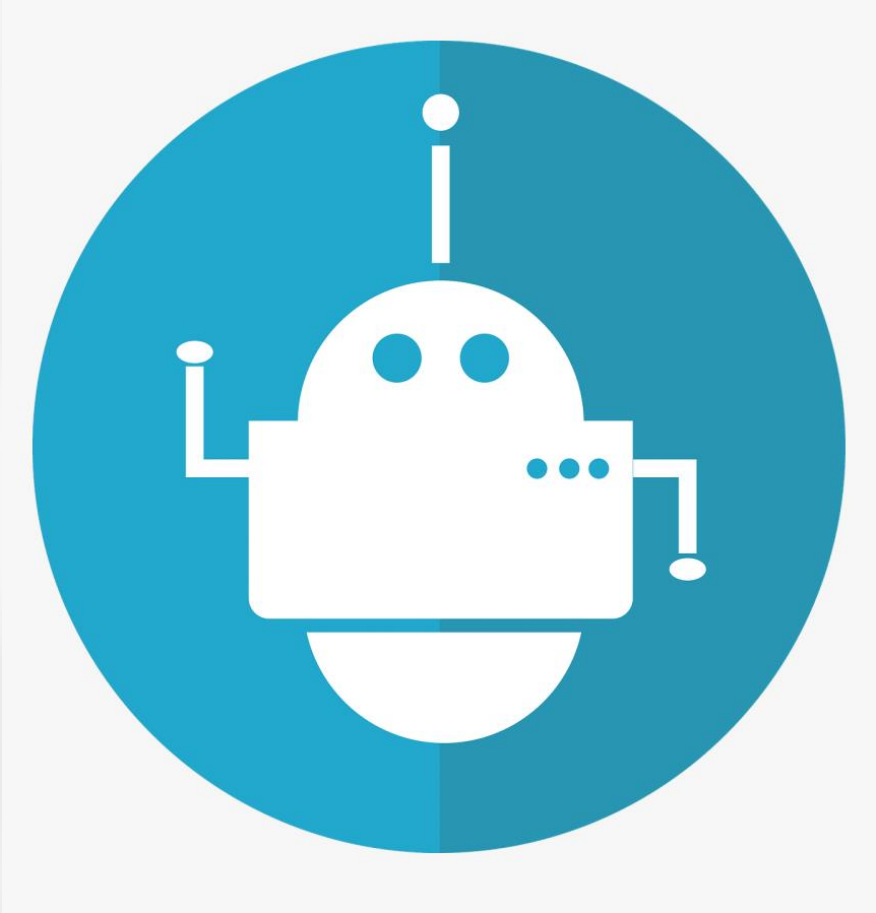
ML Engineering with Azure ML

- in a nutshell

|M|O|D|H|I| | |
Et foretak i SpareBank 1-alliansen

I had a dream….



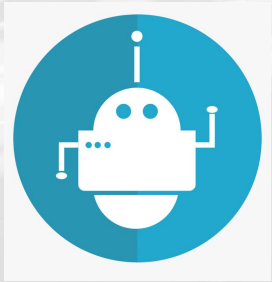
- I made a robot “model” which looks like exact as me…
- It will come to Modhi to work from 2nd Nov. instead of me
- But it uses Mac and speaks Chinese only…

Will it work in Modhi as it is meant to? Why?

Environment

- To let a model work, you need to create a porper environment, e.g.
  - Software
  - Libraries
  - Configurations..
- Just like humans…

Sigthor, I'd like to work from home



You need screen, mouse, keyboard, docking besides your laptop… remember to order the same docking as you are using in office

important to specify version

Define Environment in Azure ML

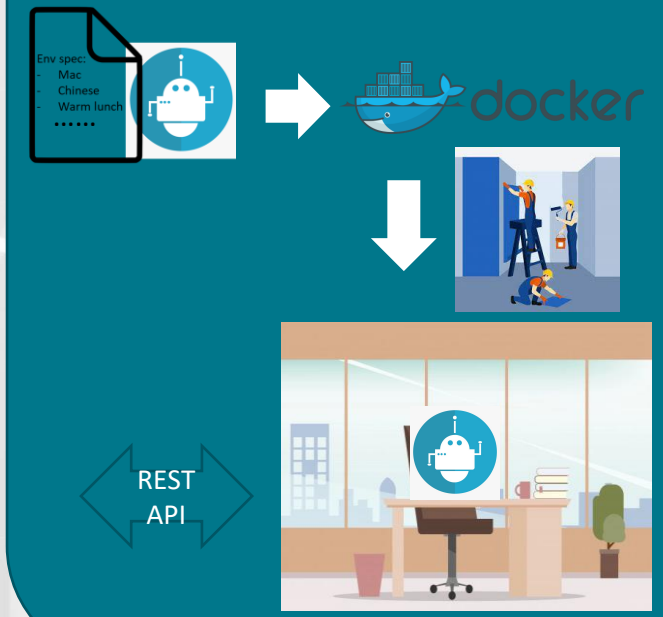
Various ways, based on Docker, Conda,e.g. via:

- Docker file
- or Conda file

(and so on…)

docker

deliver software/services in packages called containers, which are isolated from one another



CONDA

- Directory that contains a specific collection of conda packages (for data science).
- **Dependencies** required by different projects separate by creating isolated spaces.

