

CSW2LD: a Linked Data frontend for CSW

Francisco J. Lopez-Pellicer, Aneta J. Florczyk, Walter Rentería-Agualimpia,
F. Javier Nogueras-Iso, Pedro R. Muro-Medrano
IAAA, Universidad de Zaragoza

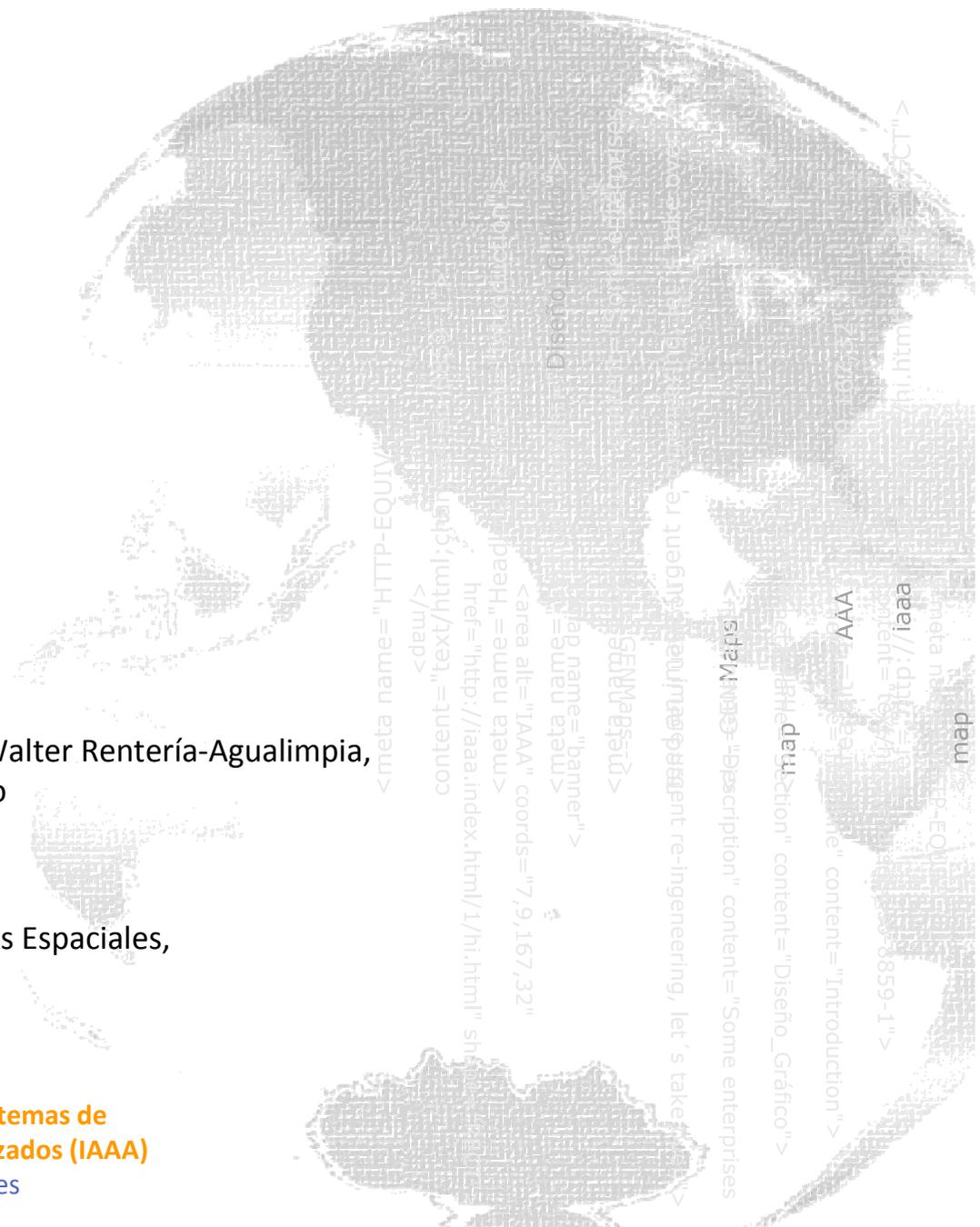
II Jornadas Ibéricas de Infraestructuras de Datos Espaciales,
Barcelona, 11 Noviembre, 2011



