PROVisualizer: the journey, the tech and a demo!

Charlie Farrugia Conal Tuohy Asa Letourneau

Public Record Office Victoria

Archives are all about complex relationships

Publications
Life Sciences
Cross-Domain
Social Networking
Geographic
Government
Media
User-Generated Content
Linguistics

Public Record Office Victoria
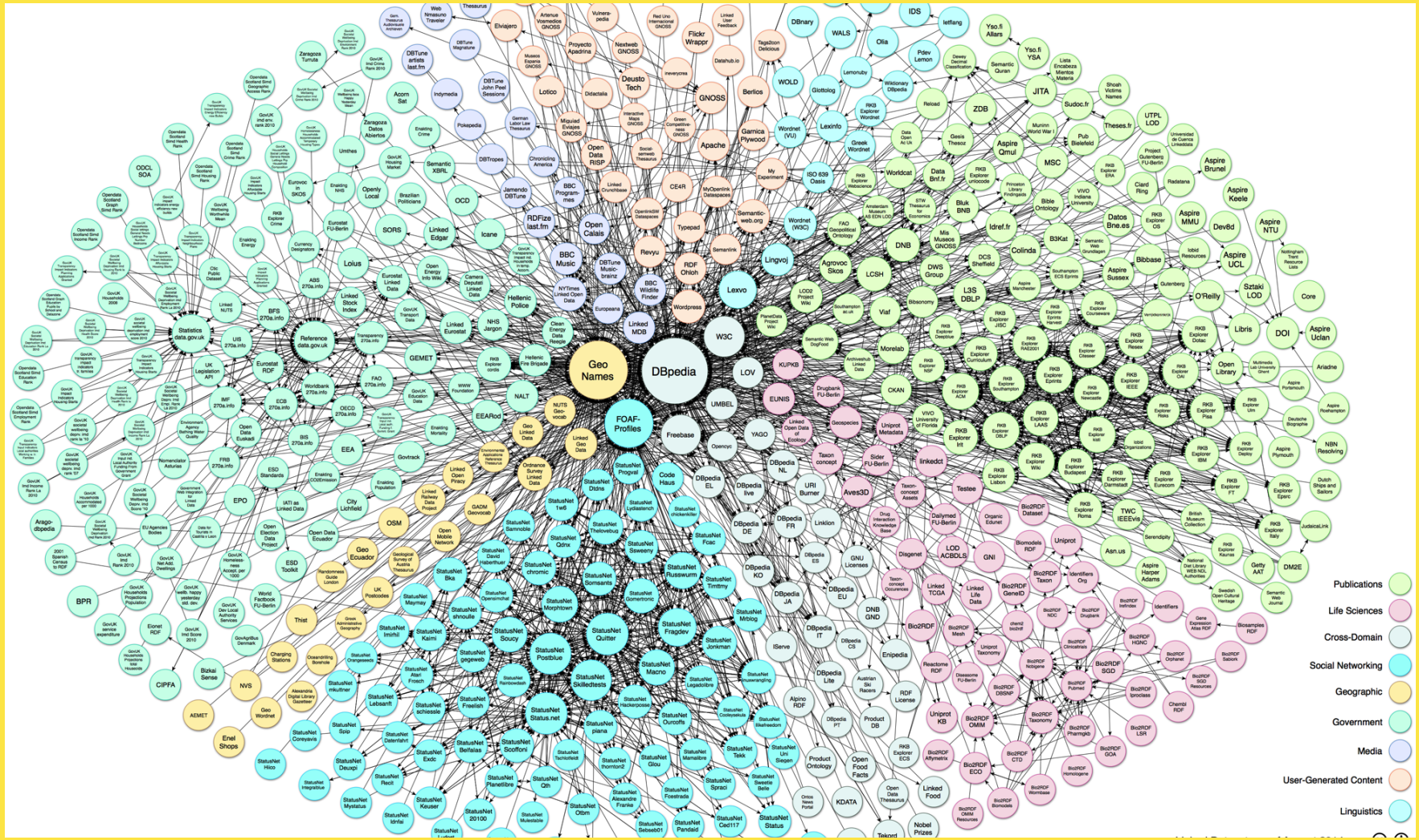
So why do we reduce them to lists?

