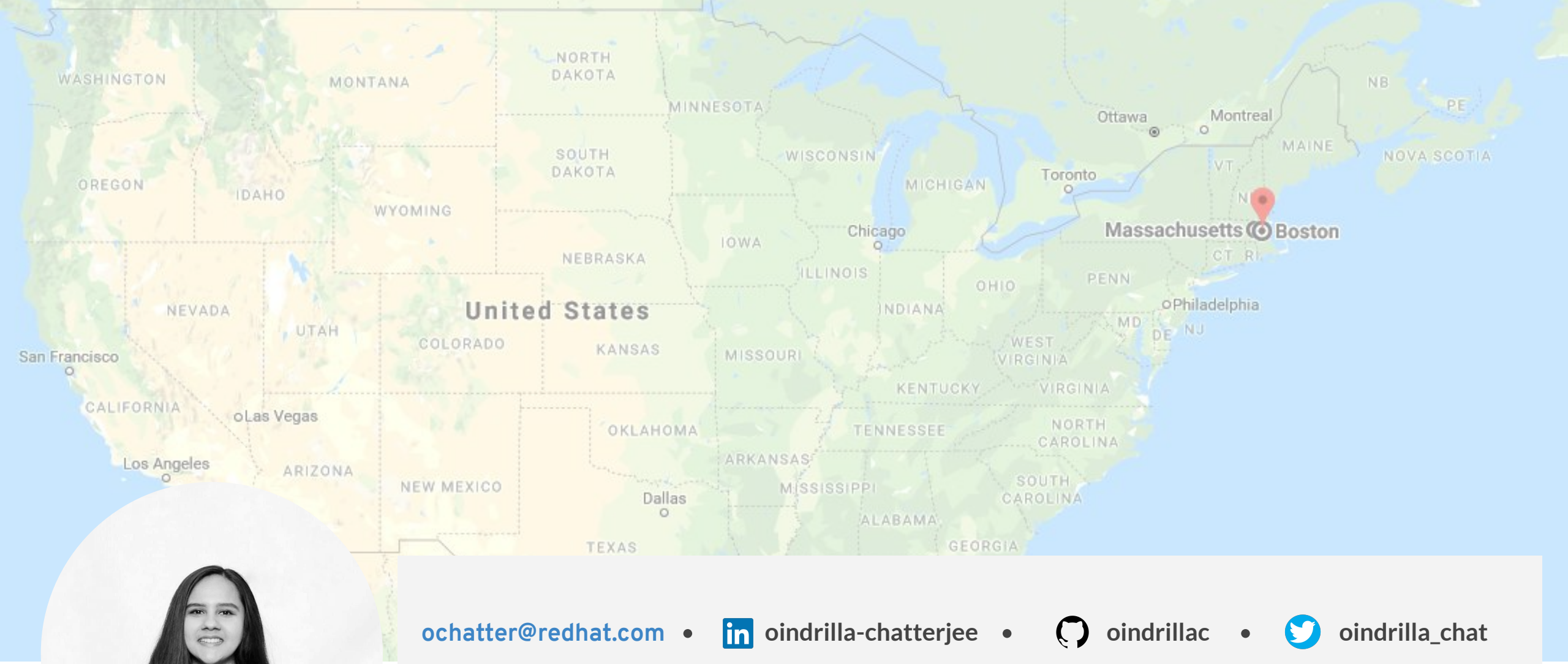
AI4CI

AI for Continuous Integration

Aakanksha Duggal, **Red Hat**
Oindrilla Chatterjee, **Red Hat**

aduggal@redhat.com • duggalaakanksha • aakankshaduggal • duggalaakanksha

Data Scientist at **Office of the CTO | AI Center of Excellence |** Red Hat **Boston, United States**



ochatter@redhat.com • oindrilla-chatterjee • oindrillac • oindrilla_chat

Data Scientist at **Office of the CTO** | **AI Center of Excellence** | Red Hat | Boston, United States

Overview

- What is AI4CI?
- Operate First Cloud
- Dive Into Open Data Sources
  - TestGrid
  - Prow
  - GitHub
- Demo
- Engage

Overview

➡ What is AI4CI?

