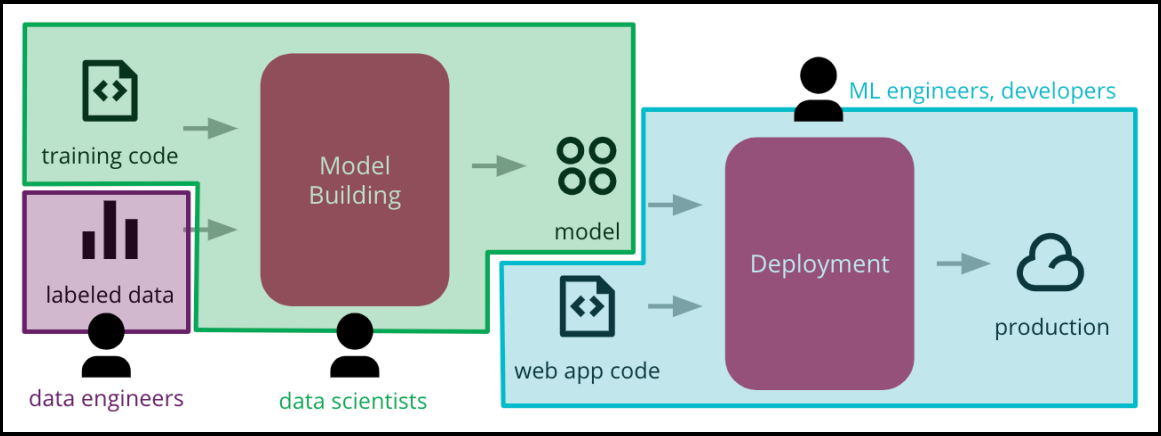
Continuous Delivery for ML

Jin Guo Oct 27th, 2020



Silos -> delays and friction, quickly becomes stale and hard to update

Hard to reproduce and audit: different tools and workflow, hard to automate end-to-end.

