



## Project Leader

**Prof G. West**, Dept of Spatial Sciences, Curtin University  
g.west@curtin.edu.au

## Research Team

**Mr J Siao Him Fa**, Curtin University; **Dr D.A. McMeekin**, Curtin University; **Dr S Moncrieff**, Curtin University; **Dr S Cox**, CSIRO.

## Project Participants

*Universities* - Curtin University, University of Canterbury  
*Private Sector* - Amristar, NGIS  
*Public sector* - DELWP, DNRM, CSIRO, PSMA, Landgate

## Objectives

To provide an automated method for seamless communication between database systems across states and territories while allowing user querying.

## Outcomes

- To Deliver a functional implementation of a federated system.
- To develop the system using open source programs.
- To adapt the system to a large area scale.

## Introduction

Authoritative spatial data is managed by Australia's jurisdictions (states and territories). Many applications require them to be joined, either because activities occur at borders, or a uniform view of data is required. The syntactic and semantic harmonization required can be done in various locations in an information architecture. Optimally data providers may conform to a community model or standard (e.g. INSPIRE model), but there is usually little incentive for agencies, as data publication for out-of-state users is not of upmost importance. Traditionally it was left to the user to merge data from multiple sources (if they could find it). Another option is for systems or services to be deployed at an intermediate point, in a 'broker' architecture. For a subset of foundation datasets the broker role is currently played in Australia by PSMA, but their processes and distribution channels depend on bulk file transfer and significant manual intervention in the process. In this project we are exploring more automated methods for federation, using semantic technologies.

## The Problem: Disparate Datasets

The picture below depicts some problems regarding the querying of differing datasets. When querying the bounding box, data needs to be fetched from various heterogeneous data sources, and hence various federation issues arise.

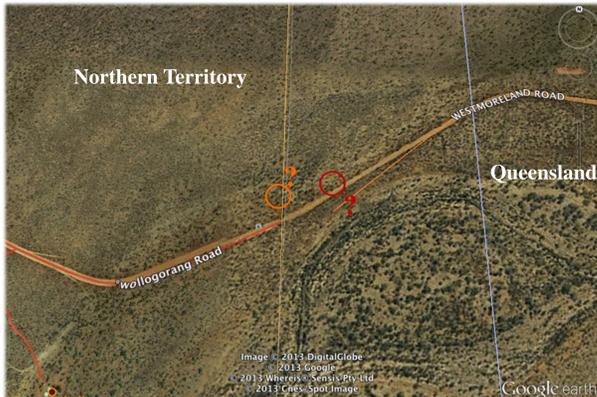


Figure 1: Disparate Datasets Problem (Google, Northern Territory, Queensland Data)

### Federation Issues

- Different road names
- Different fonts
- Different line styles and colours
- Gaps in the data
- Different locations of the state borders

## Example: Differing Representation of "LandUse"

```
<?xml encoding="ISO-8859-1"?>
<ELEMENT LandUse (LandParcel)>
<ELEMENT LandParcel (AREA, BROAD, LU1,
LU2, LU3, ..., JurisType, JurisName)>
<ELEMENT AREA (#PCDATA)>
<ELEMENT BROAD (#PCDATA)>
<ELEMENT LU1 (#PCDATA)>
.....
<ELEMENT JurisType (#PCDATA)>
<ELEMENT JurisName (#PCDATA)>
</LandParcel>
</LandUse>
```

a) Local XML data source S<sub>1</sub> of Eau Claire County.

```
<?xml encoding="ISO-8859-1"?>
<ELEMENT LandUse (LandParcel)>
<ELEMENT LandParcel (AREA, LAND_USE,
PARCEL_ID, ..., JurisType, JurisName)>
<ELEMENT AREA (#PCDATA)>
<ELEMENT LAND_USE (#PCDATA)>
<ELEMENT PARCEL_ID (#PCDATA)>
.....
<ELEMENT JurisType (#PCDATA)>
<ELEMENT JurisName (#PCDATA)>
</LandParcel>
</LandUse>
```

b) Local XML data source S<sub>2</sub> of the City of Madison.

### Differences:

- XML tags
- Names
- Information details
- Letters vs Numbers

Figure 2: LandUse Representation (Isabel Cruz – Ontology Alignment for the Semantic Integration of Heterogeneous Geospatial Data Set)

## Possible Federated System

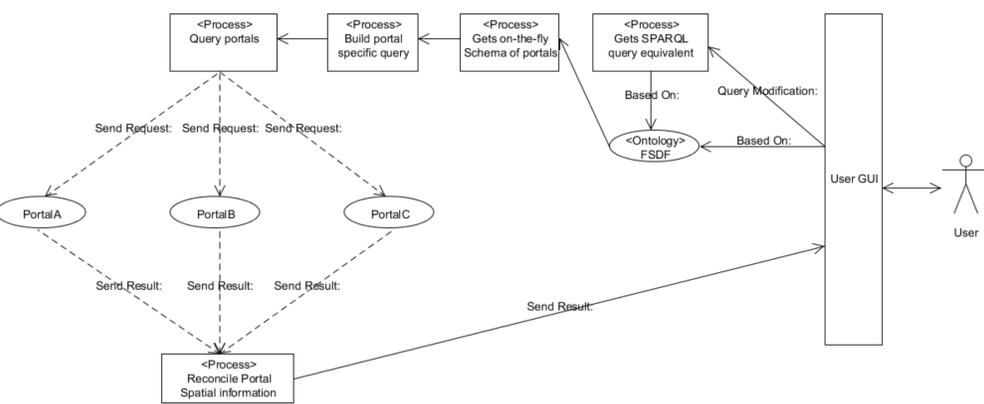


Figure 3: Sample Federated System

## Landgate LGATE-069 Graph

Figure 4 is an ontological representation of the LGATE-069 type (StateElectoralUnit). The ontology was automatically generated from Landgate's XML schema, using EXtensible Stylesheet Language Transformations(XSLT).

The ontology obtained can be integrated onto third-party ontologies via the "AbstractFeature" class, which is common in geographic ontologies.

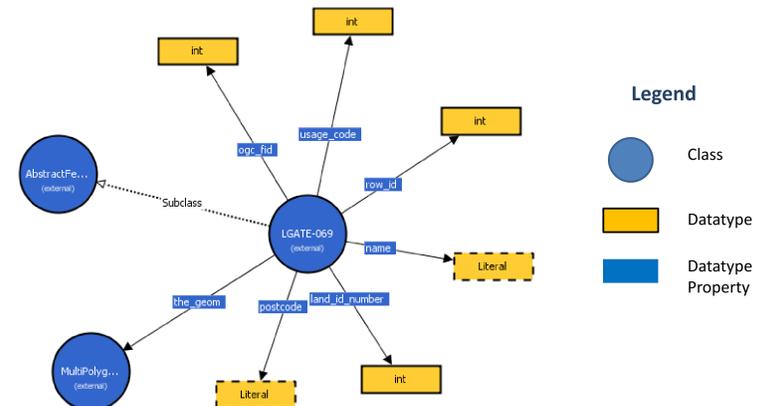


Figure 4: LGATE-069 Graph

## Summary

Due to the jurisdictions having different spatial data schemas and formats, database interoperability is an issue. By finding ways to federate Australia's spatial data automatically using semantic web techniques, it would allow the unification of all the disparate datasets to be done for the user. Hence, this would lead to easier access to nation wide spatial data, while lifting semantic burdens on the user.