Use AI in DevSecOps

Harshad Reddy Nalla
Senior Software Engineer

Red Hat

Introduction

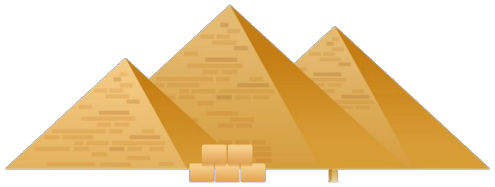
- Self introduction
- Introduction what we are going to understand from the presentation

Red Hat

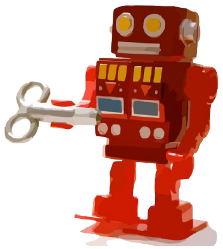
Introduction

- Sr. Software Engineer, AI Centre of Excellence, Red Hat Boston.
- Primarily part of AI DevSecOps team,
  working on Project Thoth: AI Stacks recommendation system
- Currently focusing on DevOps and Thoth bots

Red Hat

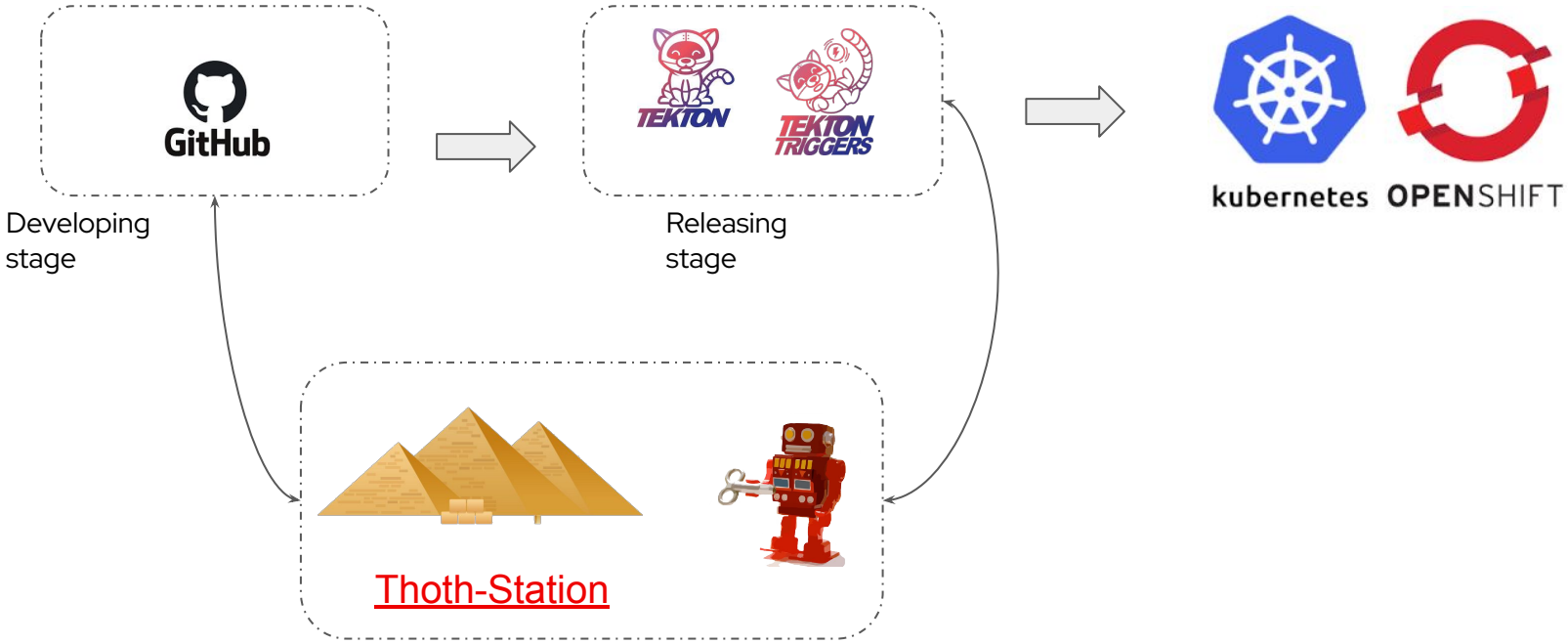


Thoth-Station

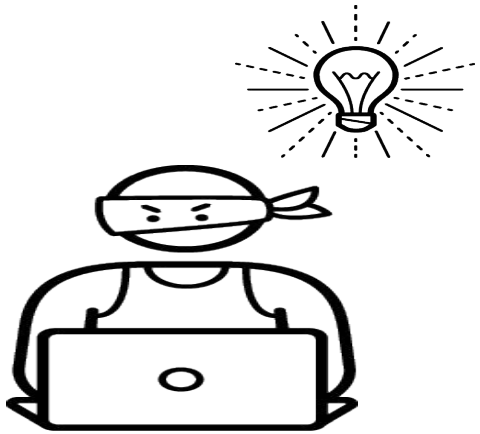


Goal

- Share the learning about the security in pipeline using AI tools



What we'll discuss today



- Why care about security in the DevOps ?
- Project Thoth
- Integration in CI/CD
- Secure Supply Chain
- More advancements

Need of Security

- Explain about the security in the code side
- Issue to be faced

Red Hat

Why care about security in the DevOps?

