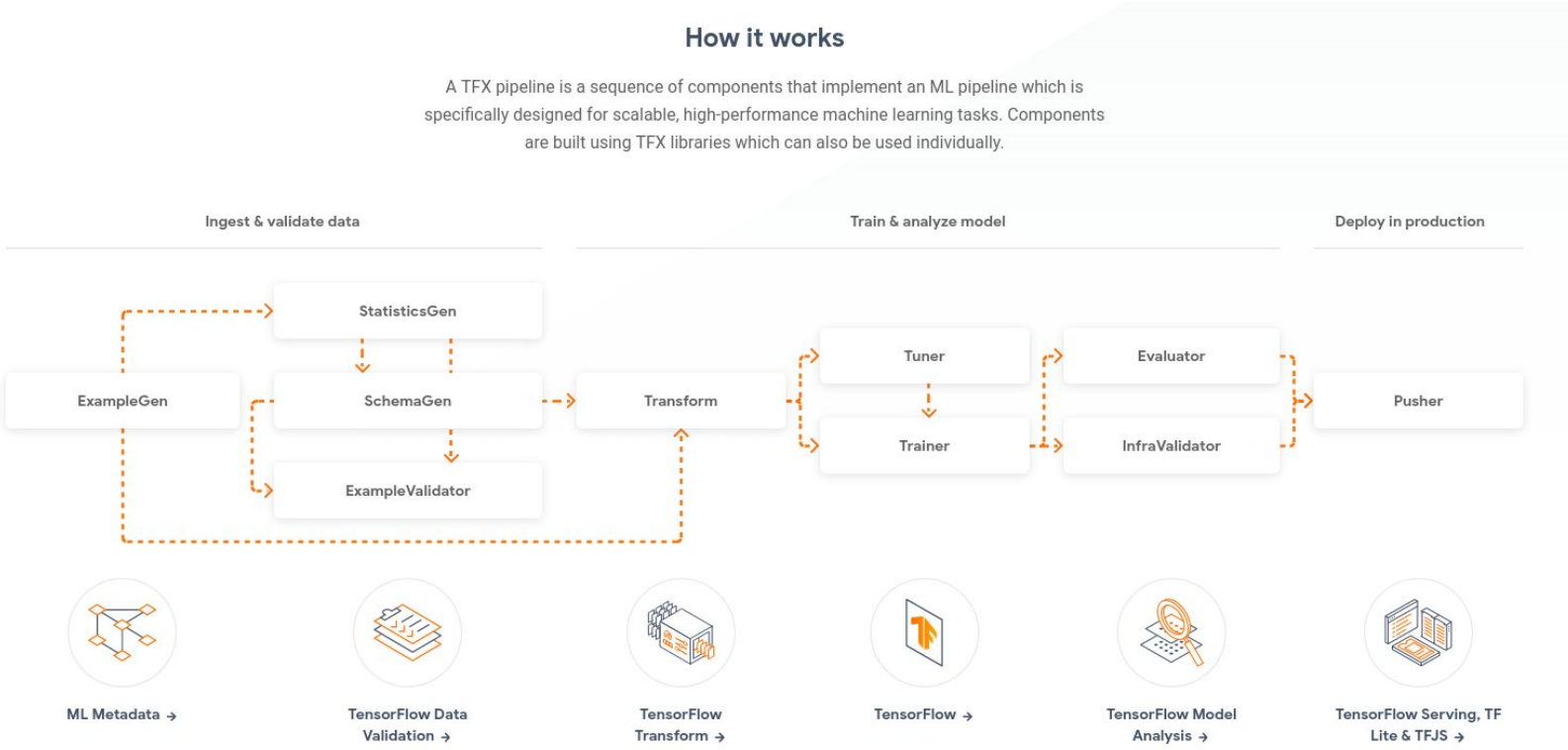
ADALTAS

https://valohai.com/mlops-platforms-compared/

TensorFlow Extended (TFX)

