

Progettare e rilasciare flussi MLOps su AWS



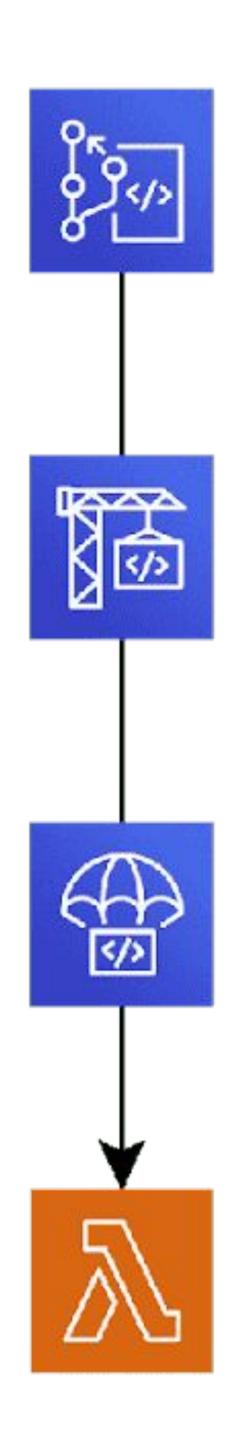
Matteo Moroni

Technical Account Manager @ beSharp spa

beSharp is an Italian Premier AWS Partner, whose core business consists in migrating its customers on AWS