Visualisations express spatial relationships and complex networks

I've always seen archives as repositories of relationships
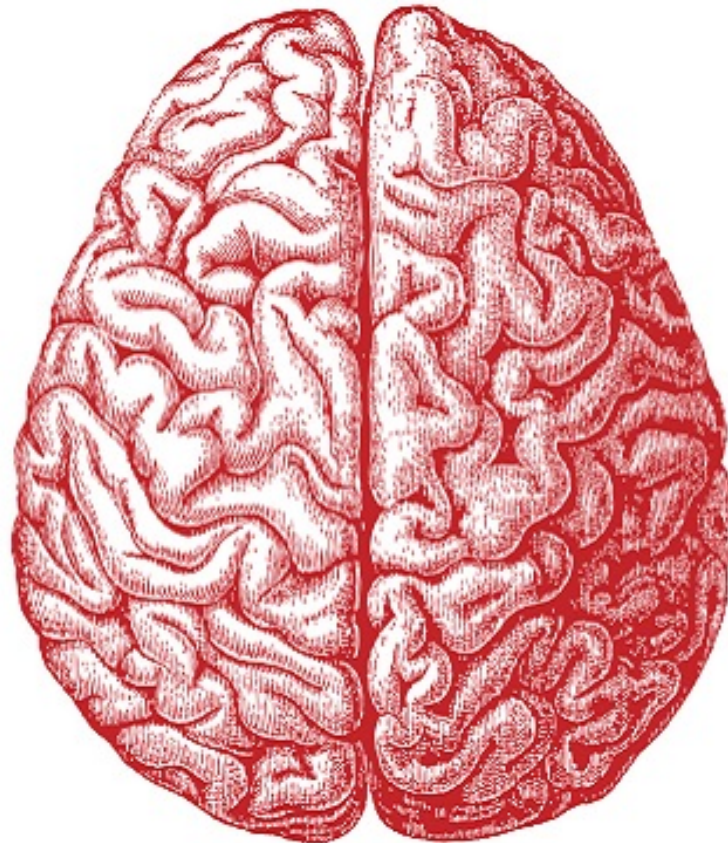
Asa's Journey

Nov 2010 While on secondment to Deloiite , I started experimenting with an Excel template developed by NodeXL to create a network chart and produced the following visualisation showing connections between the Chief Secretary's Dept VA 475 and the 200 Series it created as well as a number of items within those series

https://drive.google.com/file/d/0ByVdr3cge40gZWE2OWExYzctMTAzOS00YTZjLTg5ZDUtZmE1ZjA3YmQ2ZDVh/view?usp=sharing&authkey=CKjytYYM

Nov 2014 Experimenting with Google Fusion Tables, I produced this network chart showing the connections between the Chief Secretary's Dept VA 475 and the 200 Series it created.

https://www.google.com/fusiontables/DataSource?docid=1cMU8B61ne8I-dfkr9p3lLGxiVZKyJ4wA9Mpyw2uR