Azure

Example of Docker and Conda file

Dockerfile

spf-users$ > a50e00r > Documents > Azure ML slides > Dockerfile > ...

```
FROM python:3
COPY requirements.txt /tmp/
RUN pip install -r /tmp/requirements.txt
COPY . /src/
```

requirements.txt

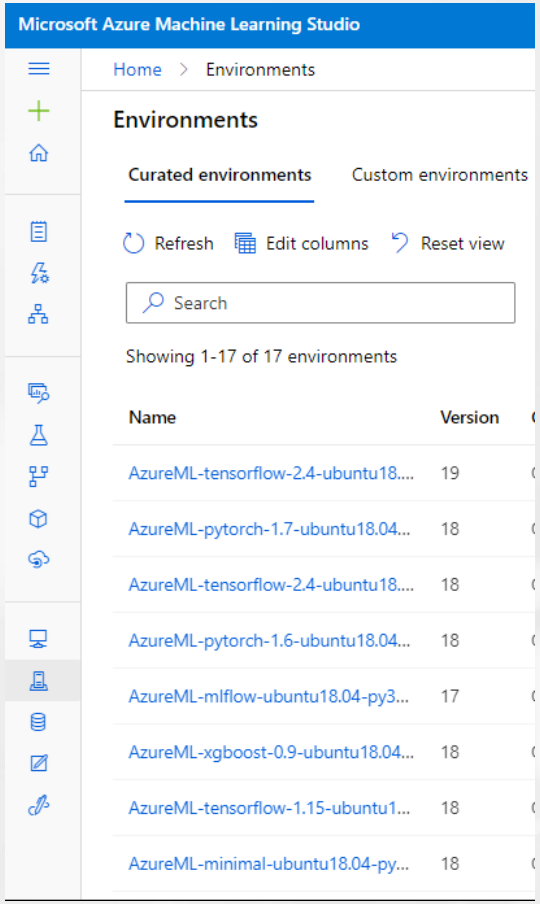
spf-users$ > a50e00r > Documents > Azure ML

```
absl-py==0.9.0
alabaster==0.7.12
astor==0.8.1
astroid==2.4.2
atomicwrites==1.4.0
attrs==19.3.0
autograd==1.3
autograd-gamma==0.5.0
autopep8==1.5.4
Babel==2.8.0
backcall==0.2.0
bleach==3.1.5
cachetools==4.0.0
certifi==2020.6.20
cffi==1.14.1
chardet==3.0.4
cloudpickle==1.5.0
cryptography==3.0
cycler==0.10.0
decorator==4.4.2
defusedxml==0.6.0
diff-match-patch==20200713
docutils==0.16
entrypoints==0.3
flake8==3.8.3
future==0.18.2
gast==0.2.2
```

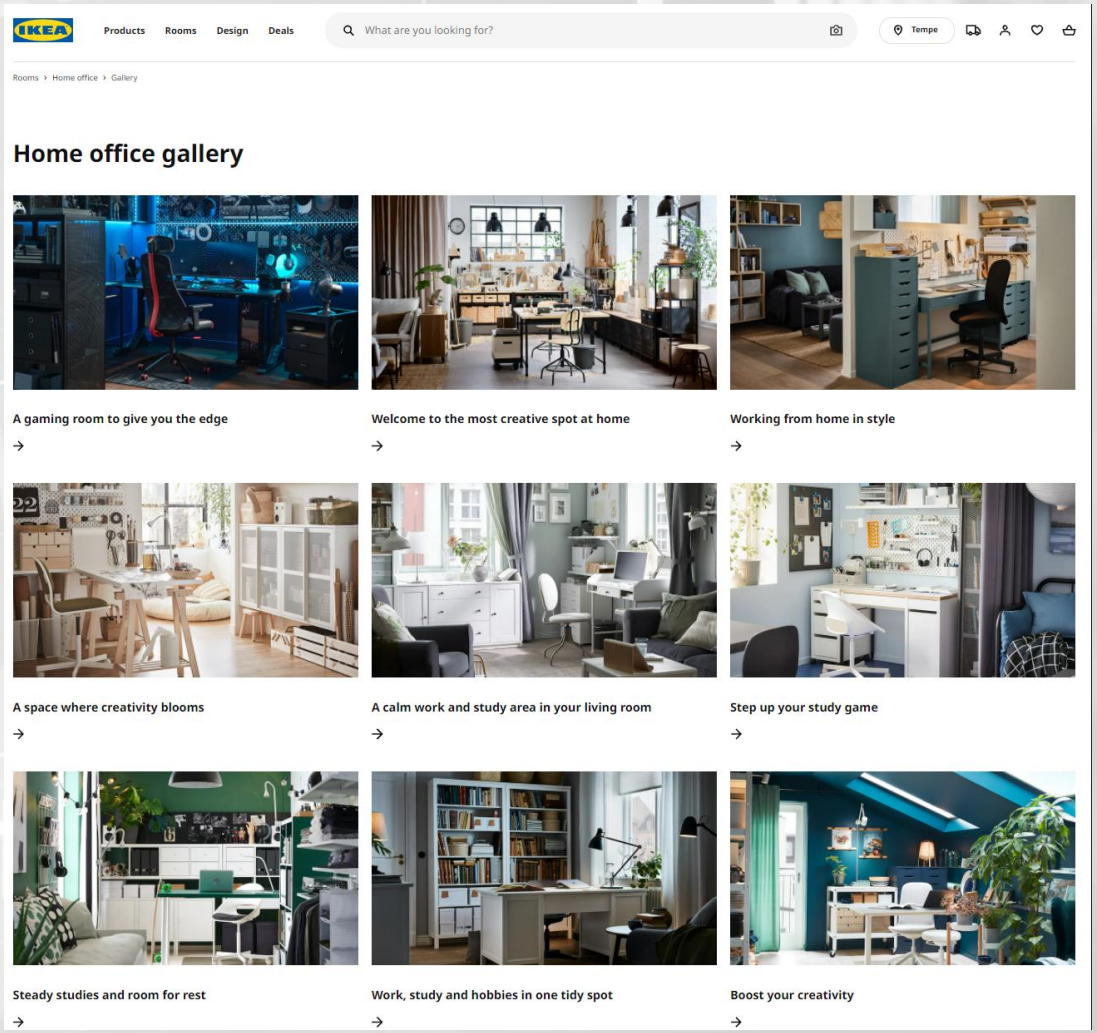
azure_ml.yml

```
name: azureML
dependencies:
  - python=3.7.11
  - jupyterlab=2.1.4
  - numpy=1.15.1
  - pandas=0.23.4
  - scikit-learn=0.19.2
  - joblib=1.0.1
  - matplotlib=2.2.3
  - cryptography=3.4.7
  - pyarrow=3.0.0
  - requests=2.26.0
  - urllib3=1.26.6
  - pip
```