Developers make sure that source code is security without any data leaks, however it significantly harder to have this kinda check on the imported packages.

- provide critical software dependencies in a secure infrastructure
- pin down dependencies based on deep developer knowledge
- provide interactive guidance services to data scientists
- CI/CD pipelines gating based on guidance services

Key Challenges in Creating Secure Software Supply Chains

- **Extensive release of dependent packages**: Latest is not always the greatest. Dependent package being released often, put more work on the security side
- **Inconsistent industry adoption**: varying approaches to identity & trust models across popular language & packaging frameworks; minimal traction within cloud-native ecosystem to date
- **Integration challenges**: Significant number of technologies in use for CI/CD, application frameworks, hybrid infrastructure.
- **Increased complexity**: shift-left puts more responsibility for security on the developer, where skills gaps & fragmented ecosystems impede progress

Thoth Introduction

- Introduction to Thoth
- Explain persona
- Explain little bit on adviser, thamos
- Triggers

Red Hat

Project Thoth

- Help developers in the selection of dependencies for their applications depending on their requirements
- Deliver optimized images for your applications
- Use bots to automate mundane work to offload humans work

Personae we are targeting

Data Scientist

responsible for releasing a machine-learning model

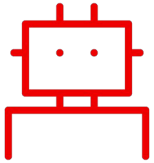
(AI) Developer

maintains an application wrapping the model

DevOps Engineer

operates the model service

Persona we are providing



(AI) DevSecOps Cyborg

maintaining application dependencies
and deployments
based on the knowledge we leaned

How Thoth can help developers?

Create Machine-generated and apply Machine-learned knowledge via well-known DevSecOps tools so that teams win a junior developer

Knowledge generation and acquisition/learning

A system to aggregating knowledge across communities and customers to derive aggregate value automatically.

Build and run time analysis for application stacks

Knowledge application (aka Services)

A project to introduce Analytics / AI-based automation into CI/CD processes

Provide guidance via Thoth Services (OpenShift Pipelines, GitHub apps, etc.)

Optimize build artifacts, and deployment configurations

Provide optimized AI Stacks, like TensorFlow

Pipelines to deploy optimized

and customized application configuration

How do we use this knowledge?

- Recommender system is called **Adviser** in Thoth.
- It uses Reinforcement Learning (RL).

Project Thoth

Check the video

What Thoth stores within it’s knowledge graph

Application Stack related

Build time and runtime environment

Dependencies graphs

Performances and Security Indicators

Software Package

Application Binary Interfaces (ABI)

Security (CVE, Prescriptions, analyzers)

Source Code Meta Information

Project features (TTR, TTCI, etc,..) from different software development platform

Thoth Recommendation types

- Latest
- Stable
- Security
- Performance

```yaml
host: {THOTH_SERVICE_HOST}
tls_verify: true
requirements_format: {requirements_format}

runtime_environments:
  - name: '{os_name}:{os_version}'
    operating_system:
      name: {os_name}
      version: '{os_version}'
    hardware:
      cpu_family: {cpu_family}
      cpu_model: {cpu_model}
    python_version: '{python_version}'
    cuda_version: {cuda_version}
    recommendation_type: stable
    platform: '{platform}'
```

Asking for an advise - CLI

- Thamos CLI
  - https://pypi.org/project/thamos/
  - .thoth.yaml configuration file already generated

```
$ pip install thamos
$ cd projects/my-project/
$ thamos advise
```

  - Hands On:
    - https://github.com/thoth-station/cli-examples
    - https://katacoda.com/aicoe/courses/ai-machine-learning/thoth-cli

Thoth Integrations

Thamos

Command line tool (developer)

Jupyter Tools (data scientist browser)

Source-to-Image (container builder)

Kebechet
Cyborg (pull request/issues creator)

Optimizing Deployment Pipeline

Security Introduction in thoth

- Change of environments
- CVE data, prescriptions
- Scorecard from ossf
- Results

Red Hat

Thoth Security Recommendations

Thoth uses three main sources for security-based advisories that affect Python packages.

- Ingests CVE information for each package from Python Packaging Advisory Database
- Computes security indicator via workflow using bandit tool from PyCQA
- Security scorecards for open source projects by the Open Source security Foundation (OpenSSF)

Provide Security to Source code

Ingested CVE

Thoth data aggregation periodically fetches the database of known vulnerabilities and ingests them in knowledge base.

- Automatically blocks the resolution of software package versions that are prone to security vulnerabilities.
- The vulnerabilities in open source Python libraries is available at public database provided by Python Packaging Advisory Database

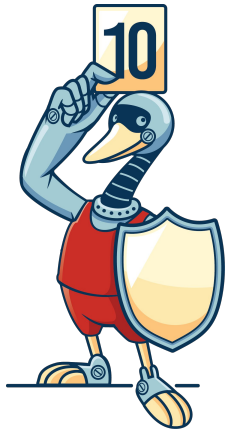
Security Indicators

Thoth engineers created the second source of data for security-based advisories.

- Each package imported by the application is statically scanned for possible issues using the open source Bandit tool.

