

Hands-on Guide To Linked Data Applications

Tutorial at ISWC 13

Liliana Cabral, Heiko Mueller, Ahsan Morshed

22 October 2013

CSIRO COMPUTATIONAL INFORMATICS, HOBART, TAS, AUSTRALIA
www.csiro.au



Tutorial Goal

Provide a practical introduction into building Linked Open data (LOD) applications using open data sets and existing tools.

Content

- LOD Basics
- LOD Application Development Workflow
 - Generic steps and tools to build an application
- Hands-on example of LOD Application
 - Linking Tasmania Geo-spatial Environmental Features

Tutorial Assumptions

- Attendees will have a basic knowledge of RDF
- Attendees will have a basic knowledge of the Linked Open Data (LOD) principles and the LOD cloud
- Only a incomplete number of (free to use) tools will be suggested for application development. Example will be based in a few selected ones.

LOD Basics

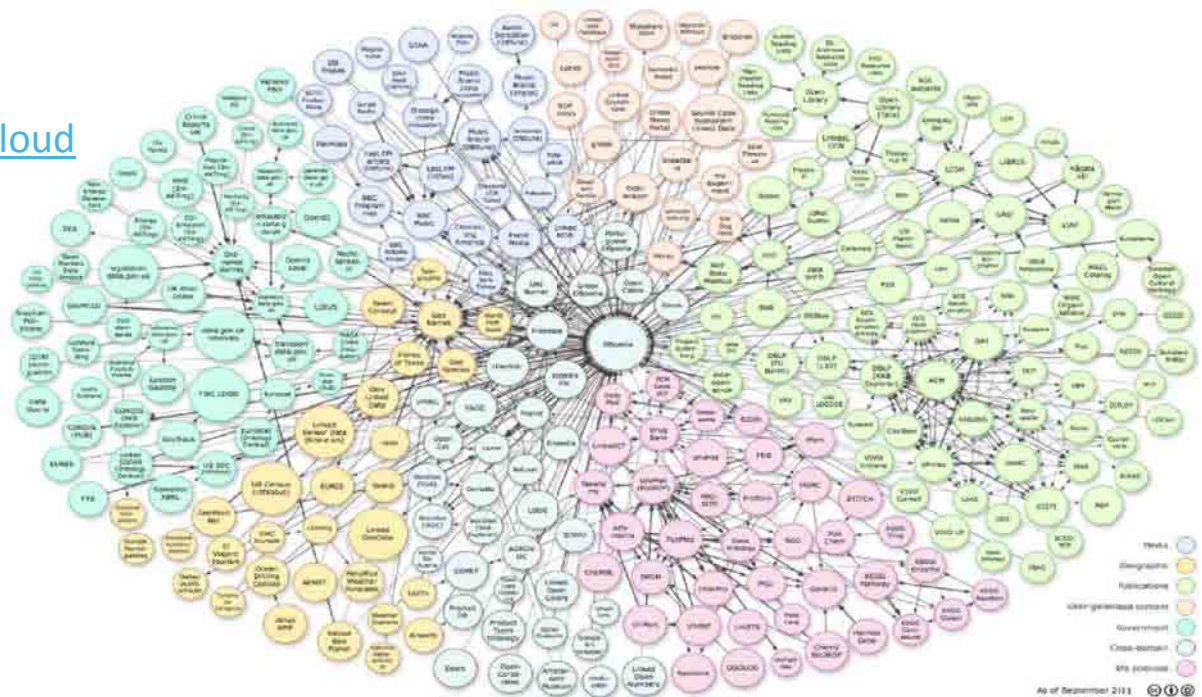
Linked Open Data

*It is about linking data on the Web. A term used to describe a recommended best practice for exposing, sharing, and connecting pieces of **data**, **information**, and **knowledge** on the Semantic Web using **URIs** and **RDF**. [Wikipedia]*

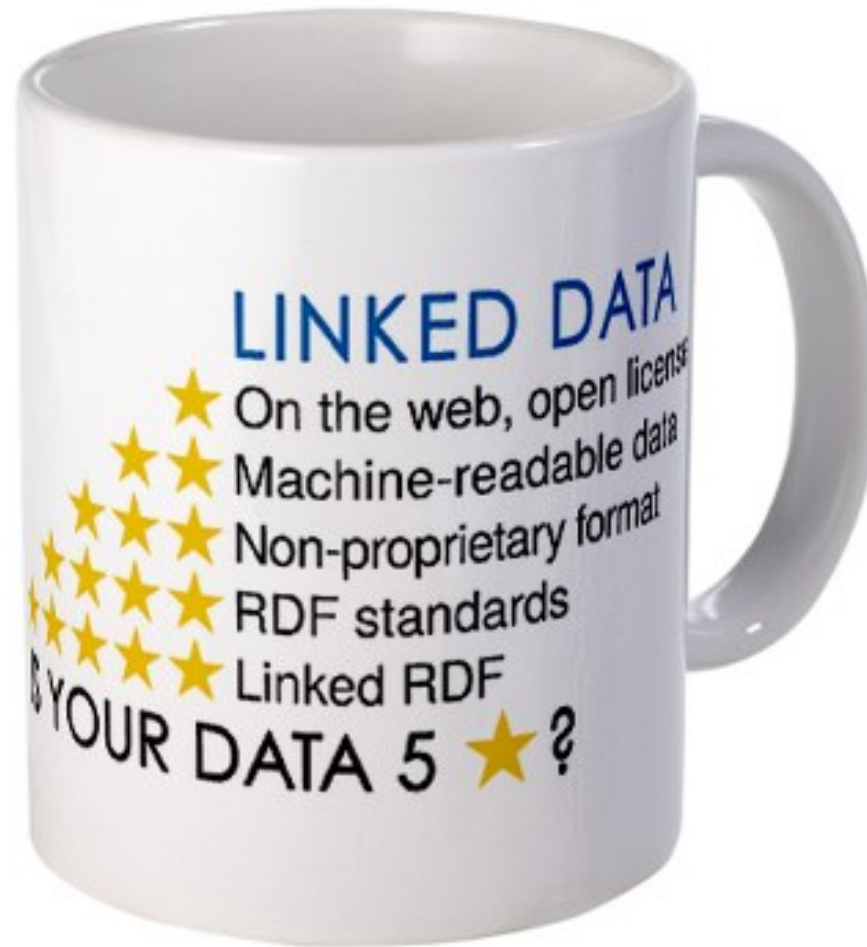
The Linked Open Data Cloud

<http://linkeddata.org>

<http://datahub.io/group/lodcloud>



Linked Data Principles



RDF

<http://www.w3.org/TR/2004/REC-rdf-primer-20040210/>

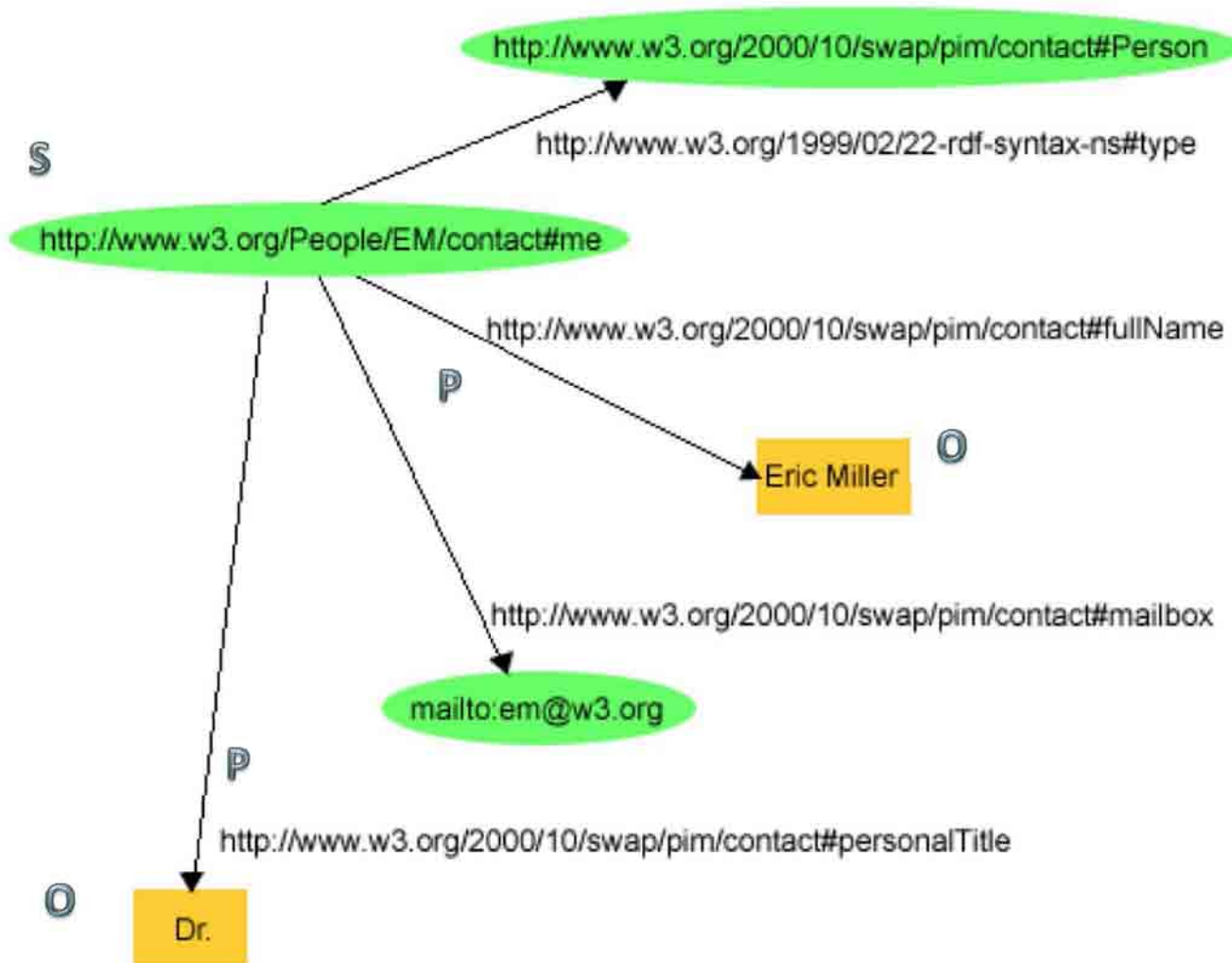


Figure 1: An RDF Graph Describing Eric Miller

RDF

Triple: <Subject, Predicate, Object>

Figure 1 illustrates that RDF uses URIs to identify:

