NASA Earthdata Knowledge Base

WGISS-48 October 9th 2019

Doug Newman

NASA EED-2 Data Use Architect

douglas.j.newman@nasa.gov

Dr. Christopher Lynnes

NASA EOSDIS System Architect

christopher.s.lynnes@nasa.gov

INTRODUCTION

Rationale

'Connect together the main elements of Earth Observation knowledge AND context in a way that is: machine-readable, human-usable and curatable'

Chris Lynnes, WGISS-47

Use Cases

- As an earth data scientist, I want to determine the datasets used in an article/report/paper
- As an earth data scientist, I want to determine the articles/reports/papers that cite a dataset
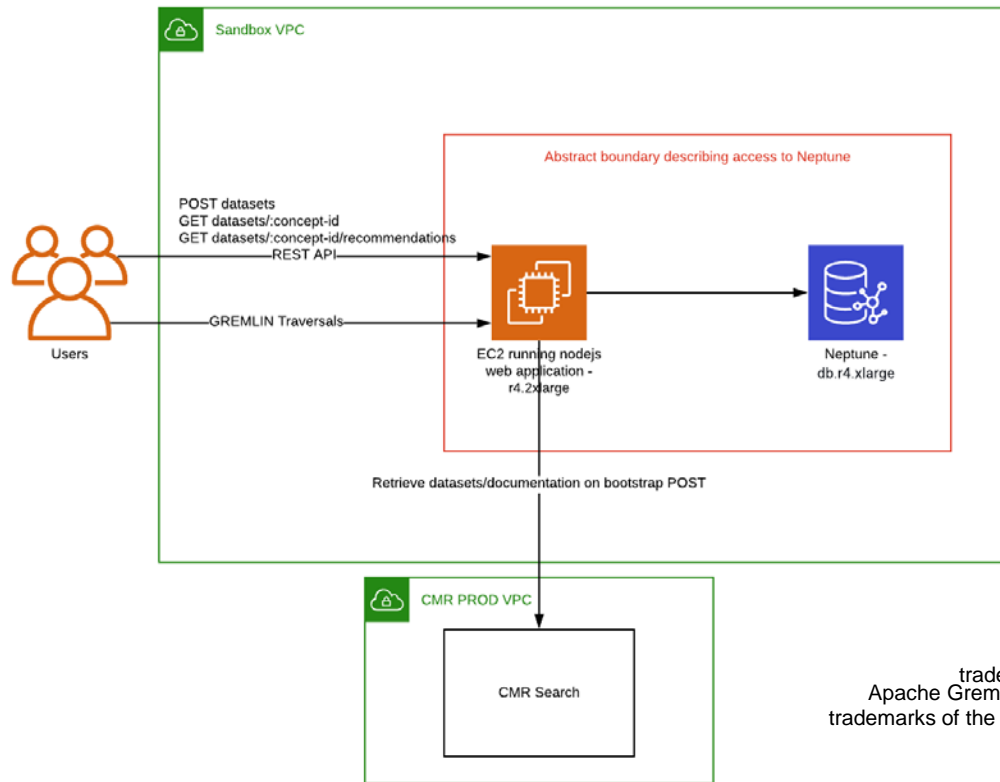
APPROACH

Graph Technologies

Implementations
- AWS Neptune
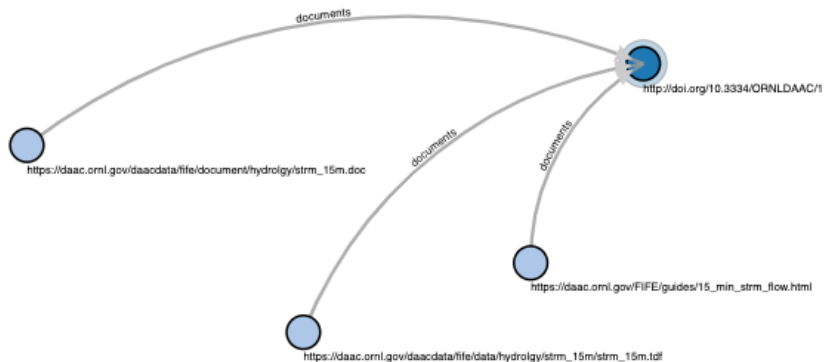- Neo4J

APIs
- SPARQL
- Gremlin
- Cyper

AWS Neptune + Gremlin

Collections and documentation

The common metadata repository contains dataset metadata records and references to their respective documentation. We used this inventory to create a graph of datasets and documentation vertices and 'documents' edges:

Recommendations

Recommendations for C1238517289-GES_DISC

- AIRS/Aqua L3 8-day Standard Physical Retrieval (AIRS-only) 1 degree X 1 degree V006 (AIRS3ST8) at GES DISC
- AIRS/Aqua L3 Daily Standard Physical Retrieval (AIRS+AMSU+HSB) 1 degree x 1 degree V006 (AIRH3STD) at GES DISC
- AIRS/Aqua L3 8-day Standard Physical Retrieval (AIRS+AMSU+HSB) 1 degree x 1 degree V006 (AIRH3ST8) at GES DISC
- AIRS/Aqua L3 Daily Standard Physical Retrieval (AIRS+AMSU) 1 degree x 1 degree V006 (AIRX3STD) at GES DISC
- AIRS/Aqua L3 Monthly Standard Physical Retrieval (AIRS+AMSU+HSB) 1 degree x 1 degree V006 (AIRH3STM) at GES DISC
- AIRS/Aqua L3 Monthly Standard Physical Retrieval (AIRS-only) 1 degree x 1 degree V006 (AIRS3STM) at GES DISC
- AIRS/Aqua L3 5-day Quantization in Physical Units (AIRS+AMSU) 5 degrees x 5 degrees V006 (AIRX3QP5) at GES DISC
- AIRS/Aqua L3 8-day Standard Physical Retrieval (AIRS+AMSU) 1 degree x 1 degree V006 (AIRX3ST8) at GES DISC
- AIRS/Aqua L3 Monthly Standard Physical Retrieval (AIRS+AMSU) 1 degree x 1 degree V006 (AIRX3STM) at GES DISC

.hasLabel('dataset')   9

CONCLUSION

If you build it...

- 10 minutes looking at a visualization of two CMR concepts gave me the inspiration for the recommendation engine. Imagine what an expert could do?
- Providing a read-only traversal API over a knowledge base like this could open the floodgates for novel discovery techniques

NEXT STEPS

Articles

- The motivation for dataset -> documentation edges was to get to articles. A lot of articles reference dataset documentation. The next step is to bridge that gap.
  - Google scholar
  - OpenCitations

QUESTIONS?

This work was supported by NASA/GSFC under Raytheon Co. contract number NNG15HZ39C.