



Institute of Information Science
Academia Sinica

中央研究院 資訊科學研究所

An Ontology of Place name for Information Retrieval and Reasoning

Dongpo Deng

dongpo@iis.sinica.edu.tw

Agenda

- Motivation
- What's an Ontology?
- Ontology and GIScience
- An Ontology of Place-name
- RDF and Jena
- An Example
- Conclusion



Motivation

- A frequently repeated factoid is 80% of all digital data generated today includes geospatial reference.
 - geographic coordinates → experts
 - place-names, address, postal codes → folks
- Formal model enable interoperability.
- Development of Human-Computer Interface
- There is a need to specify the place name to canonical and interchangeable geospatial knowledge



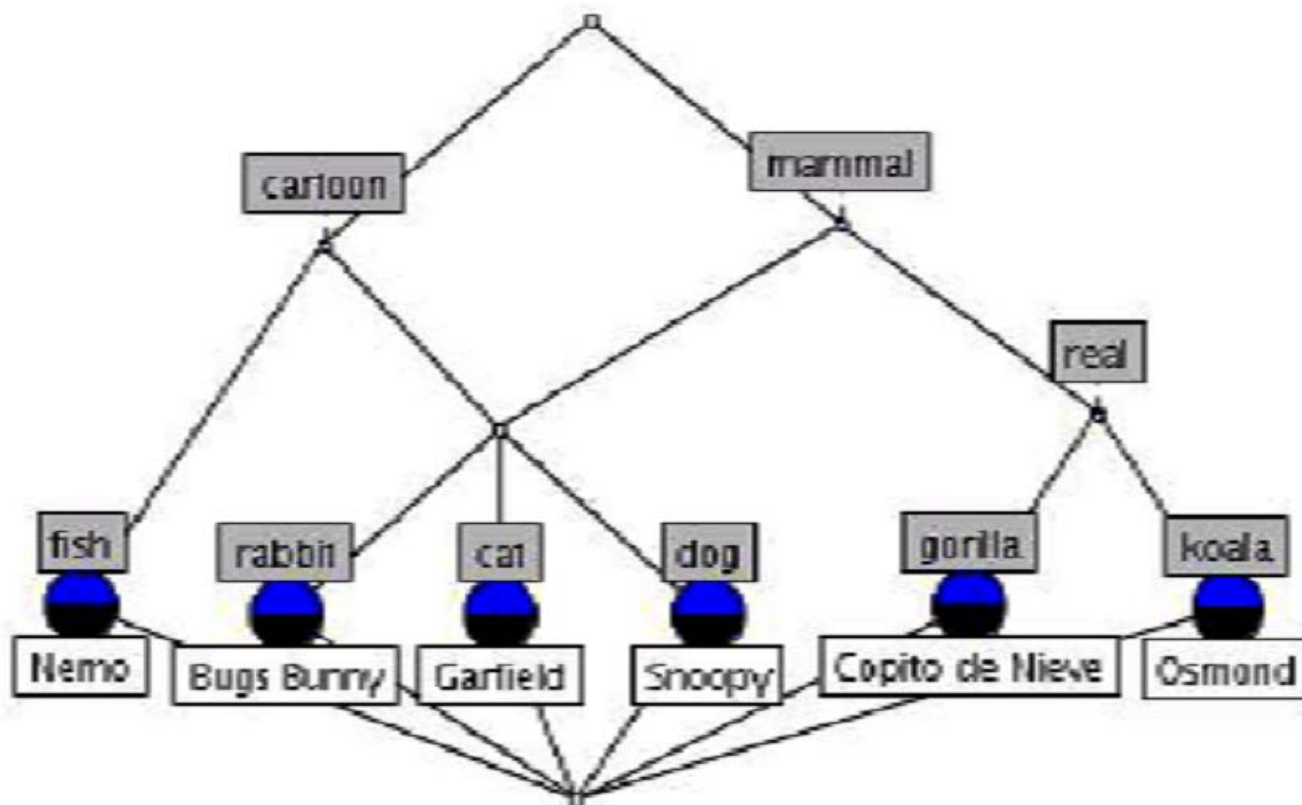
An Example of Formal Concept Analysis

	real	cartoon	cat	dog	rabbit	fish	gorilla	koala	mammal
Garfield		x	x						x
Snoopy		x		x					x
Bugs Bunny		x			x				x
Nemo		x				x			
Copito	x						x		x
Osmond	x							x	x