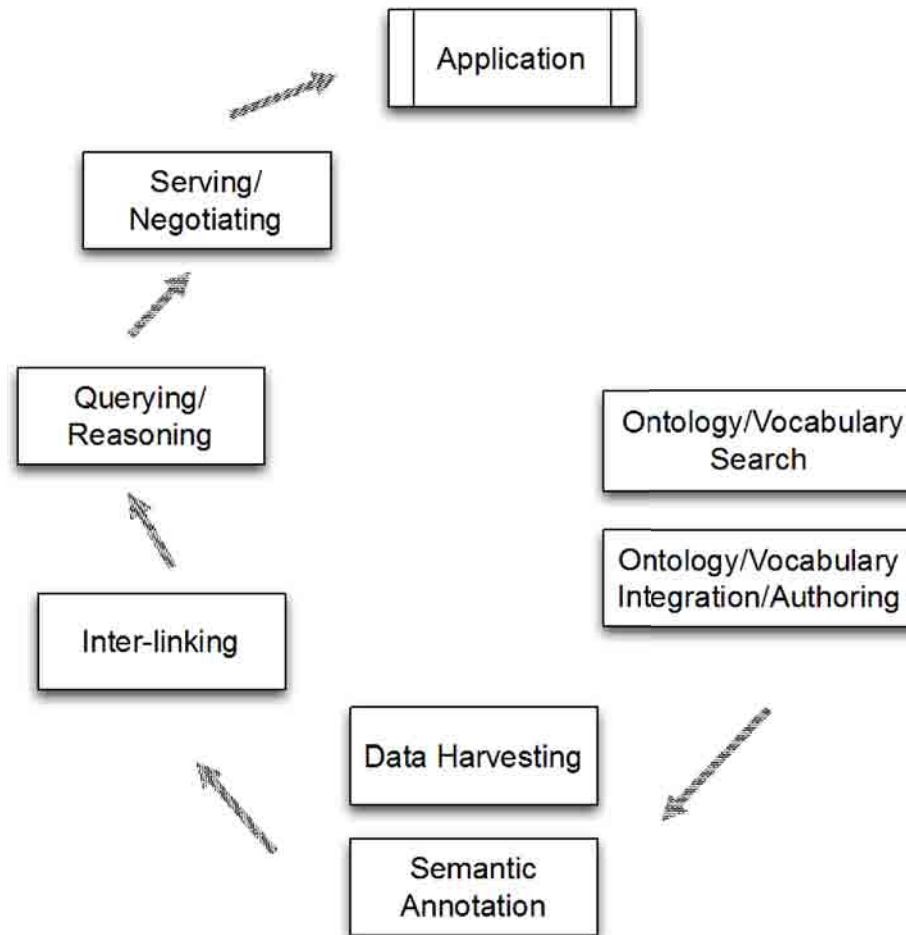
- S** • individuals, e.g., Eric Miller, identified by <http://www.w3.org/People/EM/contact#me>
- P** • kinds of things, e.g., Person, identified by <http://www.w3.org/2000/10/swap/pim/contact#Person> **O**
- P** • properties of those things, e.g., mailbox, identified by <http://www.w3.org/2000/10/swap/pim/contact#mailbox>
- O** • values of those properties, e.g. <mailto:em@w3.org> as the value of the mailbox property (RDF also uses character strings such as "Eric Miller", and values from other datatypes such as integers and dates, as the values of properties)

Ontology examples

- **Ontology for DBpedia (Wikipedia)**
 - <http://wiki.dbpedia.org/Ontology>
 - <http://dbpedia.org/ontology/>
 - <http://mappings.dbpedia.org/server/ontology/classes/>
 - <http://dbpedia.org/sparql>
- **Friend-of-a-friend (FOAF) Vocabulary**
 - <http://xmlns.com/foaf/spec/>
- **Ontology for E-commerce**
 - <http://www.heppnetz.de/ontologies/goodrelations/v1>
- **Ontology for Academic conferences**
 - <http://data.semanticweb.org/ns/swc/ontology>
- **Ontologies from Semantic.org**
 - <http://semanticweb.org/wiki/Ontology>

LOD Application Development Workflow

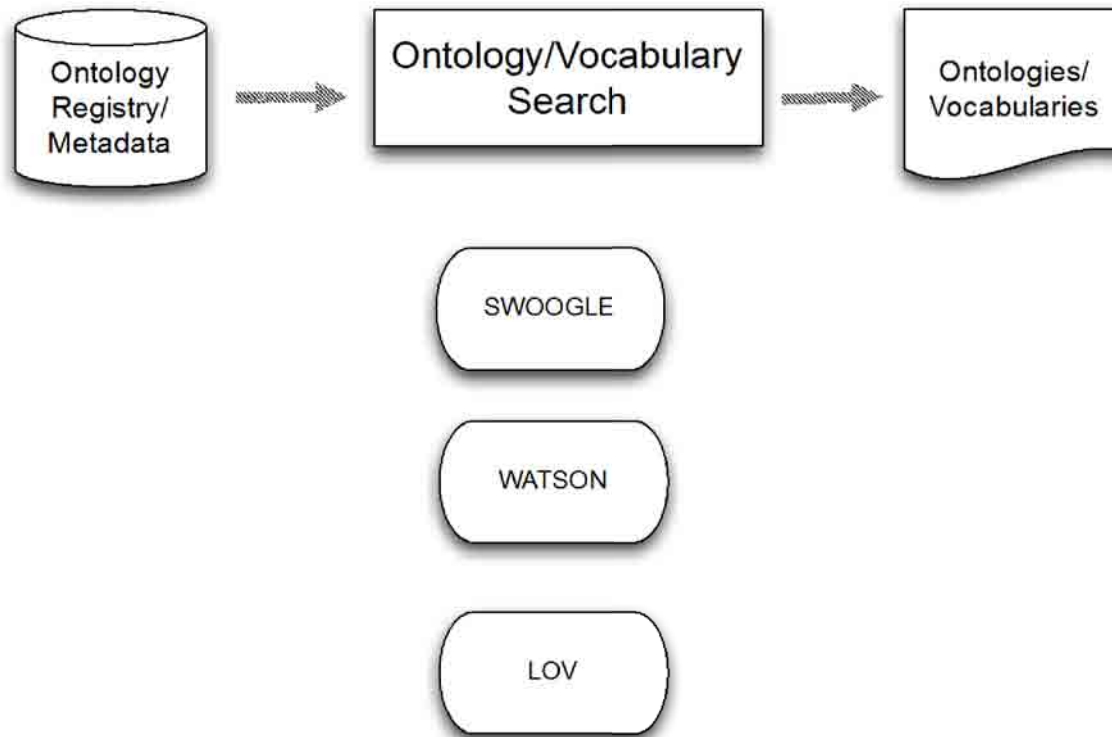
LOD Application Development Workflow



LOD Application Development Workflow

- We present a loose set of activities that a software developer will perform, not necessarily in sequence, and sometimes more than once to build an application.
- For each activity, we present inputs and outputs along with a list of suggested tools.
- The tools are used to complete part of an activity. Further setup, cleaning or programming may be needed.
- Some tools (suite) can be used in more than one activity
- The use of tools depend on the nature of the application and complexity of source data sets.

Ontology/Vocabulary Search



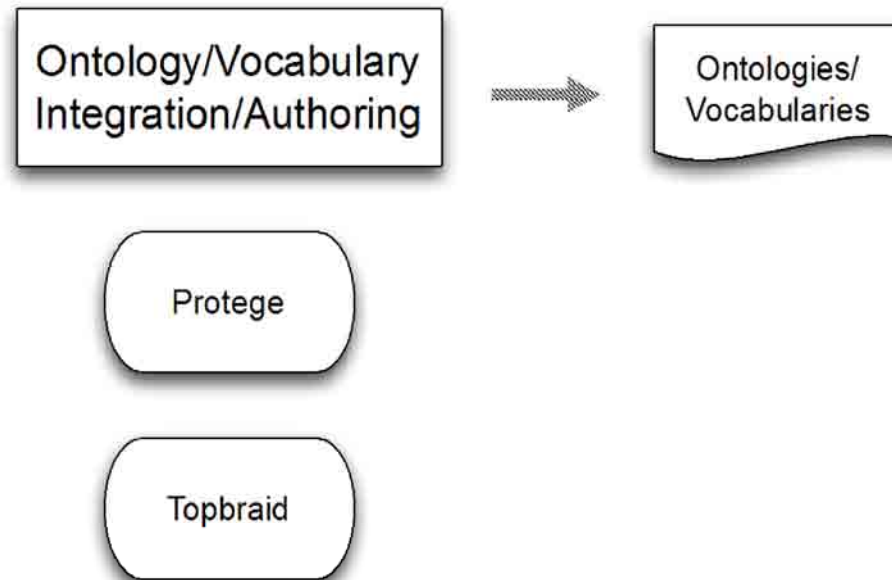
Ontology/Vocabulary Search

- Tools for ontology/vocabulary search provide:
 - A registry of vocabulary metadata
 - A search mechanism
 - Output is a list of metadata and URLs to ontologies or related concepts, which match the search input against registered ontologies.

Ontology/Vocabulary Search

- SWOOGLE
 - <http://swoogle.umbc.edu>
- Watson
 - <http://watson.kmi.open.ac.uk/WatsonWUI/>
- LOV
 - <http://lov.okfn.org/dataset/lov/>

Ontology/Vocabulary Integration/Authoring



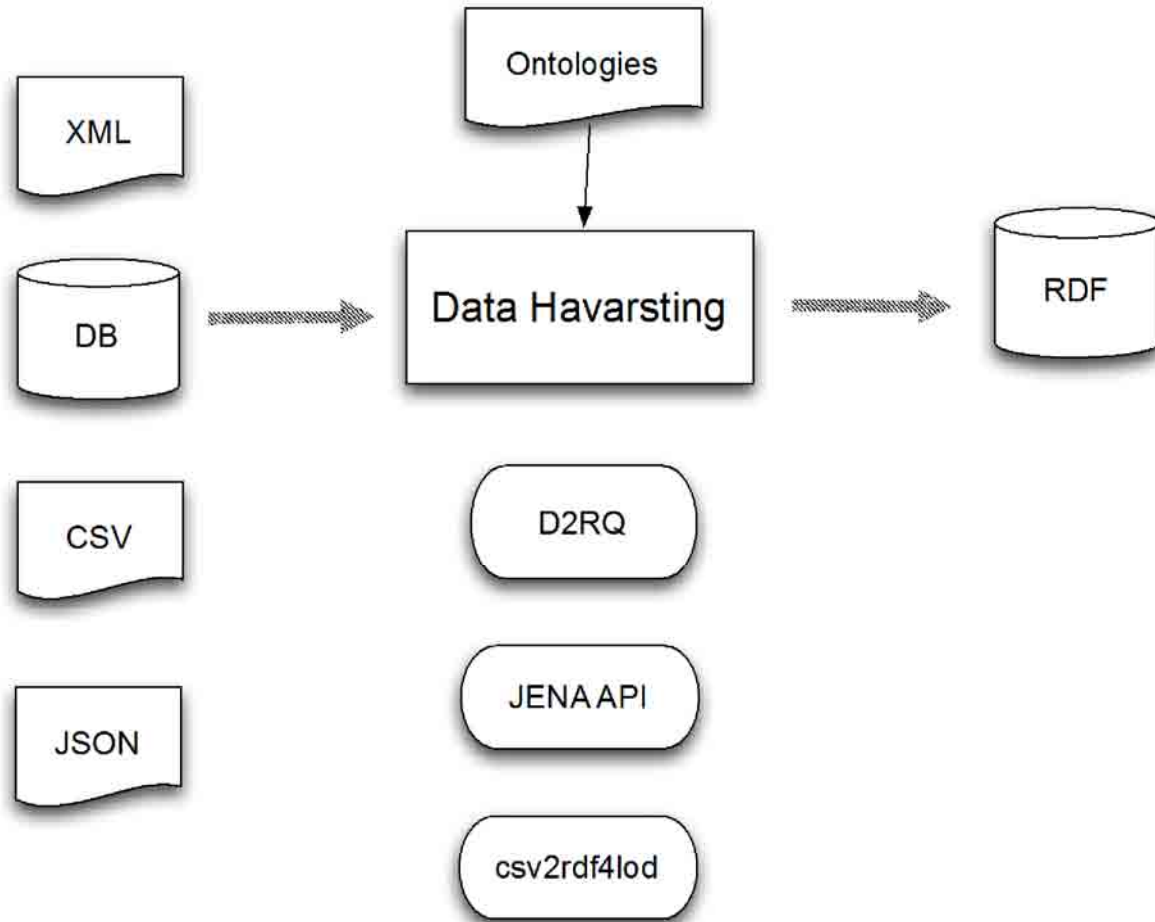
Ontology/Vocabulary Integration/Authoring