Materiale in parte da https://github.com/xtreamsrl/amld22-mlops-on-aws

Collaborazione con Xstream srl



Outline

ML ed MLOps

AWS introduzione

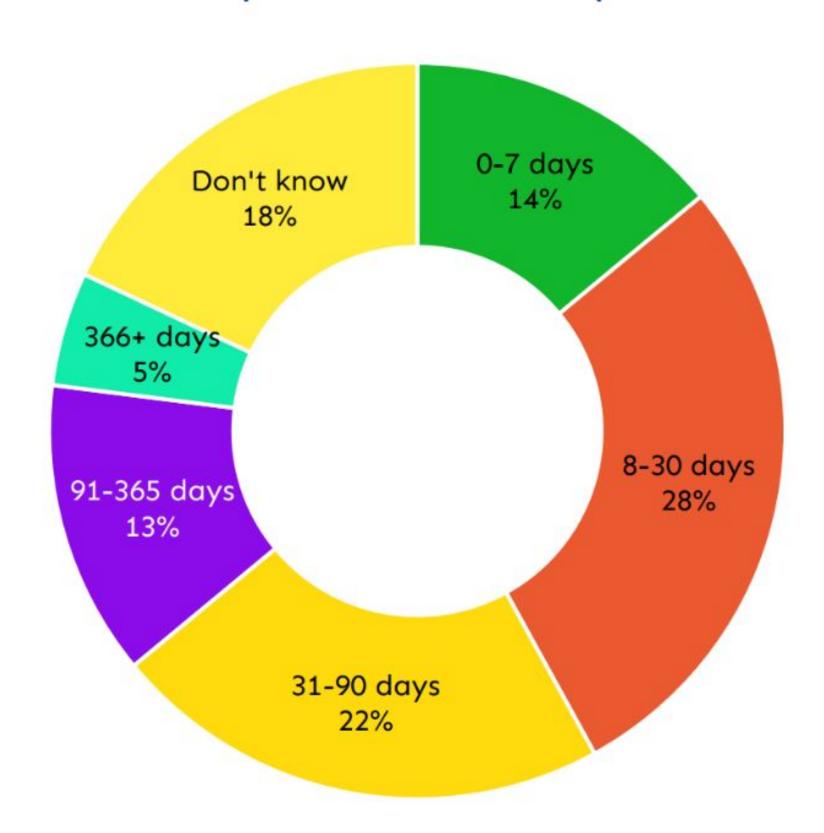
MLOps su AWS

Un esempio pratico

Conclusioni

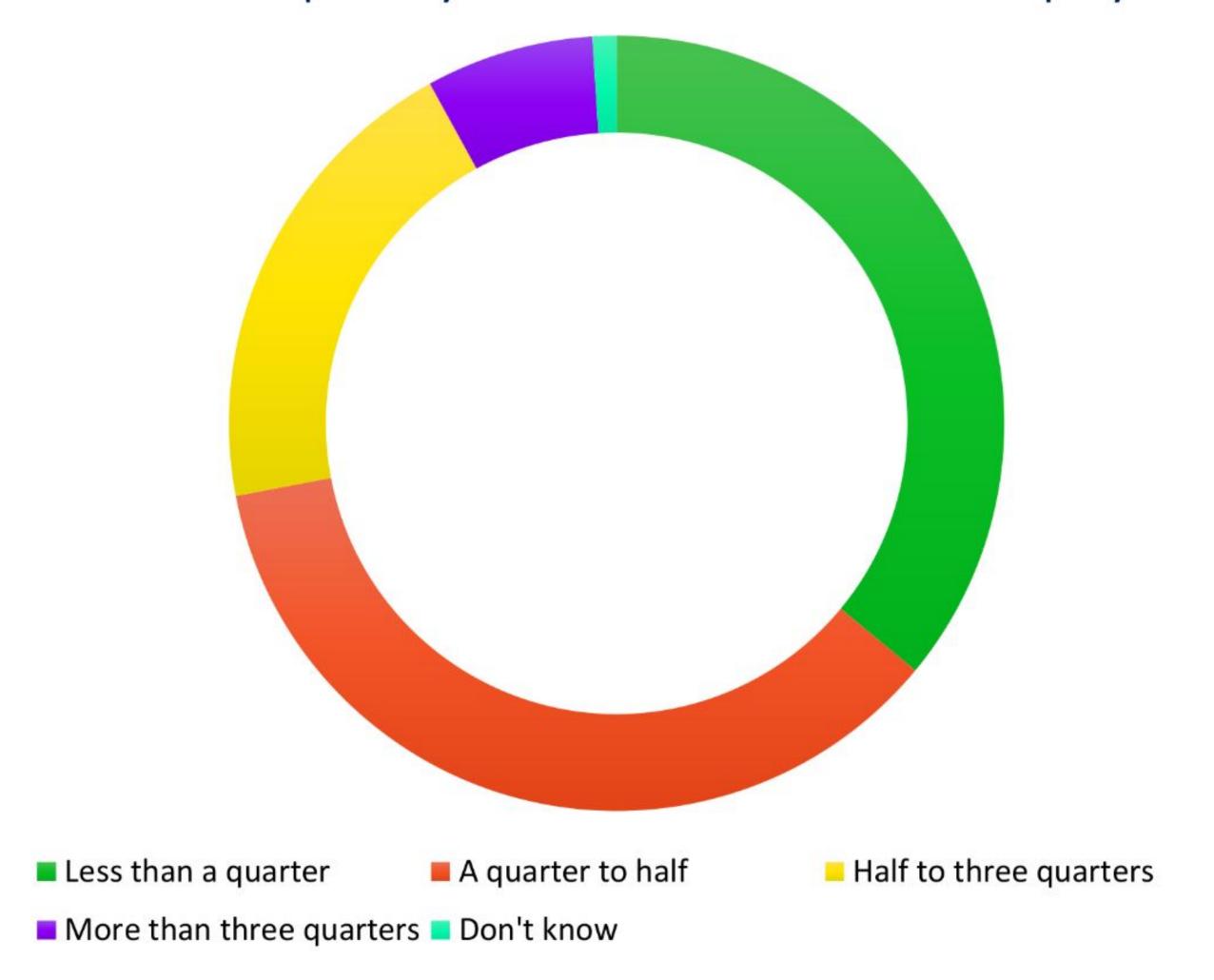
ML Development is slow

Time to ship a new model to production

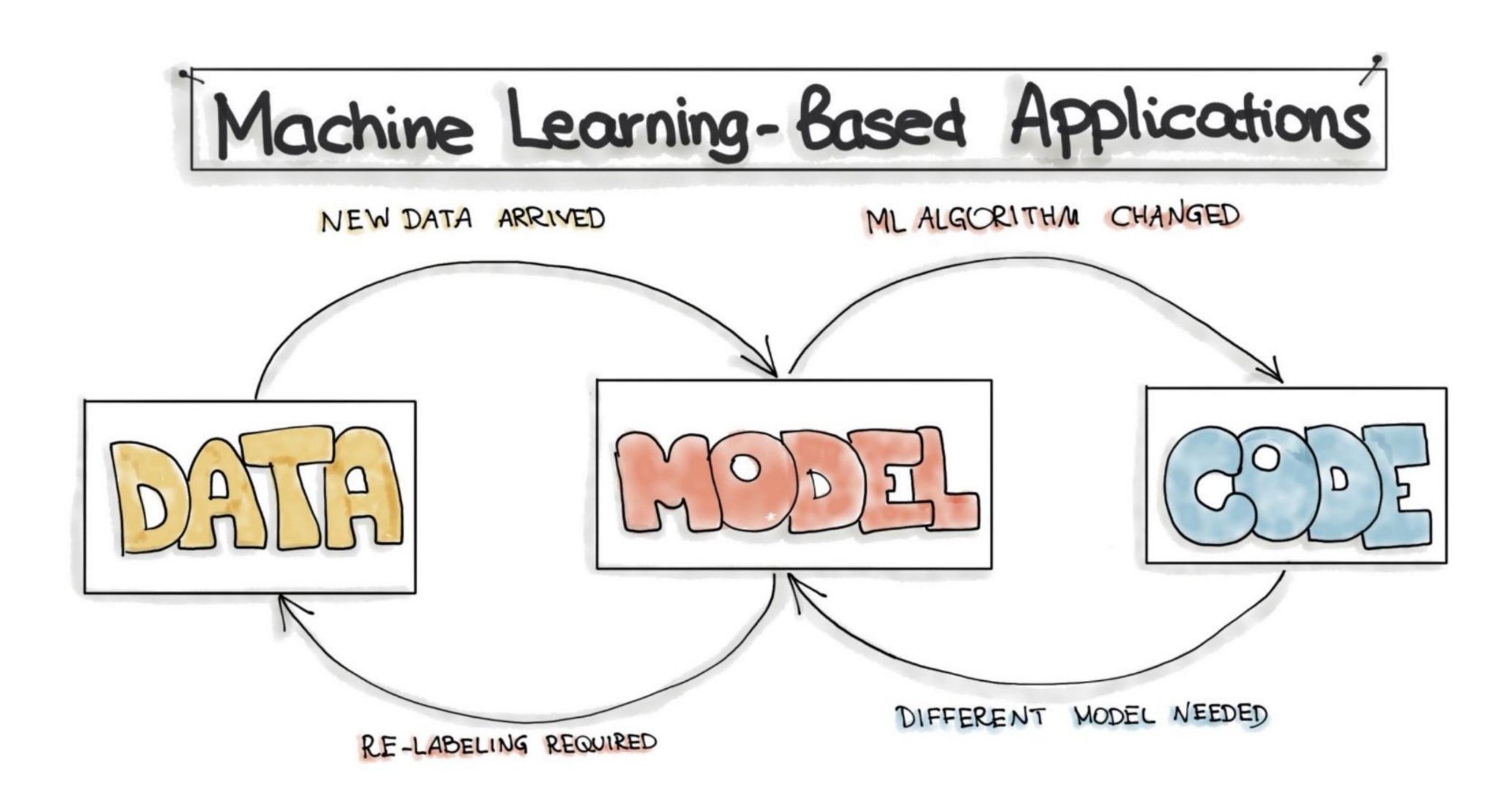


ML Deployment is time-consuming

Share of time spent by data scientists on model deployment



