Ashurbanipal’s clay tablet library (650BC) had index tablets pointing to content. Both indexes and content still exist!

How long will I URIs last?

Persistent URIs for the Australian Govt.

Simon Cox, Nicholas Car

CSIRO Research Scientist & Member of the Australian Government Linked Data Working Group
Outline

- The Aust Govt Linked Data WG
- AGLDWG PIDs
- Initial arrangements
- New arrangements
Australian Government Linked Data WG

Community of Commonwealth Government experts and champions, with invited non-voting participation of individuals, corporations and other entities.

- Established in August 2012, with strong growth in membership since the Government released the outcomes of an inquiry on Data Availability and Use in the Australian Government.
- No official Government mandate or related legislation, but a community of practice that promotes and represents a series of federal Government entities who seek to implement and use Linked Data technologies for the betterment of Australian Government data sharing.
- Several members have signed an MoU to support the use and persistence of linked.data.gov.au URIs.
Australian Government Linked Data WG

Community of Commonwealth Government experts and champions, with invited non-voting participation of individuals, corporations and other entities.

- Established in August 2012, with strong growth in membership since the Government released the outcomes of an inquiry on Data Availability and Use in the Australian Government.
- No official Government mandate or related legislation, but a community of practice that promotes and represents a series of federal Government entities who seek to implement and use Linked Data technologies for the betterment of Australian Government data sharing.
- Several members have signed an MoU to support the persistence of linked.data.gov.au URIs.
Things the AGLDWG Does

- Facilitate the Aust Govt LD community
- Engage with WoG data initiatives
- Manage persistent URIs for resources
  - and provide Guidelines for their use
- Starting to register Linked Data resources
  - Datasets
  - Ontologies
  - Others…
AGLDWG PIDs – Initial Arrangements

• It is seen as essential for useful Linked Data to have persistent URIs

• A PID regime started in 2013 provided for URIs within 26 subdomains of data.gov.au:
  – Top-level AGIFT terms used to segment management
    • environment
    • maritime
    • infrastructure
    • …

• Other domain segmentation vocabs considered, e.g. COFOG

• Regime has run until now

• The non-AGIFT subdomain reference.data.gov.au added in 2015 to cater for general LD resources

• Only environment.data.gov.au has been well utilised, reference a little bit
AGLDWG PIDs – Initial Arrangements

- {environment, reference, infrastructure}.data.gov.au managed with the PID Svc tool
  - A CSIRO-developed management UI for a proxy server
  - Scales to large numbers of PID mappings
  - Can be edited without low-level server access

- Individual resources (ontology files) and whole collections of resources (APIs) have been allocated persistent URIs

- Important datasets include: ACORN-SAT & Geofabric

- Important definitional resources include:
  - AGIFT
  - Dataset Ontology
  - BA Glossary

The UI of the PID Svc managing environment.data.gov.au
AGLDWG PIDs – Initial Arrangements

• Questions about how to deal with non-AGIFT terms were always anticipated

• In 2017, URIs about spatial data were requested for the Linked Data version of the Geocoded National Address File (G-NAF: http://gnafld.net)

• With ‘statistics’ being added to AGIFT, ‘spatial’ was requested and ultimately received
  – A need to do things differently to AGIFT management segmentation was perceived
  – Some members of the WG felt uncomfortable with thematic classifiers in URIs, even for management

The National Archive’s AGIFT as an online vocabulary

**Australian Governments' Interactive Functions Thesaurus (AGIFT)**

**Subject**
The Australian Governments' Interactive Functions Thesaurus (AGIFT) is a three-level hierarchical thesaurus that describes the business functions carried out across Commonwealth, state and local governments in Australia.

**Top Concept**
- BUSINESS-SUPPORT-AND-REGULATION
- CIVIC-INFRASTRUCTURE
- COMMUNICATIONS
- COMMUNITY-SERVICES
AGLDWG PIDs – New Arrangements

Aims of the New Arrangements:

- Work within a new MoU arrangement across agencies to lock down a namespace
- Avoid indicating classification of resources in URI
- Avoid constricting top-level vocabulary
- Make 1st round allocation easy
  - Anyone can easily get a dataset or definitional PID
- Provide for a 2nd-round of class-based register URIs
  - Maximally re-use top-level registers for resources of the same type
AGLDWG PIDs – New Arrangements

- New PID URI arrangements proposed in Feb, 2018. Currently under test

- Proposed:
  - linked.data.gov.au to be used for all PIDs
    - No other subdomains used
AGLDWG PIDs – New Arrangements

• New PID URI arrangements proposed in Feb, 2018. Currently under test

• Proposed:
  – linked.data.gov.au to be used for all PIDs
    • No other subdomains used

The National Archive’s AGIFT as an online vocabulary

Australian Governments' Interactive Functions Thesaurus (AGIFT)

Subject
The Australian Governments' Interactive Functions Thesaurus (AGIFT) is a three-level hierarchical thesaurus that describes the business functions carried out across Commonwealth, state and local governments in Australia.