**Universidad
Zaragoza**



**Laboratorio de Sistemas de
Información Avanzados (IAAA)**
<http://iaaa.unizar.es>

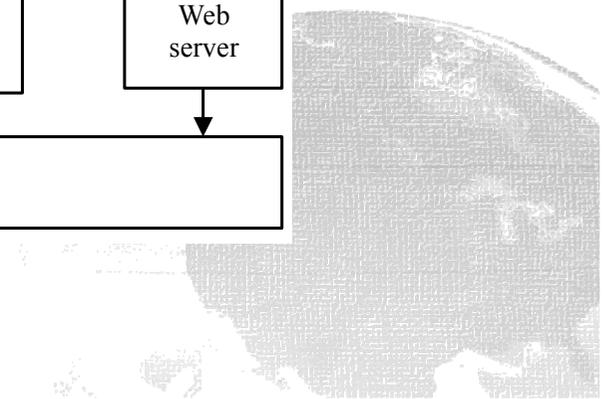
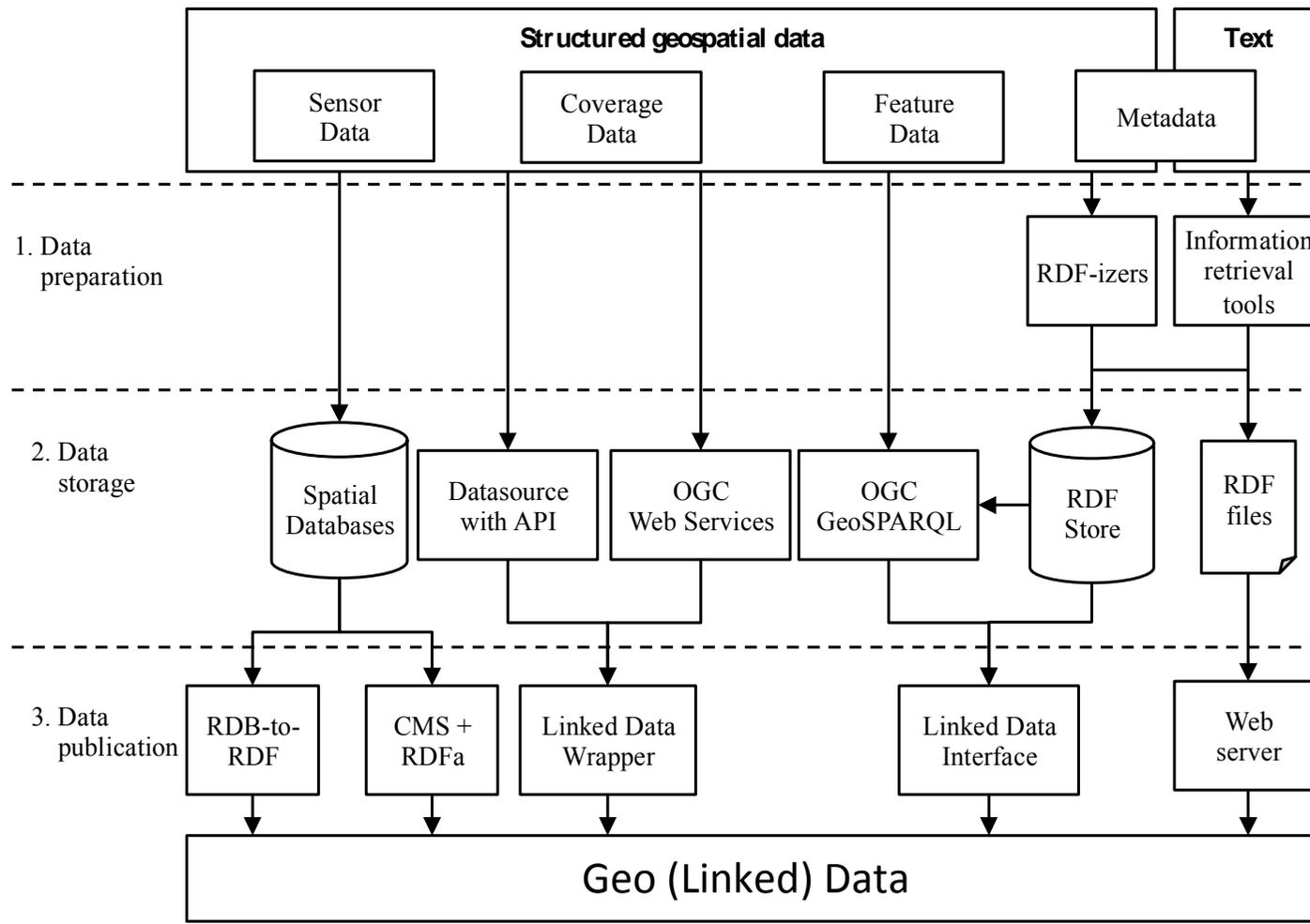


