Enterprise NeuroSystem Group Meeting

Feb 14, 2022

Sprint 1 Activities

Activity 1:

- Flesh out use-cases
- Volunteers: **Tong Zhang (Intel),** Bill Wright (RH), Ravi Sinha (Jio), Dinesh Verma (IBM), John Overton (Kove)

Activity 2:

- Study other open source catalogs
- Volunteers: Tong Zhang (Intel), Sanjay Aiyagari(RH), **Sathya Santhara (IBM)**, Theo Thomas (IBM), John Overton (Kove)

Activity 3:

- Code design and development
- Volunteers: Austin Eovito, **Josh Purcell**, Rahul Batra, Shirley Han, Janki Vora, Amir Khanof, Dinesh Verma (all from IBM), John Overton (Kove)

Activity 4: (added in 2/14/call)

- Governance Model
- Volunteers: **Audrey Reznik**, Bill Wright , Eric Erlandsen, Trevor (all from RedHat) , John Overton (Kove), Andy Poling (Kove) , Dinesh Verma (IBM)

Name is bold is leader/coordinator of the activity.

Agenda

Activity Summary

- Activity 1 – work in progress
- Activity 2 – See Slides 4-5
- Activity 3: See Slides 6-10

Objectives

Activity 2 Report

1. What are the functions supported by other Catalogs and whether the problem of 'Self describing assets" can be solved by existing catalogs?

2. If any of the existing catalogs can be modified a little (option b) to address the problem of metadata, what should be the incremental addition?

3. What additional functionalities are provided by existing catalogs like execution engines, so that they can be reused in new catalog?

4. Are there any other problems to be addressed with respect to metadata (other than the one mentioned) like automatic updation of metadata when the actual data/model changes?

Catalog – Data, Model, Pipelines

Activity 2 Report

- MLX – Machine learning exchange
- ONNX
- Acumos
- Portable format for analytics – PFA & MLEAP
- NNEF
- MAX and DAX/Datashim
- Informatica CDGC

Flask + Mongo Report

Activity 3 Report

Easy to develop SDS catalog code in python with Flask + Mongo, and package it as a container with Docker

Comments from developing the code

- Structure of the self-describing entries – keep current or change to a flat structure
- Structure constraints and input checks for different fields
  - constraints for scope and type
  - Data and metadata links – unchecked or forced to satisfy constraints such as being an URI
- CRUD interface will be inefficient for common use-cases – need to add search capability
  - What search capability to expose
- Security Incorporation for containers

Structure Questions

Activity 3 Report

Current Approach:

– {“data”: ”blob”,

“metadata”: {“owner”: slac.edu, “type”: “dataset”, “link”: ”None”}}

Decision:
Go now with flat structure,
See how it goes and revert
Quickly if problems emerges

Alternative:

{“data”: “blob”, “owner”: slac.edu, “type”: “dataset”, “metadata”: “blob2”}}

Reasons:

- Simpler representation, less manipulation
- Constraint support for “owner” and “type” (e.g. owner must be of type dns structure }
- Constraints are only for two fixed fields in the top structure, otherwise constraints supported for two links in metadata, but others are up to the user.

Constraint Support

Activity 3 Report

Scope and type may need to satisfy some constraints

- Users should only be allowed to put in some values of scope
- Only some types are valid for some users.

Catalog can allow the definition of these constraints

- e.g. define policies/constraints based on user-identity, or define global constraints based on configuration

Decision:
Important to add this capability
Keep an eye on other open source and groups that may be relevant to this type of constraint.

Derivative: 4th Activity Group on Governance

Deployment Considerations

Activity 3 Report

Deployment models

- As a container with the catalog
- Security constraints in deployment
  - Defer to front-end webserver?
  - Support inbuilt into the catalog, e.g. incorporate Apache within the catalog container?
- Other deployment considerations
  - Load-balancing?
- Leave deployment considerations externally

Decision:
Mechanisms for addressing deployment issues exist in Service Mesh and Istio. One should document the considerations but leave the capabilities out of scope for now.

A metadata describing how to size/scope should be put into the documentation.

Search Support

Activity 3 Report

Common use-case

- A published pushes digital assets into the catalog
- A subscriber retrieves the digital asset

In may cases, subscribers would not have direct links to descriptors, and would like to search for items

- e.g. search for items that are within some scope (e.g. do longest prefix search for the same)

Discussions – Deferred till next meeting.