Security Scorecards

The third source of security-related advisories consists of security scorecards that provide health metrics for open-source software



Thoth Prescriptions

- Prescriptions to heal Python applications
  - More Information: Documentation

- Prescriptions form a declarative interface to the resolution engine

  Set of YAML files that are automatically consumed by resolver in a deployment

Prescriptions Example

Adjust requirements in GPU enabled environments

- Use tensorflow-gpu as a “pseudonym” to tensorflow if gpu enabled environment is available
- Use the right tensorflow-gpu for the environment following support matrix
  tf_cuda.yaml
  tf_cudnn.yaml

Fixing library overpinning issue

- Tensorflow in version 2.1 can cause runtime errors when running with h5py>=3 caused by overpinning
  tf_s2_h5py.yaml

Block using certain library functions due to security reasons

- Mktemp is deprecated due to vulnerability to race conditions
  tempfile.yaml

Prescriptions Example

```
units:
  wraps:
    - name: TensorFlowMultipleProcessesGPUBug
      type: wrap
      should_include:
        adviser_pipeline: true
      match:
        state:
          resolved_dependencies:
            - name: tensorflow-gpu
              version: "~=2.3.0"
      run:
        justification:
          - type: WARNING
            message: "tensorflow in version 2.3 has a bug that prevents from running if multiple TensorFlow
processes are running"
            link: tf_38518
```

Thoth Security Recommendations

Security based recommendations using Thamos CLI

```
$ pip install thamos
```

```
$ thamos config
```

```
$ thamos add flask
```

```
$ thamos advise --recommendation-type security
```

Cyborgs Introduction

- Cyborgs of thoth

Red Hat

Bots helping with app development

- Reduce mundane work and focus on delivering solutions.
- Dependency management and managing life cycle of repository
  - Automatic update of libraries
  - Automatic release management

Kebechet Config

```
managers:
 - name: info
   configuration:
     enabled: false
 - name: thoth-advise
   configuration:
     enabled: false
 - name: update
   configuration:
     enabled: false
 - name: version
   configuration:
     enabled: false
```

Search or jump to... / Pull requests Issues Marketplace Explore

GitHub App

Khebhut

Configure

Manage your installation settings.

I'm Kebechet bot, goddess of freshness - I will keep your source code fresh and up-to-date.

A Cyborg team member that is enhancing your application stack by applying Thoth's knowledge.

Read more about this app on the Marketplace.

Developer

AICoE

Website

Khebhut is provided by a third-party and is governed by separate terms of service, privacy policy, and support documentation.

Report abuse

© 2021 GitHub, Inc. Terms Privacy Security Status Docs Contact GitHub Pricing API Training Blog About

Source: https://github.com/apps/khebhut

Automatic update of dependencies by kebechet. #1614

Closed **khebhut** wants to merge 1 commit into `master` from `kebechet-automatic-update`

Conversation 11 | Commits 1 | Checks 1 | Files changed 1

khebhut bot commented on Dec 8, 2020 — Contributor

Kebechet has updated the depedencies to the latest version 🚀
The direct dependencies updated in the pull request are -

Package Name	Old Version	Updated Version	Is Dev
thoth-python	0.10.2	0.11.0	False
thoth-storages	0.29.3	0.29.4	False
voluptuous	0.12.0	0.12.1	False
hypothesis	5.41.4	5.41.5	True

Kebechet Version: 1.2.2

khebhut bot requested review from **fridex**, **goern** and **sesheta** as code owners 11 months ago

khebhut bot added the bot label on Dec 8, 2020

Release of version 0.47.0 #2160

Merged **harshad16** merged 1 commit into `master` from `v0.47.0` 2 days ago

Conversation 4 | Commits 1 | Checks 0 | Files changed 2

khebhut bot commented 2 days ago — Contributor

Hey, **@fridex**!

Opening this PR to create a release in a backwards compatible manner.

Closes: #2159

```
Changelog:
* Provide information about last CVE update to users
* :arrow_up: Automatic update of dependencies by Kebechet for the rhel8 environment
* Provide a link to CVE based on info as present in PyPA/advisory-db
* :arrow_up: Automatic update of dependencies by Kebechet for the rhel8 environment
* Fix matching operating system in prescriptions
* Fix ABI sieve filtering packages when no image analysis is available
* Introduce a CVE sieve
* :arrow_up: Automatic update of dependencies by Kebechet for the rhel8 environment
```

Release of version 0.47.0 — 5bf72ee

AICoE CI/CD

- Explain of ci/cd
- Introduction to aicoe-ci
- Integration of thoth
- Run

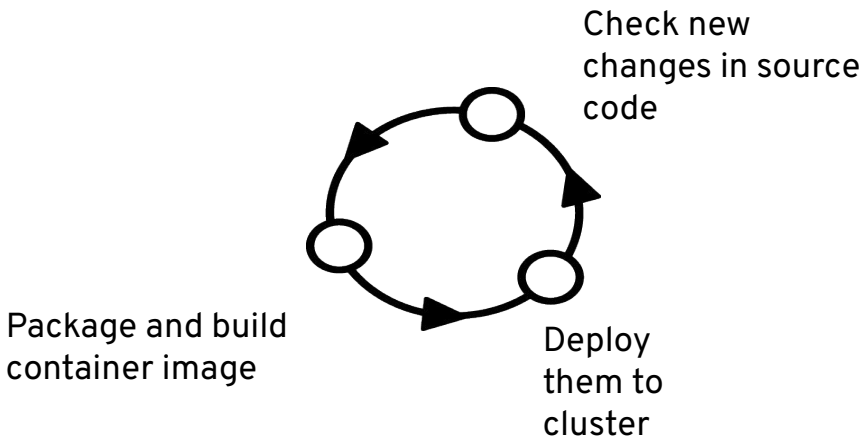
Why we care about CI/CD ?