Data pipeline

Model training pipeline

Deployment pipeline

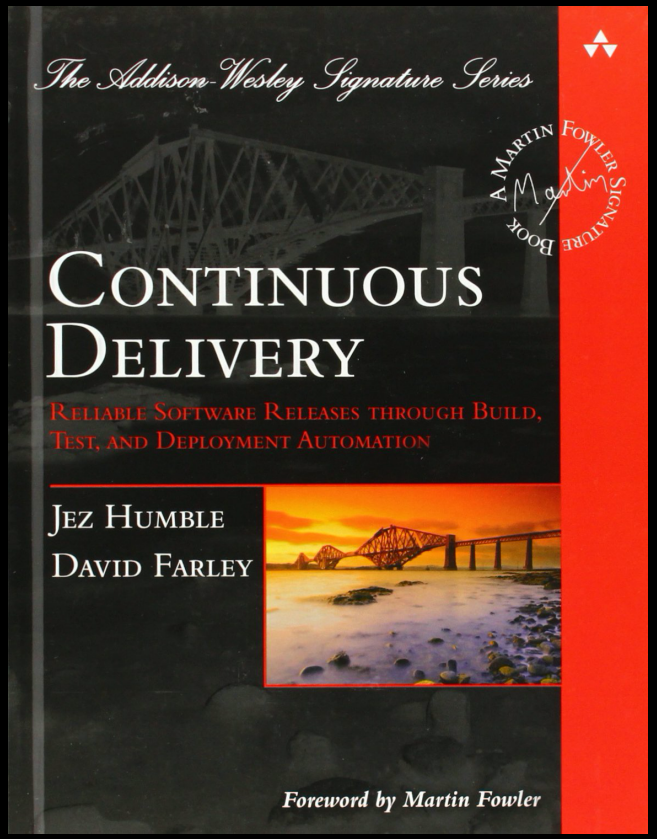
Safely Quickly Sustainably

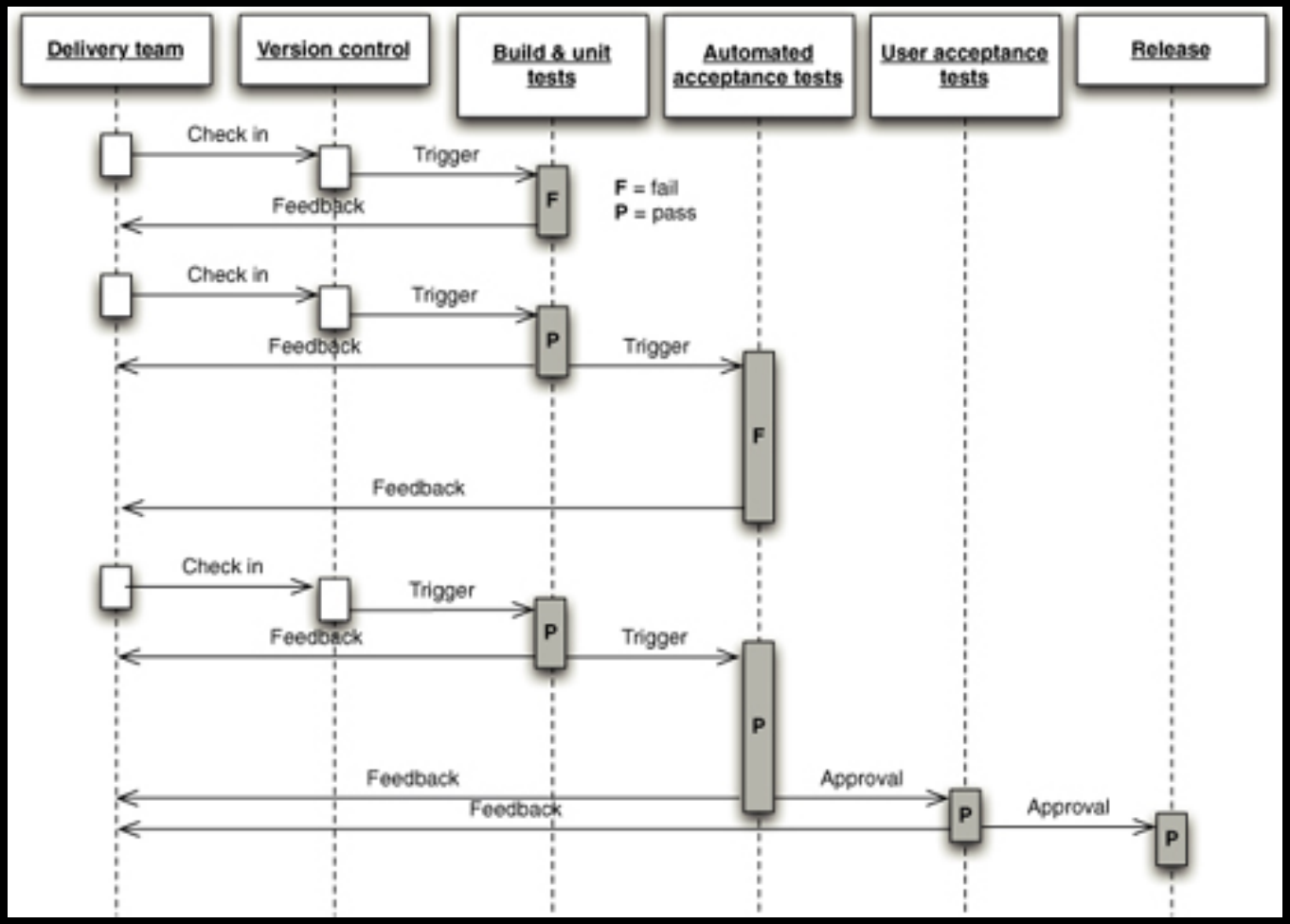
https://martinfowler.com/articles/cd4ml.html