Operate First Cloud

Dive Into Open Data Sources

- TestGrid
- Prow
- GitHub

Demo

Engage

AIOps

AI + DevOps

Artificial Intelligence for IT Operations

+

Dev & Ops : **DevOps**

AIOps

AI + DevOps

Artificial Intelligence for IT Operations

+

Dev & Ops : **DevOps**

CI/CD

Continuous Integration,
Continuous Delivery

Automating stages of app development

AIOps

AI + DevOps

+

CI/CD

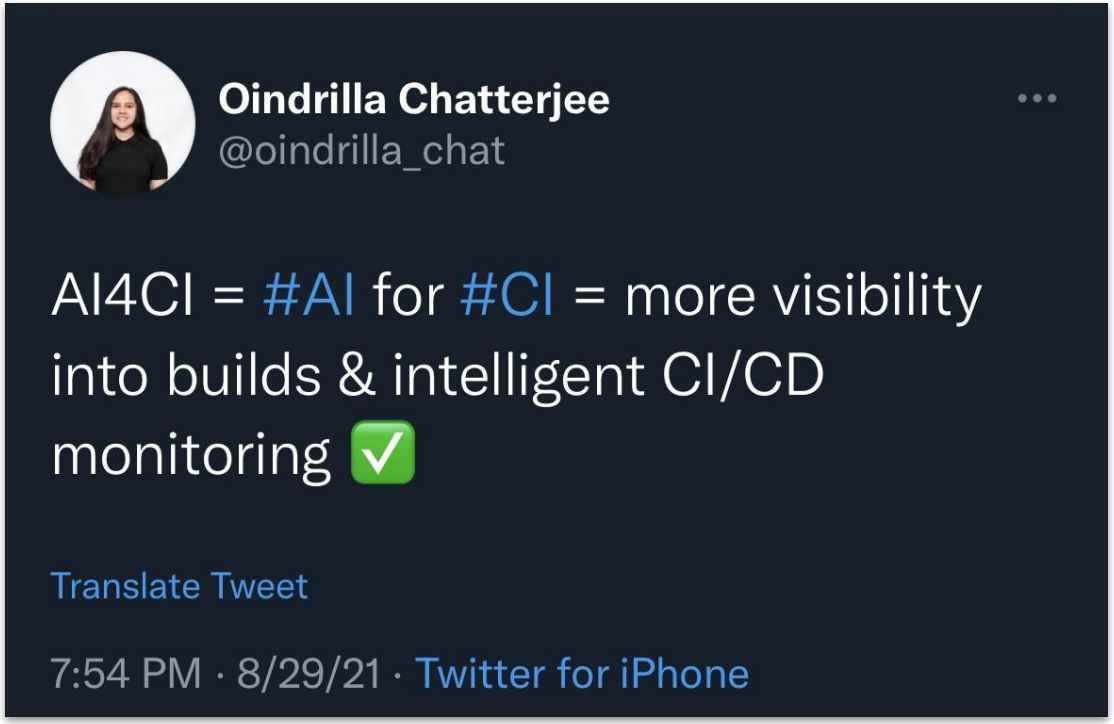
Continuous Integration,
Continuous Delivery

=

AI4CI

AI for Continuous Integration

What is AI4CI?



Our current initiatives

Our current initiatives

Collecting Data

Our current initiatives

Collecting Data

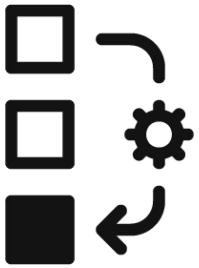
Metrics

Our current initiatives

Collecting Data → Metrics → AI/ ML

Our current initiatives

Collecting Data → Metrics → AI/ ML → E2E Workflow



Overview

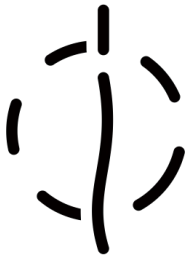
- What is AI4CI?
- ➡ Operate First Cloud
- Dive Into Open Data Sources
  - TestGrid
  - Prow
  - GitHub
- Demo
- Engage

Operate First

"A concept to incorporate operational experience into software projects."

www.operate-first.cloud

Operate First

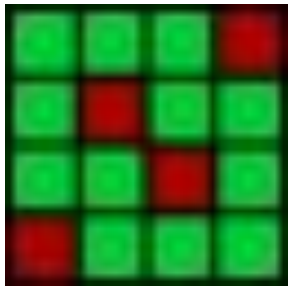


Overview

- What is AI4CI?
- Operate First Cloud
- ➡ Dive Into Open Data Sources
  - TestGrid
  - Prow
  - GitHub
- Demo
- Engage

