



Thoughts on deploying a PROV capacity

ESIP Summer

July 2018 | Tucson AZ




ESIP Lab

Overview

Very early work... *experimenting in code*

The stardate is Summer 2018, so if it's later than that and you are reading this, look to contact ESIP or others about where things have evolved too..

pres·ent·ism

/prezenˌtɪzəm/ 

noun

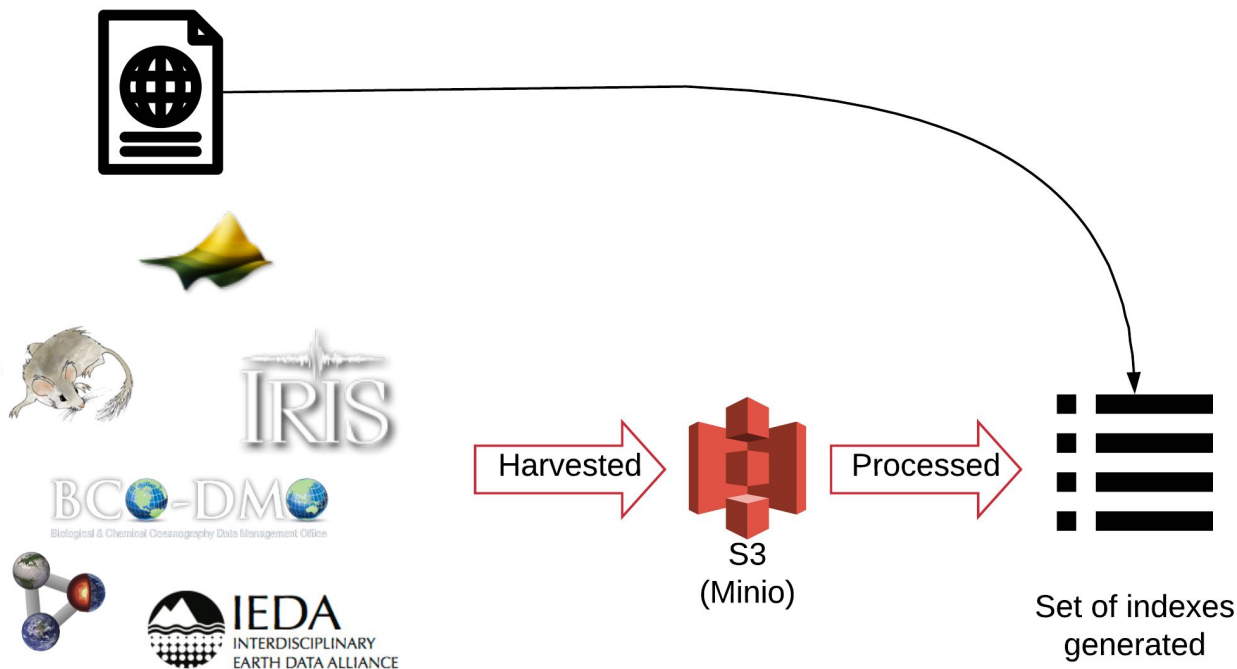
uncritical adherence to present-day attitudes, especially the tendency to interpret past events in terms of modern values and concepts.

So Many Issues in Implementation!

- What do we capture?
- How do we capture it?
- How do we represent it?
- Once we have it.. How do we store and/or send it?
- How do we provide access and search?
- How do we let others know?
- How is it updated and maintained?
- How do we know it's still valid?
(representation rot or representation uplifting)

