AIOps for community health

Project AI4CI

AI4CI: Open Source AIOps toolkit

Problem

- Need for **AIOps** - Automated monitoring, analysis, alerting with Ops (CI/CD, development processes)

- **Open Source data** originating from real world production systems is a rarity for public datasets.

- Lack of AI driven metrics for open source community health.

Opportunity

- **Open operations data** made available by running open source software and applications in production.

- Data includes CI/CD data, code, telemetry, logs, operational dashboards.

- Eg: Kubernetes testing infrastructure, Fedora make their testing data available open source.

Solution

- Collection of intelligent and open source **data science tools** to collect and analyze the CI/CD data.

- **AIOps models** like Github time-to-merge service, optimal stopping time prediction, build log classifier

- KPI and Metric dashboards

- Goal is to foster an open source AIOps community with open ops data, AI tools and services.

AI4CI supports CI/CD and software dev processes

What is AI4CI?

Collection of **Open Source AIOps tools** including scripts, notebooks, pipelines, dashboards and data sources.
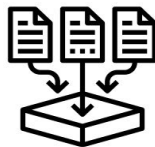
Data collection
Collection of open operations data from Kubernetes testing platforms eg: Testgrid, Github, and Prow.

Metrics
Collects metrics and **KPIs** and visualization dashboards.
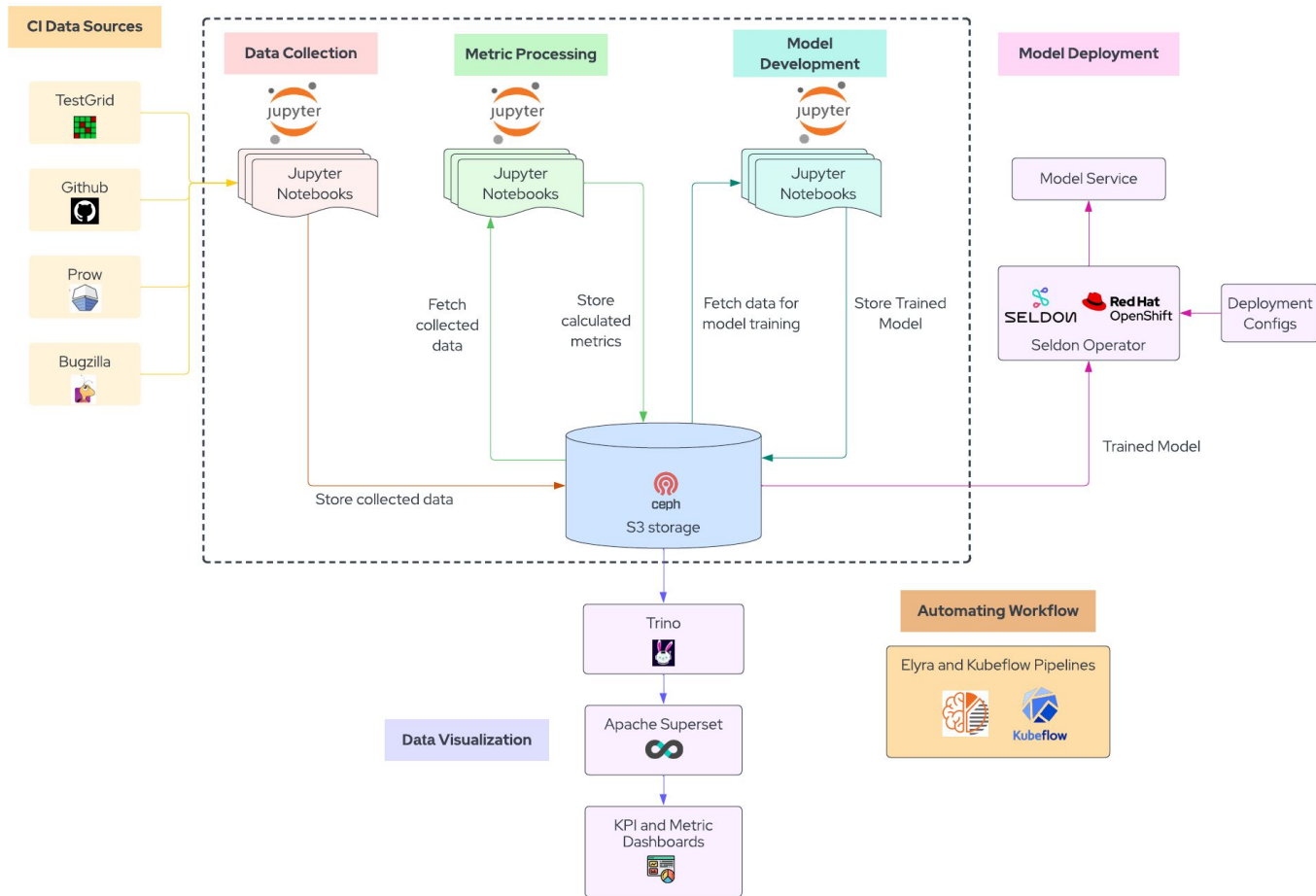
ML Services
ML services which can support CI/CD processes.
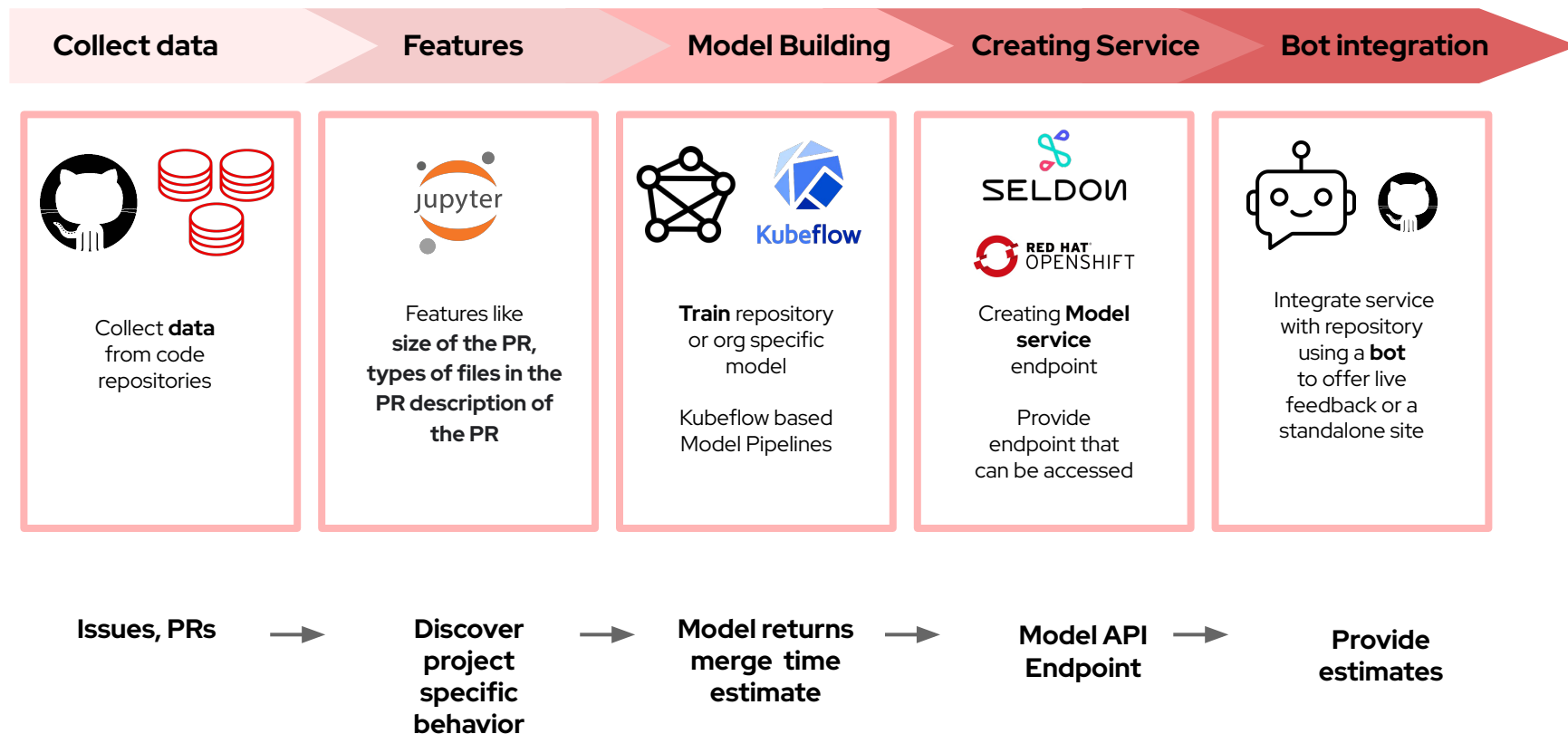
Open source AIOps template