Continuous Delivery

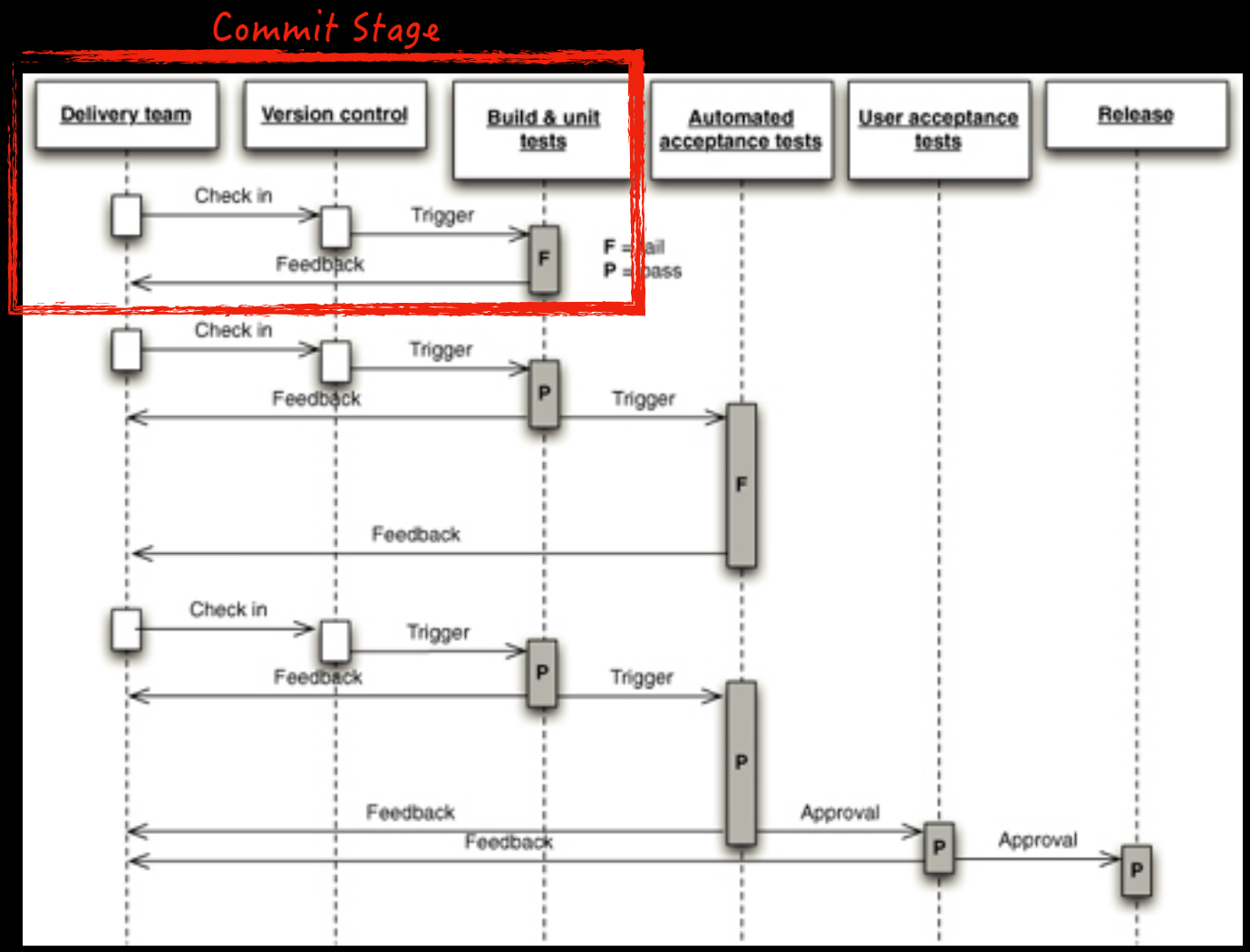
A close, collaborative working relationship between development and deployment teem