ESIP Lab

Project 418 Experience



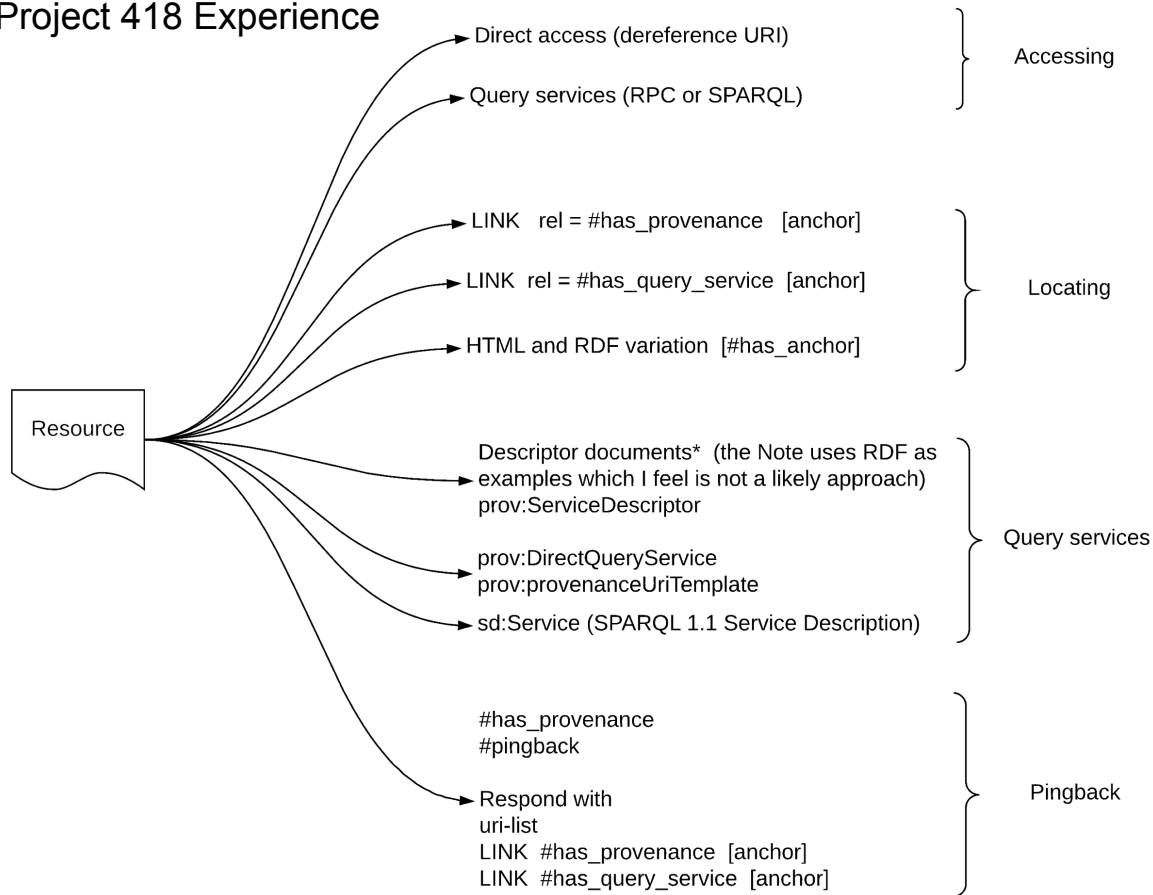
Wanted to keep track and dataset resources becoming part of indexes.

- Software
- Organizations
- Time
- Index versions
- Activity (processing)

Q: Is generating an index based on a dataset a prov event?

ESIP Lab

Project 418 Experience



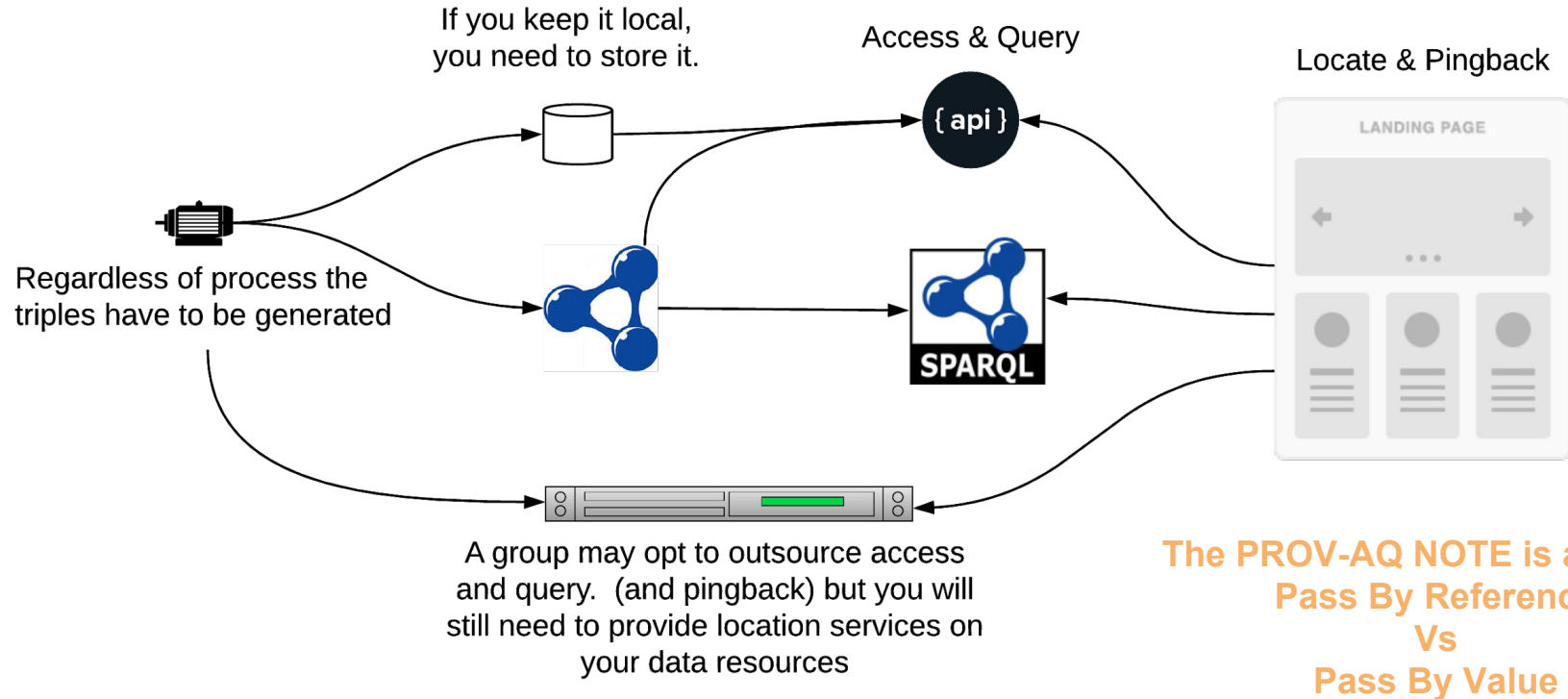
Prov introduces a range of items to keep track of for an organization. (the items left are from the PROV-AQ note)

