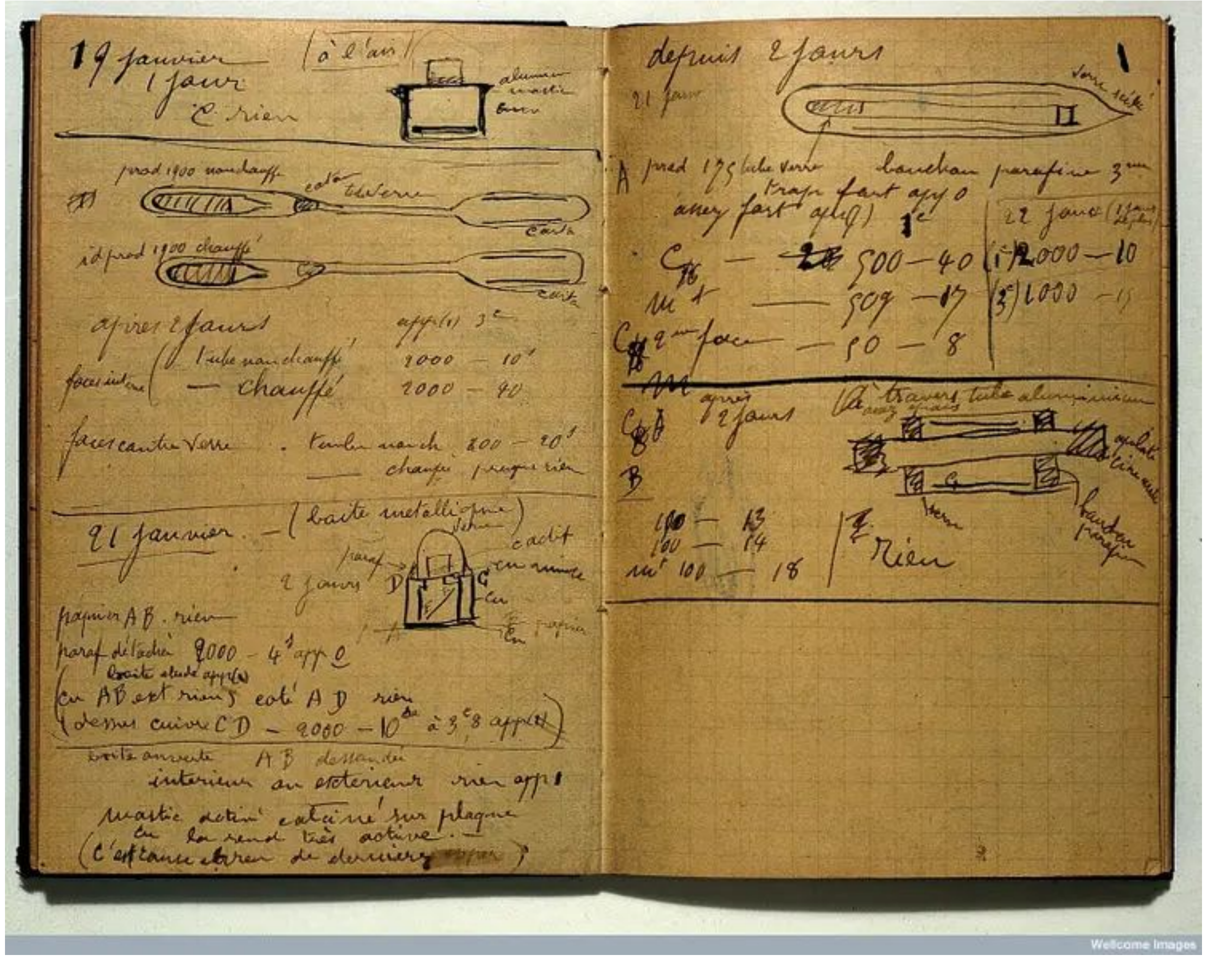
ADALTAS

3. Experiment tracking

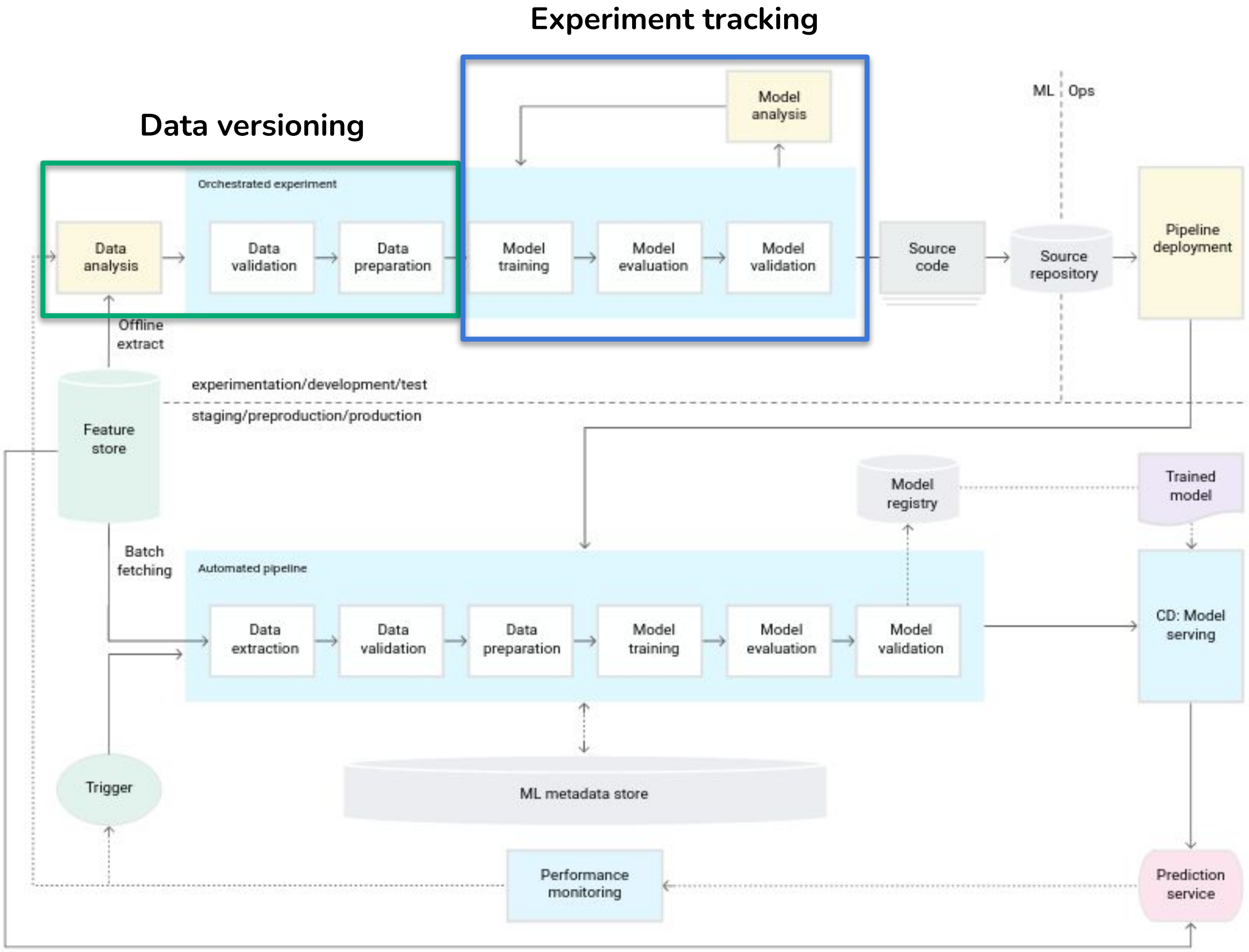
How not to forget what we already tested

Each experiment requires bookkeeping



Roadmap

ADALTAS



What is experiment tracking?

- During model development, we test many different things:
  - algorithms
  - parameters
  - datasets (split, sampling…)
  - feature

- In long term, we would like to memorize, what we tested and how
- In short term, we would like to compare results (easily) to keep the best model

What is experiment tracking?

- To be able to compare different models, we need to track (record):
  - statistical metrics
  - business metrics
  - other artefacts (plots)

- To reproduce the training/execution, we need to track the environment (libraries, versions)

- Many tools available and they mostly offer end-to-end ML lifecycle management
- We will use MLFlow

What is MLFlow?

ADALTAS

MLflow is an open source platform to manage the ML lifecycle, including experimentation, reproducibility, deployment, and a central model registry. MLflow currently offers four components:

MLflow Tracking	MLflow Projects	MLflow Models	Model Registry
Record and query experiments: code, data, config, and results	Package data science code in a format to reproduce runs on any platform	Deploy machine learning models in diverse serving environments	Store, annotate, discover, and manage models in a central repository

- We will focus on **experiment tracking**.
- We will be using managed version on Databricks.

MLFlow basics

ADALTAS

MLflow Tracking is organized around the concept of *runs*, which are executions of some piece of data science code. Each run records the following information:

Code Version

Git commit hash used for the run, if it was run from an MLflow Project.

Start & End Time

Start and end time of the run

Source

Name of the file to launch the run, or the project name and entry point for the run if run from an MLflow Project.

Parameters

Key-value input parameters of your choice. Both keys and values are strings.

Metrics

Key-value metrics, where the value is numeric. Each metric can be updated throughout the course of the run (for example, to track how your model's loss function is converging), and MLflow records and lets you visualize the metric's full history.

Artifacts

Output files in any format. For example, you can record images (for example, PNGs), models (for example, a pickled scikit-learn model), and data files (for example, a Parquet file) as artifacts.

https://www.mlflow.org/docs/latest/tracking.html

What will we do?

ADALTAS

databricks

- Track some experiments with the linear model
- Convert the linear model to Random Forest and repeat