Extensive automation of all possible parts of the delivery process



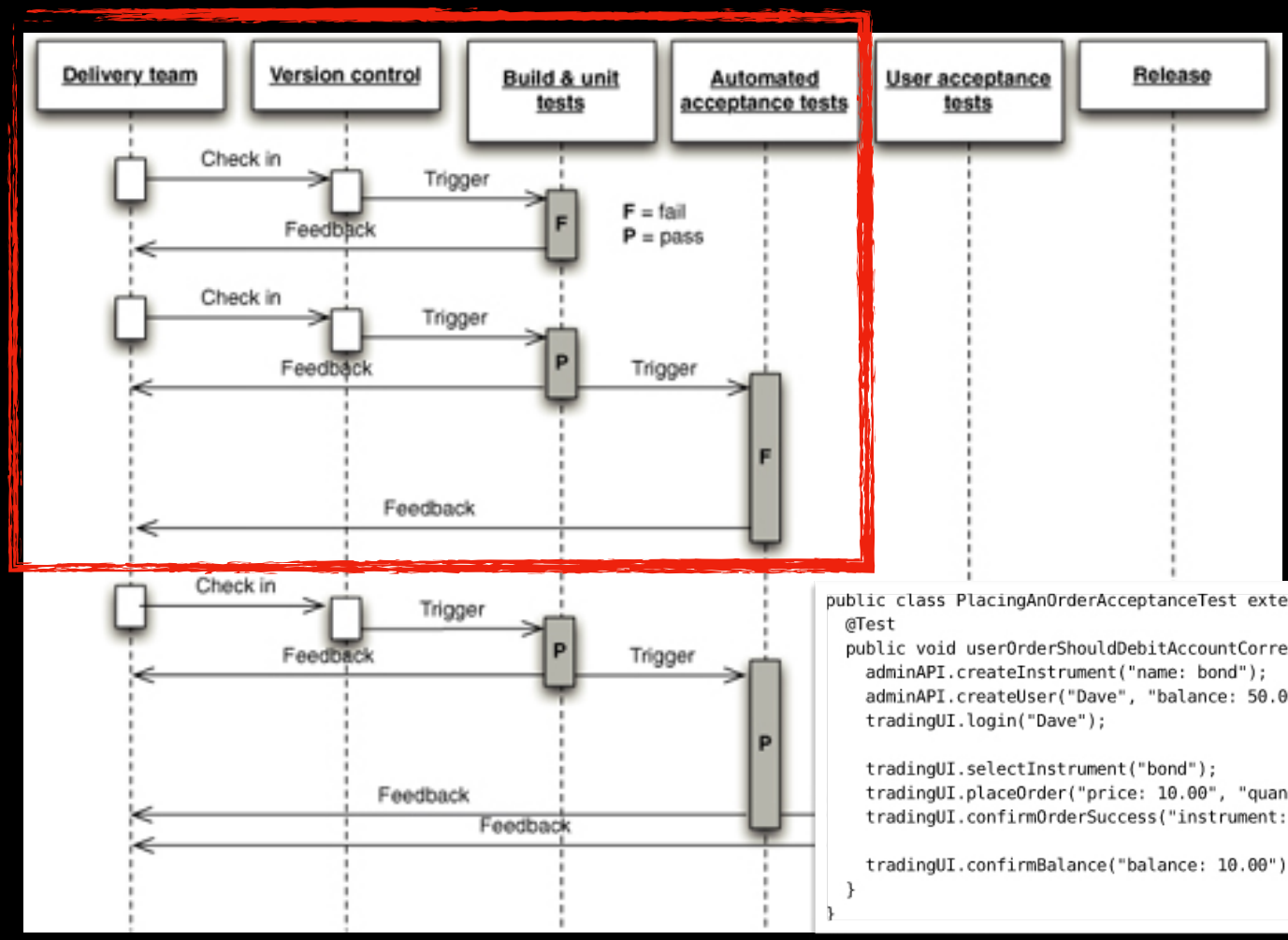


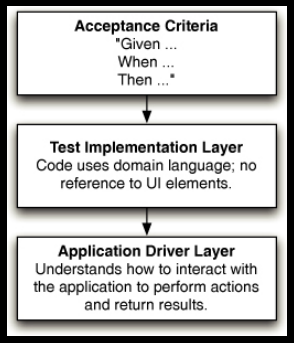
Continuous Delivery Reliable Software Releases through Build, Test, and Deployment Automation by Jez Humble and David Farley



detecting, as fast as possible, the most common failures that changes to the system may introduce, and notifying the developers so they can fix the problems quickly.

Continuous Delivery Reliable Software Releases through Build, Test, and Deployment Automation by Jez Humble and David Farley



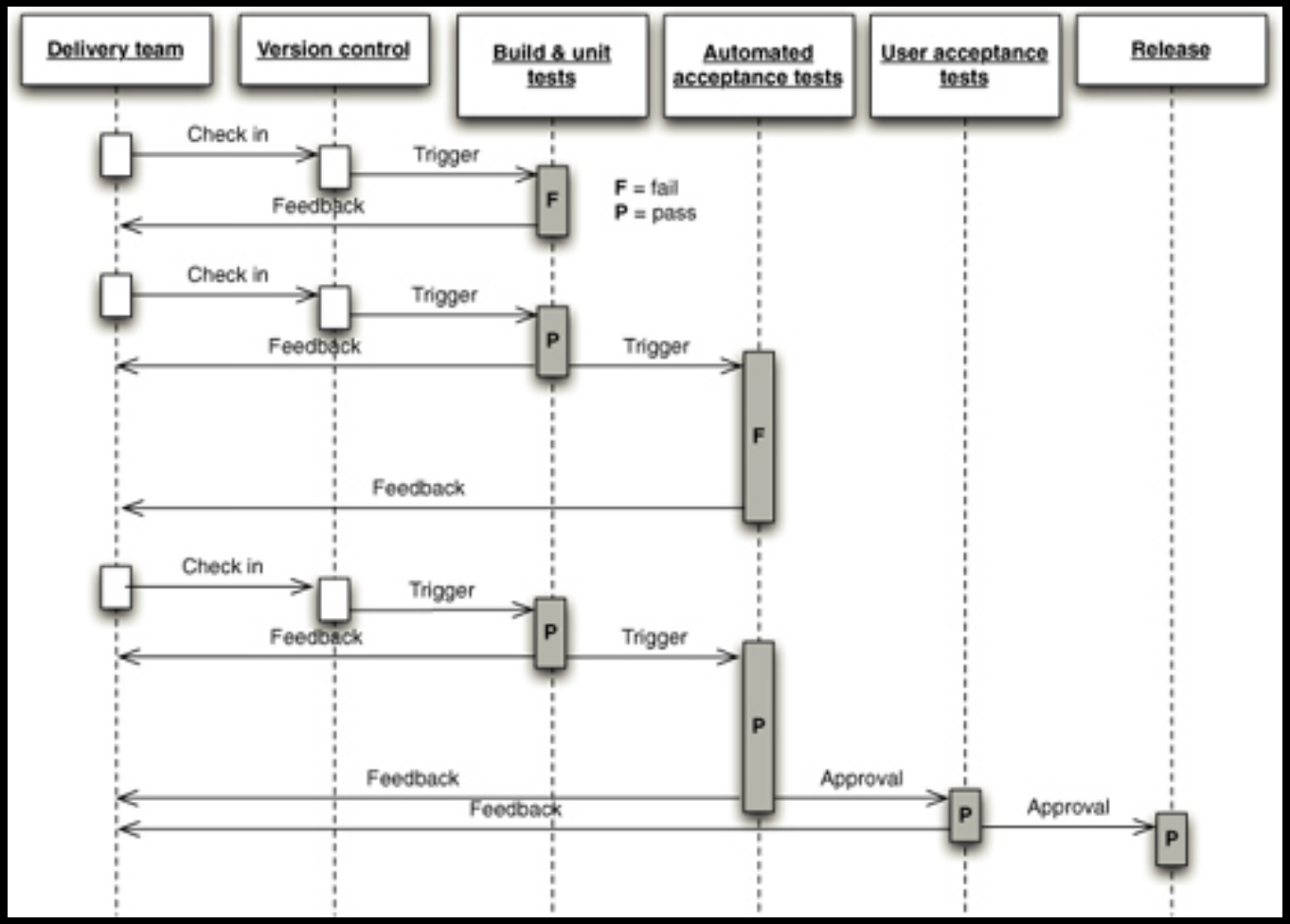



```
public class PlacingAnOrderAcceptanceTest extends DSLTestCase {
  @Test
  public void userOrderShouldDebitAccountCorrectly() {
    adminAPI.createInstrument("name: bond");
    adminAPI.createUser("Dave", "balance: 50.00");
    tradingUI.login("Dave");

    tradingUI.selectInstrument("bond");
    tradingUI.placeOrder("price: 10.00", "quantity: 4");
    tradingUI.confirmOrderSuccess("instrument: bond", "price: 10.00", "quantity: 4");

    tradingUI.confirmBalance("balance: 10.00");
  }
}
```