2015 PROVisualizer Phase 1 http://metadata.prov.vic.gov.au/provisualizer (Asa,Charlie, Conal)
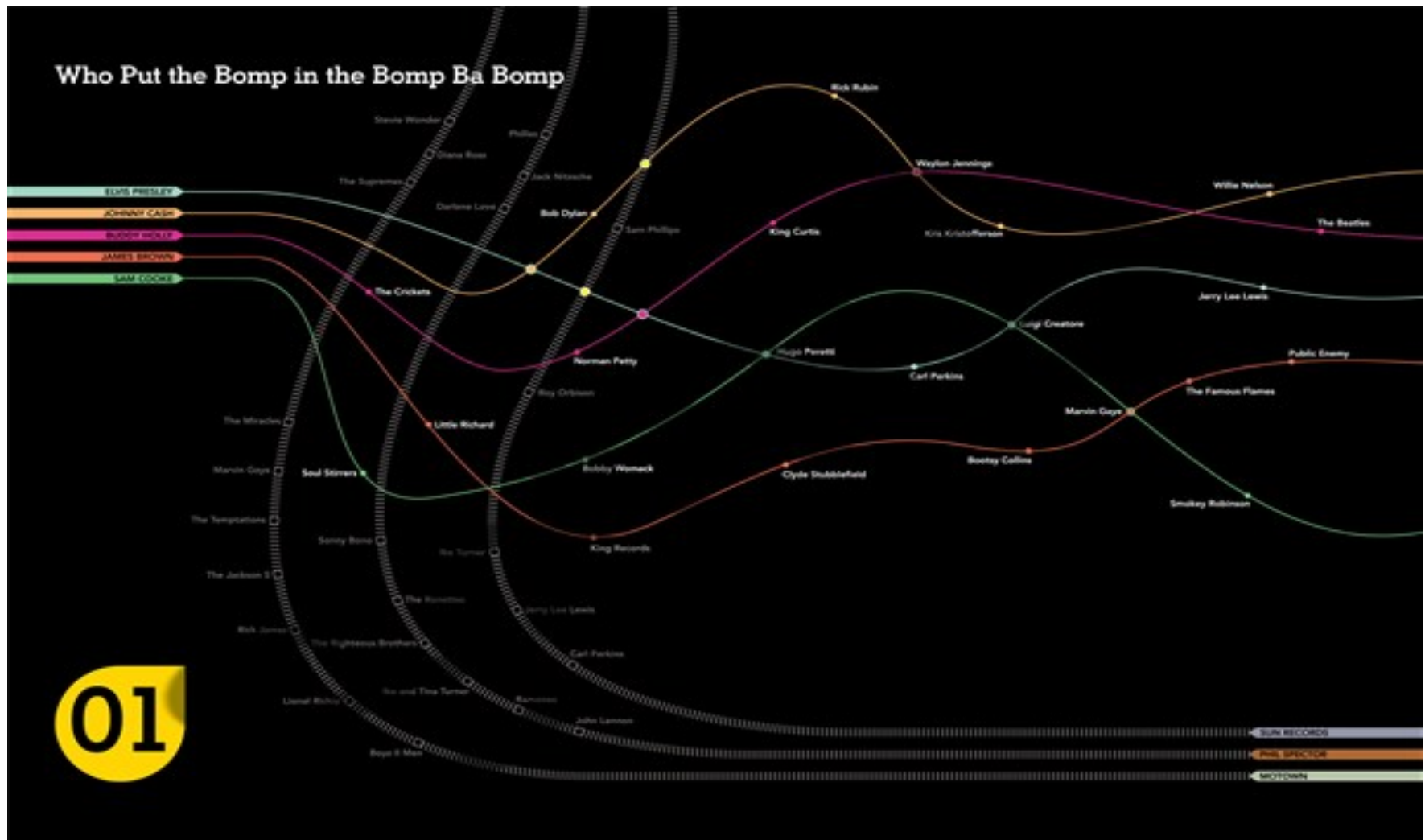
Essential feedback? Hard to use, no context, visually confusing

2016 After a series of UAT sessions and further iterations we developed the final tool we will demo today after hearing from Charlie and Conal.