- Tools for vocabulary authoring provide an environment for:
 - Editing, validating, instantiating ontologies
 - Importing existing ontologies
 - Transforming schemas (e.g. XML) and other formalisms into ontologies
- Output is an ontology with a defined namespace

Ontology/Vocabulary Integration/Authoring

- TopBraid Composer
 - http://www.topquadrant.com/products/TB_Composer.html
- Protege
 - <http://protege.stanford.edu>

Data Harvesting



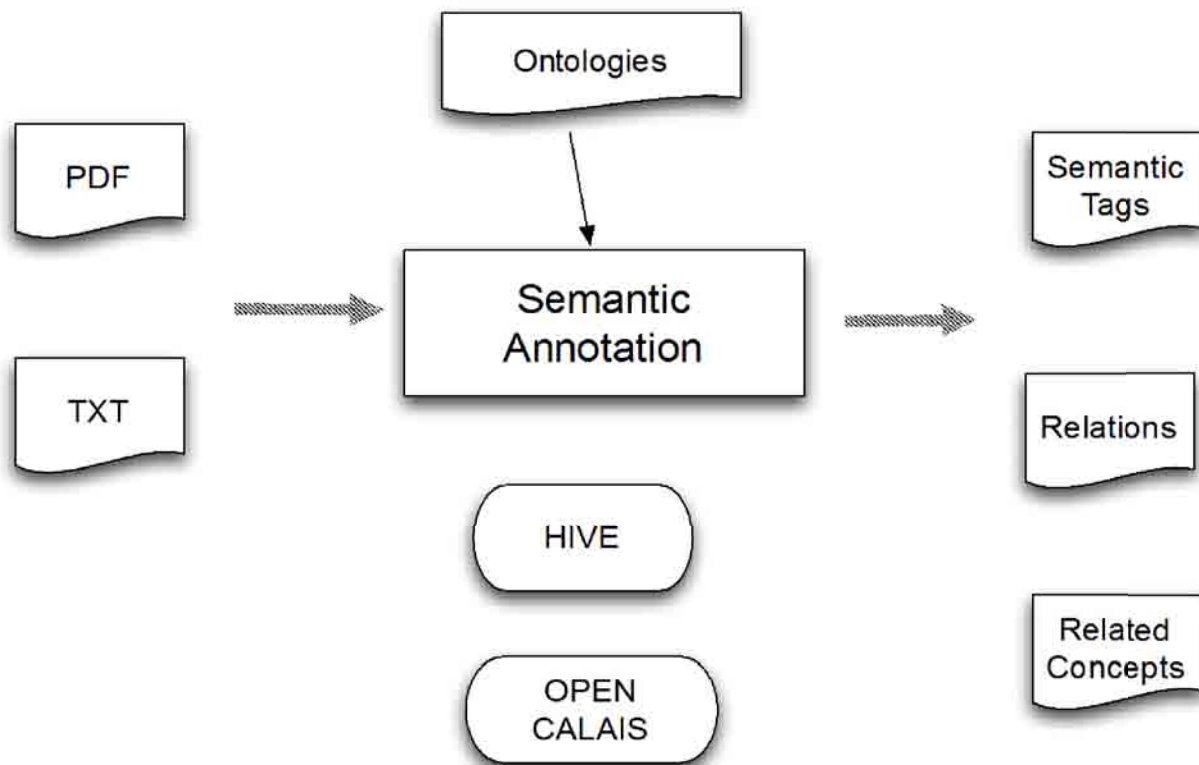
Data Harvesting

- Data harvesting tools :
 - Have as input structured documents, whose structure can be mapped to ontology concepts and relations, and data can be instantiated according to given ontologies
 - Provide a mapping mechanism
 - Can be based on a standard language such as R2RML
 - <http://www.w3.org/TR/2012/REC-r2rml-20120927/#overview>
 - Can be part of a Database plugin for automated RDF generation and query rewriting, for example as in
 - Virtuoso (<http://virtuoso.openlinksw.com/rdf-quad-store/>),
 - AllegroGraph (<http://www.franz.com/agraph/allegrograph/>).
- Output is RDF data according to input documents, which can be stored/dumped into a RDF triple store (e.g. SESAME)

Data Harvesting

- D2RQ
 - <http://d2rq.org>
- JENA RDF API
 - <http://jena.apache.org>
- CSV2RDF4LOD
 - <https://github.com/timrdf/csv2rdf4lod-automation/wiki>

Semantic Annotation



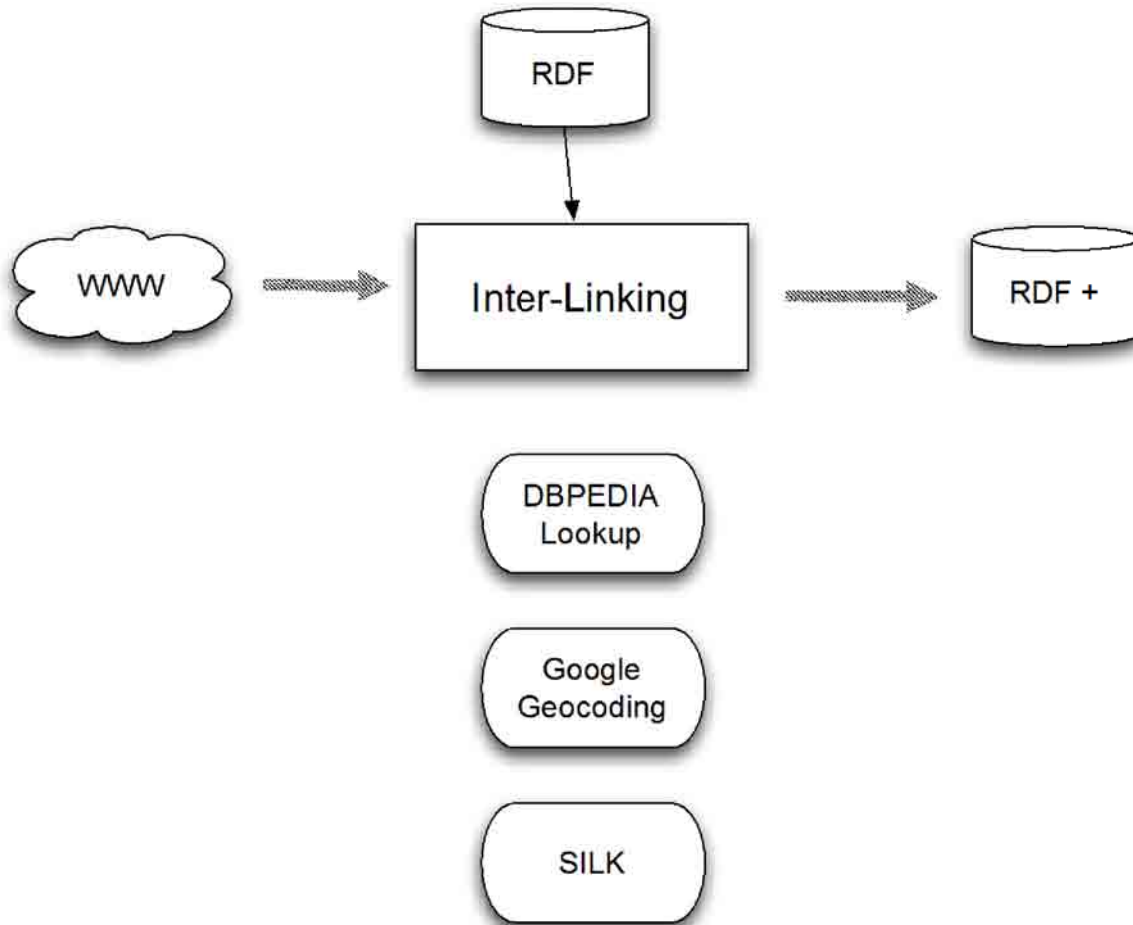
Semantic Annotation

- Semantic annotation:
 - Have as input un-structured documents.
 - Provide a similarity matching algorithm, which must be trained with specific ontologies, for example
 - KEA (<http://www.nzdl.org/Kea/description.html>)
- Output is a list of semantic tags or semantic relationships expressed in terms of a given ontology (e.g. SKOS).

Semantic Annotation

- HIVE
 - <https://code.google.com/p/hive-mrc/>
 - <http://hive.nescent.org/home.html>
- Open Calais
 - <http://www.opencalais.com/about>

Inter-linking



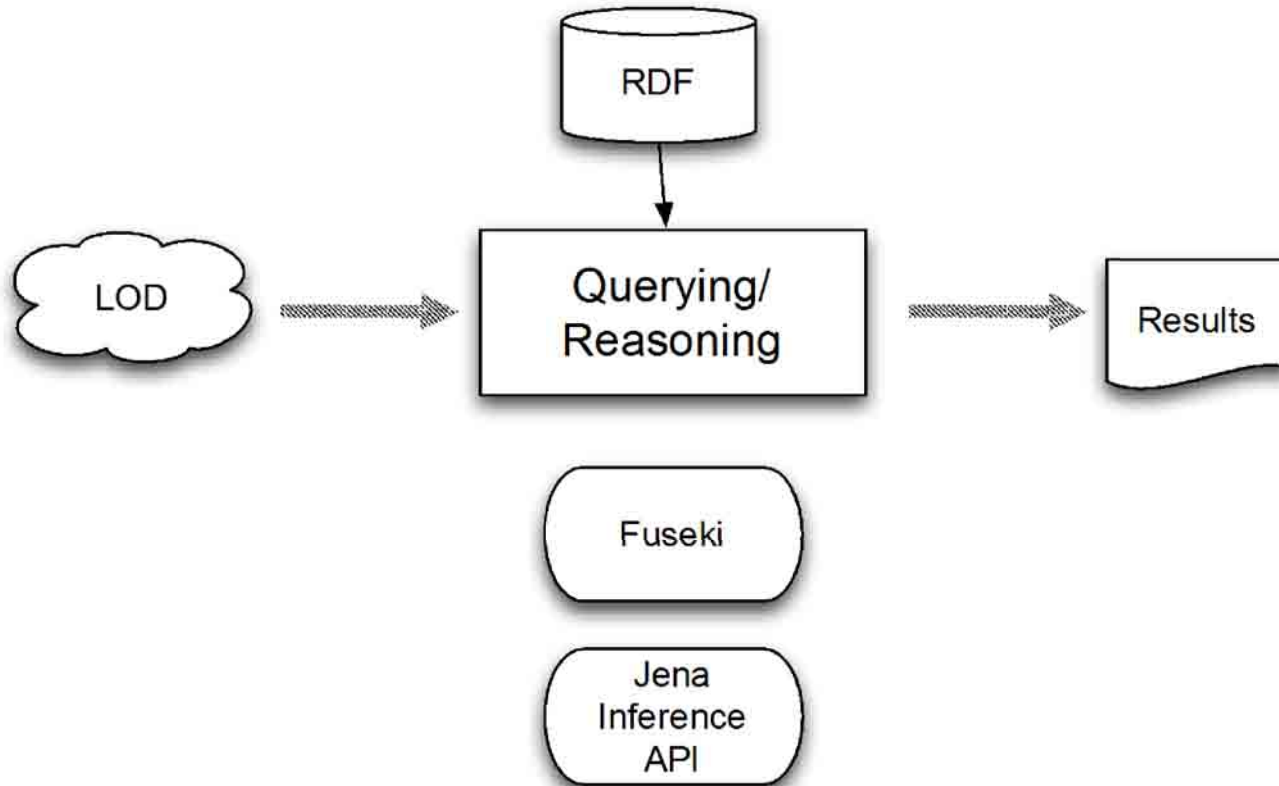
Inter-linking

- Interlinking tools can provide:
 - Lookup functions over a dataset.
 - Similarity matching algorithms to find degrees of relationship such as equality (e.g. same as), inclusion (e.g. broader) between resources
- Output is a set of new relations, which link existing subjects to external resources.