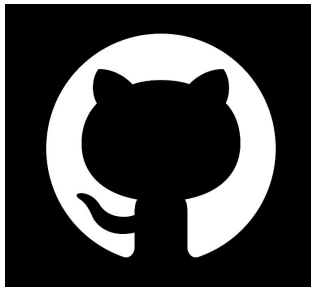
Open Data Sources

TestGrid



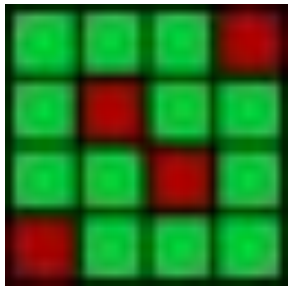
Prow

Github



Open Data Sources

TestGrid



Prow

Github

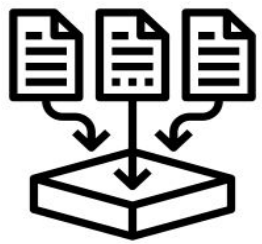
TestGrid

prowjob_name: e2e-kops-aws-distro-imagecentos7 prowjob_config_url: https://git.k8s.io/test-infra/config/jobs/kubernetes/kops/kops-periodics-distros.yaml
08-12 09:51 PDT @1425862280918077440 -- 05-14 17:50 PDT @1393368069975314432; Served 223344 results in 2.74 seconds
About
Size
Options
Graph
Local Time: OFF
Display Clustered Failures List
08-12
08-11
08-10
09:51 PDT
01:51 PDT
17:51 PDT
09:50 PDT
01:50 PDT
17:50 PDT
09:50 PDT
1425862280911425741482621425620686551425499817761425379020671425258223521425137369
1dae5cda1
adc21d1dd
a4af85bd9
0e01893de
6eab49e05
4a0fa2d10
c2ff89a4a
Show 29 stale tests (no results in last 10+ runs)
Overall Changes
kubetest2.Test Changes
Kubernetes e2e suite.[sig-apps] StatefulSet Basic StatefulSet functionality [StatefulSetBasic] should perform rolling updates and
Kubernetes e2e suite.[sig-storage] In-tree Volumes [Driver: local][LocalVolumeType: dir-bindmounted] [Testpattern: Pre-provis
Kubernetes e2e suite.[sig-storage] In-tree Volumes [Driver: local][LocalVolumeType: tmpfs] [Testpattern: Pre-provisioned PV (
Kubernetes e2e suite.[sig-storage] In-tree Volumes [Driver: local][LocalVolumeType: dir-bindmounted] [Testpattern: Pre-provis
Kubernetes e2e suite.[sig-storage] In-tree Volumes [Driver: local][LocalVolumeType: dir-bindmounted] [Testpattern: Pre-provis
Kubernetes e2e suite.[sig-storage] In-tree Volumes [Driver: local][LocalVolumeType: dir-bindmounted] [Testpattern: Pre-provis
Kubernetes e2e suite.[sig-storage] In-tree Volumes [Driver: local][LocalVolumeType: tmpfs] [Testpattern: Pre-provisioned PV (f
Kubernetes e2e suite.[sig-apps] StatefulSet Basic StatefulSet functionality [StatefulSetBasic] should provide basic identity
Kubernetes e2e suite.[sig-network] Networking Granular Checks: Pods should function for node-pod communication: udp [Linux
Kubernetes e2e suite.[sig-network] Services should have session affinity timeout work for NodePort service [LinuxOnly] [Conforr
Kubernetes e2e suite.[sig-node] Pods Extended Pod Container Status should never report success for a pending container
Kubernetes e2e suite.[sig-storage] In-tree Volumes [Driver: aws] [Testpattern: Dynamic PV (default fs)] subPath should suppor
Kubernetes e2e suite.[sig-storage] In-tree Volumes [Driver: aws] [Testpattern: Dynamic PV (delayed binding)] topology should
Kubernetes e2e suite.[sig-storage] In-tree Volumes [Driver: aws] [Testpattern: Dynamic PV (ext4)] volumes should store data
Kubernetes e2e suite.[sig-storage] In-tree Volumes [Driver: local][LocalVolumeType: blockfs] [Testpattern: Pre-provisioned PV
Kubernetes e2e suite.[sig-storage] In-tree Volumes [Driver: local][LocalVolumeType: blockfs] [Testpattern: Pre-provisioned PV
Kubernetes e2e suite.[sig-storage] In-tree Volumes [Driver: local][LocalVolumeType: tmpfs] [Testpattern: Pre-provisioned PV (
Kubernetes e2e suite.[sig-storage] In-tree Volumes [Driver: local][LocalVolumeType: dir-link-bindmounted] [Testpattern: Pre-p
Kubernetes e2e suite.[sig-storage] In-tree Volumes [Driver: local][LocalVolumeType: tmpfs] [Testpattern: Pre-provisioned PV (
Kubernetes e2e suite.[sig-storage] In-tree Volumes [Driver: local][LocalVolumeType: tmpfs] [Testpattern: Pre-provisioned PV (
Kubernetes e2e suite.[sig-storage] PersistentVolumes-local [Volume type: dir-link-bindmounted] Two pods mounting a local volu
Kubernetes e2e suite.[sig-apps] Deployment should not disrupt a cloud load-balancer's connectivity during rollout
Kubernetes e2e suite.[sig-apps] StatefulSet Basic StatefulSet functionality [StatefulSetBasic] should not deadlock when a pod's
Kubernetes e2e suite.[sig-cli] Kubectl Port forwarding With a server listening on 0.0.0.0 should support forwarding over websock
Kubernetes e2e suite.[sig-storage] CSI mock volume CSI Volume expansion should not expand volume if resizingOnDriver=off,
Kubernetes e2e suite.[sig-storage] ConfigMap updates should be reflected in volume [NodeConformance] [Conformance]
Kubernetes e2e suite.[sig-storage] Ephemeralstorage When pod refers to non-existent ephemeral storage should allow deletion