(Azure) Curated and Registered Env.



- Like Ikea sample rooms… «Shop the look»



- You can also register custom envrionments yourself, and reuse them

Now we know: evironment is important for model to perform.

Time to let it doing some real business!

Infereing: real-time(on-demand)

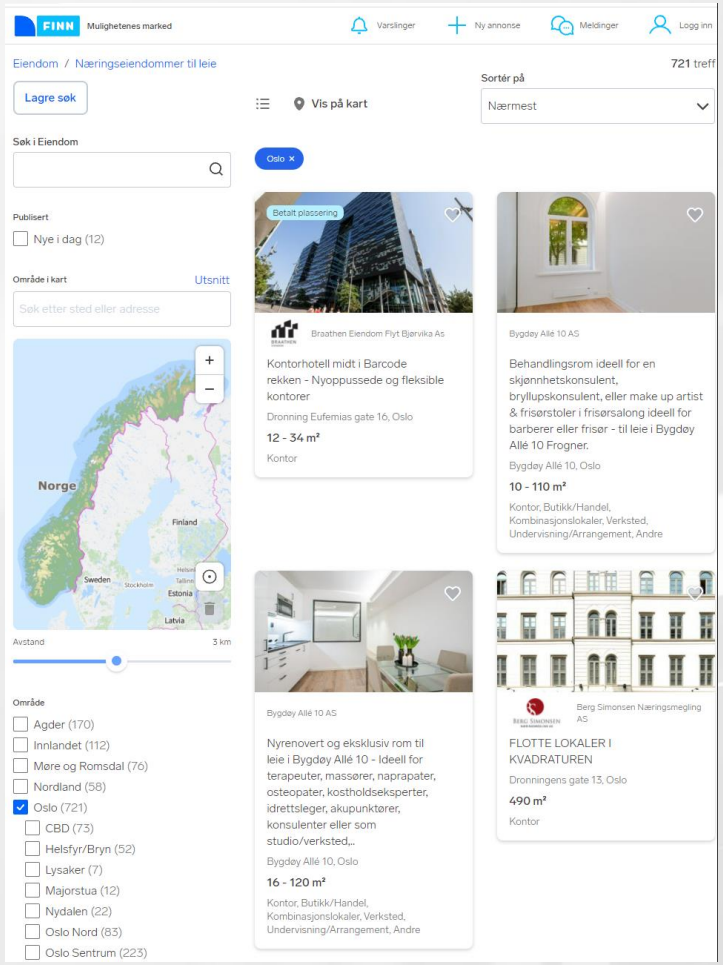
- After received the model and its office setup specs, Modhi DreamWorks decided to launch a new business:

Surprise!
Horoscope astrology!

- Business mode: prepaid customers can request the prediction of their luck at anytime, they will receive the response immediately (sounds like REST API?)

What is the next?

Rent a office place!