Inter-linking

- DBPedia Lookup
 - <https://github.com/dbpedia/lookup>
- Google Geocoding
 - <https://developers.google.com/maps/documentation/geocoding/>
- SILK
 - <http://wifo5-03.informatik.uni-mannheim.de/bizer/silk/>

Querying/Reasoning



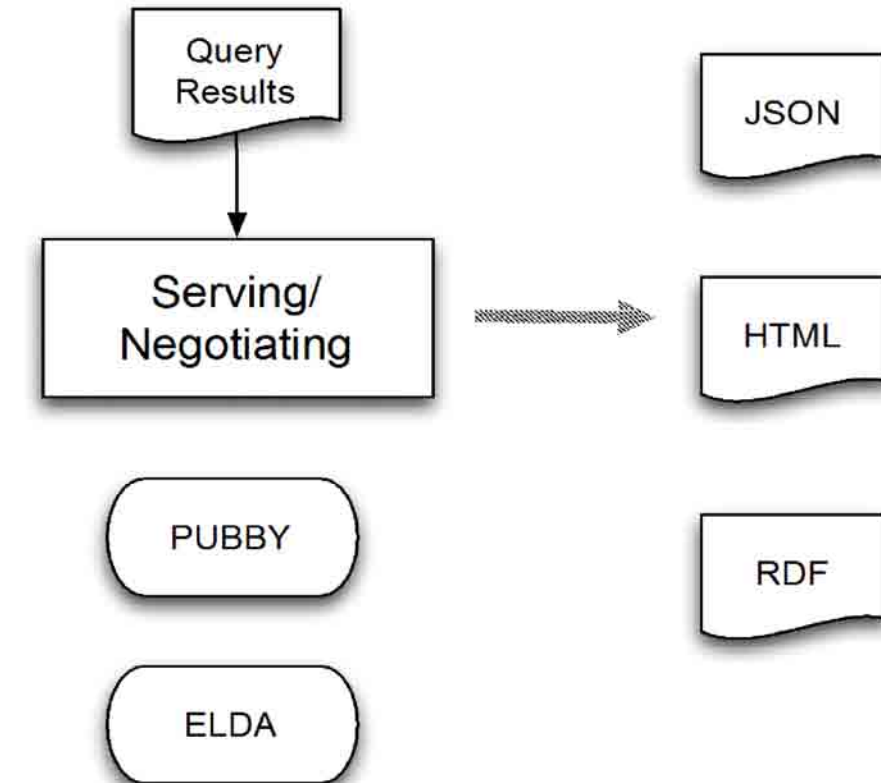
Querying/Reasoning

- Any tool that uses SPARQL
 - It can be part of another tool, e.g. Snorql over D2RQ
- Any tool that uses OWL/RDF reasoners
- Output is a table with results of the query, which can be converted to different formats (e.g. JSON)

Querying/Reasoning

- Fuseki
 - http://jena.apache.org/documentation/serving_data/
- Jena Inference API
 - <http://jena.apache.org/documentation/inference/index.html>

Serving/Negotiating



Serving/Negotiating

- Tools that serve or negotiate RDF provide:
 - A Web Server
 - Content negotiation provides access to data in different formats
 - Defined namespace
 - SPARQL endpoint
 - A dataset that can be exposed to the LOD Cloud following LOD principles
 - All elements have a URL that can be dereferenced
 - It can be part of another tool, e.g. D2RQ

Serving/Negotiating

- Pubby
 - <http://wifo5-03.informatik.uni-mannheim.de/pubby/>
- ELDA
 - <http://code.google.com/p/elda/>

LOD Application Example

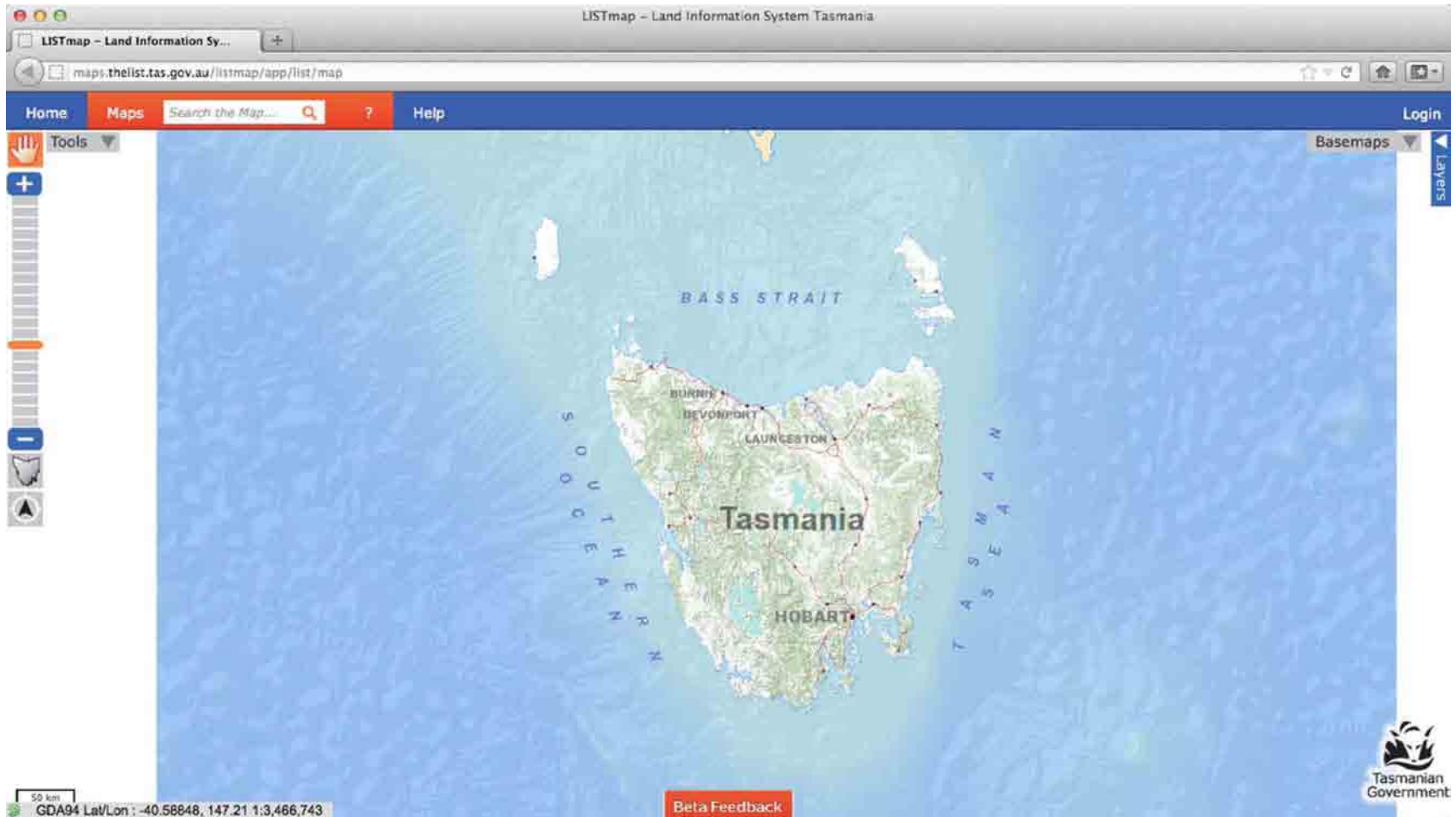
Linking Tasmania Geo-spatial Environmental Features

- Goal:
 - Access environmental features about Tasmania from Geo-spatial infrastructures and create an LOD dataset
 - Create an application that will provide integrated information
- Geo-spatial infrastructures provide:
 - Access to a registry of geographic content (e.g. ArcGIS)
 - Mapping services – Maps according to geo location
 - Feature services – Features according to geo location
 - Standards for naming and interoperability

Linking Tasmania Geo-spatial Environmental Features

- Examples of geo-spatial applications:
 - The List – Tasmania Government
 - <http://www.dpiw.tas.gov.au/inter-nsf/WebPages/LBUN-59H3DY?open>
 - <http://maps.thelist.tas.gov.au/listmap/app/list/map>
 - <http://services.thelist.tas.gov.au/geoserver/web/?wicket:bookmarkablePage=:org.geoserver.web.demo.MapPreviewPage>
 - Gazetteer Australia
 - <http://www.ga.gov.au/search/index.html#/>
 - <http://www.ga.gov.au/darwin-gazetteer/index.xhtml>

The List – Tasmania Maps and Features



The List – Tasmania Maps and Features

Identify Results

One feature found in one layer

~ Rivers Streams and Creeks (one feature)

Feature	
Name	Coal River
Nomenclature Register Number	1389X
Hydro Class	River
Hyd Line Type1	Watercourse
Hyd Line Type2	Definite
Existing	N/A
MHWM Type	N/A
In Use	N/A
Hyd Line ID	626115
RELGRND	N/A