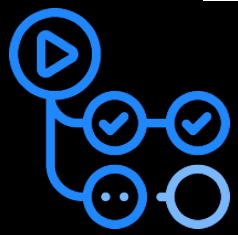
Continuous Delivery Reliable Software Releases through Build, Test, and Deployment Automation by Jez Humble and David Farley

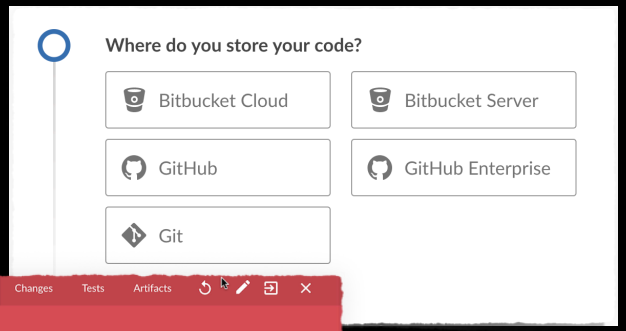


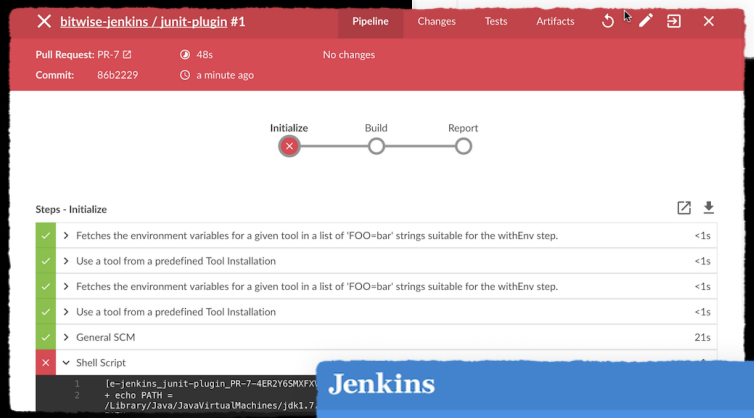
Continuous Delivery Reliable Software Releases through Build, Test, and Deployment Automation by Jez Humble and David Farley