From **testgrid.k8s.io** to **dashboards**

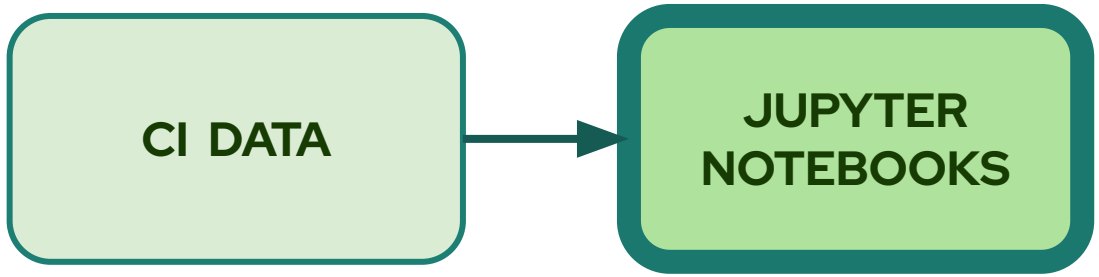
How?

CI DATA



From **testgrid.k8s.io** to **dashboards**

How?

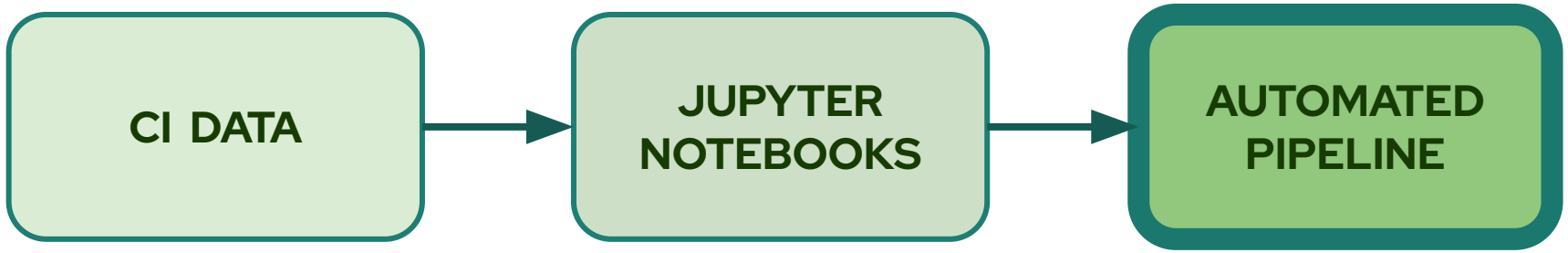


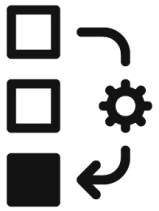
Metrics

ML/ AI

From **testgrid.k8s.io** to **dashboards**

How?