Updates cascade



The Machine Learning process

3. Data Exploration

- Explorative analysis
- Data cleaning



1. Business Understanding

- Problem statement
- Evaluation metrics
- Literature review

2. Data Gathering

- Source discovery
- Data preparation
- Quality assessment



5. Evaluation

- Performance evaluation
- Time evaluation

4. Modelling

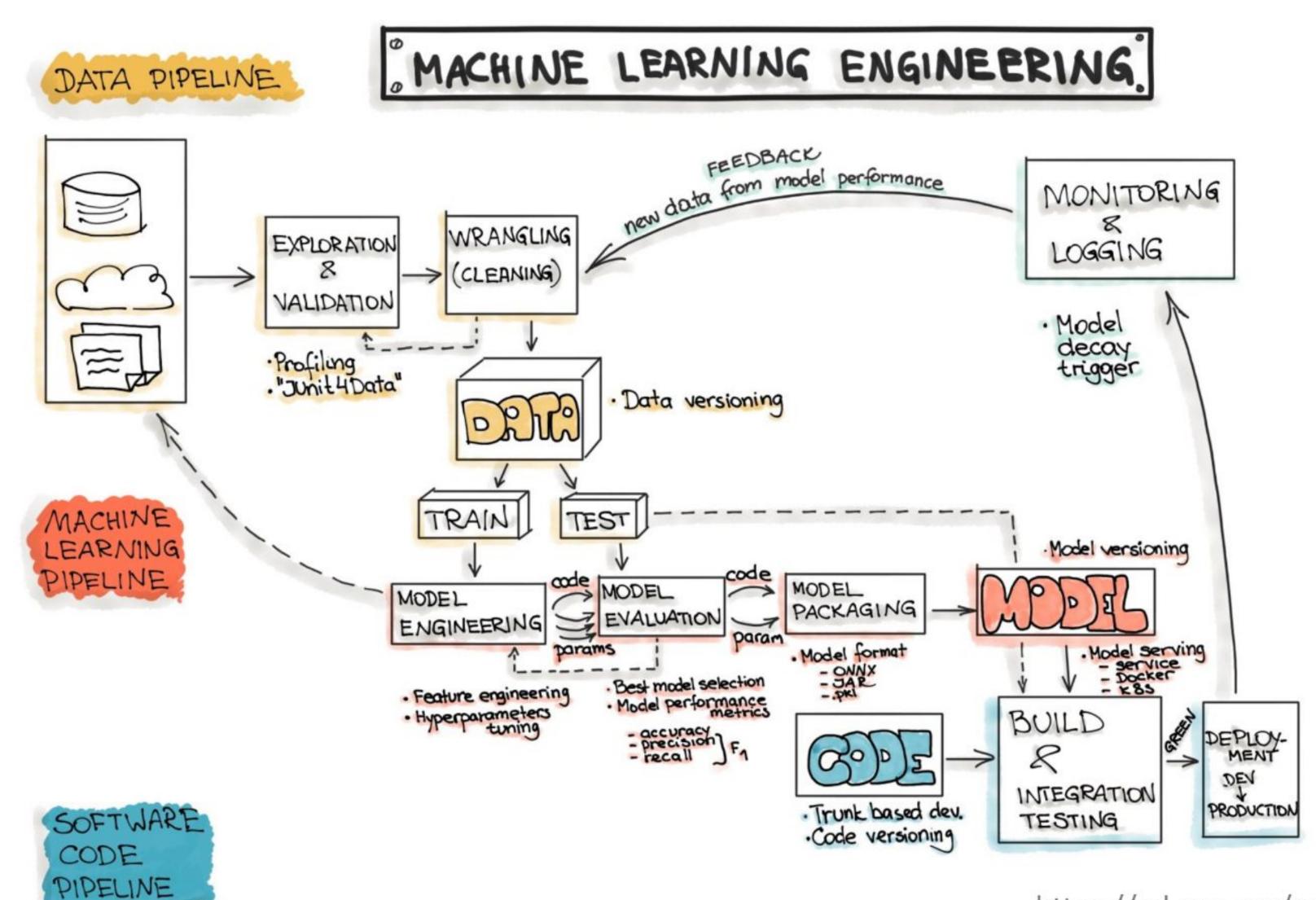
- Feature engineering
- Model design
- Model implementation