From Uta Priss, <http://www.upriss.org.uk/>



Organizing the Objects as a Lattice



From Uta Priss, <http://www.upriss.org.uk/>

What's an ontology

- domain-specific and user-dependent view (knowledge engineer perspective) (Gruber, 1993)
 - An Ontology is a
 - formal specification ⇒ Executable, Discussable
 - of a shared ⇒ Group of persons
 - conceptualization ⇒ About concepts
 - of a domain of interest ⇒ Between application and “unique truth“



What's a geospatial ontology

- UCGIS defined
 - “the totality of geospatial concepts, categories, relations, and processes”
 - “conceptual systems that people use “
 - “canonical description of knowledge domains”
 - “associated classificatory theories”
 - “the formal modeling of the geospatial world as this is experienced and conceptualized by non-experts”

Naïve Geography



Why develop an Ontology?

- To make domain assumptions explicit
 - Easier to change domain assumptions
 - Easier to understand and update legacy data
- To separate domain knowledge from operational knowledge
 - Re-use domain and operational knowledge separately.
- A community reference for applications.
- To share a consistent understanding of what information means.

From Steffen Staab, 2006, <http://isweb.uni-koblenz.de>



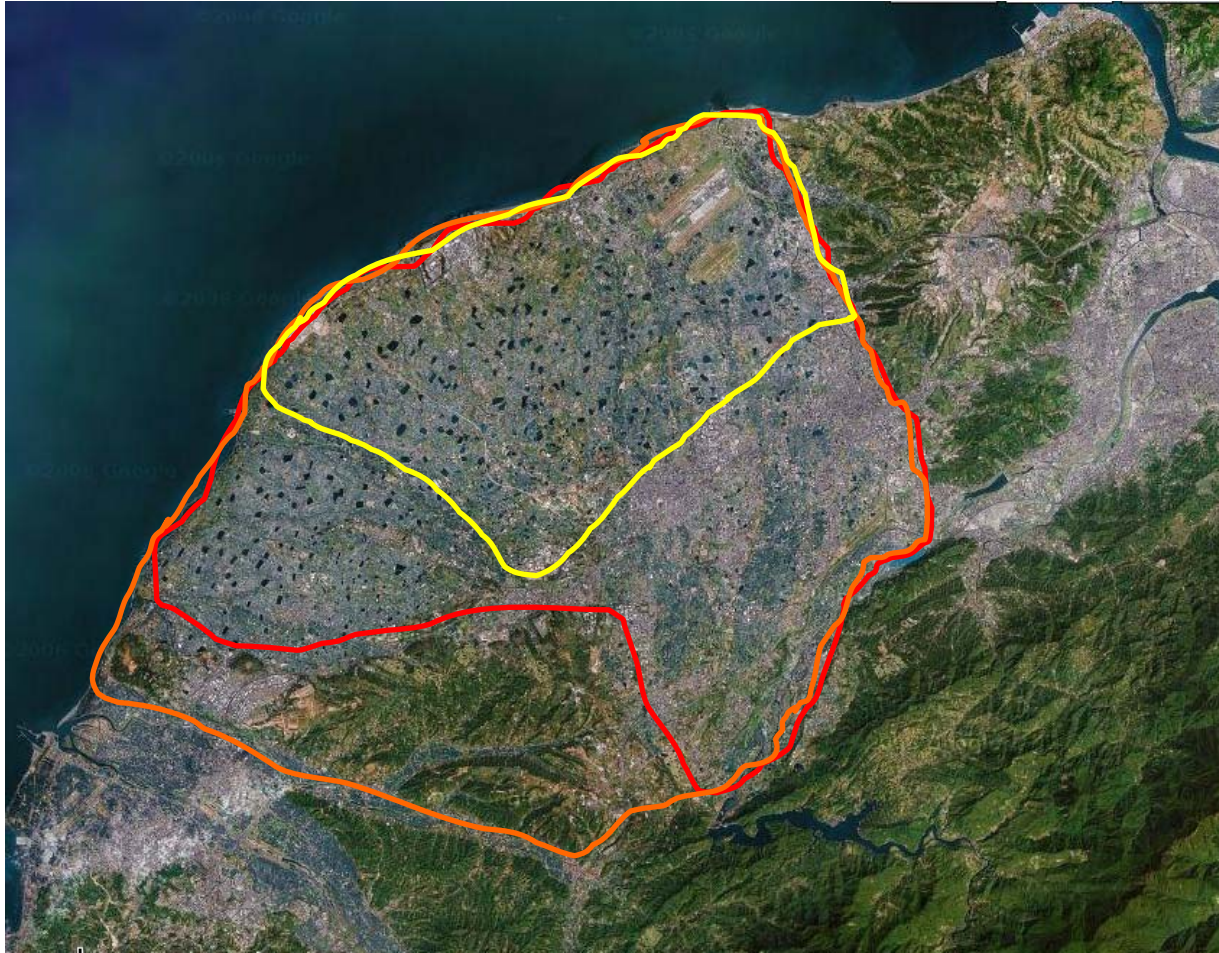
What can the Ontology do?