Automation | Continuous Validation | Secure Delivery

- Automate test checks on new changes for source code.
- Validation checks to ensure changes are fit for source code.
- Constant packaging and delivery of the container image to deployment.
- Cloud native CI/CD system which could run on-premise OpenShift.

Challenges with the CI/CD system available:

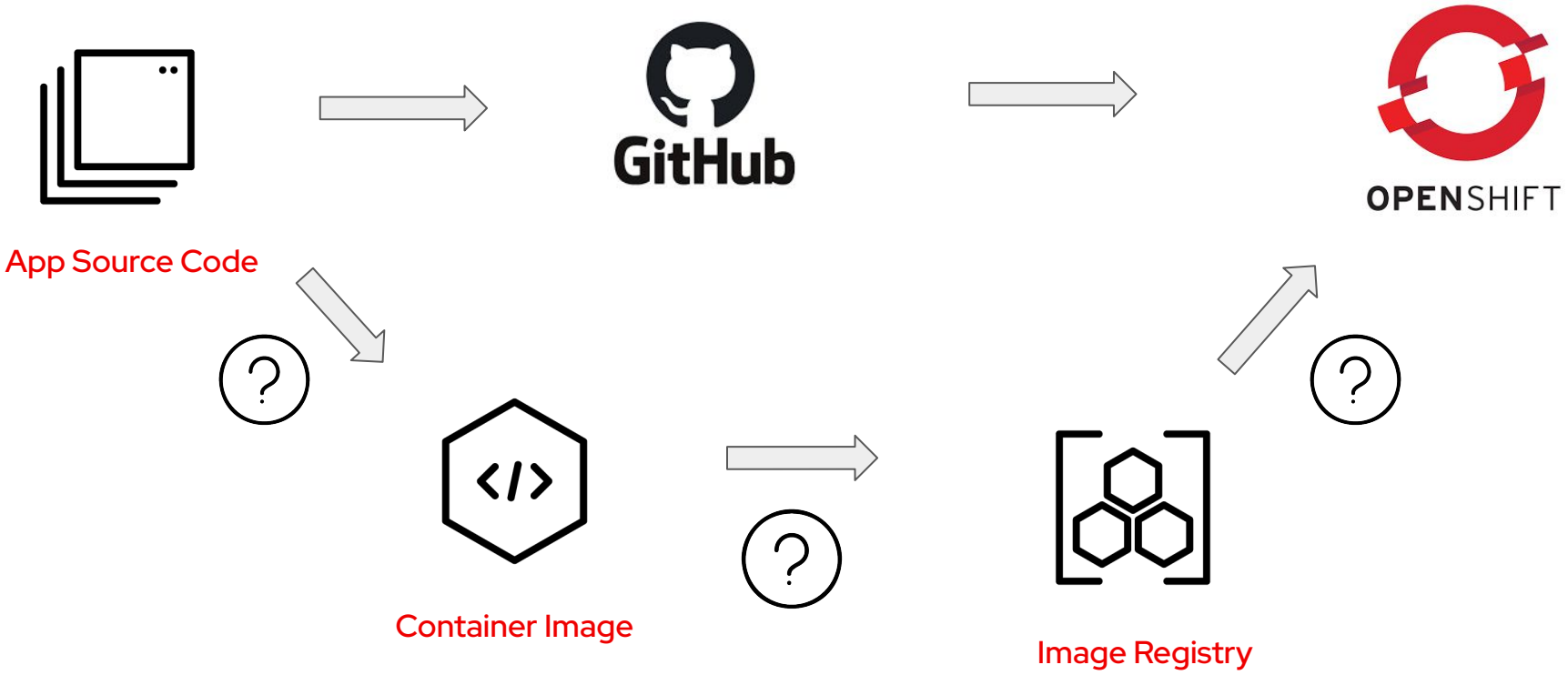
- **No one size fits all**



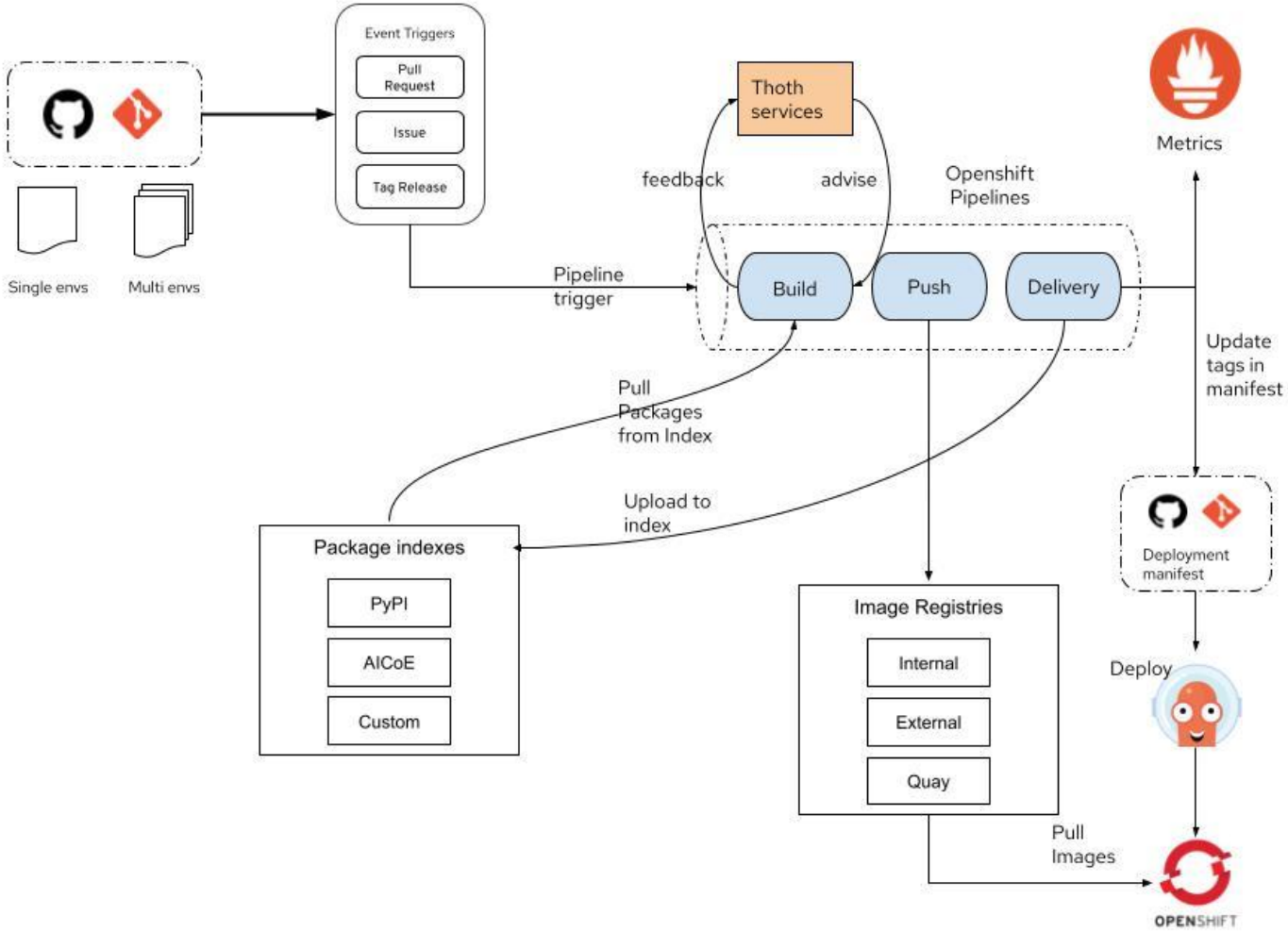
Develop CI app with Tekton

- Tekton is a flexible open-source framework for creating CI/CD systems.
- Allows developers to build, test, and deploy across cloud providers and on-premise systems.
- Now part of CD Foundations .

Information Flow



CI Architecture



AICoE CI

GitHub App

aicoe-ci

This is the Continuous Integration Cyborg maintained by Thoth Station.

Configure

Manage your installation settings.

Developer

AICoE

Website

aicoe-ci is provided by a third-party and is governed by separate terms of service, privacy policy, and support documentation.

Report abuse

Source: https://github.com/apps/aicoe-ci

AICoE CI

```
- name: inference
  build:
    base-image: "quay.io/thoth-station/s2i-thoth-ubi8-py38:v0.28.0"
    build-stratergy: Source
    registry: quay.io
    registry-org: thoth-station
    registry-project: elyra-aidevsecops-tutorial
    registry-secret: thoth-station-thoth-pusher-secret
  deploy:
    project-org: "thoth-station"
    project-name: "elyra-aidevsecops-tutorial"
    image-name: "elyra-aidevsecops-tutorial"
    overlay-contextpath: "manifests/overlays/test/imagestreamtag.yaml"
```

Source: https://github.com/AICoE/elyra-aidevsecops-tutorial/blob/master/docs/source/build-images.md

Build Pipeline

- Source-to-Image container images are used to build the image with thoth induced.

Container image with s2i

- Set the environment variables to start Thoth advise

THAMOS_RUNTIME_ENVIRONMENT=""

THOTH_ADVISE="1"

THOTH_ERROR_FALLBACK="1"

THOTH_DRY_RUN="1"

THAMOS_DEBUG="0"