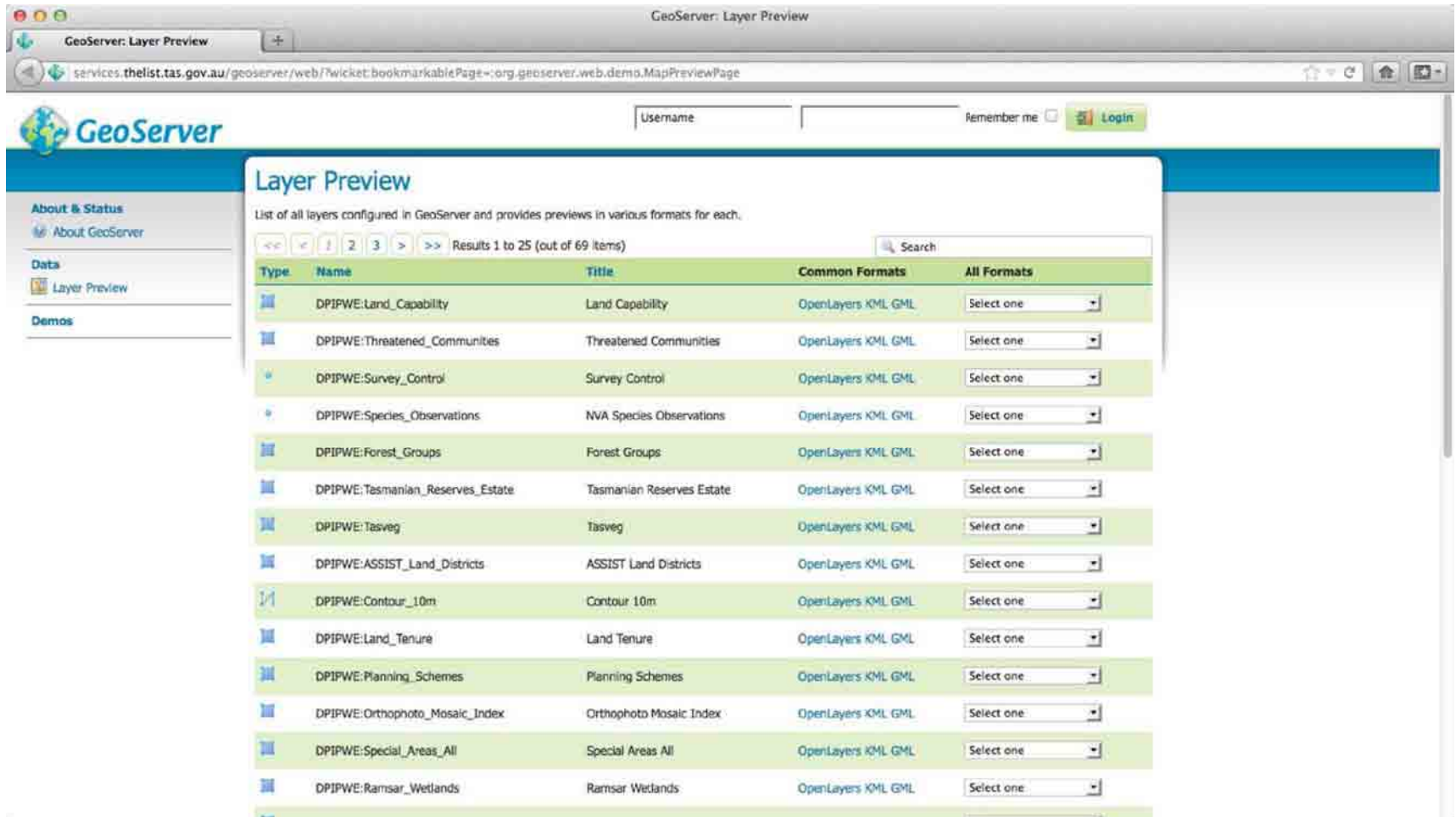
Search Results for 'vineyard'

142 features found

- Clemens Hill Vineyard, Cambridge
- Clover Hill Vineyard, Lebrina
- Clover Hill Vineyard, Lebrina
- Clover Hill Vineyard, Lebrina
- Coal River Rise Vineyard, Campania
- Coal Valley Vineyard, Cambridge
- Coal Valley Vineyard, Cambridge
- Coolinda Vale Vineyard / Pooley Wines, Campania
- Craigow Vineyard, Cambridge
- Craigow Vineyard, Cambridge

Previous 1 2 3 ... 15 Next

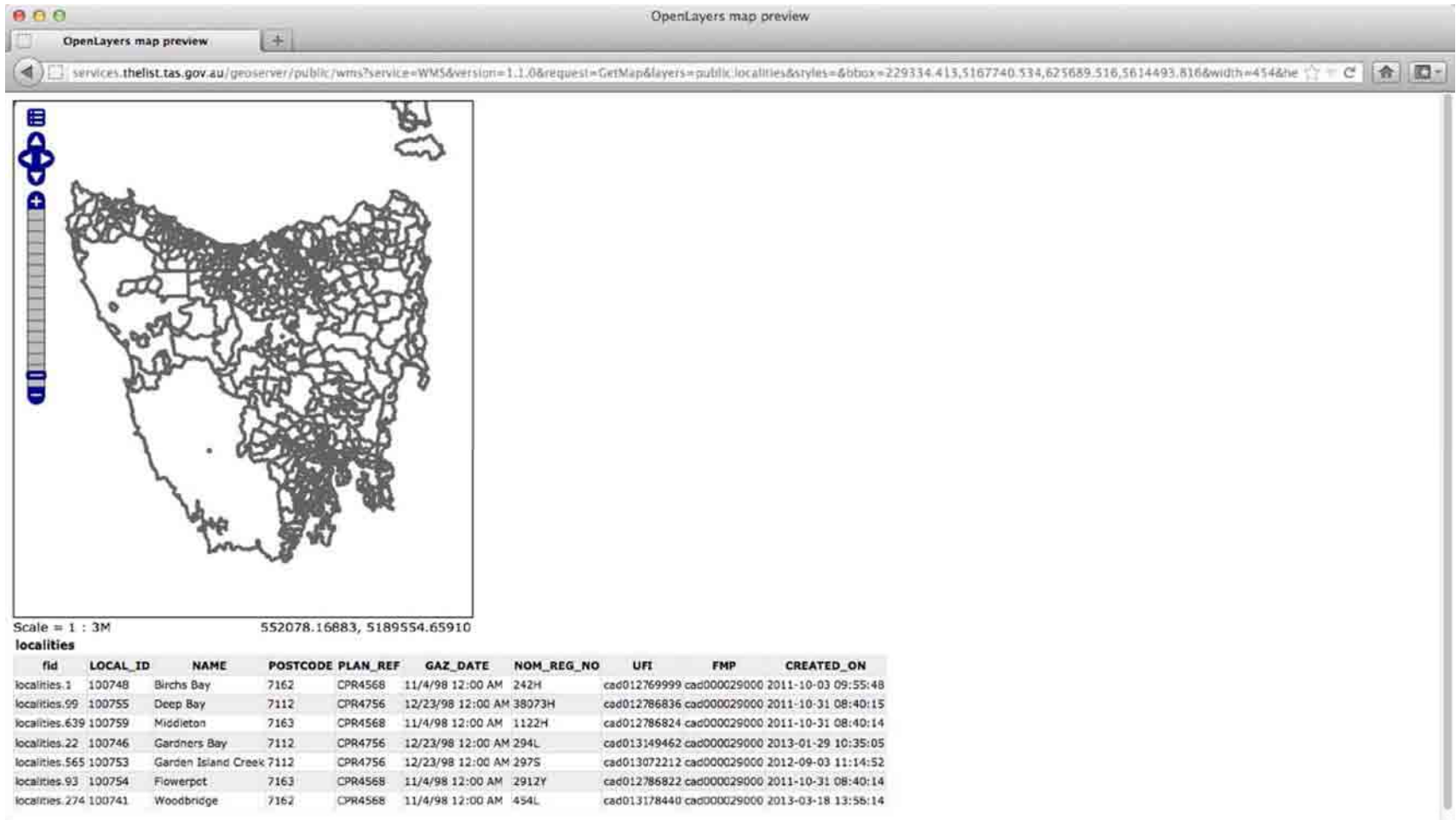
The List services – Tasmania Maps and Features



The screenshot shows the GeoServer Layer Preview interface. The browser address bar displays the URL: `services.thelist.tas.gov.au/geoserver/web/?wicket:bookmarkablePage=:org.geoserver.web.demos.MapPreviewPage`. The page title is "GeoServer: Layer Preview". The main content area is titled "Layer Preview" and contains a list of layers configured in GeoServer. The list includes a search bar and pagination controls showing "Results 1 to 25 (out of 69 items)". The layers are listed in a table with columns: Type, Name, Title, Common Formats, and All Formats. The layers are:

Type	Name	Title	Common Formats	All Formats
DPIPWE:Land_Capability	Land Capability	OpenLayers KML GML	Select one	
DPIPWE:Threatened_Communities	Threatened Communities	OpenLayers KML GML	Select one	
DPIPWE:Survey_Control	Survey Control	OpenLayers KML GML	Select one	
DPIPWE:Species_Observations	NVA Species Observations	OpenLayers KML GML	Select one	
DPIPWE:Forest_Groups	Forest Groups	OpenLayers KML GML	Select one	
DPIPWE:Tasmanian_Reserves_Estate	Tasmanian Reserves Estate	OpenLayers KML GML	Select one	
DPIPWE:Tasveg	Tasveg	OpenLayers KML GML	Select one	
DPIPWE:ASSIST_Land_Districts	ASSIST Land Districts	OpenLayers KML GML	Select one	
DPIPWE:Contour_10m	Contour 10m	OpenLayers KML GML	Select one	
DPIPWE:Land_Tenure	Land Tenure	OpenLayers KML GML	Select one	
DPIPWE:Planning_Schemes	Planning Schemes	OpenLayers KML GML	Select one	
DPIPWE:Orthophoto_Mosaic_Index	Orthophoto Mosaic Index	OpenLayers KML GML	Select one	
DPIPWE:Special_Areas_All	Special Areas All	OpenLayers KML GML	Select one	
DPIPWE:Ramsar_Wetlands	Ramsar Wetlands	OpenLayers KML GML	Select one	

The List services – Tasmania Maps and Features



Gazetteer Australia – Feature discovery

The screenshot displays the 'Discovery and Delivery System' web application. The browser address bar shows the URL www.ga.gov.au/search/index.html#/bounds. The page header includes the Australian Government Geoscience Australia logo and navigation links: Home, Discovery and Delivery System Home, and Discovery and Delivery System. A search bar contains the text 'tasmania', and a 'Quick Links' dropdown is visible. Below the header, there are tabs for 'Map View', 'Split View', and 'List View'. The 'Map View' tab is active, showing a map of Australia with various regions labeled. The map is centered on Tasmania, with coordinates 'Lat -56.166°, Long -178.951°' displayed. A scale bar indicates 500 km. On the left side, the 'SEARCH CRITERIA' section shows 'What: tasmania' and '100 of 331 results displayed'. Below this, a list of search results is shown, each with a title, a brief description, and a download icon. The results include:

- Crustal architecture of Tasmania based on deep seismic reflection...**
This paper was presented at the 17th Australian Geological Convention, Geological Society of Australia
- NW Tasmania, TAS, 1984 (P494)**
Linespacing for the survey is 500 metres
- Apatite fission track thermochronology of Tasmania**
Legacy product - no abstract available
- The basement elements of Tasmania**
Legacy product - no abstract available
- Modelling Gold Deposition in North-eastern Tasmania Using the El...**
Modelling Gold Deposition in Northeastern Tasmania Using the Elf Geochemical Modelling Program
- Project T3 Tasmania Mineral Province Geoscientific database, 3D ...**
Project T3 Tasmania Mineral Province Geoscientific database, 3D Geological Modelling, Mines and Mineral Prospectivity
- Arthur-Pieman Area. NW Tasmania, TAS, 1996 (P652)**
Linespacing for the survey is 100/200 metres
- North Central Tasmania, Tas, 1999 (P699) - magnetic, radiometric ...**
Linespacing for the survey is 200 metres

At the bottom of the page, there are links for Privacy, Accessibility, Freedom of Information, and a disclaimer. The footer also includes the Creative Commons Attribution (CC BY) license logo and the Department of Resources, Energy and Tourism logo.

Gazetteer Australia – Search - Tasmania (caves)

Discovery and Delivery System – Place Names Search

Discovery and Delivery System ...

www.ga.gov.au/darwin-gazetteer/index.xhtml#

Australian Government
Geoscience Australia

Place Names Search

Quick Links

Map view Split View

Name Where

Lat -42.521, Long 149.587

Retrieved 100 of 182 features

Clear Results

▼ Westmorland Cave (TAS) Lat:-41.610001 Lon:146.389999

State: Tasmania

Feature Code: CAVE (Cave, Blowhole, Cavern, Grotto)

Status: Official

Latitude: 41°36'36"S [Decimal Degrees -41.610001°]

Longitude: 146°23'24"E [Decimal Degrees 146.389999°]

Record ID: TAS09218

Authority: Tasmania

100K MAP No.: 8114

Variant Name(s):

► Leech Pot (TAS) Lat:-41.610001 Lon:146.389999

► Kellys Pot (TAS) Lat:-41.599998 Lon:146.369995

► Prohibition Cave (TAS) Lat:-41.59 Lon:146.320007

► Flowers Pot (TAS) Lat:-41.59 Lon:146.330002

► Hidden Cave (TAS) Lat:-41.59 Lon:146.25

► Waterworks Cave (TAS) Lat:-41.59 Lon:146.350005

► Dangerous Hole (TAS) Lat:-41.599998

Export Results list CSV

Department of Resources, Energy and Tourism

Map data ©2013 Google Imagery ©2013 NASA, TerraMetrics

Privacy Accessibility Freedom of Information

Geoscience Australia

CC BY

Unless otherwise noted, all Geoscience Australia material on this website is licensed under the Creative Commons Attribution 3.0 Australia License.