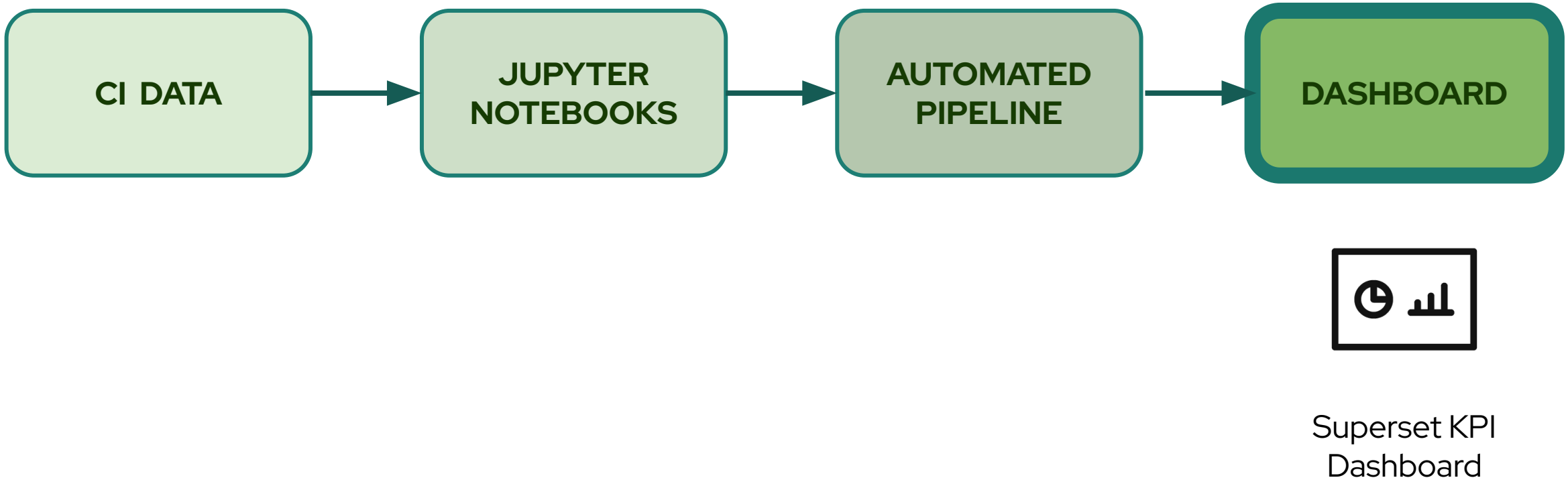




Elyra and Kubeflow Pipelines

From **testgrid.k8s.io** to **dashboards**

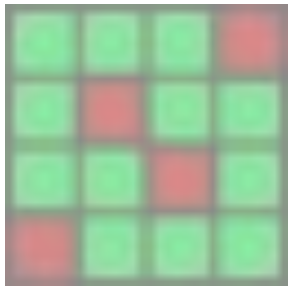
How?