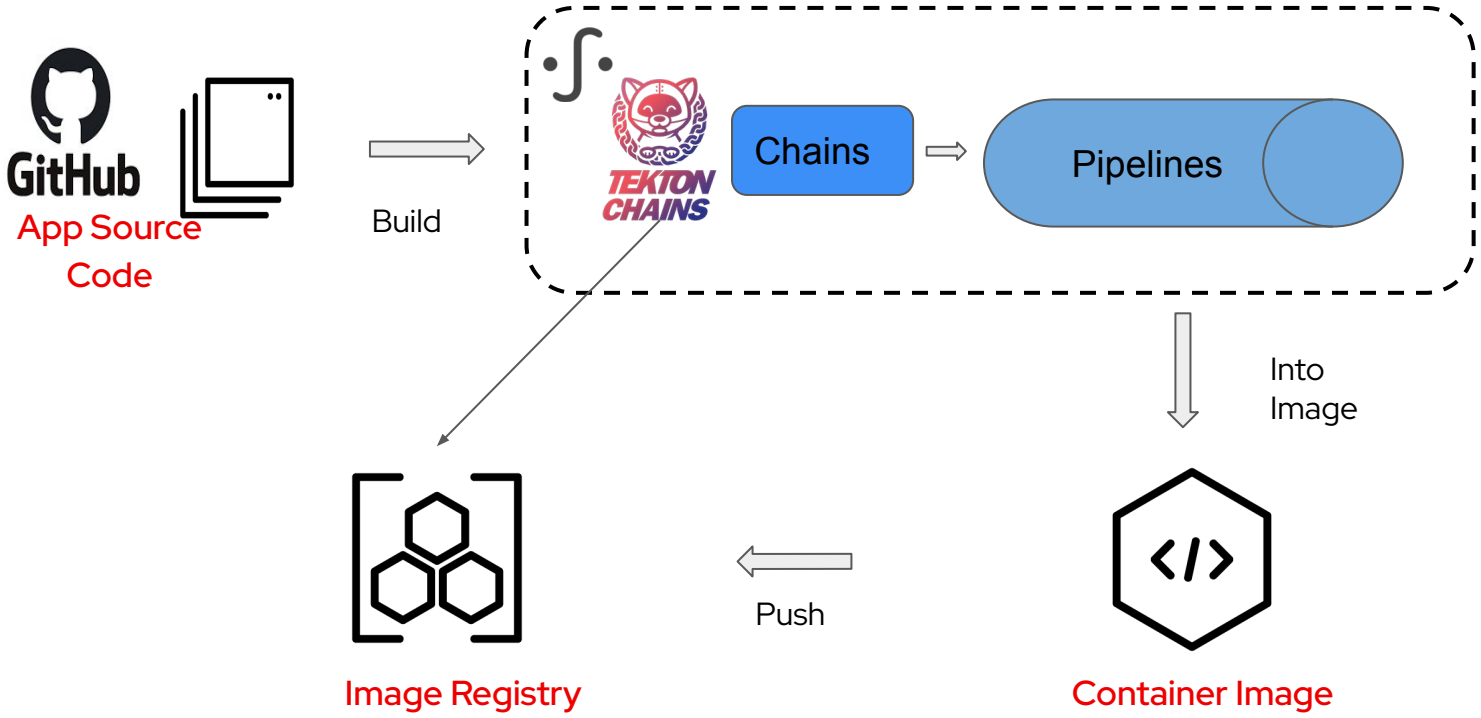
THAMOS_VERBOSE="1"

Integrated Tekton chain

- Secure the supply chain
- Integration
- Signature and finishing the end step

Red Hat

Secure the supply build pipeline



Signed Image for security

Build

Build the container image for an application

x509,KMS,Cosign

Signing with a variety of cryptographic key types and services

Sign Image

Configuring Tekton Chains to generate and sign images

Push signed Image

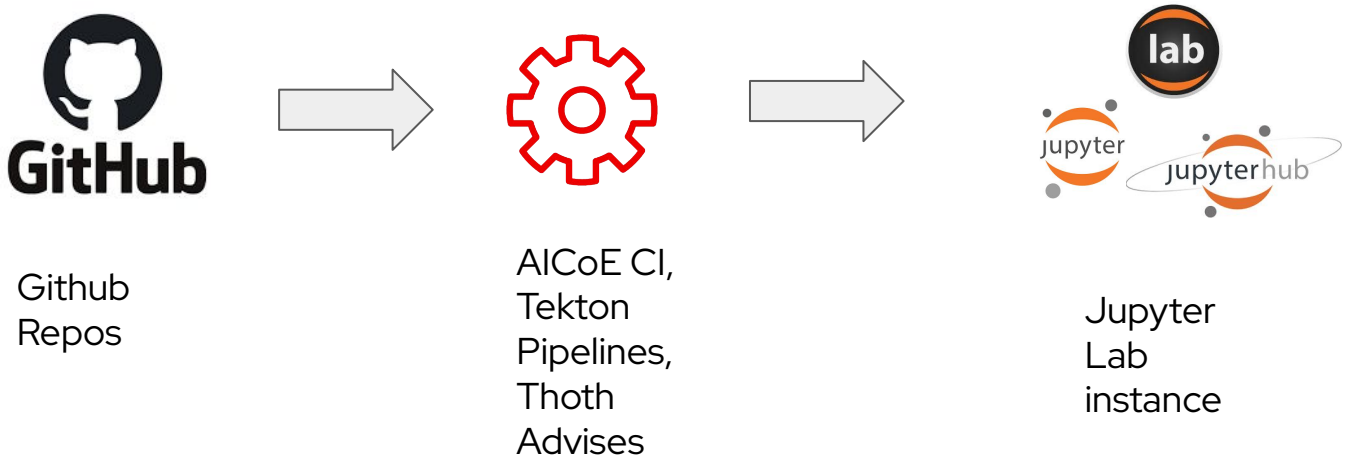
Pushing signatures to an OCI registry after signing an image