- No need to go to finn.no…
- Azure provides dedicated office places for real-time use:

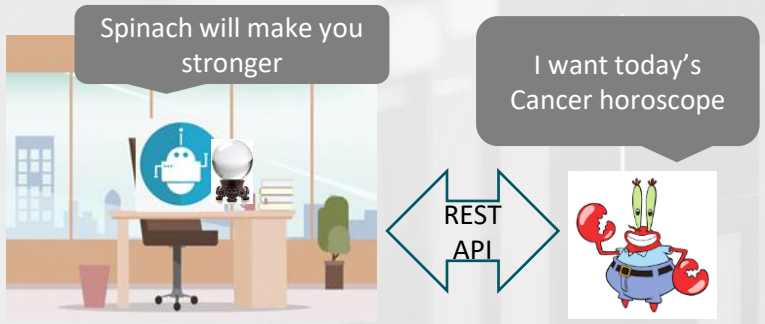
Azure Container Instance (ACI)

Azure Kubernetes Service (AKS)

Azure Container Instance (ACI)

- for one ~~man~~ robot team

- Allows you to run Docker containers without having to set up VMs yourself
- But you need to specify its size, rent price varies with it (of course, which model to use, and env.) – then Azure fix everything include Rest API



Linux-operativsystem

Varigheten til beholdergruppe

Ressurser	Pris per sekund	Pris per time	Pris per måned
Minne	**kr0,0000123** per GB	**kr0,04425** per GB	**kr32,3003** per GB
vCPU	**kr0,0001121** per prosessorenhet for virtuell maskin	**kr0,40351** per prosessorenhet for virtuell maskin	**kr294,5585** per prosessorenhet for virtuell maskin

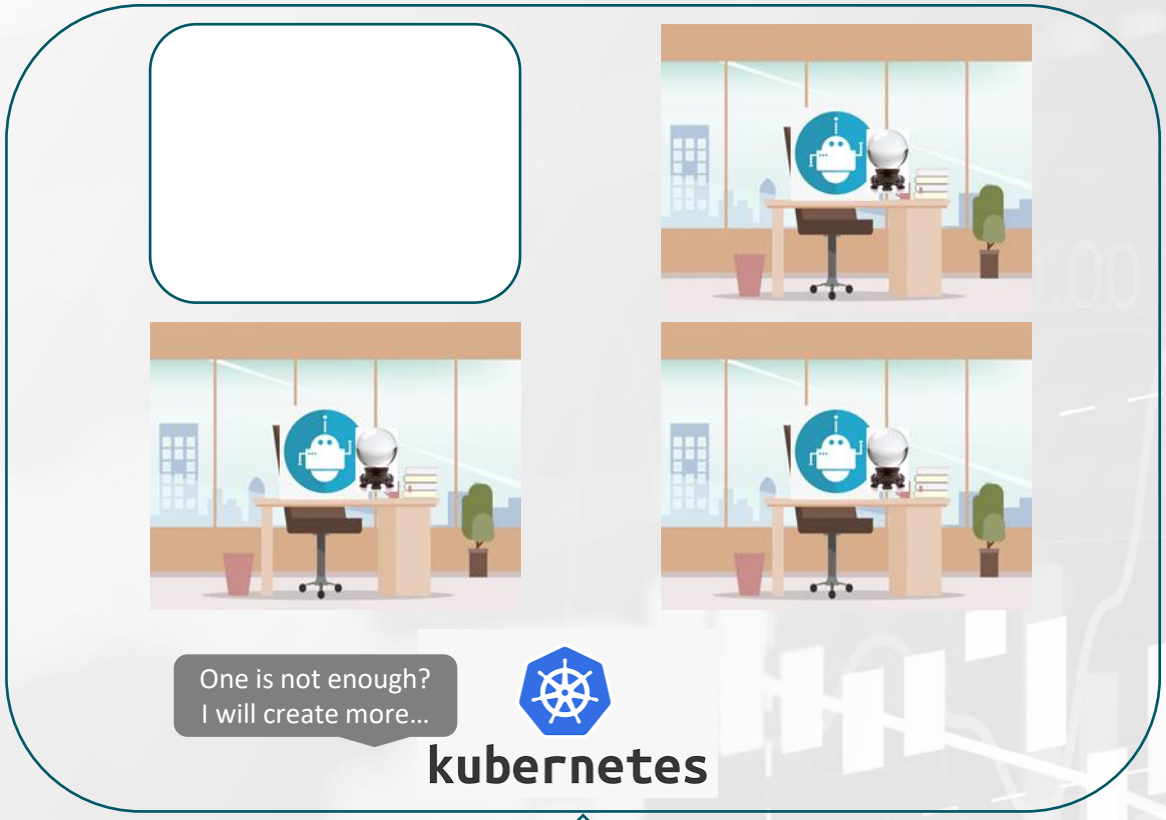
Varigheten til GPU-beholdergruppe

Ressurser	Pris per sekund	Pris per time	Pris per måned
Minne	**kr0,00001334** per GB	**kr0,04801** per GB	**kr35,0437** per GB
vCPU	**kr0,0000996** per prosessorenhet for virtuell maskin	**kr0,35848** per prosessorenhet for virtuell maskin	**kr261,6893** per prosessorenhet for virtuell maskin
K80	**kr0,0011257** per grafikkprosessor for virtuell maskin	**kr4,0523652** per grafikkprosessor for virtuell maskin	**kr2 958,2265960** per grafikkprosessor for virtuell maskin
P100	**kr0,0027449** per grafikkprosessor for virtuell maskin	**kr9,8815367** per grafikkprosessor for virtuell maskin	**kr7 213,5217764** per grafikkprosessor for virtuell maskin
V100	**kr0,0089187** per grafikkprosessor for virtuell maskin	**kr32,1072012** per grafikkprosessor for virtuell maskin	**kr23 438,2568759** per grafikkprosessor for virtuell maskin

When business growing rapidly
- ~~Hiring?~~ Clone!