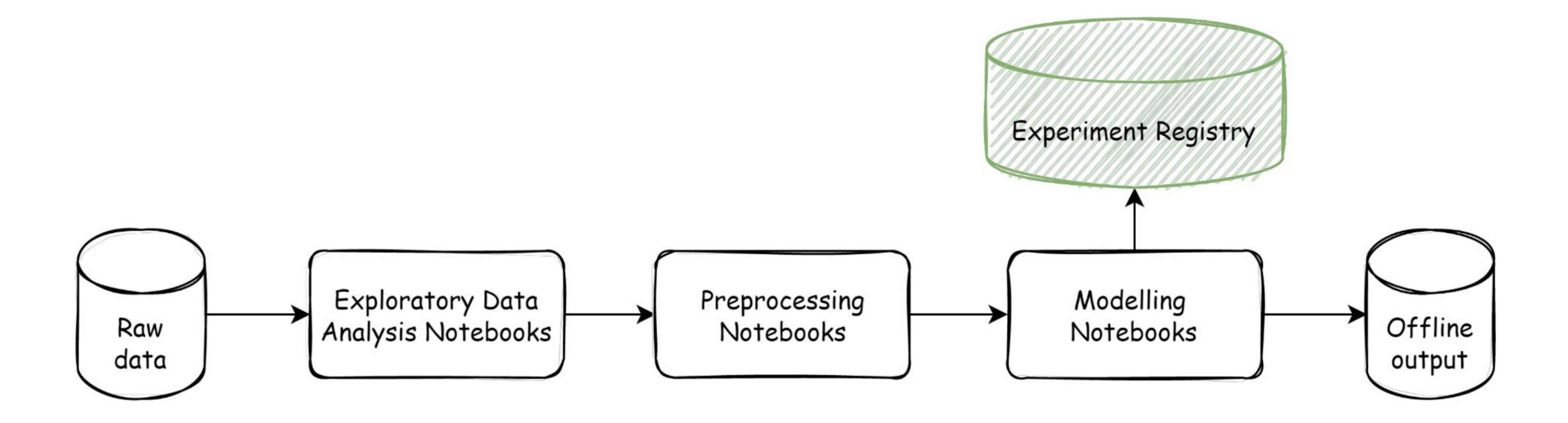
6. Deployment

- Architecture design
- Execution scheduling
- Performance tuning
- Integration

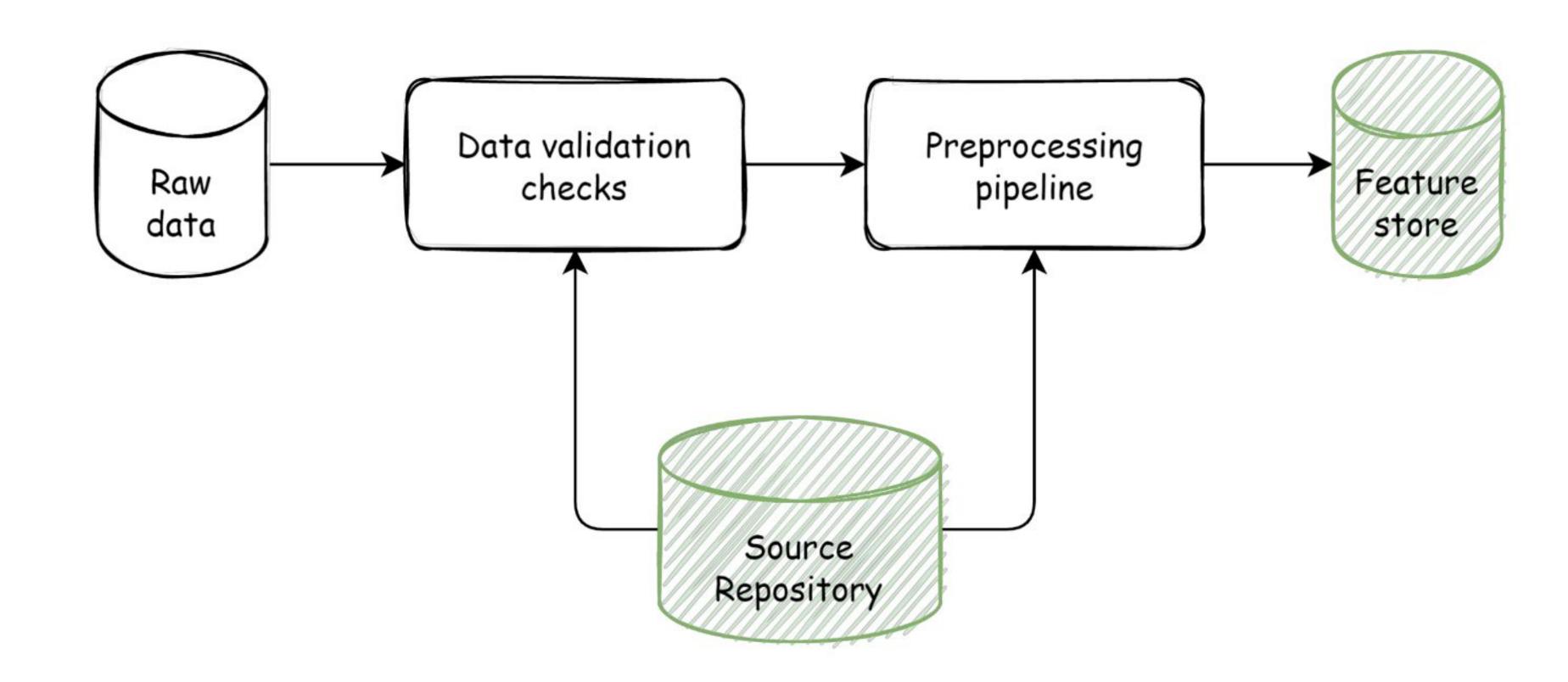