Linking Tasmania Geo-spatial Environmental Features

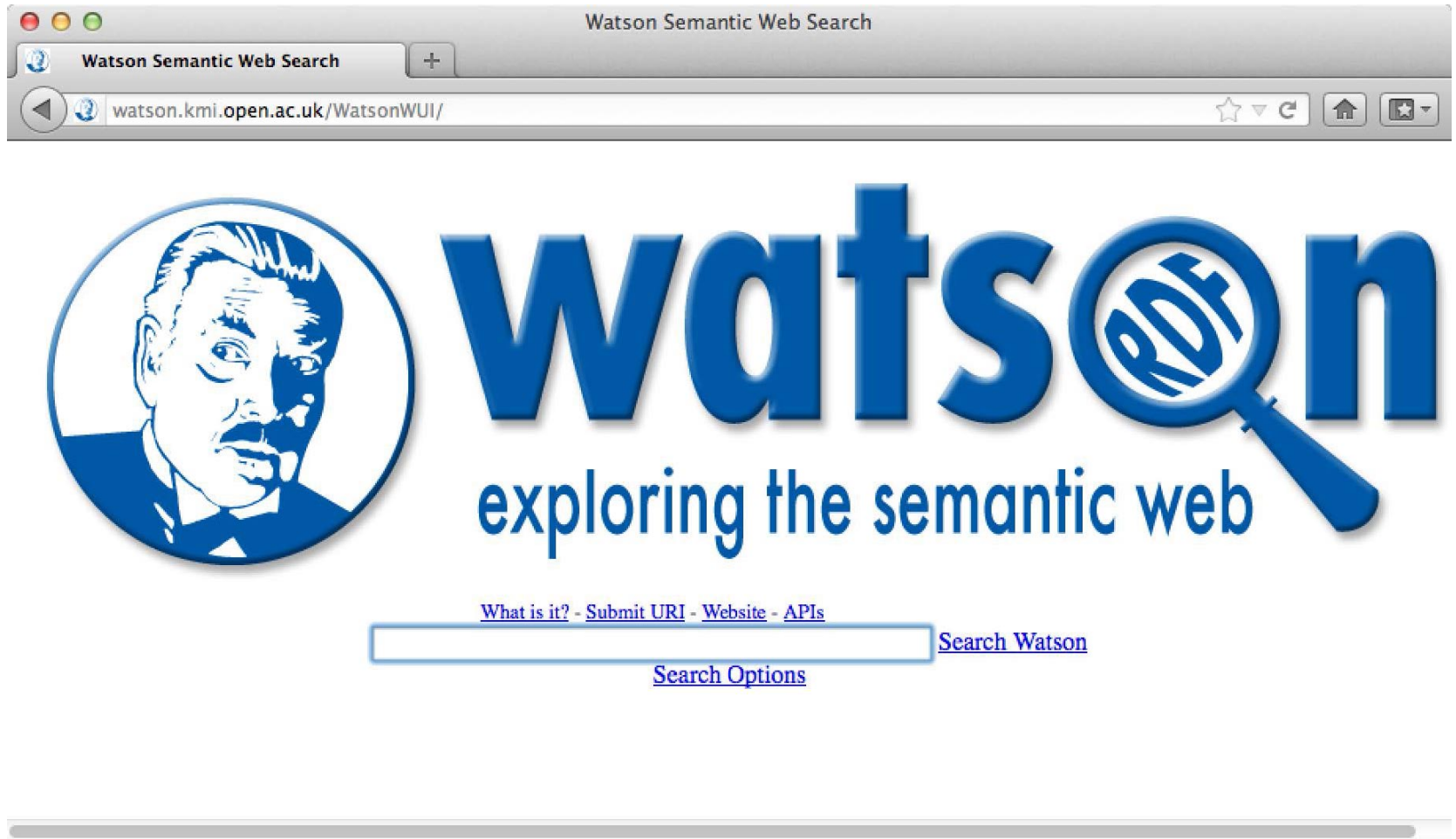
- Dataset Sources:
 - Gazetteer Australia
 - <http://www.ga.gov.au/darwin-gazetteer/index.xhtml>
 - Parks & Reserves; Caves; Landmarks
 - Public Toilets
 - <http://data.gov.au/dataset/national-public-toilet-map>
 - Tasmanian Devil Sightings
 - http://biocache.ala.org.au/occurrences/search?taxa=Sarcophilus+harrisii#tab_recordsView



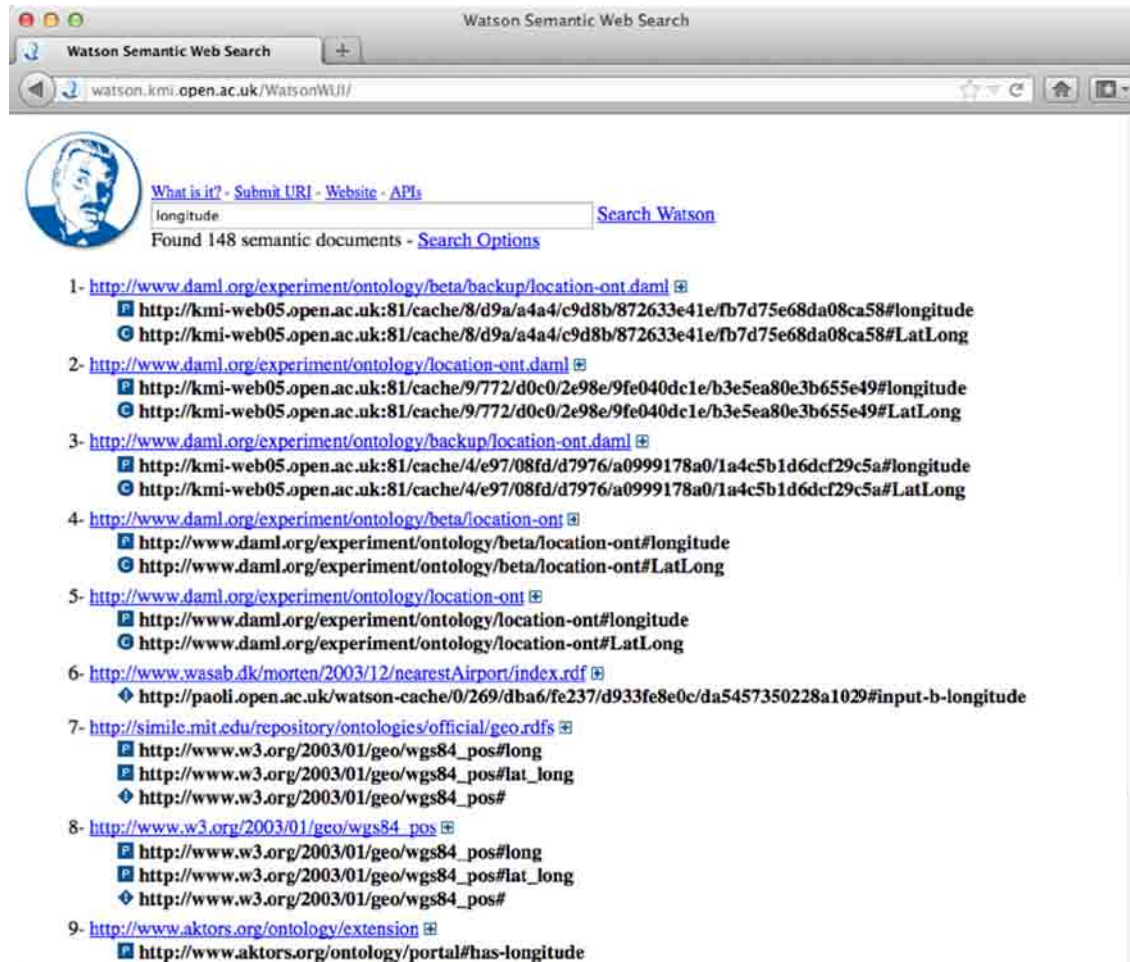
Development Steps – Ontology Search

- From Watson page
 - <http://watson.kmi.open.ac.uk/WatsonWUI/>
 - Type 'geonames'
 - Select <http://www.geonames.org/ontology>
 - http://www.geonames.org/ontology/ontology_v2.0_Lite.rdf
 - Type 'longitude'
 - Select http://www.w3.org/2003/01/geo/wgs84_pos

Development Steps – Ontology Search



Development Steps – Ontology Search



The screenshot displays the Watson Semantic Web Search interface in a web browser. The browser's address bar shows the URL `watson.kmi.open.ac.uk/WatsonWUI/`. The page features a search bar with the text "longitude" and a "Search Watson" button. Below the search bar, it indicates "Found 148 semantic documents - Search Options". A list of 9 search results is displayed, each with a link icon and a URL. The results include links to DAML.org, KMI-web05.open.ac.uk, and W3.org, all related to ontology searches for "longitude".

Watson Semantic Web Search

What is it? - Submit URI - Website - APIs


longitude Search Watson

Found 148 semantic documents - Search Options

- 1- <http://www.daml.org/experiment/ontology/beta/backup/location-ont.daml>
<http://kmi-web05.open.ac.uk:81/cache/8/d9a/a4a4/c9d8b/872633e41e/fb7d75e68da08ca58#longitude>
<http://kmi-web05.open.ac.uk:81/cache/8/d9a/a4a4/c9d8b/872633e41e/fb7d75e68da08ca58#LatLong>
- 2- <http://www.daml.org/experiment/ontology/location-ont.daml>
<http://kmi-web05.open.ac.uk:81/cache/9/772/d0c0/2e98e/9fe040dc1e/b3e5ea80e3b655e49#longitude>
<http://kmi-web05.open.ac.uk:81/cache/9/772/d0c0/2e98e/9fe040dc1e/b3e5ea80e3b655e49#LatLong>
- 3- <http://www.daml.org/experiment/ontology/backup/location-ont.daml>
<http://kmi-web05.open.ac.uk:81/cache/4/e97/08fd/d7976/a0999178a0/1a4c5b1d6dcf29c5a#longitude>
<http://kmi-web05.open.ac.uk:81/cache/4/e97/08fd/d7976/a0999178a0/1a4c5b1d6dcf29c5a#LatLong>
- 4- <http://www.daml.org/experiment/ontology/beta/location-ont>
<http://www.daml.org/experiment/ontology/beta/location-ont#longitude>
<http://www.daml.org/experiment/ontology/beta/location-ont#LatLong>
- 5- <http://www.daml.org/experiment/ontology/location-ont>
<http://www.daml.org/experiment/ontology/location-ont#longitude>
<http://www.daml.org/experiment/ontology/location-ont#LatLong>
- 6- <http://www.wasab.dk/morten/2003/12/nearestAirport/index.rdf>
<http://paoli.open.ac.uk/watson-cache/0/269/dba6/fe237/d933fe8e0c/da5457350228a1029#input-b-longitude>
- 7- <http://simile.mit.edu/repository/ontologies/official/geo.rdfs>
http://www.w3.org/2003/01/geo/wgs84_pos#long
http://www.w3.org/2003/01/geo/wgs84_pos#lat_long
http://www.w3.org/2003/01/geo/wgs84_pos#
- 8- http://www.w3.org/2003/01/geo/wgs84_pos
http://www.w3.org/2003/01/geo/wgs84_pos#long
http://www.w3.org/2003/01/geo/wgs84_pos#lat_long
http://www.w3.org/2003/01/geo/wgs84_pos#
- 9- <http://www.aktors.org/ontology/extension>
<http://www.aktors.org/ontology/portal#has-longitude>