ADALTAS

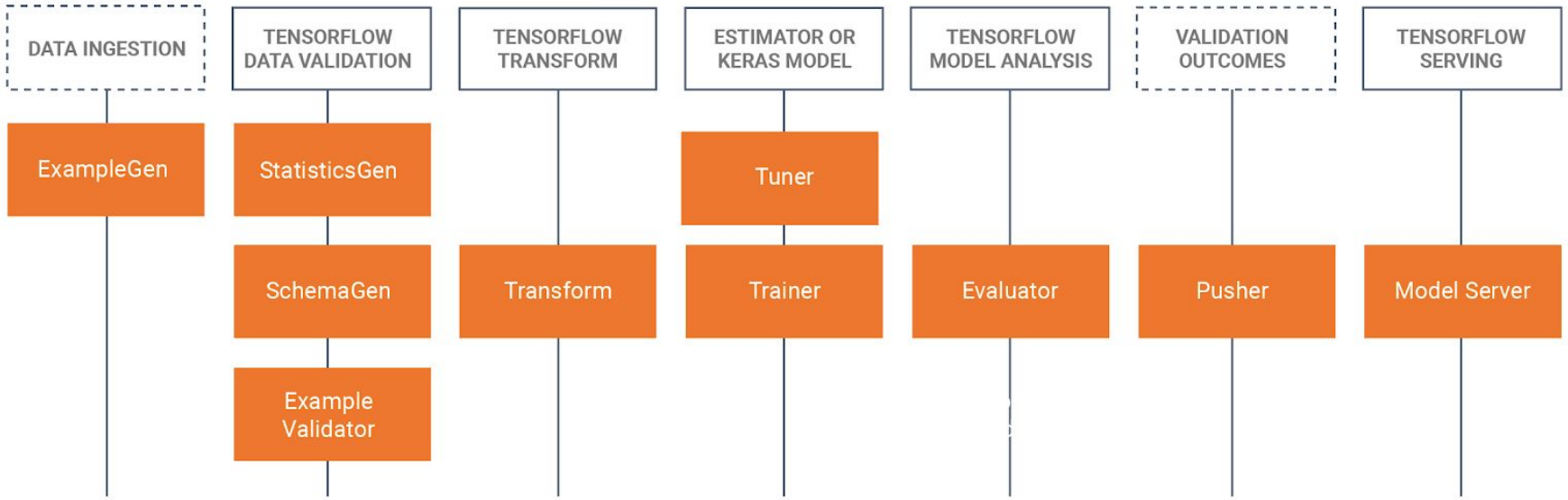
- Google's solution for deploying their proper models
- Robust and scalable
- Open-sourced in 2019