The Machine Learning process



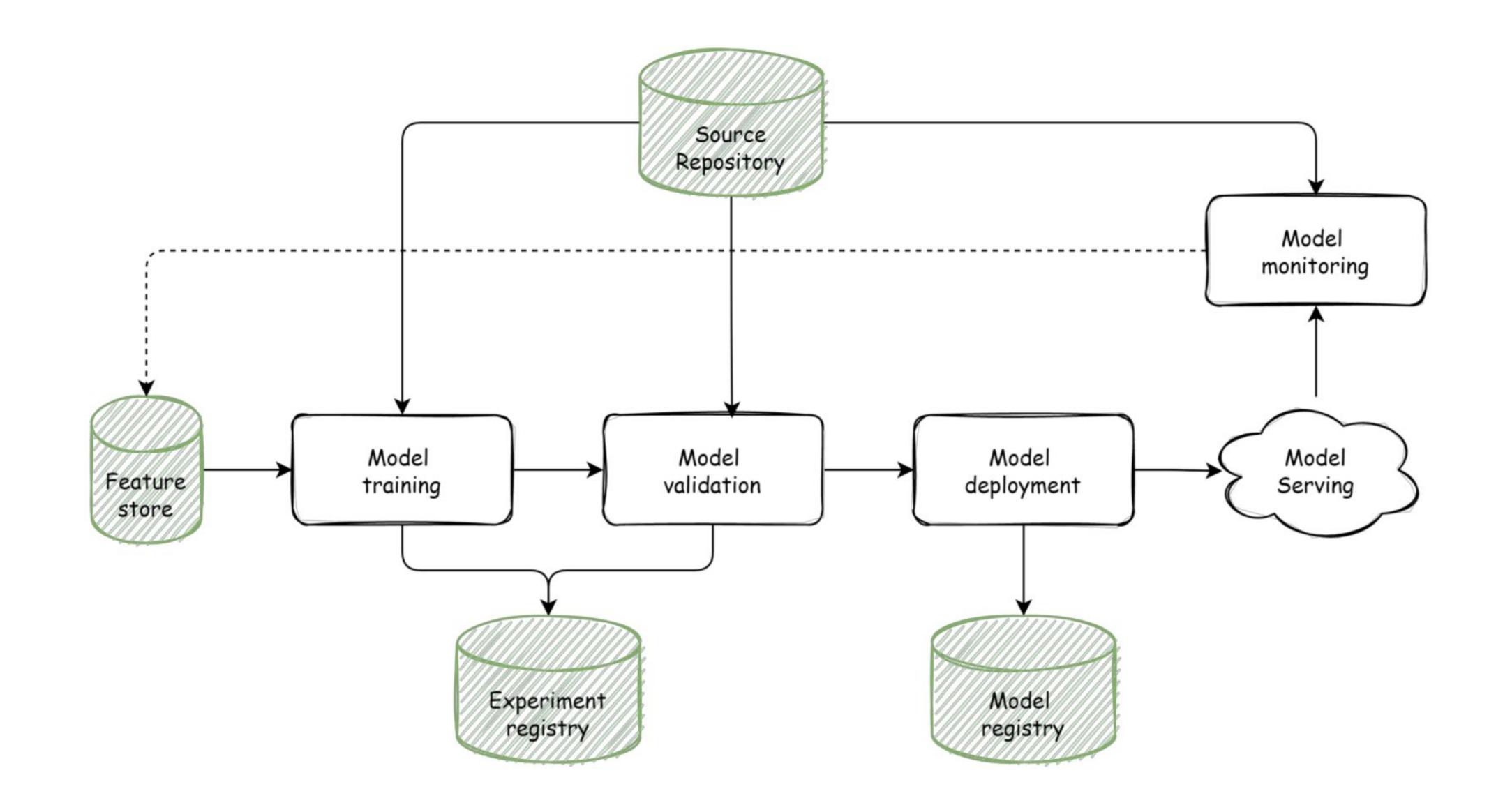
Offline development



Automated data pre-processing

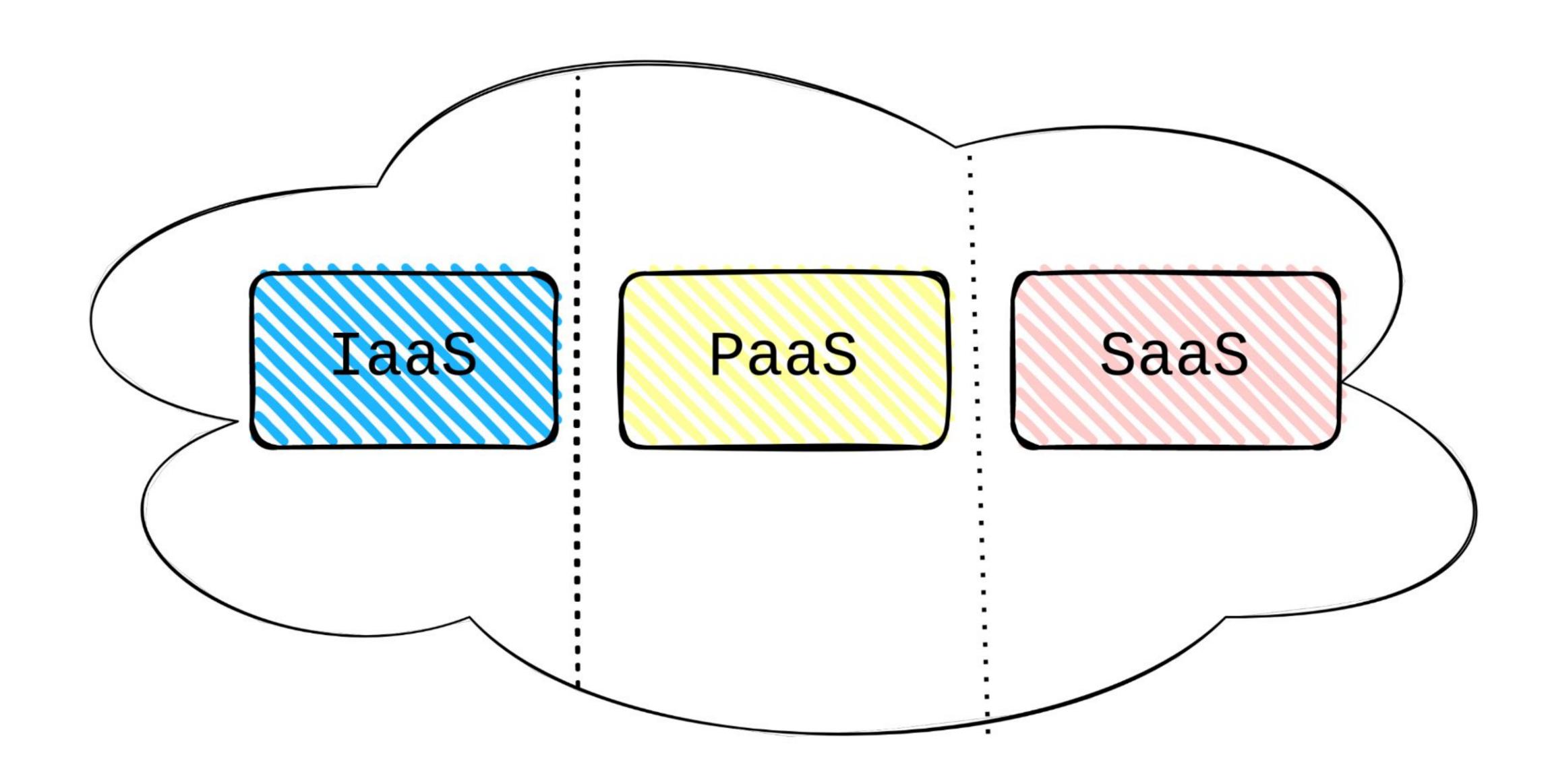