Superset KPI Dashboard

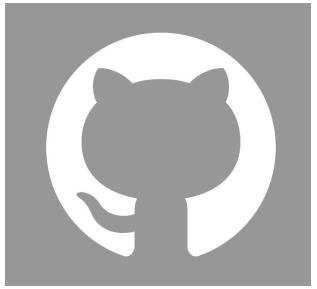
Open Data Sources

TestGrid

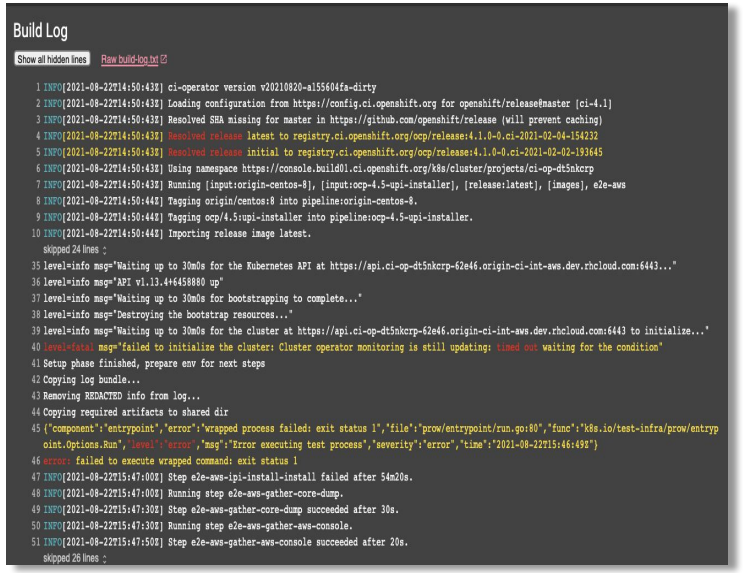


Prow

Github

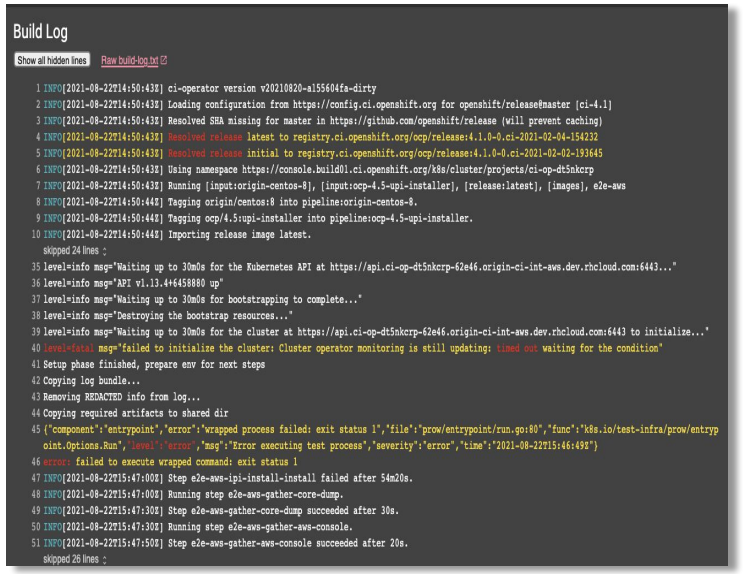


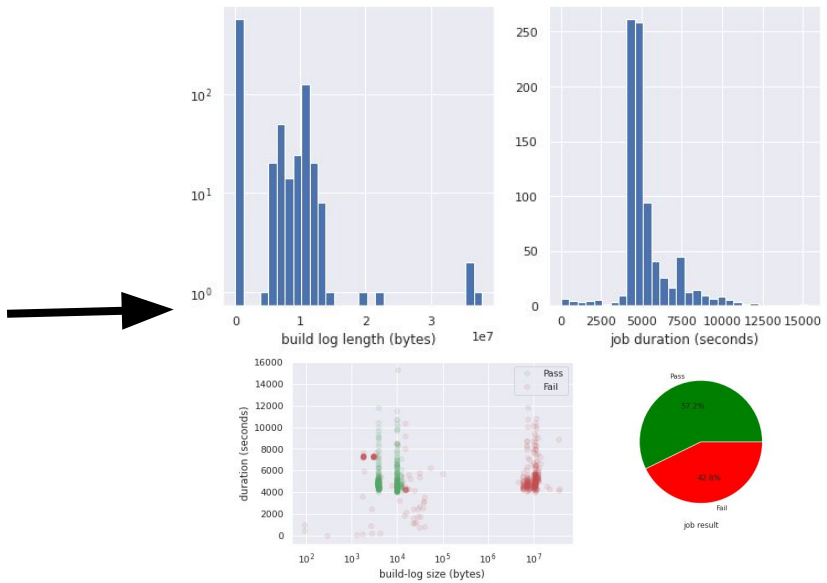
Prow



Build Logs

Prow

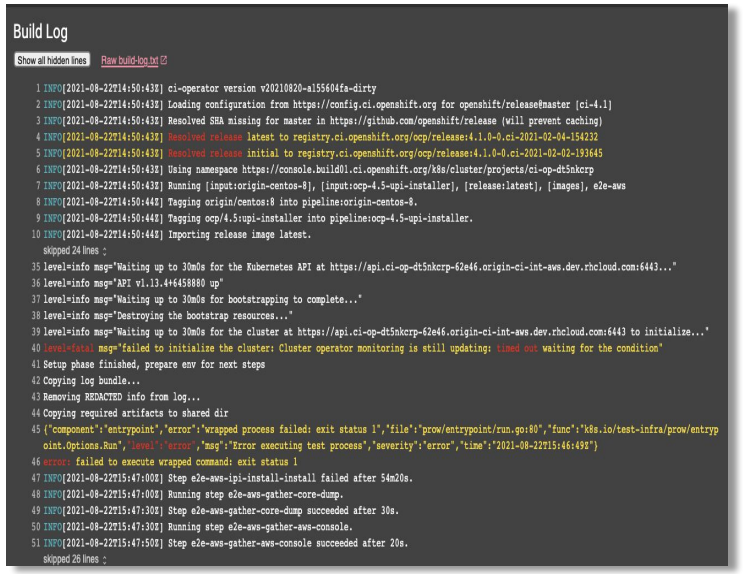