TFX: how it works

ADALTAS

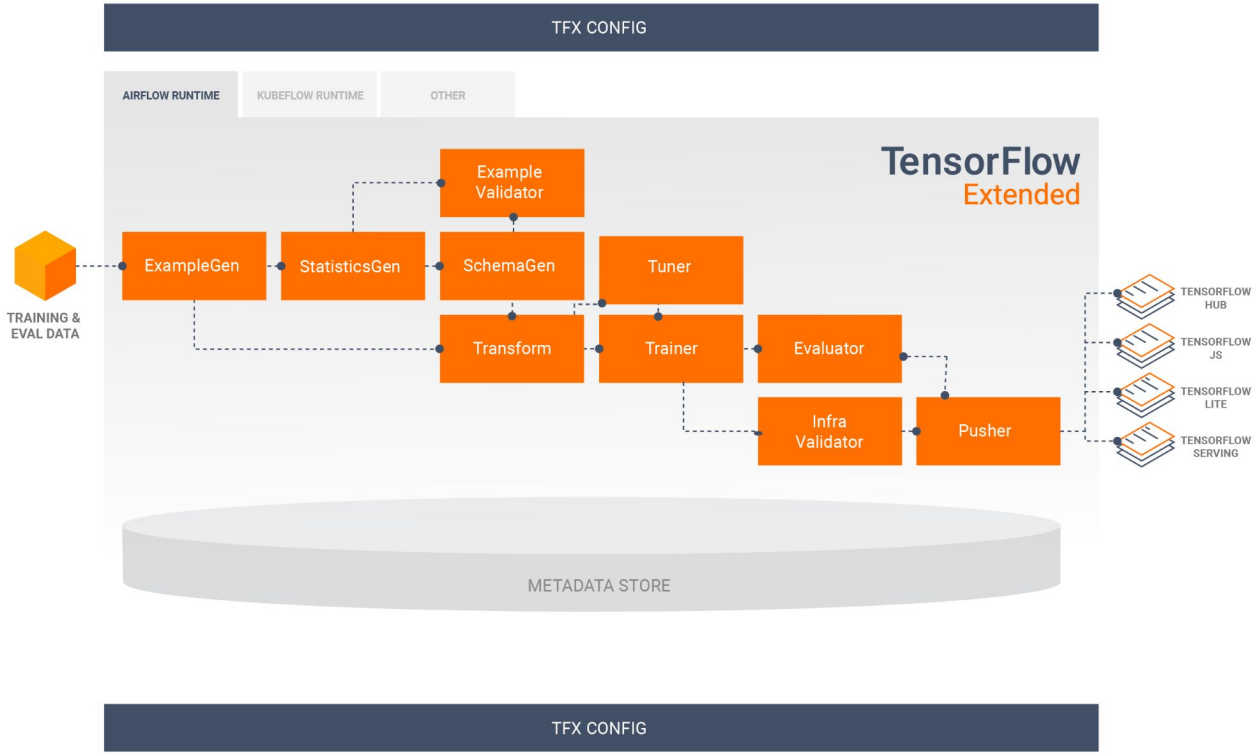
- It is based on TensorFlow libraries
- Many libraries: potential problems with compatibility



TFX: how it works

ADALTAS

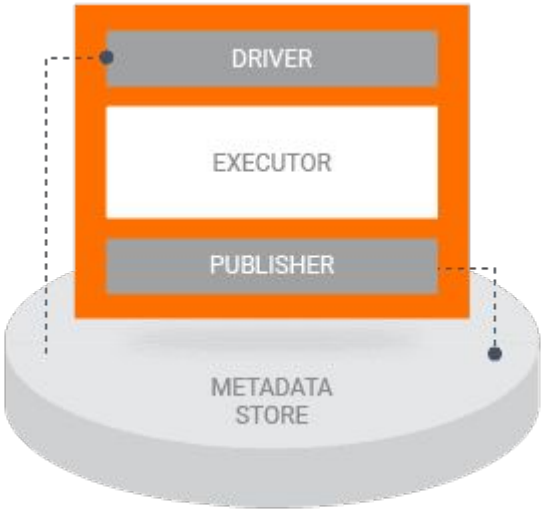
- Modular: modules can be used independently or to build pipelines
- Output from one module is input for the next one
- Reduction of boilerplate code to connect modules and make the pipeline work