Automated model lifecycle

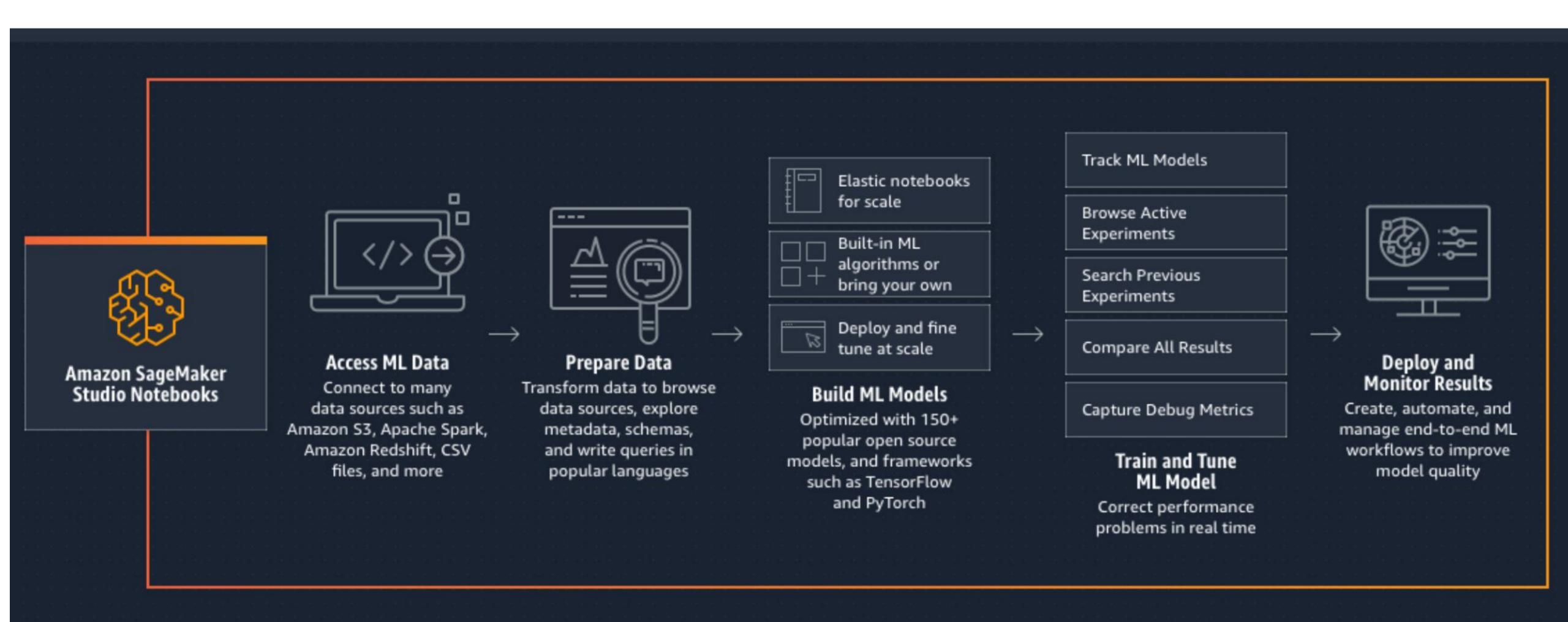


AWS Cloud

AWS Cloud basics



Sagemaker Components



Sagemaker Training