Advancements

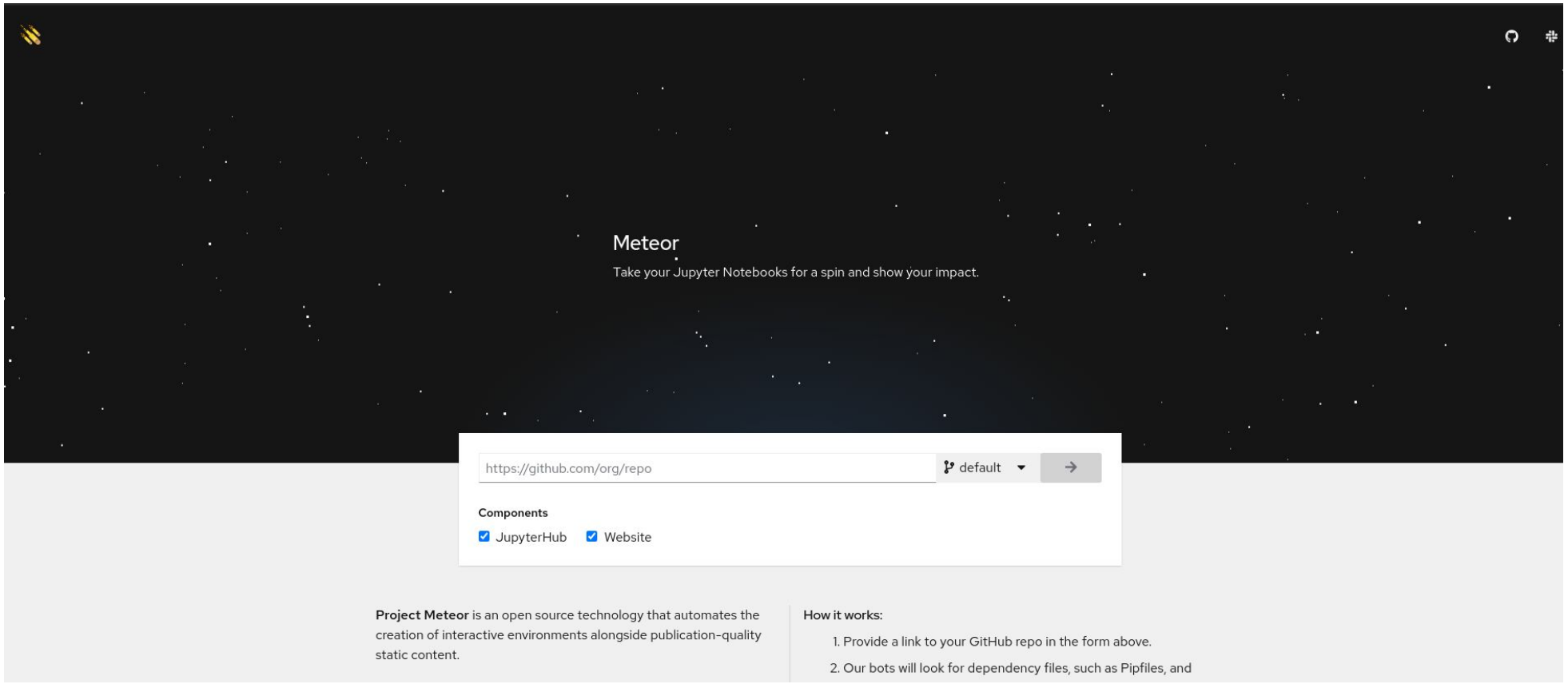
- Predictable pipeline
- Feedback
- Metrics mlops

Project Meteor

Generating AI/ML Projects into consumable jupyter notebooks.