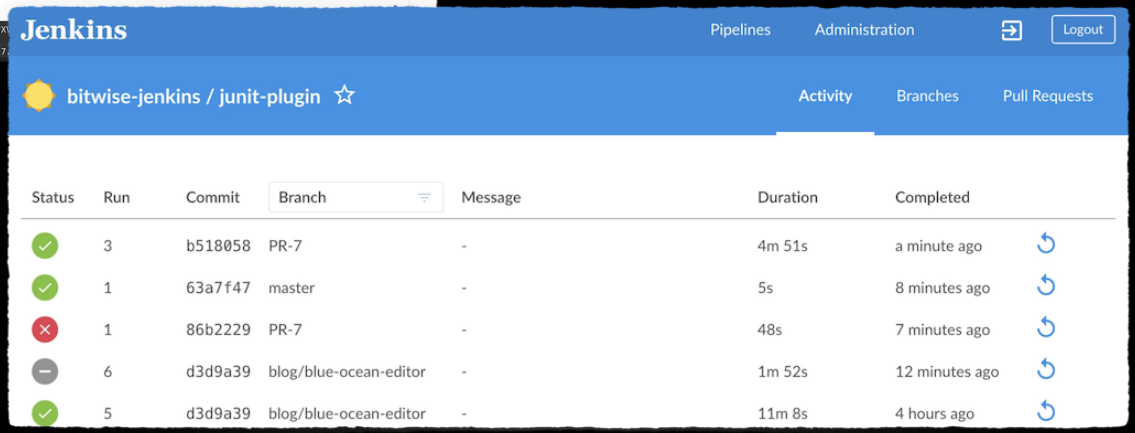
CI/CD Tools


go



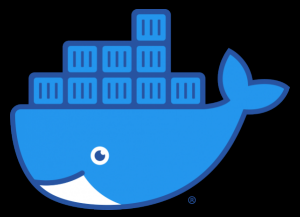

Where do you store your code?
Bitbucket Cloud
Bitbucket Server
GitHub
GitHub Enterprise
Git


bitwise-jenkins / junit-plugin #1
Pipeline
Changes
Tests
Artifacts
Pull Request: PR-7
48s
No changes
Commit: 86b2229
a minute ago
Initialize
Build
Report
Steps - Initialize
Fetches the environment variables for a given tool in a list of 'FOO=bar' strings suitable for the withEnv step.
<1s
Use a tool from a predefined Tool Installation
<1s
Fetches the environment variables for a given tool in a list of 'FOO=bar' strings suitable for the withEnv step.
<1s
Use a tool from a predefined Tool Installation
<1s
General SCM
21s
Shell Script


Jenkins
Pipelines
Administration
Logout
bitwise-jenkins / junit-plugin
Activity
Branches
Pull Requests
Status
Run
Commit
Branch
Message
Duration
Completed
3 b518058 PR-7 - 4m 51s a minute ago
1 63a7f47 master - 5s 8 minutes ago
1 86b2229 PR-7 - 48s 7 minutes ago
6 d3d9a39 blog/blue-ocean-editor - 1m 52s 12 minutes ago
5 d3d9a39 blog/blue-ocean-editor - 11m 8s 4 hours ago

ML Specific Concerns

Serve Model in Production



- **Embedded model:** treat the model artifact as a dependency that is built and packaged within the consuming application.
- **Model deployed as a separate service:** the model is wrapped in a service that can be deployed independently of the consuming applications.
- **Model published as data:** the model is treated and published independently, but the consuming application will ingest it as data at runtime.