- An ontology will:
 - enable general, though not necessarily efficient, reasoning about spatial.
 - link with more efficient specialized reasoning engines for spatial reasoning.
 - link with the numerous database that exist containing a wealth of specific e.g. geographical, spatial information.
 - support convenient query capability for spatial information.

From Jerry Hobbs, Keynote “*Interoperability among Geospatial Ontologies*” GEOS 2005.



Geographical Entities are Indeterminate and Ambiguous Objects



Vagueness and Ambiguity

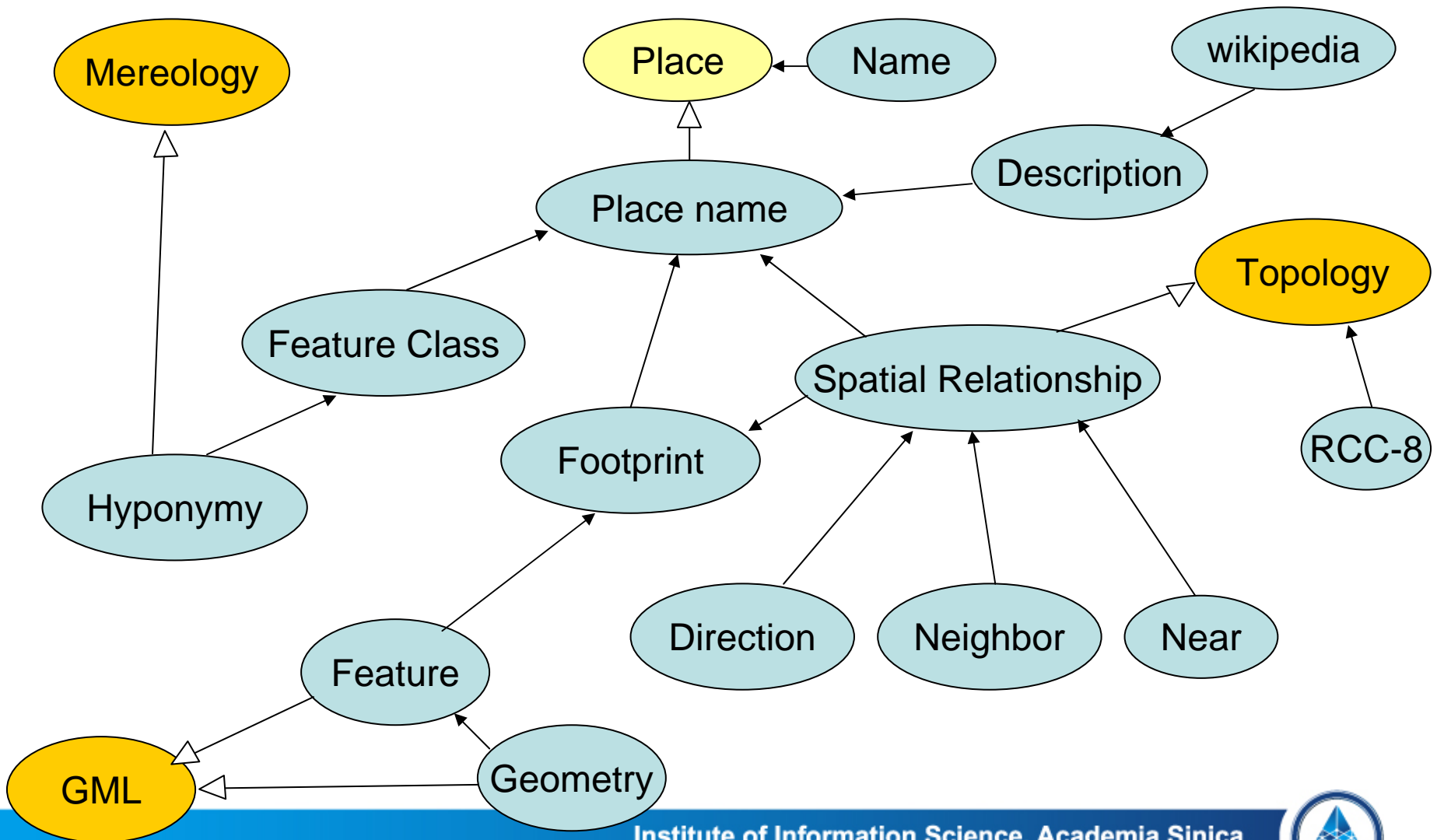
- Geographic categorization and classification are scale- and size-dependent, ⇒ relations and property inheritance are not yet clear
- Regionalization in space and time is human-dependent, ⇒ nomenclature can have different meanings in different application contexts.
- Location and structure of boundaries shape many geographical categories. ⇒ spatial granularity and complicated spatial relationship



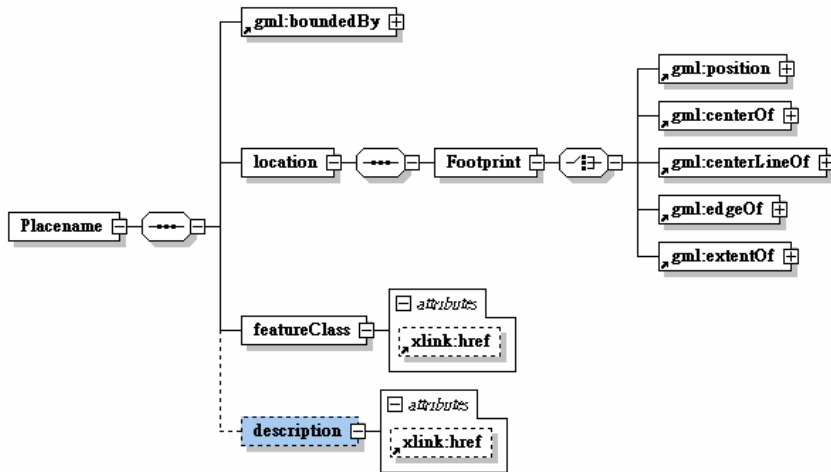
- *fiat* (tennis court) and *bona fide* (shoreline)
- open (bay) and closed (lake)