Charlie's journey

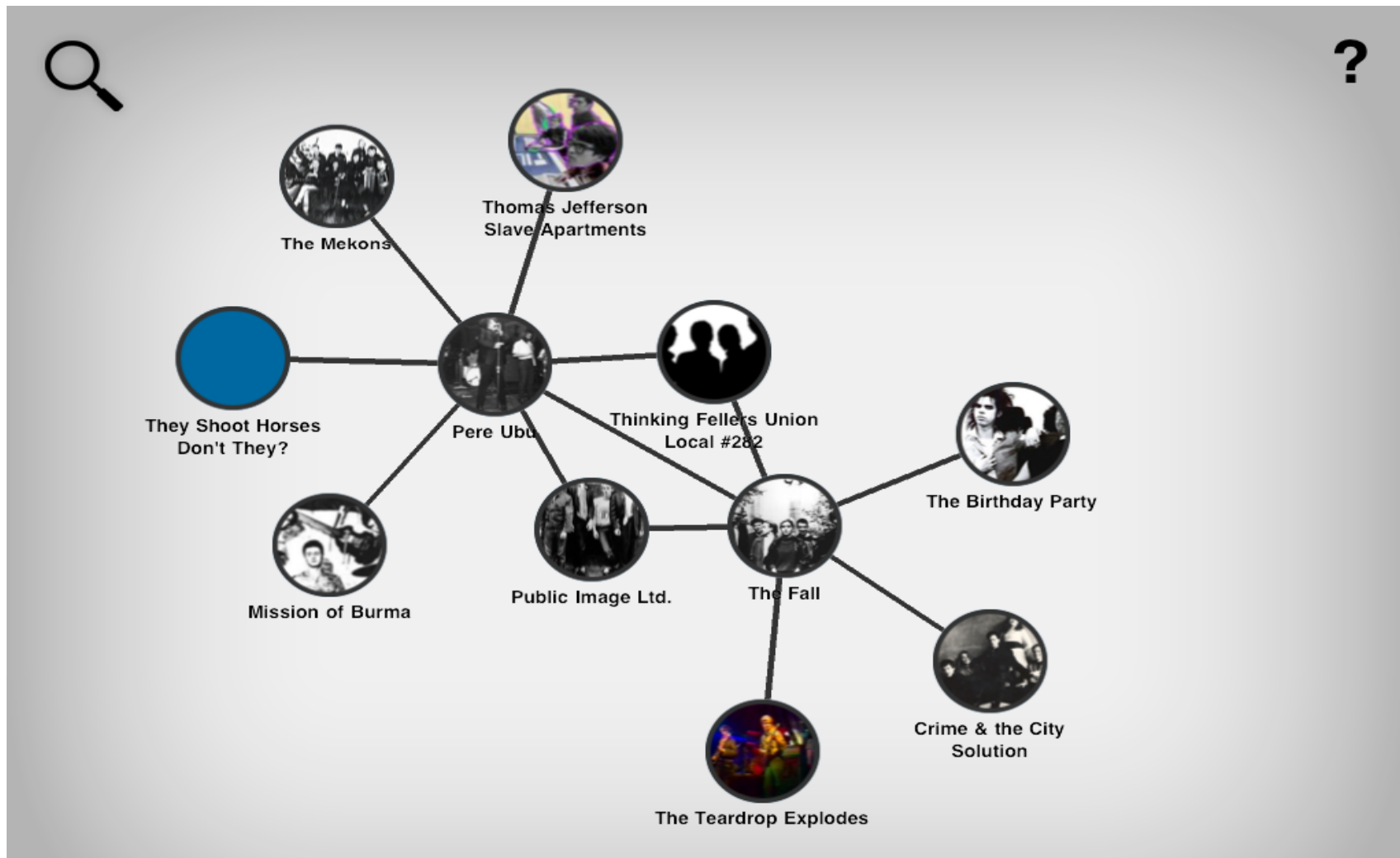Started with a challenge to visually represent administrative change

Public Record Office Victoria

VICTORIA State Government

Rock Connections by Bruno MacDonald

DiscovR Music

A DiscovR Search Result

Charlie's Non Tech Savvy Attempt

PROV's public data feed

PROVisualizer's data pipeline

Open Archives Metadata Repository (OAI-PMH)

metadata harvest via the internet

RIF-CS XML files

PROVisualizer's harvester / transformer converts RIF-CS to RDF

RDF graphs

download over the web

CSV Files

SPARQL queries

PROVisualizer (web page widget)