Azure Kubernetes Service (AKS)

- managing things when demand grows (recommended for production)



Kubernetes – for container orchestration:

- Autoscaling
- Load balancing
- and much more…

Similar to setting up a ACI, you only need to name model, env, and addtioanally, autoscaling parameters (min, max nodes, throus)

Like a manager

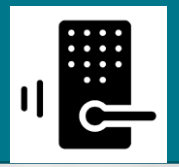
Authentication

Azure Container Instance (ACI)

key-based:
static, unless refresh the key

Azure Kubernetes Service (AKS)

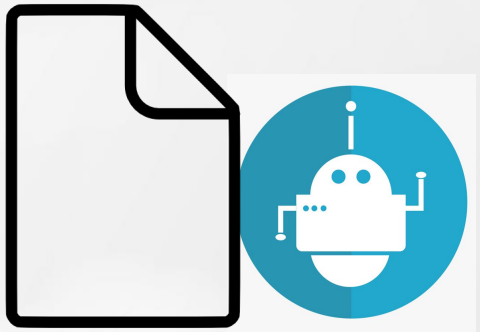
or



key-based

token-based:
token changes houly

Recap:real-time inference



Ready for your business

Tell Azure:

- the env and the model (and a little bit more) you are going to use
- the type of infrastructure (ACI or AKS) and its size
- authentication method

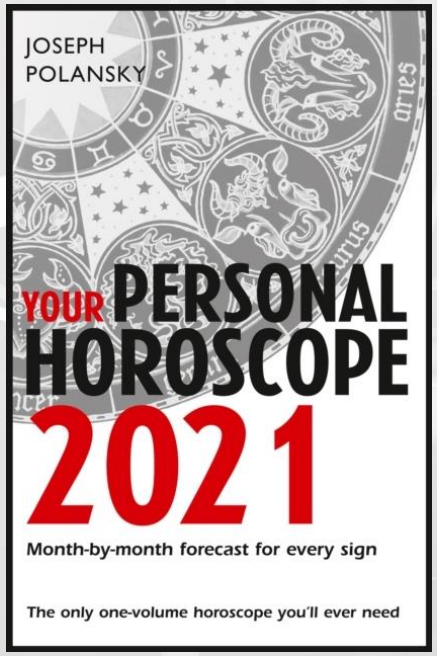
Azure will set it up for you

We haven't finished yet…

Another business mode…

Inferecing: Batch

- After the success of the real-time horoscope astrology service, Modhi DreamWorks decided to add another business mode: print all the prediction out at once!



Batch with ML pipeline

A little bit history: On December 1, 1913, Henry Ford installs the first moving assembly line for the mass production of an entire automobile. His innovation reduced the time it took to build a car from more than 12 hours to one hour and 33 minutes.

Use similar line of thinking when mass produce predictions

DataTransferStep

PythonScriptStep

ParallelRunStep

Azure ML Pipeline

Data Factory

VM

AML Cluster

AdlaStep: Runs a U-SQL job in Azure Data Lake Analytics

DatabricksStep: Runs a notebook, script, or compiled JAR on a databricks cluster.

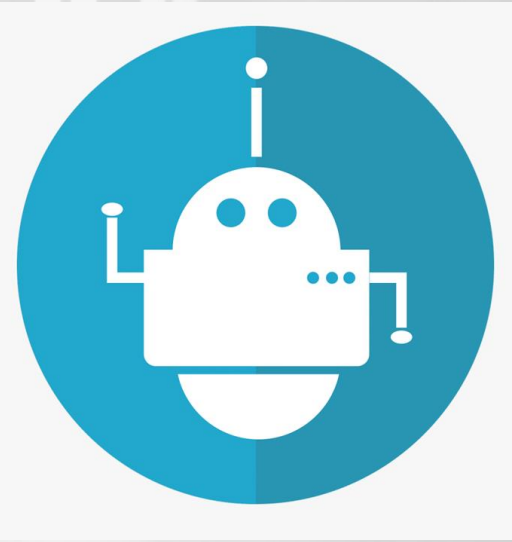
(Pipeline steps are building blocks for the pipeline. You can use them to build a pipleline as you want.)

ML Pipeline Publish and Schedule

- Publish it:
  - Get a REST Endpoint (a URL give you access to it)
- Schedule it:
  - Time-based
  - Change-based (source data)

Now your Horoscope book printing line is created!