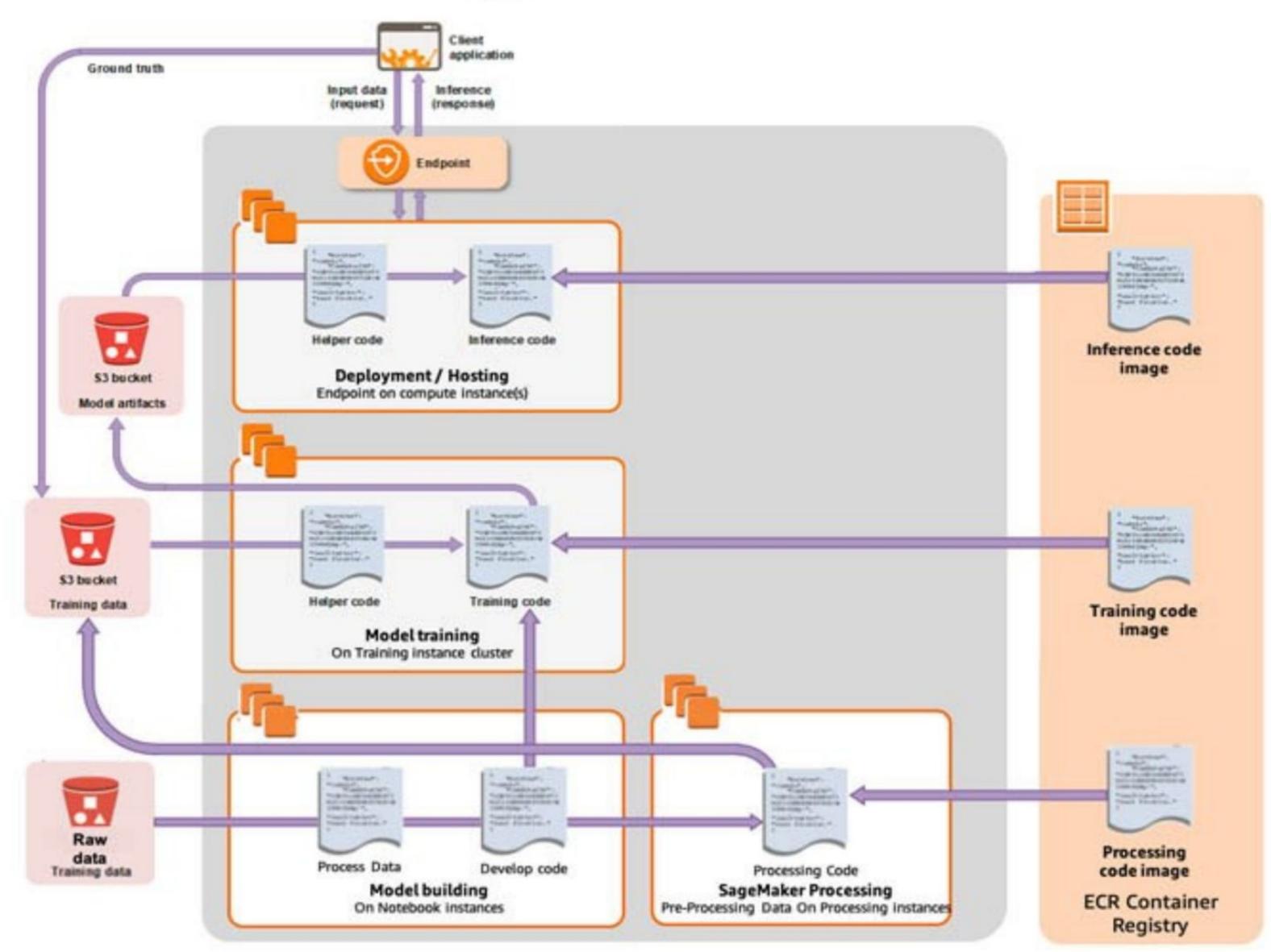
And then you deploy your model to production for realtime inference (or you use Batch Transform for batch inference*)