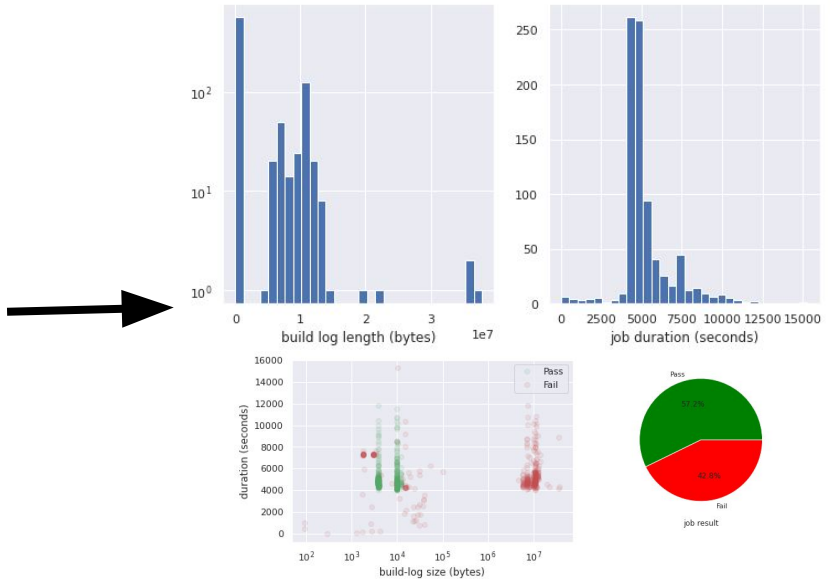


Build Logs

Log Analyses

Prow



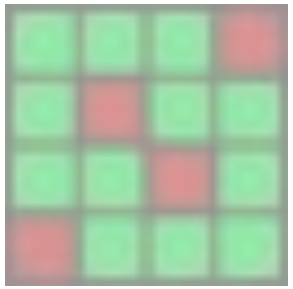


- successful builds without provider obfuscation
- failures in CI infrastructure (e.g. lack of capacity)
- Build failures due to some tests
- failures during preparation steps, before tests can be executed