Development Steps – Ontology Search

watson.kml.open.ac.uk/WatsonWUI/onto_check.html?q=http%3A//www.w3.org/2003/01/geo/wgs84_pos



Details for http://www.w3.org/2003/01/geo/wgs84_pos
[Back](#)

[Get cached file - Query with SPARQL](#)

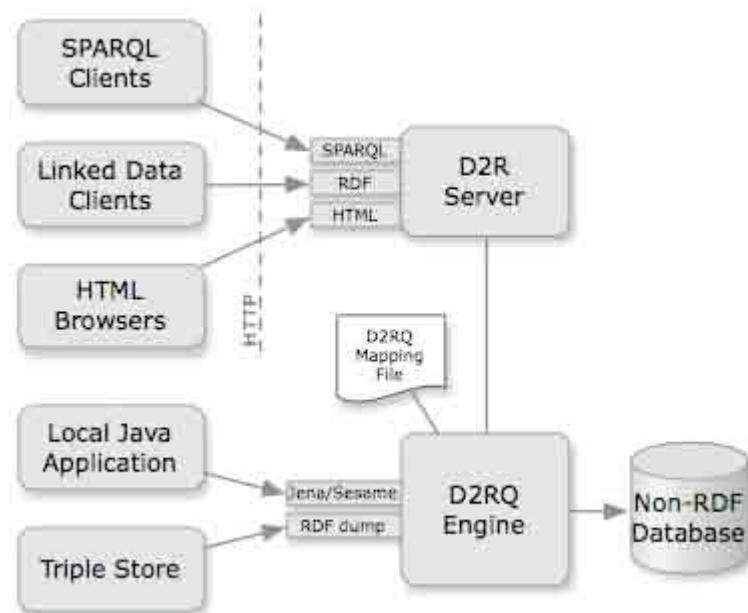
<i>Size of the file</i>	1102 KB
<i>Number of statements</i>	6061
<i>Representation languages</i>	RDF,OWL
<i>Label</i>	geo
<i>Comment</i>	<p>Recent changes to this namespace: \$Log: wgs84_pos.rdf,v \$ Revision 1.18 2006/02/01 22:01:04 danbri Clarified that I reference ellipsoid Revision 1.17 2004/02/06 17:38:12 danbri Fixed a bad commit screwup Revision 1.15 2003/04/19 1 danbri fixed a typo Revision 1.13 2003/02/19 22:27:27 connolly relaxed domain constraints on lat/long/alt from Point to XSLT doc. Revision 1.11 2003/01/12 01:20:18 danbri added a link to morten's xslt rdfls viewer. Revision 1.10 2003/01/ they would have required each occurrence of the property to mention the datatype. Revision 1.9 2003/01/11 11:41:31 da 2003/01/11 11:05:02 danbri Added an rdfs:range for each lat/long/alt property, http://www.w3.org/2001/XMLSchema# trying to be Earth-centric and neutral about coordinate system(s) at the same time. Feedback welcomed. Revision 1.6 2 rdfs:comment property of the vocabulary. Note that this is not common practice (but seems both harmless and potential lines: +16 -5 Updated schema: Added a dc:date, added url for more info. Changed the rdfs:label of the namespace from XML comment on the lat_long property suggesting that we might not need it (based on #rdfig commentary from imple +6 -5 Fixed typo; several rdfs:about attributes are now rdf:about. Thanks to MortenF in #rdfig for catching this error. re buglet in vocab, added more wgs links revision 1.2 date: 2003/01/10 11:01:11; author: danbri; state: Exp; lines: +4 -4 F to comma separated. revision 1.1 date: 2003/01/10 10:53:23; author: danbri; state: Exp; basic geo vocab</p>
<i>Employed DL</i>	ALR+HI
<i>Number of classes</i>	85
<i>Number of properties</i>	43
<i>Number of individuals</i>	1152

Development Steps – Data Harvesting

- From the Gazetteer Australia (Place names search)
 - <http://www.ga.gov.au/darwin-gazetteer/index.xhtml>
 - Go to 'Where' link and select 'Tas' as location
 - Go to 'Name' link (leave name blank)
 - Select features (parks&reserves, caves, landmarks)
 - Download CSV (by feature)
- Setup a relational database (Postgres)
 - Import CSV into tables
- Setup D2RQ
 - Generate default mappings
 - Modify mappings using selected ontologies

D2RQ

- System for accessing relational databases as RDF (query or dump).
- D2RQ platform consist of:
 - Language mapping database schema to RDFS vocabularies or OWL ontologies.
 - Engine to rewrite Jena API calls to SQL queries.
 - HTTP server that provides a Linked Data view and a SPARQL endpoint.



<http://d2rq.org>

Load data into DB (Postgres) - example

```
CREATE TABLE landmarks (  
  Record_ID VARCHAR(16) NOT NULL,  
  Name VARCHAR(64) NOT NULL,  
  Authority VARCHAR(64) NOT NULL,  
  State_ID CHAR(3) NOT NULL,  
  Feature_code VARCHAR(32) NOT NULL,  
  Status CHAR(1) NOT NULL,  
  Longitude DOUBLE PRECISION NOT NULL,  
  Latitude DOUBLE PRECISION NOT NULL,  
  Class_code CHAR(1) NOT NULL,  
  Classification VARCHAR(32) NOT NULL,  
  PRIMARY KEY(Record_ID)  
);  
  
COPY landmarks (  
  Record_ID, Name, Authority, State_ID, Feature_code, Status,  
  Longitude, Latitude, Class_code, Classification  
) FROM 'LANDMARKS-TAS.csv' WITH DELIMITER ',';
```

D2RQ – Mapping Language

- Declarative language for describing the relation between a relational database schema and RDFS vocabularies or OWL ontologies.
- Mappings are RDF documents written in Turtle syntax.
- Generate manually or automatically.

```
generate-mapping -u <user> -p <pwd> -o <out-file>  
-d <driver, e.g., org.postgresql.Driver> url (e.g.  
jdbc:postgresql://localhost:5432/tutorial)
```

<http://d2rq.org/d2rq-language>

D2RQ – Mapping Language (cont.)

LANDMARKS

Record ID	Name	State	Feature	...
TAS19510	Octopus Tree	TAS	TREE	...
TAS09412	Richards Monument	TAS	MONU	...
...



```
map:landmarks a d2rq:ClassMap;  
    d2rq:dataStorage map:database;  
    d2rq:uriPattern "landmarks/@@landmarks.record_id|urlify@";  
    d2rq:class vocab:landmarks;  
    d2rq:classDefinitionLabel "landmarks";  
    .  
map:landmarks_record_id a d2rq:PropertyBridge;  
    d2rq:belongsToClassMap map:landmarks;  
    d2rq:property vocab:landmarks_record_id;  
    d2rq:propertyDefinitionLabel "landmarks record_id";  
    d2rq:column "landmarks.record_id";
```

D2RQ – Mapping Language (cont.)

LANDMARKS

Record ID	Name	State			
TAS19510	Octopus Tree	TAS			
TAS09412	Richards Monument	TAS			
...

Represents a class or a group of similar classes of an OWL ontology or RDFS schema. A class map defines how instances of the class are identified.

```
map:landmarks a d2rq:ClassMap;  
    d2rq:dataStorage map:database;  
    d2rq:uriPattern "landmarks/@@landmarks.record_id|urlify@@";  
    d2rq:class vocab:landmarks;  
    d2rq:classDefinitionLabel "landmarks";  
.  
map:landmarks_record_id a d2rq:PropertyBridge;  
    d2rq:belongsToClassMap map:landmarks;  
    d2rq:property vocab:landmarks_record_id;  
    d2rq:propertyDefinitionLabel "landmarks record_id";  
    d2rq:column "landmarks.record_id";
```

D2RQ – Mapping Language (cont.)

LANDMARKS

Record ID	Name	State	Feature	...
TAS19510	Octopus Tree	TAS		
TAS09412	Richards Monument	TAS		
...		

Relates a database column to an RDF property. Property bridges are used to attach properties to the RDF resources created by a class map.

```
map:landmarks a d2rq:ClassMap;  
    d2rq:dataStorage map:database;  
    d2rq:uriPattern "landmarks/@@landmarks.record_id|urlify@";  
    d2rq:class vocab:landmarks;  
    d2rq:classDefinitionLabel "landmarks";  
    .
```

```
map:landmarks_record_id a d2rq:PropertyBridge;  
    d2rq:belongsToClassMap map:landmarks;  
    d2rq:property vocab:landmarks_record_id;  
    d2rq:propertyDefinitionLabel "landmarks record_id";  
    d2rq:column "landmarks.record_id";
```

Mappings to selected ontologies (example 1)

```
@prefix geonames: <http://www.geonames.org/ontology#> .  
@prefix geo: <http://www.w3.org/2003/01/geo/wgs84_pos#> .
```

```
map:caves a d2rq:ClassMap;  
    d2rq:dataStorage map:database;  
    d2rq:uriPattern "caves/@@caves.record_id|urlify@@";  
    d2rq:class geonames:Feature;  
    d2rq:classDefinitionLabel "caves";
```

```
map:caves_longitude a d2rq:PropertyBridge;  
    d2rq:belongsToClassMap map:caves;  
    d2rq:property geo:long;  
    d2rq:propertyDefinitionLabel "caves longitude";  
    d2rq:column "caves.longitude";  
    d2rq:datatype xsd:double; x
```


Mappings to selected ontologies (example 2)

```
map:parks a d2rq:ClassMap;  
    d2rq:dataStorage map:database;  
    d2rq:uriPattern "parks/@@parks.record_id|urlify@@";  
    d2rq:class geonames:Feature;  
    d2rq:classDefinitionLabel "parks";
```

#Landmarks - Caves - Nearby

```
map:landmarksCavesNearby a d2rq:PropertyBridge;  
    d2rq:belongsToClassMap map:landmarks;  
    d2rq:property geonames:nearby;  
    d2rq:refersToClassMap map:caves;  
    d2rq:join "landmarks.record_id <= landmarks_caves_nearby.landmark";  
    d2rq:join "landmarks_caves_nearby.cave => caves.record_id";
```

D2RQ – Getting Started

- Start D2RQ server:

```
d2rq-server <mapping-file>
```

- Generate RDF dump:

```
dump-rdf <mapping-file> -o <out-file>
```

Snorql: Exploring <http://152.83.253.27:2020/sparql>

SPARQL:

```
PREFIX db: <http://152.83.253.27/resource/>
PREFIX rdf: <http://www.w3.org/2002/07/owl#>
PREFIX owl: <http://www.w3.org/2002/07/owl#>
PREFIX map: <http://152.83.253.27/resource/#>
PREFIX wsl: <http://www.w3.org/2001/XMLSchema#>
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX vocab: <http://152.83.253.27/resource/vocab/>

SELECT DISTINCT * WHERE {
  ?s ?p ?o
}
LIMIT 10
```

