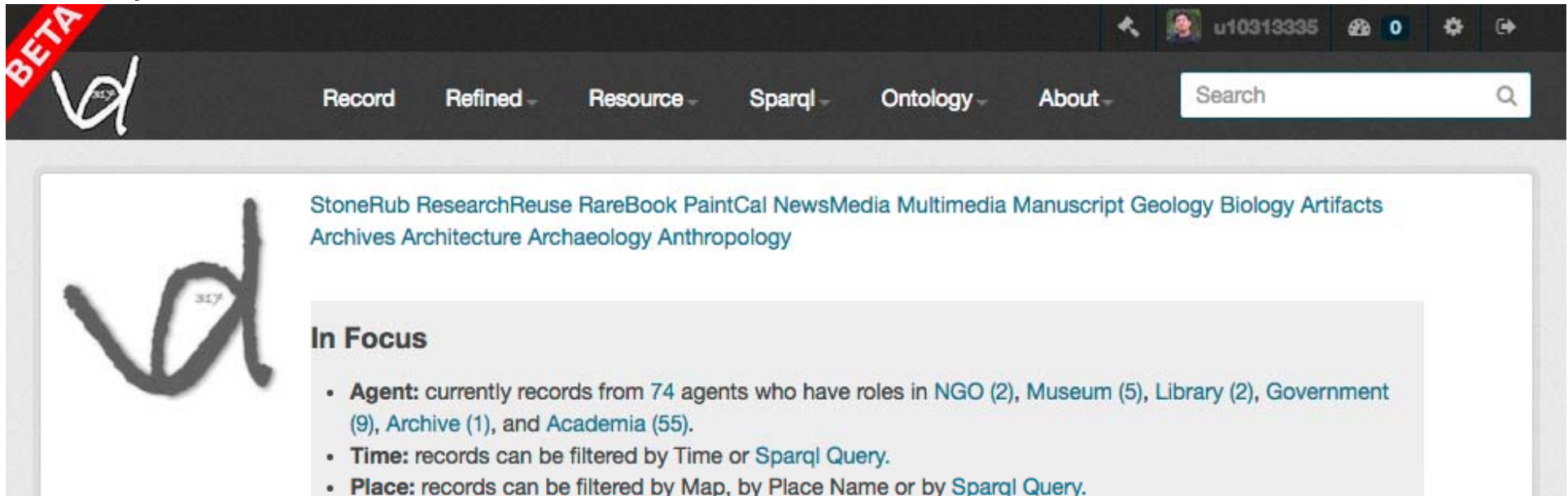


Metadata as Linked Data for Research Data Repositories

Cheng-Jen Lee, Andrea Wei-Ching Huang, Tyng-Ruey Chuang
Institute of Information Science, Academia Sinica, Taiwan

- ✓ An use case for curation, publication, and reuse of metadata as linked data
- ✓ A new method to manage data for general-use and discipline-specific repositories
- ✓ Data semantically enriched with vocabularies and knowledge bases via adaptable mechanisms



The screenshot shows the ODW interface with a navigation bar containing 'Record', 'Refined', 'Resource', 'Sparql', 'Ontology', and 'About'. A search bar is on the right. The main content area features a large 'od' logo and a list of categories: StoneRub, ResearchReuse, RareBook, PaintCal, NewsMedia, Multimedia, Manuscript, Geology, Biology, Artifacts, Archives, Architecture, Archaeology, and Anthropology. Below this is an 'In Focus' section with three bullet points:

- **Agent:** currently records from 74 agents who have roles in NGO (2), Museum (5), Library (2), Government (9), Archive (1), and Academia (55).
- **Time:** records can be filtered by Time or Sparql Query.
- **Place:** records can be filtered by Map, by Place Name or by Sparql Query.