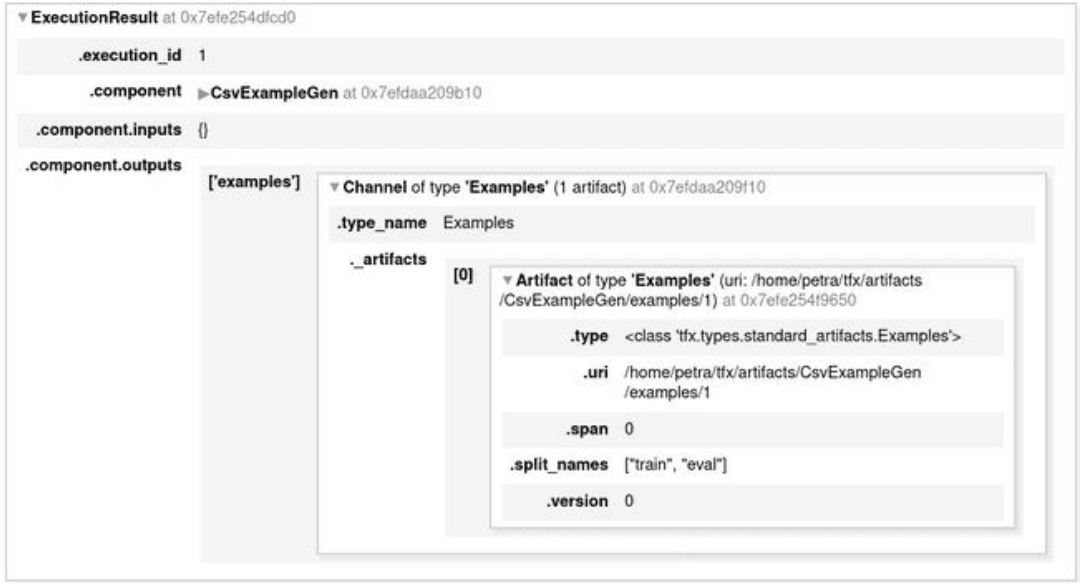


TFX: metadata store

ADALTAS

- Stores the information about the state of the system after an execution of a component
- Stores the outputs of the components:
  - information about the models, the data they were trained on, and the evaluation metrics
  - execution records for every component
  - lineage of the data objects as they flow through the pipeline.
- Makes the pipeline data and task-aware