Build Logs | **Log Analyses** | **Log Classification**

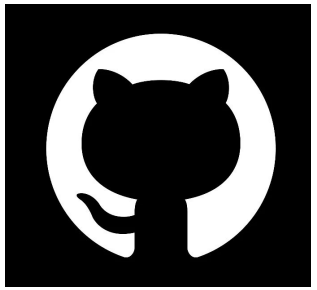
Open Data Sources

TestGrid

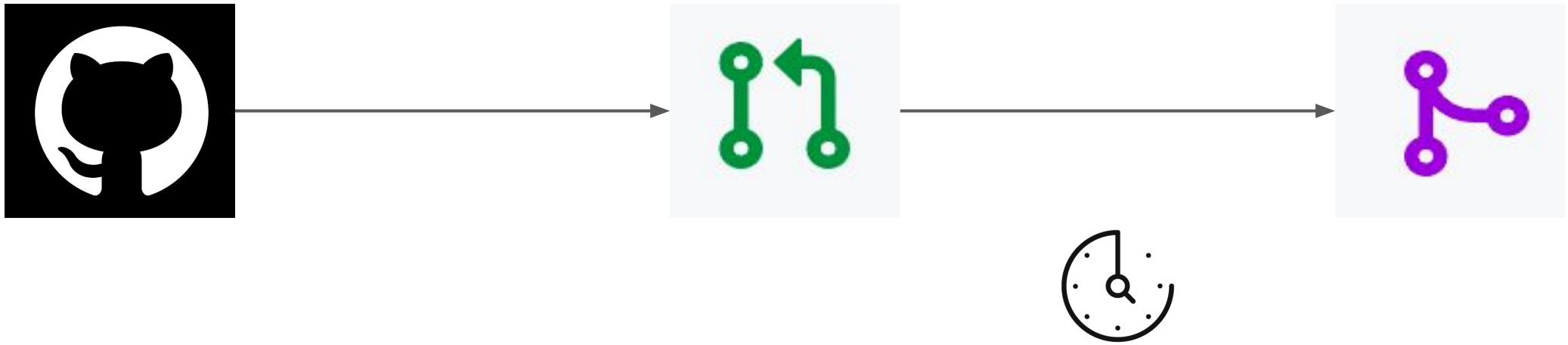


Prow

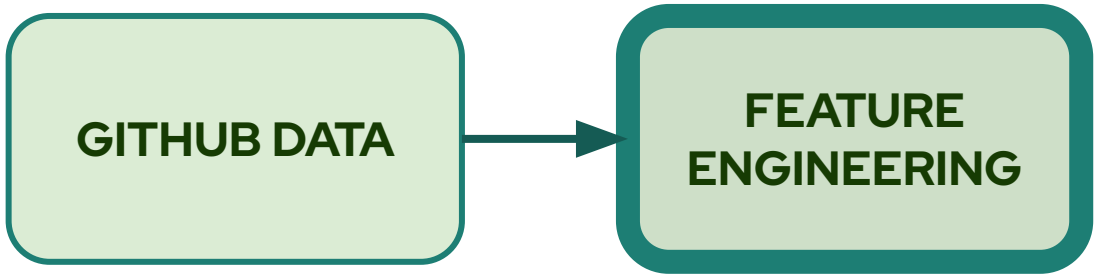
Github



Time to Merge of a PR



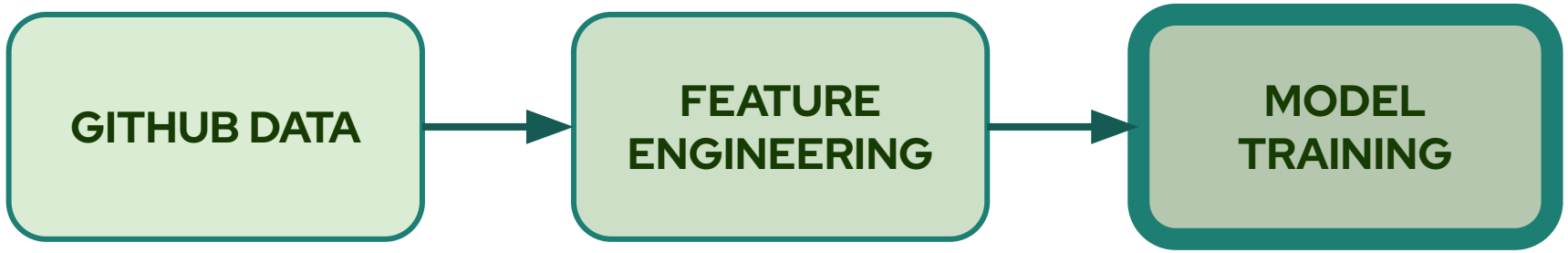
GITHUB DATA



size of the PR,

types of files in the PR

description of the PR

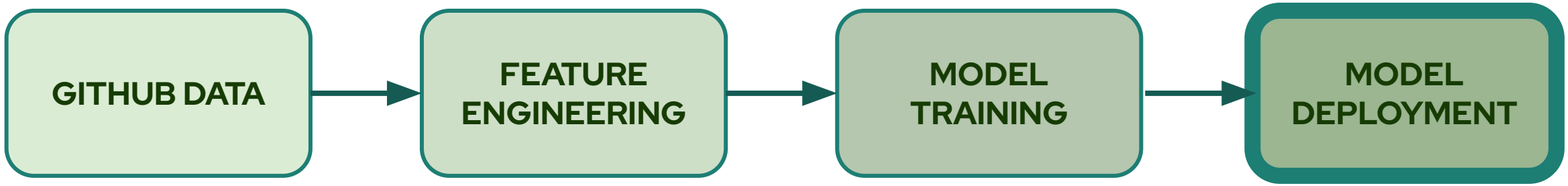


Naive Bayes

XGBoost

Random Forest

SVM


GITHUB DATA
FEATURE ENGINEERING
MODEL TRAINING
MODEL DEPLOYMENT

SELDON

Overview

What is AI4CI?

Operate First Cloud

Dive Into Open Data Sources

- TestGrid
- Prow
- GitHub

➡ Demo

Engage

Overview

- What is AI4CI?
- Operate First Cloud
- Dive Into Open Data Sources
  - TestGrid
  - Prow
  - GitHub
- Demo
- ➡ Engage

Engage

Get Started

https://www.operate-first.cloud/data-science/ai4ci/docs/get-started.md

Open Data Sources	Notebooks	Dashboards	Model Endpoints	Automated Workflows	Video Playlist

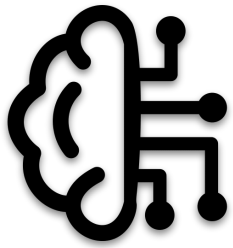
Engage

Contribute

https://www.operate-first.cloud/data-science/ai4ci/docs/how-to-contribute.md

KPI Metric

ML Analysis



Thank you, questions?

https://github.com/aicoe-aiops/ocp-ci-analysis

@OperateFirst

https://tinyurl.com/aiforci

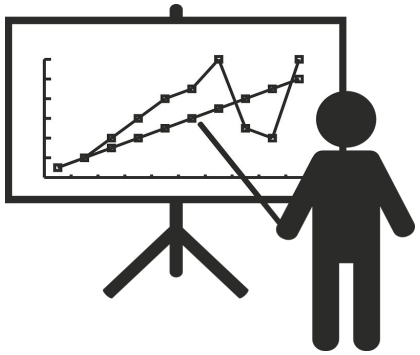
https://www.operate-first.cloud/data-science/ai4ci/

THE LINUX FOUNDATION

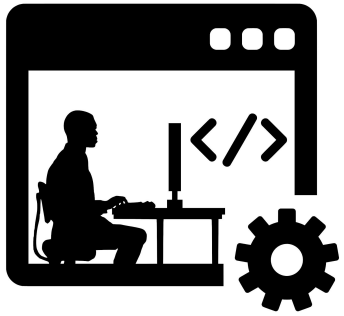
OPEN SOURCE SUMMIT

Engage!

Data Science Enthusiast



Developer



Kubernetes Contributor