An ontology of place-name



A GML of Place-name



```

<?xml version="1.0" encoding="UTF-8"?>
<TwGazetteer>
  <places>
    <PlaceName>
      <gml:boundaryBy>306248.97228,2436769.72952 306248.97228,2436769.72952 </gml:boundaryBy>
      <gml:name>紅頭村</gml:name>
      <location>
        <Footprint>
          <gml:centerOf>
            <gml:Point>
              <gml:coordinates>306248.97228,2436769.72952</gml:coordinates>
            </gml:Point>
          </gml:centerOf>
        </Footprint>
      </location>
      <featureClass xlink:href="http://w3p.org/vocab/c_17">四等城市</featureClass>
    </PlaceName>
  </places>
</TwGazetteer>
  
```

Taiwan Place-name Type Thesaurus

Hypernymy / Hyponymy

Taiwan Place Name Type Thesaurus	
行政管理區域	
= ADL_name	administrative areas
地籍區 ADL_name=cadastral areas	
軍事區 ADL_name=military areas	
政治區 ADL_name=political areas	
人口聚居區	
= ADL_name	populated places
一等城市 ADL_name=cities	
二等城市 ADL_name=	
三等城市 ADL_name=	
四等城市 ADL_name=	
郵政區 ADL_name=postal areas	
學區 ADL_name=school districts	
統計區 ADL_name=statistical areas	
原住民部落區 ADL_name=tribal areas	
水域圖徵	
= ADL_name	hydrographic features
地下含水層 ADL_name=aquifers	
灣 ADL_name=bays	
河道 ADL_name=channels	
集水區 ADL_name=drainage basins	
河口 ADL_name=estuaries	
沖積平原 ADL_name=floodplains	
湖 ADL_name=lakes	
埤塘 ADL_name=pond	
海洋 ADL_name=seas	
河流 ADL_name=streams	
地熱圖徵 ADL_name=thermal features	
土地丘塊 ADL_name=land parcels	
人工圖徵	
= ADL_name	manmade features
農業場址 ADL_name=agricultural sites	
建物 ADL_name=buildings	
墓地 ADL_name=cemeteries	