Most of this can be represented as a data graph though.

Thought: Checking data graphs with shape graphs?

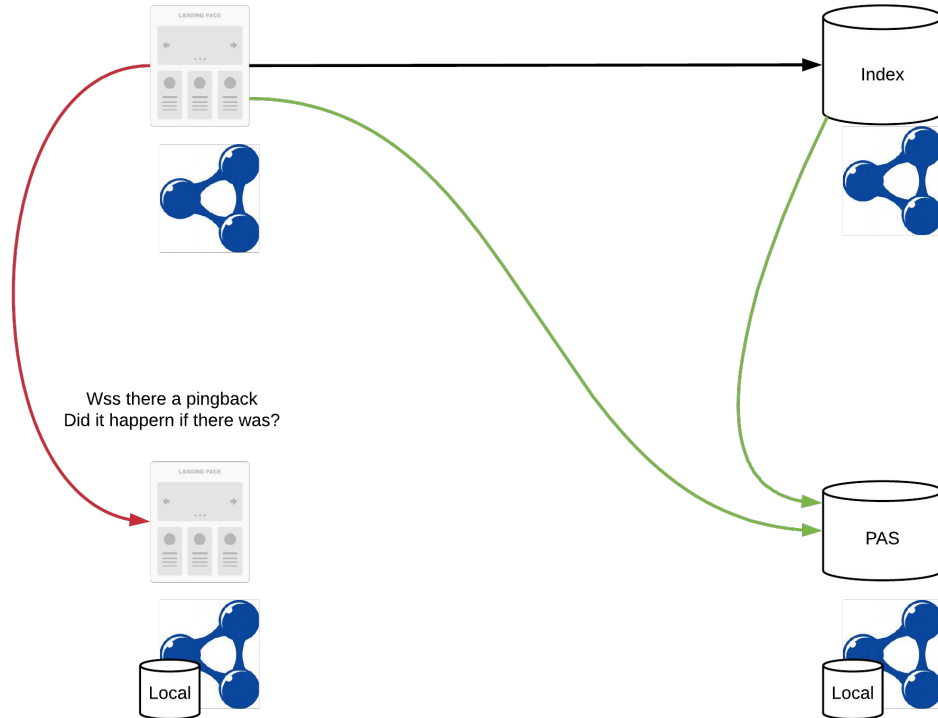
ESIP Lab

Project 418 Implementation Options



ESIP Lab

Project 418 Implementation Options



This is *Linked Open Data* and is an *Open World* (great)

Can we deal with this in the discover and index phase?

It's easy to see why people start to think about a distributed ledger.. But that is buzzword zone and is it really better? (just a monolith of a different sort)

ESIP Lab

Closing thoughts

PROV-AQ is just a note... we can influence it. (The authors want us to)

What is the compelling value add to the facilities to get them to start this process?

Is this just a special case metadata and we should advocate capturing it for all the reasons we capture metadata?

Leverage interest in schema.org to help address plumbing overhead?