<http://data.odw.tw> | <http://voc.odw.tw>

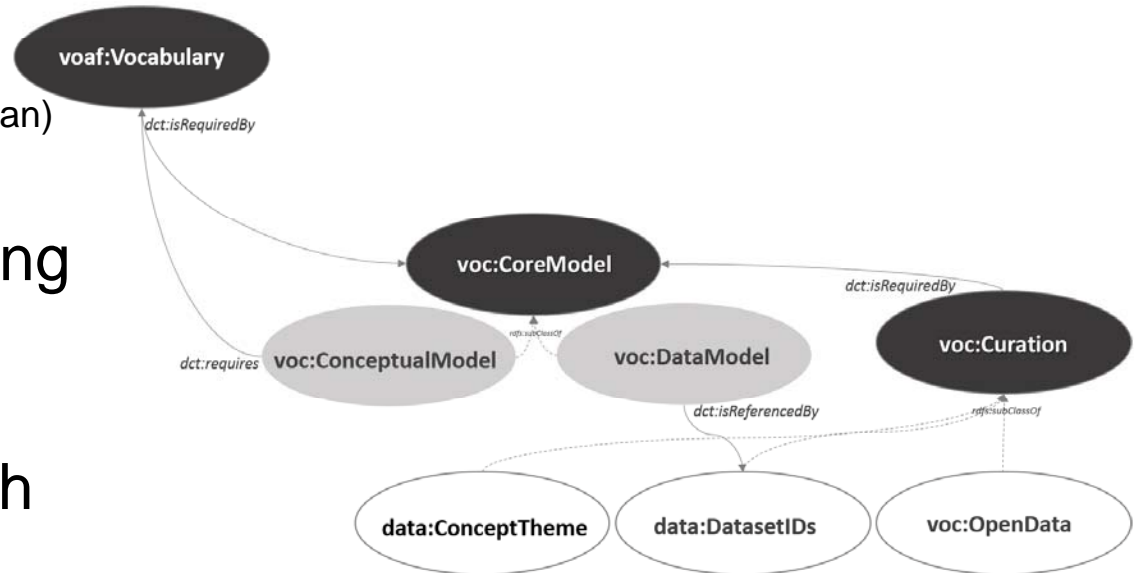


Metadata as Linked Data

- 840,000 **metadata** records
 → 44 million+ **linked data triples**

(from Union Catalog of Digital Archives Taiwan)

- A configurable, automatic **XML** → **CSV** → **RDF** mapping and conversion process
- An **ontology** to describe cross-disciplinary research data with multiple refined versions



Main Components of the Ontology for data.odw.tw (voc4odw)

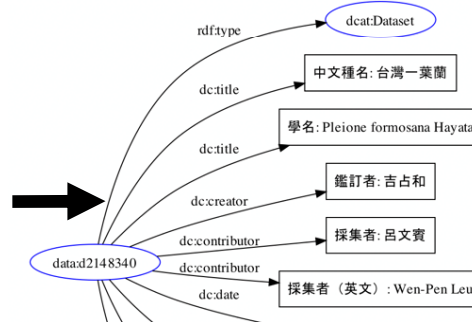
Example: *Pleione formosana* (台灣一葉蘭)

```
<?xml version="1.0" encoding="BIG5"?>
<DACatalog>
  <AdminDesc>
    <Project Creator="中研院生物多樣性中心" GenDat Project>
  </AdminDesc>
  <Catalog>
    <Record>典藏機構與計畫: 中央研究院: 生物多樣性研究
    <Record>內容主題: 生物: 植物界: 種子植物門: 單子葉植
  </Catalog>
  <DigiArchiveID>43501</DigiArchiveID>
  <Hyperlink>http://www.hast.biodiv.tw/spec
```

XML
(Original Data)



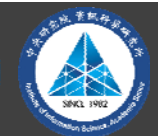
CSV with RDF properties



RDF
(Linked Data)



<http://data.odw.tw> | <http://voc.odw.tw>



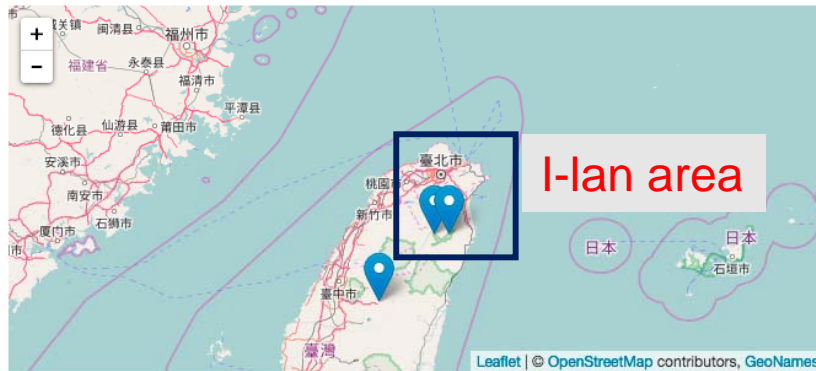
A Linked Open Research Data Repository

Browse *Pleione Formosana's* geographical distributions from a **map view**

Find other species found in 1993 in I-lan area by **spatio-temporal** search

中文種名: 台灣一葉蘭

RECORD EXTENT



I-lan area

METADATA

rdf:type	data:Refined, r4r:Data, dcat:Dataset	
r4r:locateAt	http://data.odw.tw/record/d2148340	
dcat:landingPage	http://data.odw.tw/r1/r1-r2148340	
dcat:themeTaxonomy	data:Biological	
txn:hasEOLPage	eol:1134120	
	rdf:type	dwc:Taxon
	rdfs:label	Pleione formosana

Get DC15 Records

```

36 # put your SPARQL query here.
37 SELECT ?PlaceLabel ?PlaceR ?PlaceD
38 WHERE { ?s txn:hasEOLPage eol:1134120; dct:requires ?event
39         OPTIONAL { ?event gn:parentFeature ?PlaceR. ?PlaceR rdfs:label ?PlaceLabel. }
40         OPTIONAL { ?s dc:coverage ?PlaceD. } }
41 ORDER by DESC(?PlaceR)
    
```

SPARQL Results

#	PlaceLabel	PlaceR	PlaceD
1	Chiayi	http://sws.geonames.org/1678835	國家: 台灣
2	Chiayi	http://sws.geonames.org/1678835	最低海拔: 2300
3	Chiayi	http://sws.geonames.org/1678835	行政區: 嘉義縣

Powered by open-source 

<http://data.odw.tw> | <http://voc.odw.tw>