ML Specific Concerns

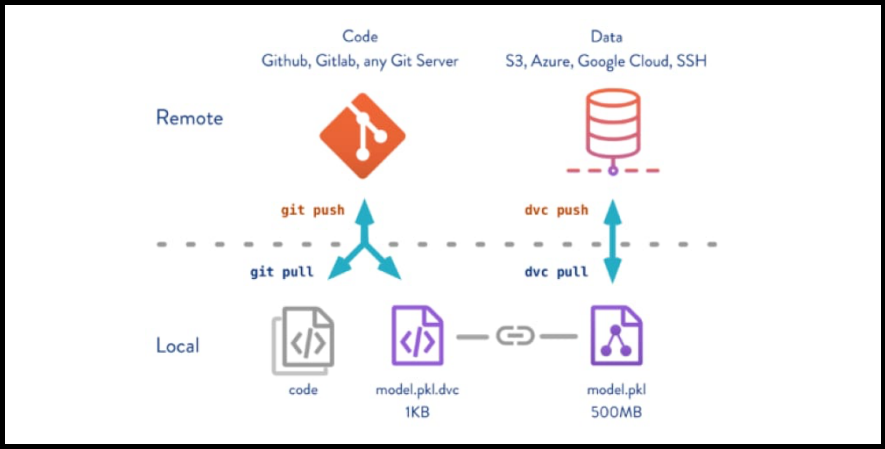
Deployment Options

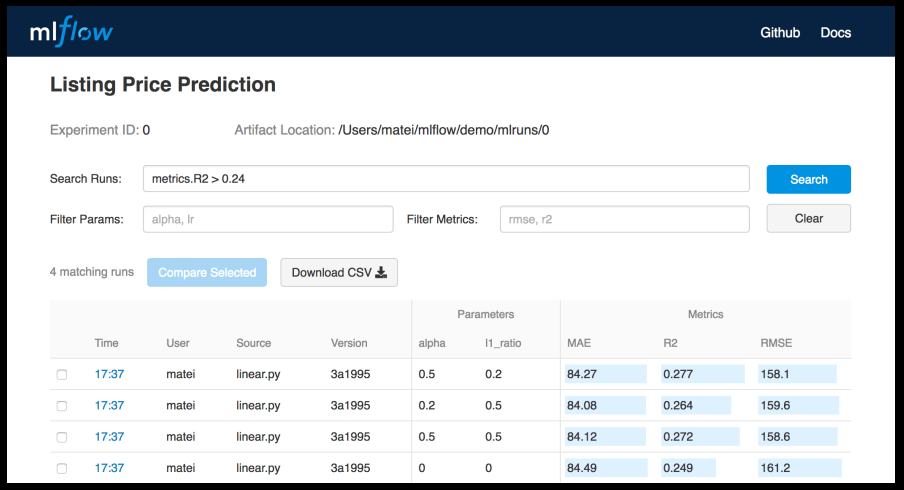
- Multiple models: release more than one model performing the same task
- Shadow models: send the same production traffic to gather data on how the shadow model performs before promoting it, similar to Blue-green deployment
- Competing models: trying multiple versions of the model in production, similar to an A/B test
- Online learning models: need to version not only the training data, but also the production data that will impact the model's performance.

ML Specific Concerns

Experiment Management

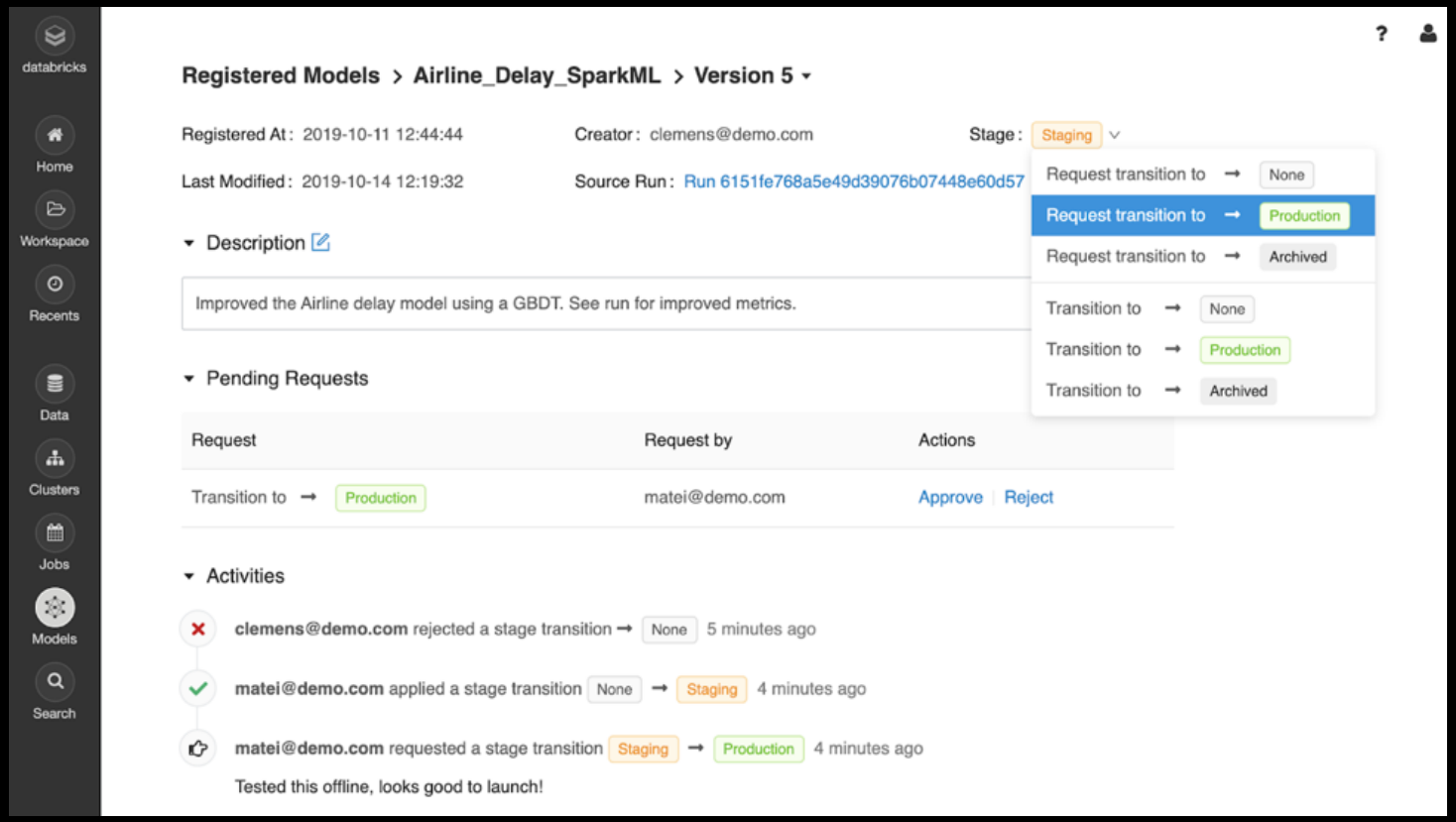
- Which configuration works best for which dataset?