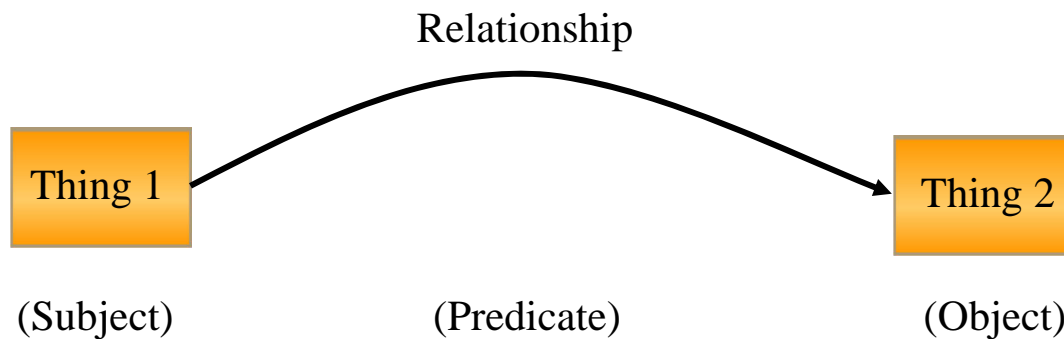


Transfer Thesaurus to RDF

```
<rdf:Description rdf:about="http://w3p.org/vocab/c_169">
  <j.0:chineseName>公園</j.0:chineseName>
  <j.0:englishName>parks</j.0:englishName>
  <j.0:hyponymOf rdf:resource="http://w3p.org/vocab/c_48"/>
</rdf:Description>
<rdf:Description rdf:about="http://w3p.org/vocab/c_17">
  <j.0:chineseName>四等城市</j.0:chineseName>
  <j.0:englishName>NA</j.0:englishName>
  <j.0:hyponymOf rdf:resource="http://w3p.org/vocab/c_12"/>
</rdf:Description>
<rdf:Description rdf:about="http://w3p.org/vocab/c_49">
  <j.0:chineseName>農業場址</j.0:chineseName>
  <j.0:englishName>agricultural sites</j.0:englishName>
  <j.0:hyponymOf rdf:resource="http://w3p.org/vocab/c_48"/>
  <j.0:hypernymOf rdf:resource="http://w3p.org/vocab/c_50"/>
  <j.0:hypernymOf rdf:resource="http://w3p.org/vocab/c_51"/>
  <j.0:hypernymOf rdf:resource="http://w3p.org/vocab/c_52"/>
  <j.0:hypernymOf rdf:resource="http://w3p.org/vocab/c_53"/>
  <j.0:hypernymOf rdf:resource="http://w3p.org/vocab/c_54"/>
  <j.0:hypernymOf rdf:resource="http://w3p.org/vocab/c_55"/>
</rdf:Description>
<rdf:Description rdf:about="http://w3p.org/vocab/c_256">
  <j.0:chineseName>洞穴</j.0:chineseName>
  <j.0:englishName>caves</j.0:englishName>
  <j.0:hyponymOf rdf:resource="http://w3p.org/vocab/c_245"/>
</rdf:Description>
```


What's the RDF?

- **Resource Description Framework**
 - represents information on Web with machine-understandable syntax and semantics
- **Framework: Subject has the predicate whose value is object**



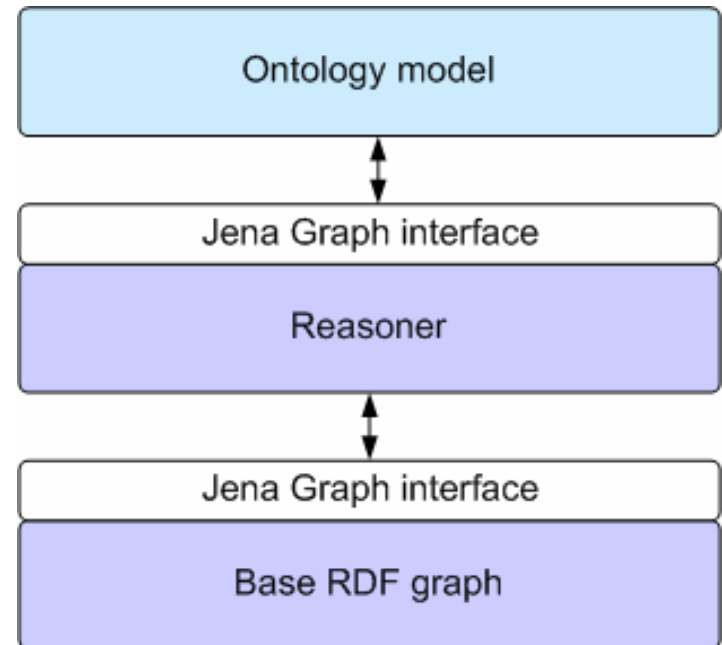
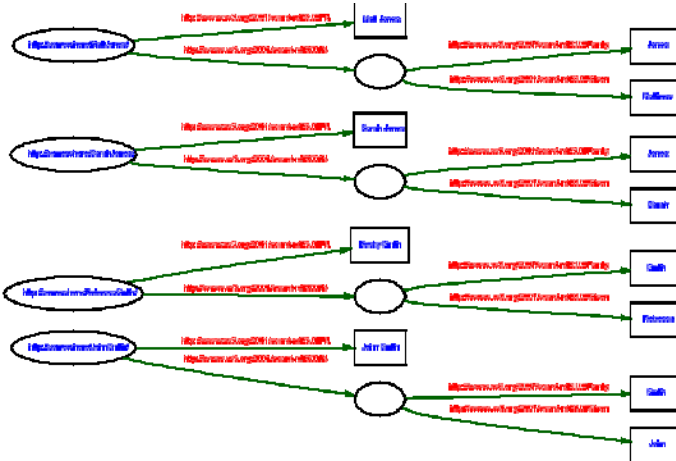
Jena