Back to the start point (when model is made)



Workspace

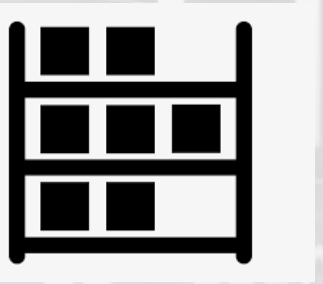
In «Ex Machina» Ava found the previous version of robots has been created.

Azure ML workspace also provides storages to keep tracks.

store, versioning, relevant: source, generated models

logs, metrics etc. outputs

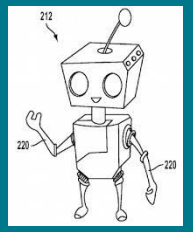
store, versioning, relevant: datasets, endpoints, artifacts etc.



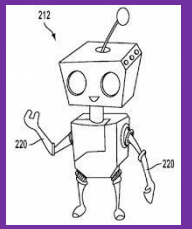
Datasets

Experiments

Models



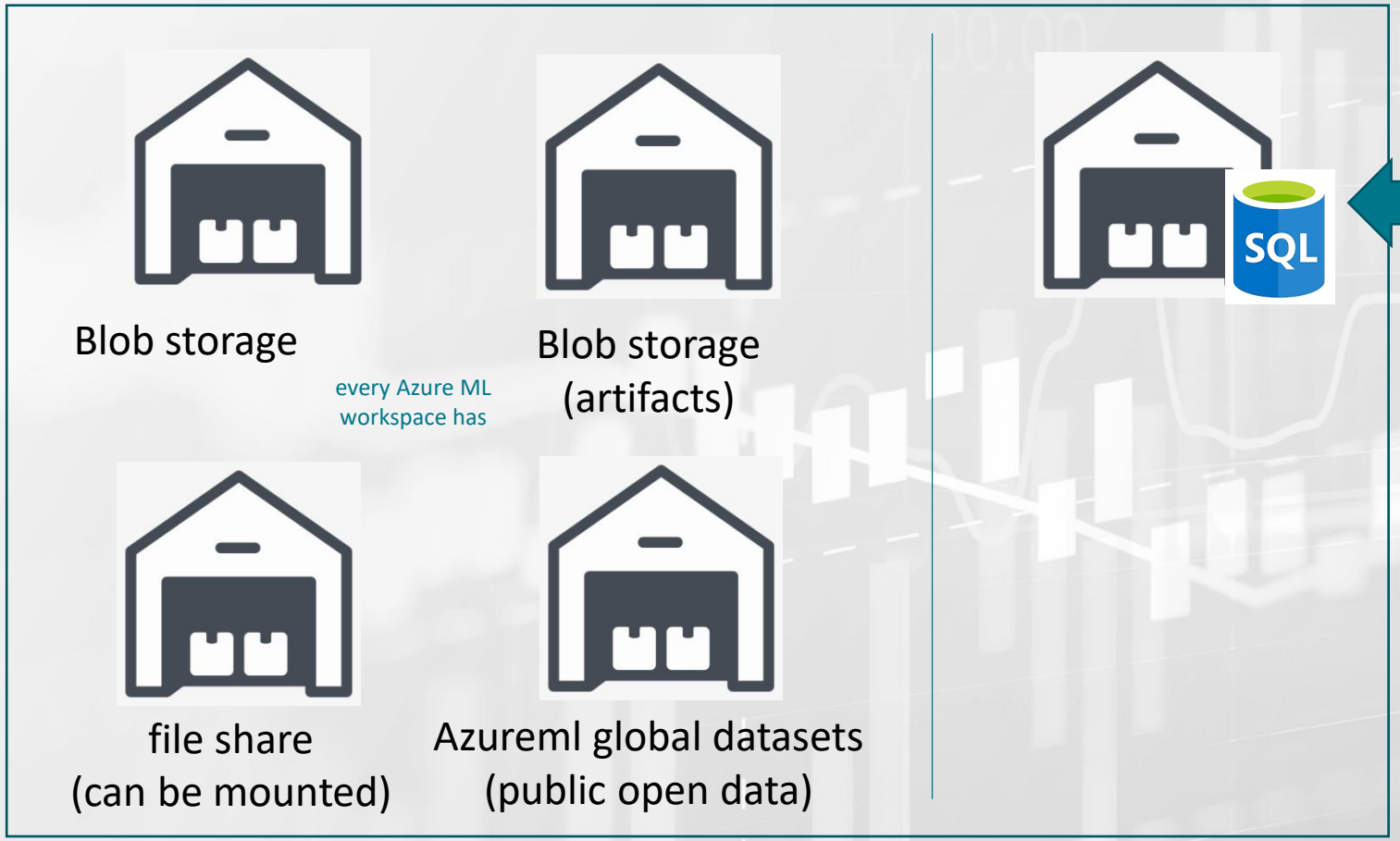
Local compute target



OEM: Remote compute target

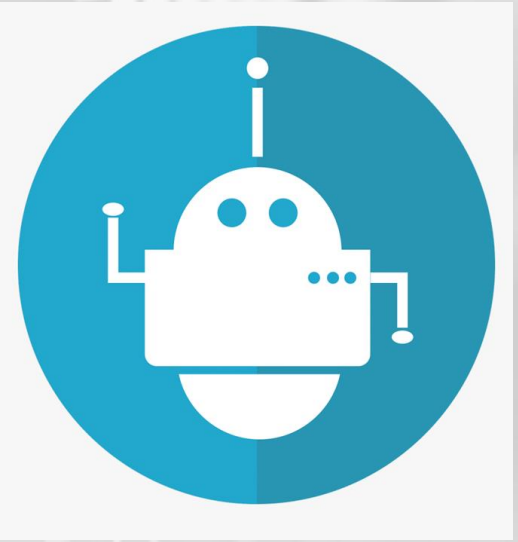
Datastores for ML workspace

- In order to be sourced, you need to be a registered supplier in ML Workspace



Storage services which can be registered:

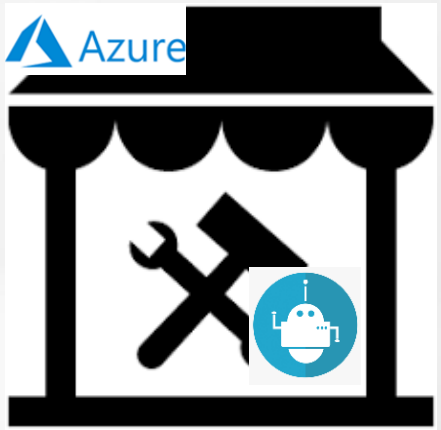
- Azure Blob Container
- Azure File Share
- Azure Data Lake
- Azure Data Lake Gen2
- Azure SQL Database
- Azure Database for PostgreSQL
- Databricks File System
- Azure Database for MySQL



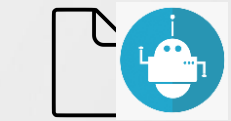
How model is made? Maybe next time.