SPARQL Graph Store

Public Record Office Victoria

VICTORIA State Government
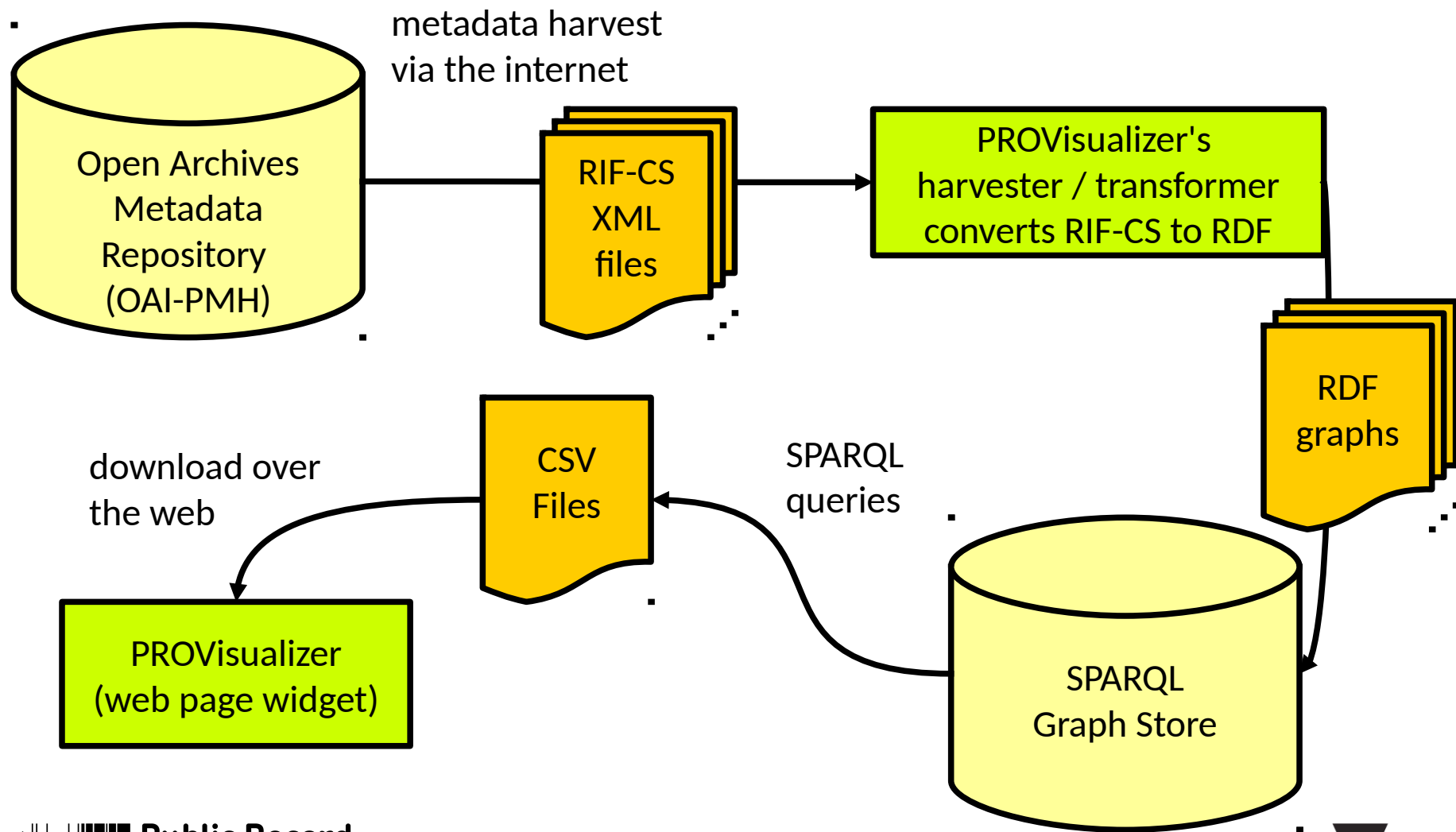
Enterprise Architecture

Factors relevant to the system architecture:

- PROV lacks in-house semantic web expertise
- PROV's strategy is to move applications to the cloud
- PROV's Archival Control System is under review and may change
- PROV's OAI-PMH repository is a long-term commitment to the National Library and Australian National Data Service

http://conaltuohy.com/blog/visualizing-government-archives-through-linked-data/

Public Record Office Victoria

VICTORIA State Government

Linked Data prospects

Using RDF and a SPARQL store was only one way to produce the data PROVisualizer needed; but RDF is an open technology that also offers:

● Interoperability with other archives and cultural heritage institutions (especially so given ICA's imminent RDF schema for Records in Context)

● Federated search across all archives

● SPARQL access for large-scale data analysis

● Enhancement of more of PROV's web real estate with relevant data drawn from the SPARQL store

● Integration of more of PROV's metadata collection

Demo

PROVisualizer can be seen at PROV's new website (which is still in beta):

http://beta.prov.vic.gov.au/explore-collection

Also viewable as a full page at:

http://prov-data.conaltuohy.com/