Open source software MLOps platform (OSS MLOps)

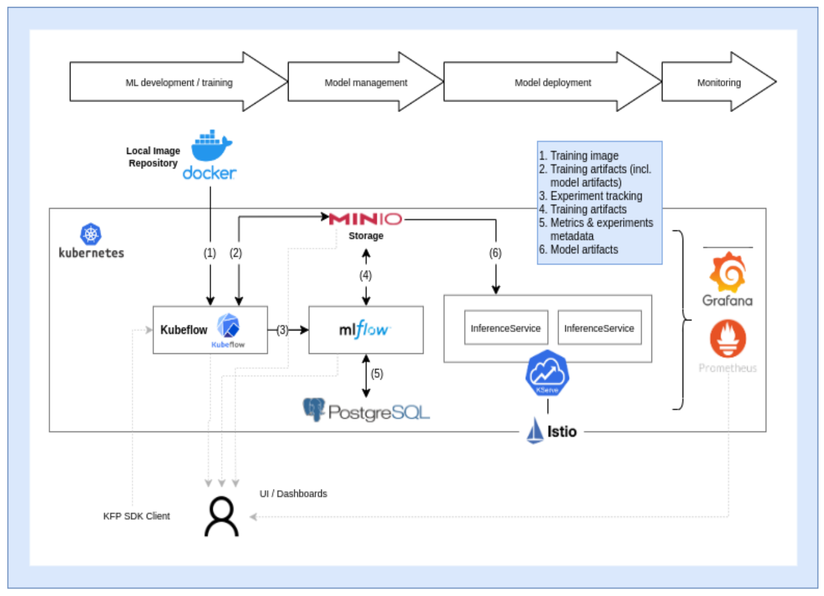
Jukka Remes[3], Niila Siilasjoki[2], Joaquín Rives[1], Harry Souris[1], Mikko Raatikainen[2], Dennis Muiruri[2], Johanna Malinen[1], Dorian Knoblauch[4]
(1) Silo AI, (2) University of Helsinki, (3) Haaga-Helia UAS, (4) Fraunhofer FOKUS

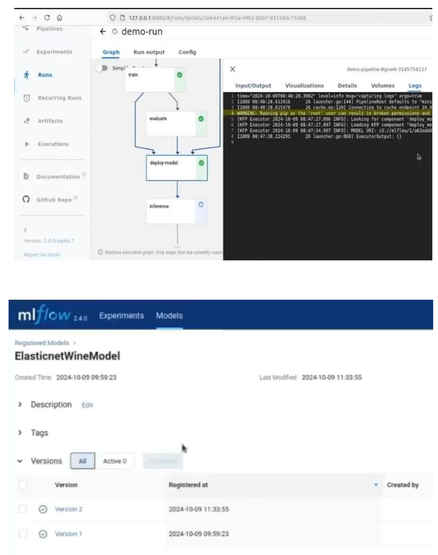
Overview:
A suite for streamlining AI/ML R&D from experimentation to deployment

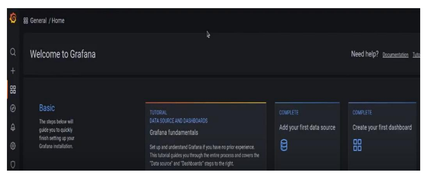
OSS MLOps platform is a Kubernetes in Docker (KinD) based platform for ML development
- Consists of industry-level and well-backed open-source tools
- Easily customizable and extensible

Requires only the following as minimum setup:
- PC with Git and Python
- Computer with Git, Docker and
  - 8 CPU
  - 16 GB RAM
  - 20 GB disk

Included services provide APIs and CLI but also web-UIs:
- **Kubeflow** for viewing and debugging e.g. pipeline executions
- **MLflow** for experiment metadata and model versioning
- **Prometheus** and **Grafana** for monitoring deployed models







Unified AI development, deployment and monitoring experience across environments

1. **AI professionals' own dedicated computer** (local, lighter deployment)
   - work laptop/desktop computer/virtual machine
   - AI development and testing for individuals (developers, researchers, students)

2. **AI teams and organizations common environments** (full deployment)
   - same workflows possible as on everybody's own computer
   - usage in cloud (Google Cloud k8s, Azure and AWS VMs, others)
   - usage in super-computing/HPC environments (Finnish IT Center of Science - CSC)

Join users and/or developers:

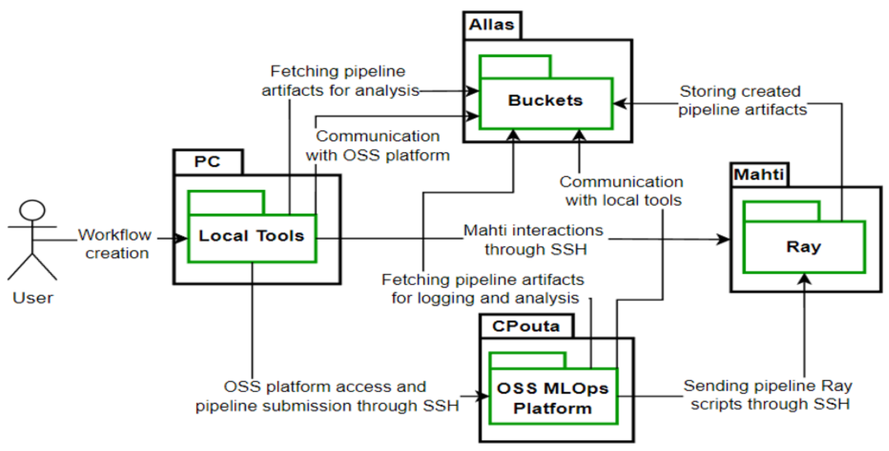
https://github.com/OSS-MLOPS-PLATFORM/oss-mlops-platform

Architectural overview of cloud-HPC integrated OSS MLOps platform at CSC

OSS MLOps platform in the CPouta virtual machine service (OpenStack)

ML pipelines use Allas (s3) object storage and Mahti supercomputer for GPUs

Customizable and scalable ML workflows using Ray in Mahti through OSS MLOps



You can use the OSS MLOps platform freely without fees in your organization.

Contributions are welcome.

Primary communication channel for issues, requests and feedback is Slack:

https://oss-mlops-platform.slack.com

New tooling coming in near future:
Streamlining platform use through git and CI/CD (software DevOps in GitHub)

Infra people
- Set up the platforms that are shared by teams
- Use the tooling to create configuration git-repository

ML developers (students, researchers, data scientists, ML engineers)
- **Use the tooling to create git-repositories** for their different
  - own ML projects
  - shared/common ML projects
- **Focus on creating ML code and configuration** (data processing, ML training, model deployments, pipelines)
- **Run pipelines simply by making git commits** to selected branches: Github Actions CI/CD automatically launches the training experiments or model deployments
  - to your local platform instance running on your own computer
  - to shared platforms (cloud, HPC/CSC)

Most of the platform details (web-API end-points, credentials etc.) can be hidden from regular ML developer making the use of platform simple