Resource for open source AIOps communities (notebooks, scripts, automated ML pipelines, dashboards, services tools)

Architecture

Current workflow: Github time to merge prediction service

Collect data	Features	Model Building	Creating Service	Bot integration

Collect **data** from code repositories

Features like **size of the PR, types of files in the PR description of the PR**

Train repository or org specific model

Kubeflow based Model Pipelines

Creating **Model service** endpoint

Provide endpoint that can be accessed

Integrate service with repository using a **bot** to offer live feedback or a standalone site

Issues, PRs → **Discover project specific behavior** → **Model returns merge time estimate** → **Model API Endpoint** → **Provide estimates**

Time to merge prediction service for community health

- Identify **bottlenecks** in development process

- Leverage the rich **historical data** of consisting of Issues, Commits, PRs

- Give **new contributors** of an estimate of when their PR will be reacted upon

- Most importantly – develop an **AI driven mindset** for community health

Optimal Stopping Point prediction service for saving resources

- Leverage the rich **historical data** of tests and build runs

- Finding an optimal stopping point for long running tests eating up resources

- Boon for developers and managers for saving resources and time

Resources

https://github.com/aicoe-aiops/ocp-ci-analysis/blob/master/docs/get-started.md

Open Data Sources	Notebooks	Dashboards	Model Endpoints	Automated Workflows	Video Playlist

https://tinyurl.com/aiforci