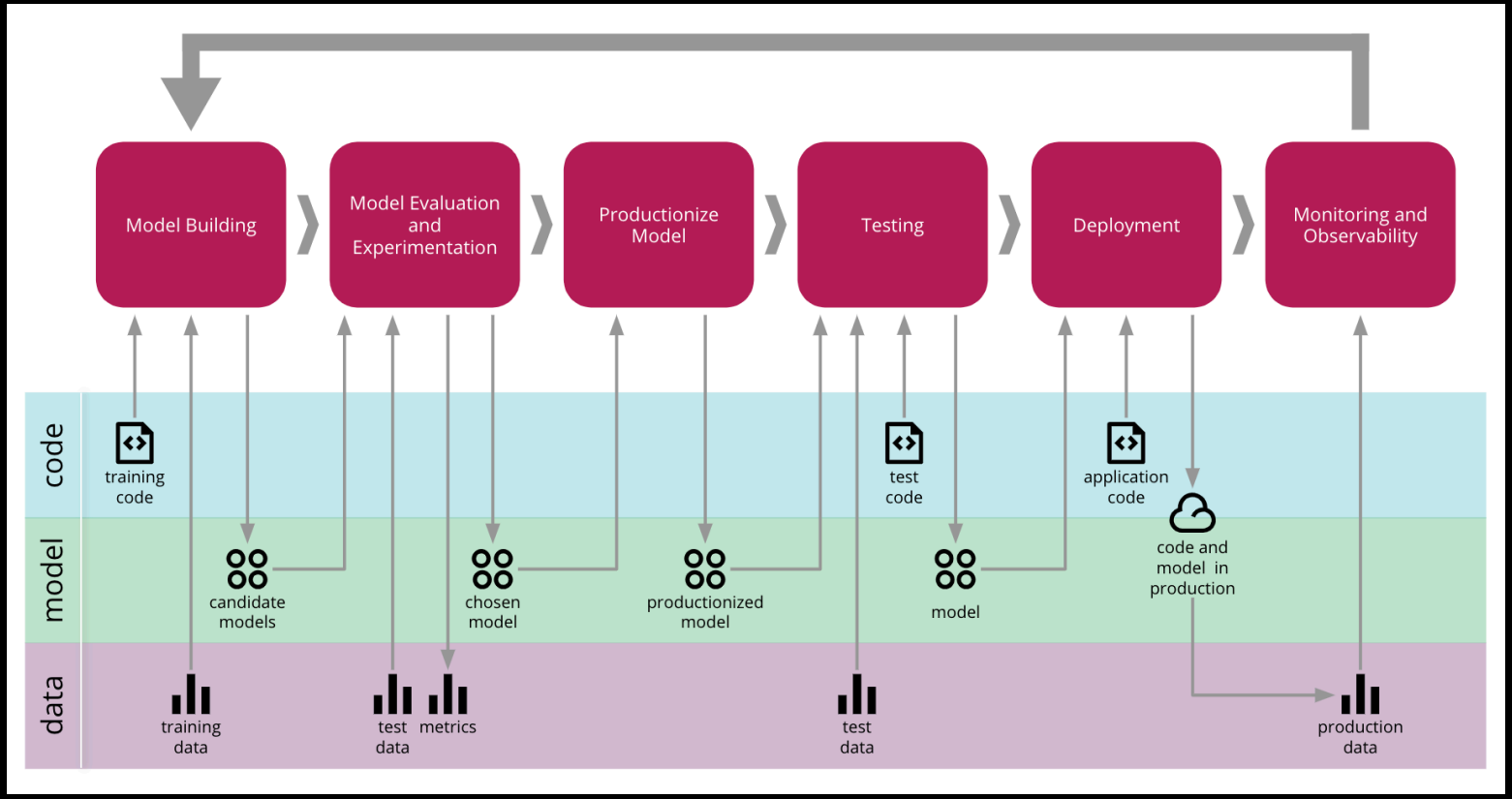
ML Specific Concerns

Register Models



https://databricks.com/blog/2019/10/17/introducing-the-mlflow-model-registry.html

Primary Objective



https://martinfowler.com/articles/cd4ml.html