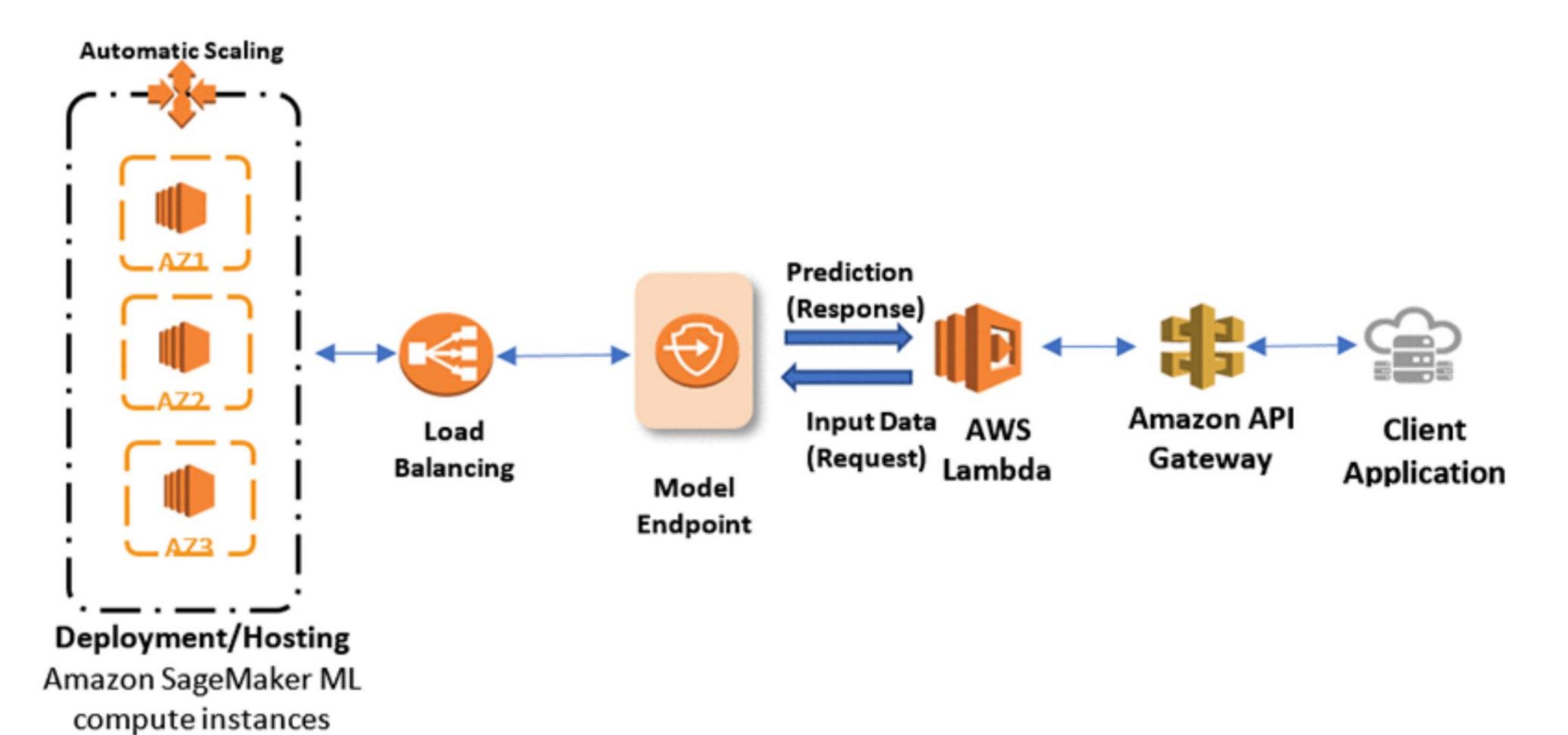
Then you train and tune your model



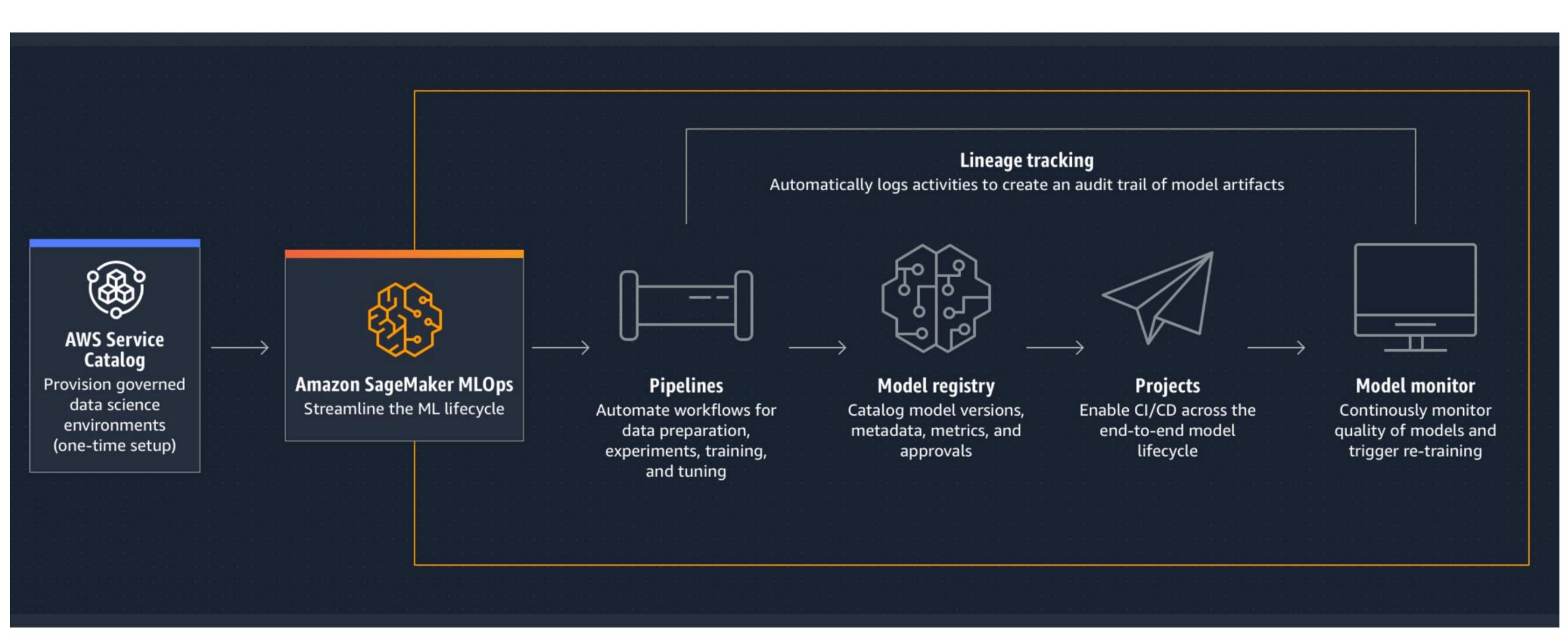
You start by building your model and processing raw data, developing your training data



Sagemaker Endpoint



AWS MLOps: Sagemaker Studio





Questions?