Top Concept
BUSINESS-SUPPORT-AND-REGULATION
CIVIC-INFRASTRUCTURE
COMMUNICATIONS
COMMUNITY-SERVICES
AGLDWG PIDs – New Arrangements

• New PID URI arrangements proposed in Feb, 2018. Currently under test

• Proposed:
  – linked.data.gov.au to be used for all PIDs
    • No other subdomains used
  – Datasets to be allocated URIs at \( \ldots/dataset/\{NAME\} \)
  – Definitional resources with URIs at \( \ldots/def/\{NAME\} \)
  – Top-level registers at \( \ldots/\{CLASS-NAME\}/ \)

The National Archive’s AGIFT as an online vocabulary

Australian Governments’ Interactive Functions Thesaurus (AGIFT)

Subject
The Australian Governments’ Interactive Functions Thesaurus (AGIFT) is a three-level hierarchical thesaurus that describes the business functions carried out across Commonwealth, state and local governments in Australia.

Top Concept
BUSINESS-SUPPORT-AND-REGULATION
CIVIC-INFRASTRUCTURE
COMMUNICATIONS
COMMUNITY-SERVICES
AGLDWG PIDs – New Arrangements

• New PID URI arrangements proposed in Feb, 2018. Currently under test

• **Proposed:**
  - linked.data.gov.au to be used for all PIDs
    • No other subdomains used
  - **Datasets** to be allocated URIs at …/dataset/{NAME}
  - **Definitional resources** with URIs at …/def/{NAME}
  - **Top-level registers** at …/CLASS-NAME/
  - Individual items either within datasets or within top-level registers:
    • …/dataset/{NAME}/{REGISTER}/{ITEM-ID}
    • …/CLASS-NAME/{ITEM-ID}

The National Archive’s AGIFT as an online vocabulary
AGLDWG PIDs – New Arrangements

• New PID URI arrangements proposed in Feb, 2018. Currently under test

• **Proposed:**
  – **linked.data.gov.au** to be used for all PIDs
    • No other subdomains used
  – **Datasets** to be allocated URIs at …/dataset/{NAME}
  – **Definitional resources** with URIs at …/def/{NAME}
  – **Top-level registers** at …/{CLASS-NAME}/

  – Individual items either within datasets or within top-level registers:
    • …/dataset/{NAME}/{REGISTER}/{ITEM-ID}
    • …/{CLASS-NAME}/{ITEM-ID}

Anyone can get these by registering a dataset

You get these by registering things with accepted definitions
AGLDWG PIDs – New Arrangements

- A dataset now in test: AU Orgs Register

AU Organisations Register

This is a demonstration system only: 'test.' in web addresses will be removed in production

This is the Linked Data version of directory.gov.au. It allows web requests to be made for directory.gov.au content (no search, direct links only!) and responds with either simple web pages (HTML) or Resource Description Framework (RDF), machine-readable, data. Internet protocol content negotiation can be used to request the different formats.
AGLDWG PIDs – New Arrangements

• A dataset now in test: AU Orgs Register
  
  
AGLDWG PIDs – New Arrangements

- A dataset now in test: AU Orgs Register

Don’t have to have an ontology per dataset, this one just does
AGLDWG PIDs – New Arrangements

- A dataset now in test: AU Orgs Register
    - Defined as org:Organization
    - Expected to share register with other orgs in the future
AGLDWG PIDs – New Arrangements

- A dataset now in test: AU Orgs Register

  - Defined as org:Organization
  - Expected to share register with other orgs in the future

Things in the register could have any URI. The /org/ here is just for the register of things of a particular class.
AGLDWG PIDs – New Arrangements

- A dataset now in test: AU Orgs Register


- Items in the AU Orgs Register dataset appear in top-level and 2nd-level registers:
  - .../org/ - org:Organisations
  - .../dataset/auorg/subStructure/
AGLDWG PIDs – New Arrangements

Timeline:

- Mar 2018: Proposed
- May 2018: demonstrated
- June 2018: New Arrangements URI Guidelines to be issued
- July 2018: AGLDWG vote to proceed with New Arrangements
- July 2018+: New Arrangement-style URIs allocated
Contacts

AGLDWG:  http://linked.data.gov.au
Simon:    simon.cox@csiro.au
Nick:      nicholas.car@csiro.au

Questions?