Summary

Summary



workspace where model is created



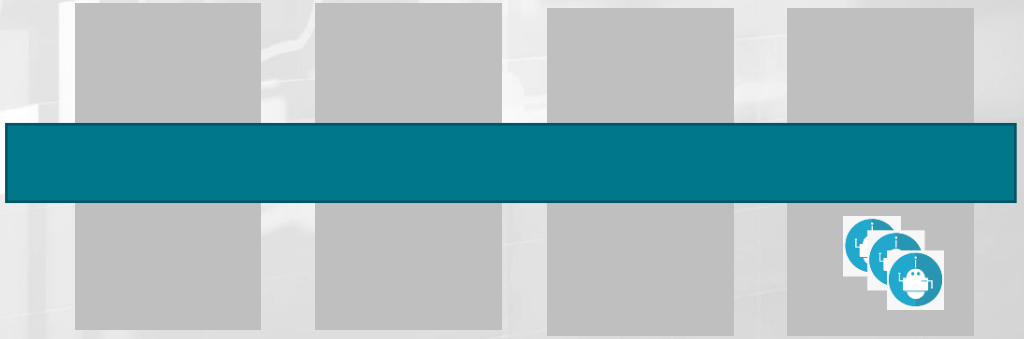
real-time service

ACI

AKS

OR

mass producing

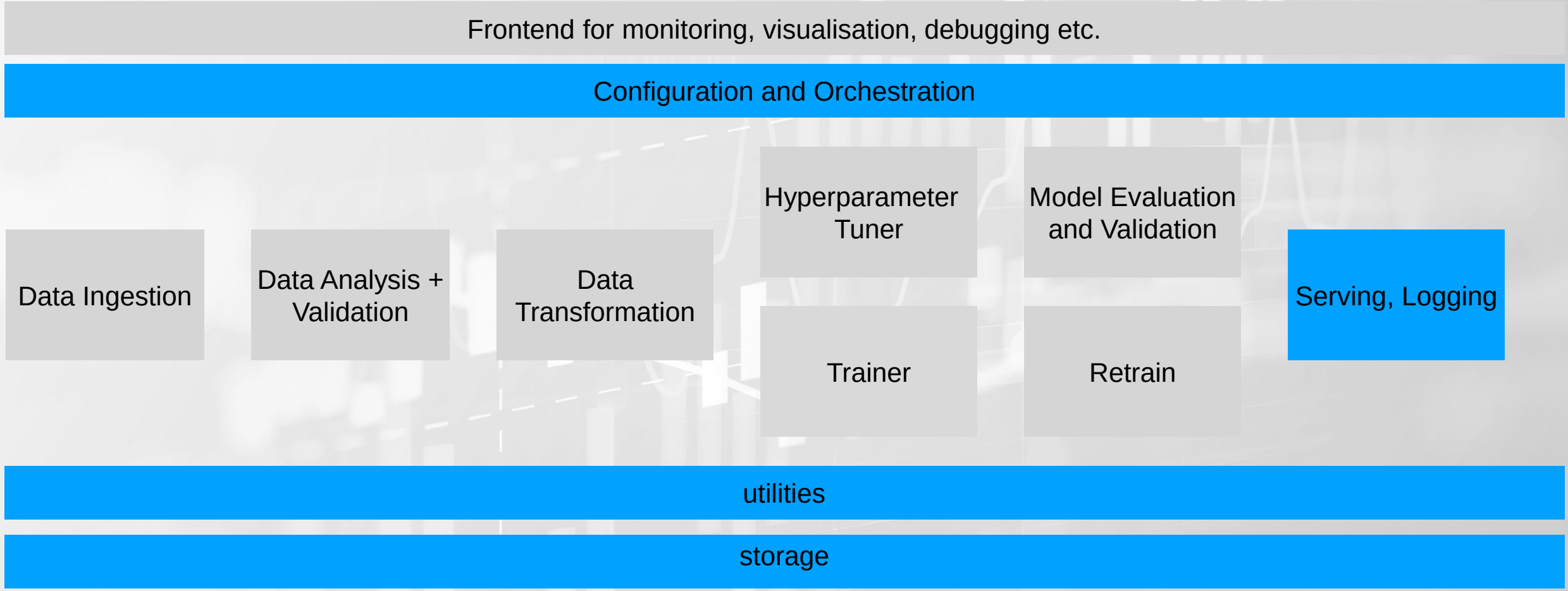


Data stores

Logs, registered datasets, models, enviroments Etc.