- A Java API for RDF
 - Developed by Brian McBride of HP
 - Derived from SiRPAC API
 - Can parse, create, and search RDF models
- Jena packages
 - jena.model
 - Key package for application developer. It contains interfaces for model, resource, ...
 - jena.mem
 - Contains an implementation of Jena API which stores all model state in main memory
 - Jena.common
 - Contains implementation classes



Jena's reasoner

- The Jena Reasoner need
 - your Ontology model → Jena Graph Interface
 - Base RDF graph → Jena Graph Interface



An OWL of Thesaurus

```
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  xmlns:owl="http://www.w3.org/2002/07/owl#"
  xmlns:daml="http://www.daml.org/2001/03/daml+oil#" >
  <rdf:Description rdf:about="http://w3p.org/vocab/hyponymOf">
    <rdf:type rdf:resource="http://www.w3.org/1999/02/22-rdf-syntax-ns#Property"/>
    <owl:inverseOf rdf:resource="http://w3p.org/vocab/hyponymOf"/>
    <rdf:type rdf:resource="http://www.w3.org/2002/07/owl#TransitiveProperty"/>
  </rdf:Description>
  <rdf:Description rdf:about="http://w3p.org/vocab/hyponymOf">
    <rdf:type rdf:resource="http://www.w3.org/1999/02/22-rdf-syntax-ns#Property"/>
    <owl:inverseOf rdf:resource="http://w3p.org/vocab/hyponymOf"/>
    <rdf:type rdf:resource="http://www.w3.org/2002/07/owl#TransitiveProperty"/>
  </rdf:Description>
</rdf:RDF>
```

A RDQL query

```
"SELECT ?first, ?firstChinese, ?second, ?secondChinese WHERE " +  
" (?first <w3p:chineseName> ?firstChinese) " +  
" (?first <w3p:hypernymOf> ?second) " +  
" (?second <w3p:chineseName> ?secondChinese) " +  
" (?third <w3p:instanceOf> ?second) " +  
" (?third <w3p:chineseName> ?thirdChinese) " + "USING w3p FOR  
<http://w3p.org/vocab/>" ;
```



The result of the RDQL query

```
C:\WINDOWS\system32\cmd.exe

Literal: 紅頭

紅頭 是一個 四等城市 也是一個 Taiwan_Place_Name_Type_Thesaurus
  Why? [Reasoning Process]: <first, http://w3p.org/vocab/c_1> <firstChinese, Taiwan_Place_Name_Type_Thesaurus> <second, http://w3p.org/vocab/c_17> <secondChinese, 四等城市> <third, http://w3p.org/vocab/meta-list.php?id=6> <thirdChinese, 紅頭>

紅頭 是一個 四等城市 也是一個 行政管理區域
  Why? [Reasoning Process]: <first, http://w3p.org/vocab/c_2> <firstChinese, 行政管理區域> <second, http://w3p.org/vocab/c_17> <secondChinese, 四等城市> <third, http://w3p.org/vocab/meta-list.php?id=6> <thirdChinese, 紅頭>

紅頭 是一個 四等城市 也是一個 人口聚居區
  Why? [Reasoning Process]: <first, http://w3p.org/vocab/c_12> <firstChinese, 人口聚居區> <second, http://w3p.org/vocab/c_17> <secondChinese, 四等城市> <third, http://w3p.org/vocab/meta-list.php?id=6> <thirdChinese, 紅頭>

C:\java\w3pontoImpl>
```

Conclusion

- An ontology of place-name is a specification of domain knowledge that enable spatial reasoning and information retrieval.
- Create a more interoperability environment.
- In the future, we will expand the ontology of place-name to include spatial relationships, like direction, neighbor, near.



Question?



- Acknowledge
- I would like to thank Mr. Hao-Chung Wang who helped me to deal with the technique support.
- I also would like to thank the Open Geospatial Information Team who kindly gave me useful advice.