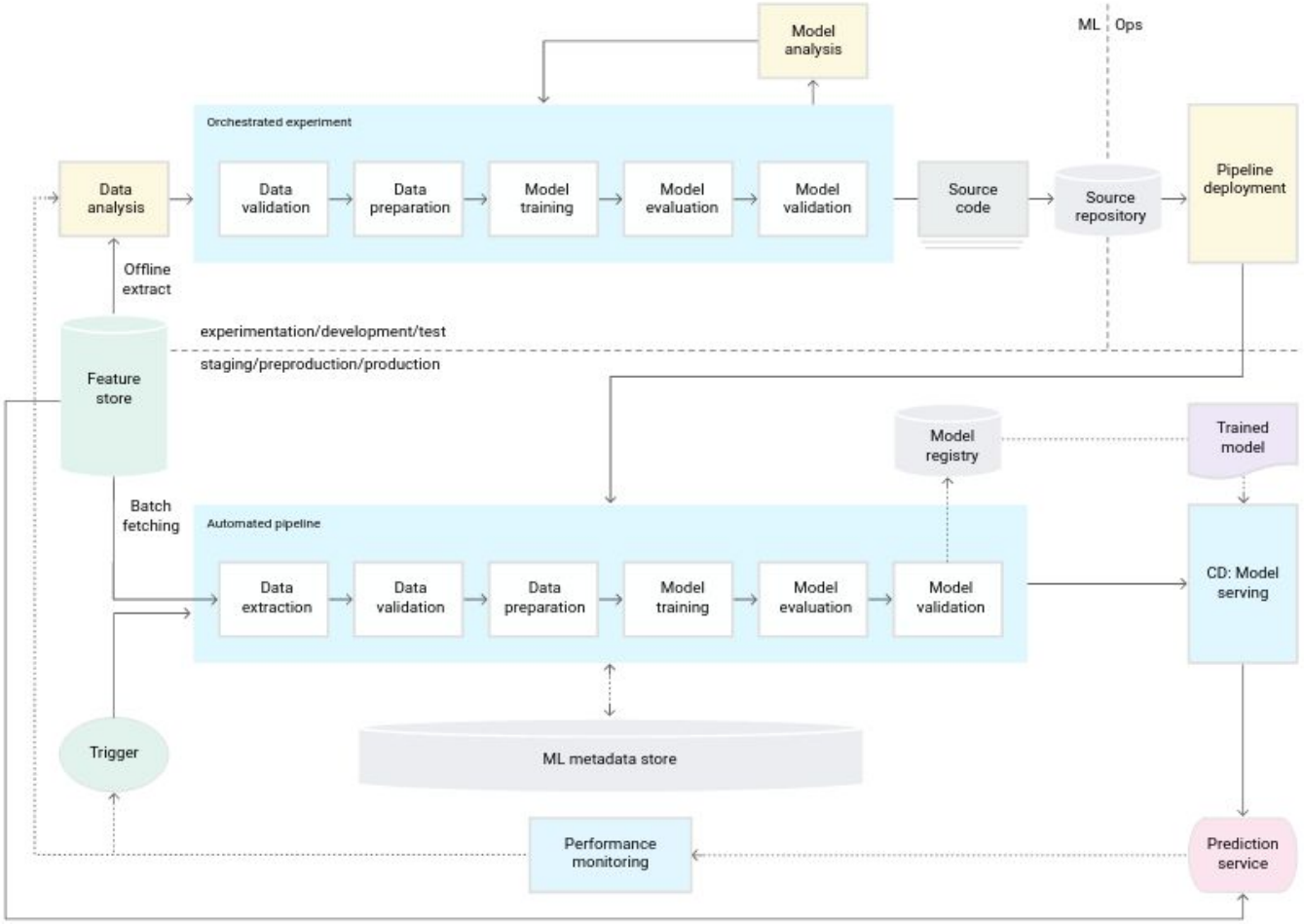
What will we do

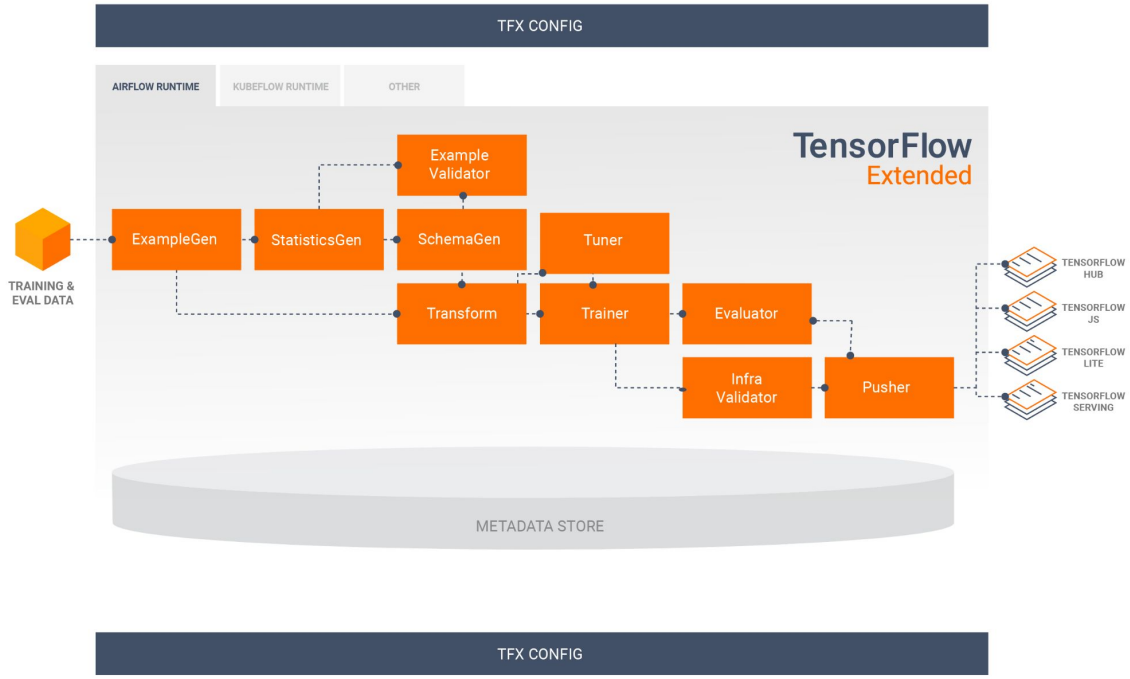
ADALTAS

- Exercise with TFX in colab: https://colab.research.google.com/github/tensorflow/tfx/blob/master/docs/tutorials/tfx/components_keras.ipynb
- Understand the functioning of the individual components of TFX

Comparison of the initial pipeline and TFX

ADALTAS





To go from here

ADALTAS

- Learn TFX at Coursera: https://www.coursera.org/learn/ml-pipelines-google-cloud
- To install and try it locally: https://www.adaltas.com/en/2021/03/05/tfx-overview/
- Look into one of the platforms: https://valohai.com/mlops-platforms-compared/