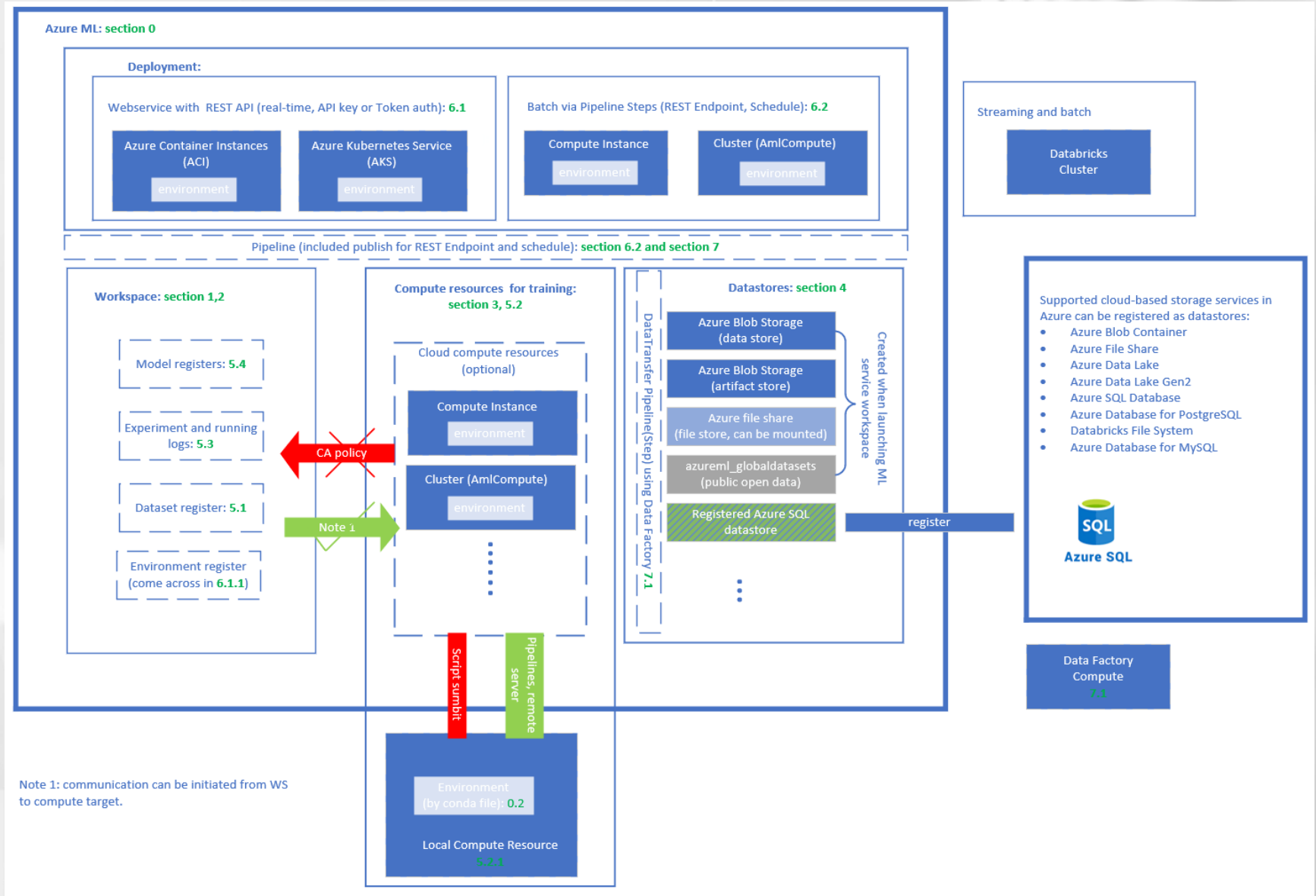
We touched

Machine Learning Model in Production



How?

- Read the tutorials
- Feel free to try and explore in lab environment



Covers only part of what Azure ML offers

Thank you!