Results:

SPARQL results:

s	p	o
db:wdsightings/f5de9c70-b88b-4d66-bbe8-4f130ab1e86c	db:vocab/wdsightings_sciname	"Megaptera novaeangliae"
db:wdsightings/f5de9c70-b88b-4d66-bbe8-4f130ab1e86c	db:vocab/wdsightings_day	10
db:wdsightings/f5de9c70-b88b-4d66-bbe8-4f130ab1e86c	db:vocab/wdsightings_year	1990
db:wdsightings/f5de9c70-b88b-4d66-bbe8-4f130ab1e86c	db:vocab/wdsightings_recordid	"f5de9c70-b88b-4d66-bbe8-4f130ab1e86c"
db:wdsightings/f5de9c70-b88b-4d66-bbe8-4f130ab1e86c	db:vocab/wdsightings_lon	115.6E0
db:wdsightings/f5de9c70-b88b-4d66-bbe8-4f130ab1e86c	db:vocab/wdsightings_month	9
db:wdsightings/f5de9c70-b88b-4d66-bbe8-4f130ab1e86c	db:vocab/wdsightings_lat	-20.7E0
db:wdsightings/f5de9c70-b88b-4d66-bbe8-4f130ab1e86c	db:vocab/wdsightings_catalognr	"NSSD-3534"
db:wdsightings/f5de9c70-b88b-4d66-bbe8-4f130ab1e86c	db:vocab/wdsightings_taxo	"urn:lsid:biodiversity.org.au:afd:taxon:8472dd02-e960-4c15-8b06-5e3156c7b0c0"
db:wdsightings/f5de9c70-b88b-4d66-bbe8-4f130ab1e86c	rdfs:label	"wdsightings #f5de9c70-b88b-4d66-bbe8-4f130ab1e86c"

Powered by [D2RQ Server](#)

<http://d2rq.org/getting-started>

D2RQ – Linked Data Support

Resolvable URIs

Following the Linked Data principles, D2R Server assigns a URI to each entity that is described in the database, and makes those URIs resolvable – that is, an RDF description can be retrieved simply by accessing the entity's URI over the Web.

<http://d2rq.org/d2r-server>

Development Steps – Interlinking

- Linked features:
 - Parks nearby Landmarks
 - Parks nearby Caves
 - Landmarks nearby Parks
 - Caves nearby Parks
 - Parks seeAlso towns nearby
 - Caves seeAlso towns nearby
 - Parks seeAlso devils nearby
 - Parks seeAlso toilets nearby
 - Landmarks seeAlso toilets nearby
 - Caves seeAlso toilets nearby

Development Steps – Interlinking

- ‘geoname:nearby’ relation linking
 - Use geo:lat and geo:long of features to calculate the vector distance between features (e.g. SQL join).
 - Link feature to nearby feature using the ‘geoname:nearby’ relation
- ‘rdfs:seeAlso’ relation linking:
 - Use Google Geocoding to find nearest place names of features given their latitude and longitude
 - Query Dbpedia for towns in Tasmania with retrieved place name
 - Link feature to place name using ‘rdfs:seeAlso’ relation
- Link feature to nearby devil (URL) using ‘rdfs:seeAlso’ relation
- Link feature to nearby toilet (URL) using ‘rdfs:seeAlso’ relation
- Homework: Which relations could replace rdfs:seeAlso?

Development Steps – Interlinking

- Reverse Geocoding

- Use Latitude and Longitude to retrieve nearest place name:
 - <http://maps.googleapis.com/maps/api/geocode/json?latlng=-42.77000046,147.0700073&sensor=false>
- Parse JSON document to retrieve place name
- Query DBPedia for towns in Tasmania with given name

```
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX schema:<http://dbpedia.org/class/yago/>

SELECT DISTINCT ?uri WHERE{
    ?uri rdf:type schema:TownsInTasmania;
    rdfs:label ?label .
    FILTER (regex(?label, 'New Norfolk'))
}
```

Development Steps – Querying

- We use SPARQL to query the data
- D2RQ offers a SPARQL tool (Snorql)
- The RDF Dataset produced from D2RQ was dumped into a Sesame repository
 - Namespace created: <http://www.csiro.au/issl/holdt/>
- We query the data using SPARQL from the Sesame Workbench
- In the example (next) we ask for all features nearby to another feature and their latitude and longitude


Development Steps – Querying

- SPARQL Query example (Sesame repository)

```
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX geo: <http://www.w3.org/2003/01/geo/wgs84_pos#>
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX vocab: <http://www.csiro.au/issl/holdt/vocab/>
PREFIX geonames: <http://www.geonames.org/ontology#>
```

```
SELECT DISTINCT ?feature1 ?feature2 ?lat1 ?lon1 ?lat2 ?lon2 WHERE {
  ?feature1 a geonames:Feature .
  ?feature1 geonames:nearby ?feature2 .
  ?feature1 geo:lat ?lat1 .
  ?feature1 geo:long ?lon1 .
  ?feature2 geo:lat ?lat2 .
  ?feature2 geo:long ?lon2 .
}
```


Development Steps – Querying

 **Workbench**

Sesame server

Repositories

New repository

Delete repository

Explore

Summary

Namespaces

Contexts

Types

Explore

Query

Saved Queries

Export

Modify

SPARQL Update

Add

Remove

Clear

System

Information

Current Selections:

Sesame server: <http://152.83.253.27:8080/openrdf-sesame> [\[change\]](#)

Repository: ISWC 2013 Linked Data Tutorial (TUTORIAL) [\[change\]](#)

User (optional): > none < [\[change\]](#)

Query Result (1-100 of 34461)

Download format: [Download](#)

Results per page: The results shown maybe truncated.

Results offset: [Previous 100](#) [Next 100](#)

Show data types & language tags: ☒

Feature1	Feature2	Lat1	Lon1	Lat2	Lon2
http://www.csiro.au/issl/holdt/landmarks/TAS19510	http://www.csiro.au/issl/holdt/parks/TAS19861	-42.900000153E0	147.25E0	-42.790000092E0	147.24000055E0
http://www.csiro.au/issl/holdt/landmarks/TAS19510	http://www.csiro.au/issl/holdt/parks/TAS23704	-42.900000153E0	147.25E0	-42.860000061E0	147.27000043E0
http://www.csiro.au/issl/holdt/landmarks/TAS19510	http://www.csiro.au/issl/holdt/parks/TAS23707	-42.900000153E0	147.25E0	-42.840000015E0	147.30999976E0
http://www.csiro.au/issl/holdt/landmarks/TAS19510	http://www.csiro.au/issl/holdt/parks/TAS21200	-42.900000153E0	147.25E0	-42.810000137E0	147.33999954E0
http://www.csiro.au/issl/holdt/landmarks/TAS19510	http://www.csiro.au/issl/holdt/parks/TAS38333	-42.900000153E0	147.25E0	-42.81999969E0	147.32000073E0
http://www.csiro.au/issl/holdt/landmarks/TAS19510	http://www.csiro.au/issl/holdt/parks/TAS19696	-42.900000153E0	147.25E0	-42.770000046E0	147.05999976E0
http://www.csiro.au/issl/holdt/landmarks/TAS19510	http://www.csiro.au/issl/holdt/parks/TAS16980	-42.900000153E0	147.25E0	-42.840000015E0	147.28999933E0
http://www.csiro.au/issl/holdt/landmarks/TAS19510	http://www.csiro.au/issl/holdt/parks/TAS16944	-42.900000153E0	147.25E0	-42.810000137E0	147.25E0
http://www.csiro.au/issl/holdt/landmarks/TAS19510	http://www.csiro.au/issl/holdt/parks/TAS03578	-42.900000153E0	147.25E0	-42.810000137E0	147.32000073E0
http://www.csiro.au/issl/holdt/landmarks/TAS19510	http://www.csiro.au/issl/holdt/parks/TAS19698	-42.900000153E0	147.25E0	-42.770000046E0	147.05000031E0
http://www.csiro.au/issl/holdt/landmarks/TAS19510	http://www.csiro.au/issl/holdt/parks/TAS19575	-42.900000153E0	147.25E0	-42.790000092E0	147.25999945E0
http://www.csiro.au/issl/holdt/landmarks/TAS19510	http://www.csiro.au/issl/holdt/parks/TAS19887	-42.900000153E0	147.25E0	-42.790000092E0	147.25E0

Development Steps: Serving

- The dataset (Sesame endpoint) has not been published (it is just for the tutorial!)
 - Thus, elements cannot be dereferenced on the Web
- Homework: How to publish the dataset in the LOD Cloud?

Development Steps: Application

- Linking Tasmania Geo-spatial Environmental Features
- Available at <http://lmd.it.csiro.au:8080/holdt/gui>
- The purpose is to explore the linked features for Tasmania
- The features (caves, landmarks, parks) are natural environments, which are frequently visited.
- We link the nearby features to each other.
- We link these features to the nearest cities/towns.
- We link these features to sites where Tasmanian devils have been seen and to toilets! (data is available with lat, long!)
- Application uses Sesame API (Java) to display the data

Application: <http://lmd.it.csiro.au:8080/holdt/gui>

ISWC 2013 - Hands on Linked Data Tutorial

Linked Features for Tasmania

Click one of the features below to explore nearby features and places

Caves Landmarks Parks

PARK <http://www.csiro.au/issl/holdt/parks/TAS38312>

http://www.csiro.au/issl/holdt/vocab/authority	Tasmania
http://www.csiro.au/issl/holdt/vocab/classCode	P
http://www.csiro.au/issl/holdt/vocab/classification	Parks & Reserves
http://www.csiro.au/issl/holdt/vocab/featureCode	RESV
http://www.csiro.au/issl/holdt/vocab/recordId	TAS38312
http://www.csiro.au/issl/holdt/vocab/state	TAS
http://www.csiro.au/issl/holdt/vocab/status	O
http://www.geonames.org/ontology#name	Devils Guliet State Reserve
http://www.geonames.org/ontology#nearby	http://www.csiro.au/issl/holdt/caves/TAS02478
http://www.geonames.org/ontology#nearby	http://www.csiro.au/issl/holdt/caves/TAS02479
http://www.geonames.org/ontology#nearby	http://www.csiro.au/issl/holdt/caves/TAS02480
http://www.geonames.org/ontology#nearby	http://www.csiro.au/issl/holdt/caves/TAS02481
http://www.geonames.org/ontology#nearby	http://www.csiro.au/issl/holdt/caves/TAS02482
http://www.geonames.org/ontology#nearby	http://www.csiro.au/issl/holdt/caves/TAS02484
http://www.geonames.org/ontology#nearby	http://www.csiro.au/issl/holdt/caves/TAS07127
http://www.geonames.org/ontology#nearby	http://www.csiro.au/issl/holdt/caves/TAS07128
http://www.geonames.org/ontology#nearby	http://www.csiro.au/issl/holdt/caves/TAS07129
http://www.geonames.org/ontology#nearby	http://www.csiro.au/issl/holdt/caves/TAS09173

Other Considerations

- Size of datasets
- Access to datasets (e.g. Rest)
- Frequency of updates
- Types of relationships

Thank you

Intelligent Sensing and Systems Laboratory
Hobart, Tasmania, Australia

Liliana Cabral

Research Fellow

E [liliana.silvacabral @ csiro.au](mailto:liliana.silvacabral@csiro.au)

CSIRO COMPUTATIONAL INFORMATICS
www.csiro.au