Project Meteor

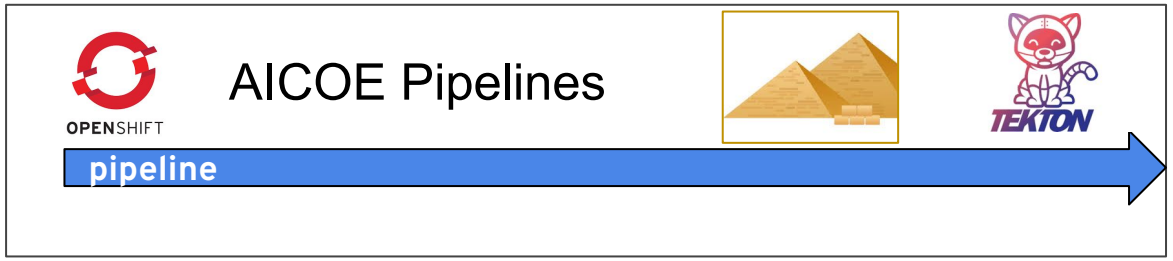


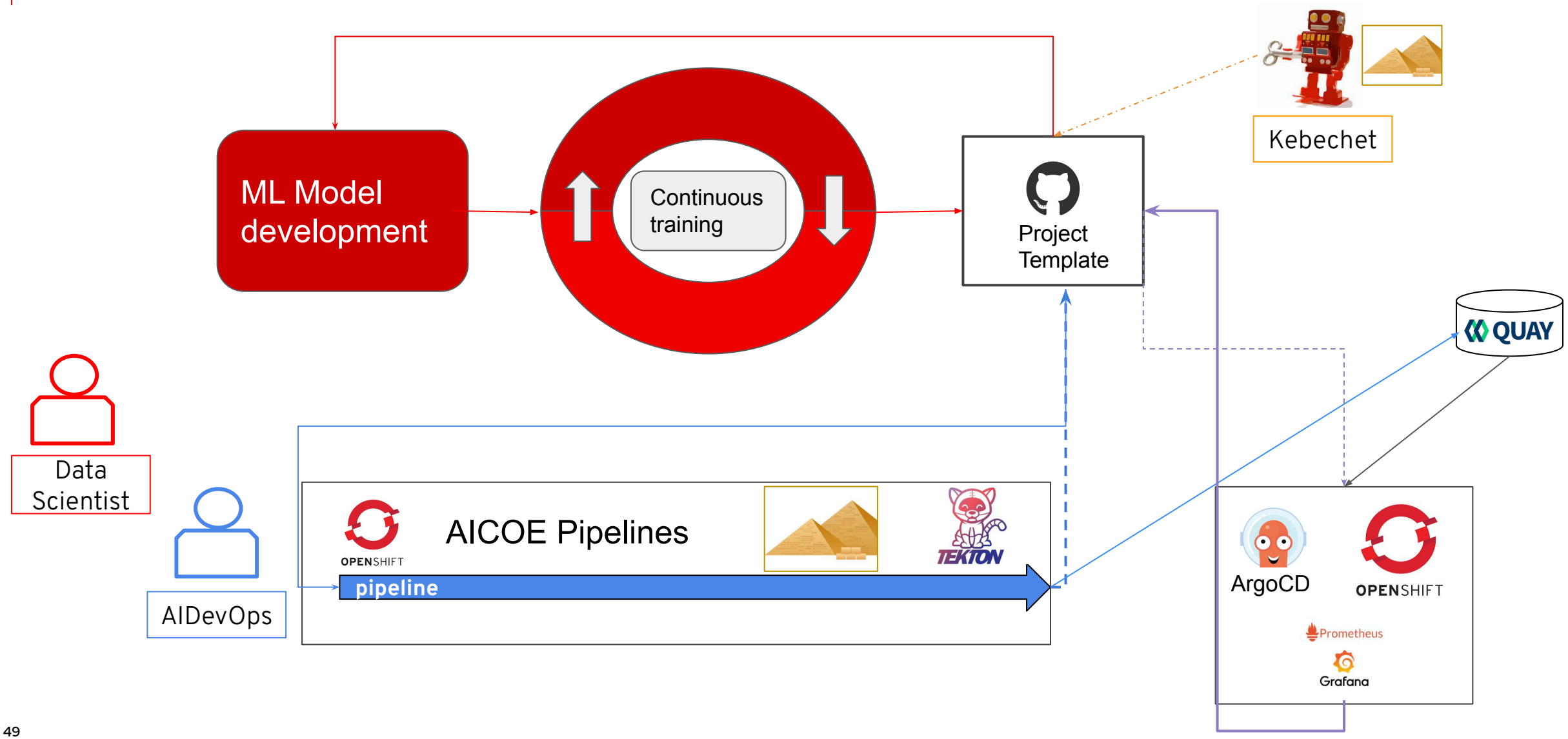
Source: https://shower.meteor.zone/

CI pipeline in Model Development

We have introduced the Feedback and reproducible builds for fast Model Development

- Model requires multiple testing in various environments
- Feedback loop to incurs from the knowledge from previous iterations
- Delivery and deploy of the each release





References

Red Hat

Website: https://thoth-station.ninja/

Twitter: https://twitter.com/thothstation

Github: https://github.com/thoth-station

Youtube: Thoth Station

Email: aicoe-thoth@redhat.com

Project Thoth

All Talks: https://github.com/thoth-station/talks

Blogs Post:

Elyra AIDevSecOps Tutorial
Secure your python applications with thoth recommendation Resolve python dependencies
Thoth Prescriptions resolving Python dependencies

Important Links:

Operate-First: https://www.operate-first.cloud/

AICoE CI: https://github.com/AICoE/aicoe-ci
Kebechet: https://github.com/apps/khebhut

Shout Out to Team Thoth

Christoph Goern
Leads the Thoth Technical Team

Frido Pokorny
Contributed most to the Adviser, knowledge Graph, prescriptions

Francesco Murdaca
Contributed most to the knowledge Graph, jupyterlab-requirements

Kevin Postlethwait
Contributed most to the Security Indicators, Kebechet

Project Thoth

aicoe-thoth@redhat.com

Thank you

Have wonderful time at Red Hat Next!

Follow the Red Hat Office of the CTO at **next.redhat.com**

Harshad Reddy Nalla
Senior Software Engineer

https://www.linkedin.com/in/harshad-reddy-nalla-2b64b2104

hnalla@redhat.com

https://github.com/harshad16

twitter.com/hnalla16

Red Hat