Setting the context: RDF wrappers for the Web of Data

- For live data and data repositories
- In open structured formats
 - A. Langegger and W. Wöß (2009), XLWrap – Querying and Integrating Arbitrary Spreadsheets with SPARQL
 - P. Coetzee, T Heath and E. Motta (2009), SparqPlug: Generating Linked Data from Legacy HTML, SPARQL and the DOM
- Or accesible through standard protocols
 - R. Cyganiak and C. Bizer (2007), Pubby – a linked data frontend for SPARQL endpoints
 - B. Haslhofer and B. Schandl (2008), The OAI2LOD Server: Exposing OAI-PMH Metadata as Linked Data



Linked Data and the Geospatial Web



Short intro to CSW

- OGC abstract (geospatial) catalogue model
 - D. Nebert, A. Whiteside, and P. Vretanos: Open GIS Catalogue Services Specification, version 2.0.2.. OpenGIS Publicly Available Standard OGC-07-006r1, Open GIS Consortium Inc., (2007).
- CSW = Catalogue Service for the Web
 - HTTP binding GET+POST
 - GetCapabilities = service metadata
 - GetRecords = query for metadata records
 - GetRecordById = present a metadata record
 - Minimum shared model derived from the Dublin Core vocabulary
- Less than 150 CSW servers in the world but they give access to huge quantities of free geospatial resources!!!



Typical CSW conversation

Support GET requests

Request:

```
GET/csw/servlet/cswservlet?request=GetRecordById&id=ESIGNMAPASRELIEVESERIE200701180000&elementSetName=full&outputSchema=http://www.opengis.net/cat/csw/2.0.2 HTTP/1.1
Host: www.ideo.es
```

Explicit request data view and format

XML based responses

Response:

```
HTTP/1.x 200 OK
Content-Type: application/xml;charset=ISO-8859-1
...
```

Dublin Core supported

Unique local identifiers

```
<?xml version = '1.0' encoding = 'ISO-8859-1'?>
<GetRecordByIdResponse
  xmlns="http://www.opengis.net/cat/csw/2.0.2"
  xmlns:dc="http://purl.org/dc/elements/1.1/"
  xmlns:dcterms="http://purl.org/dc/terms/"
  ...>
  <SummaryRecord>
    <dc:title>Mapas en Relieve</dc:title>
    <dc:identifier>
      ESIGNMAPASRELIEVESERIE200701180000
    </dc:identifier>
    ...
    <dc:format>PVC</dc:format>
    <dc:subject>elevation</dc:subject>
    <dc:subject>imageryBaseMapsEarthCover</dc:subject>
    <dc:type>dataset</dc:type>
    ...
    <dcterms:spatial>COUNTRIES.SPAIN</dcterms:spatial>
    ...
    <dcterms:spatial>
      northlimit=43.8;
      southlimit=37.83;
      westlimit=-9.32;
      eastlimit=0.72;
    </dcterms:spatial>
  </SummaryRecord>
</GetRecordByIdResponse>
```

Controlled terms (often)

Spatial content



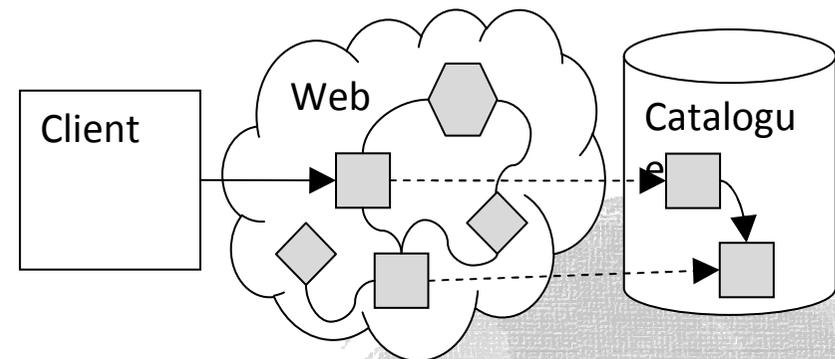
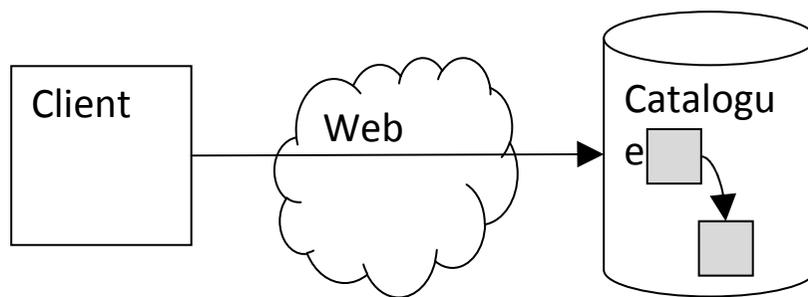
Universidad
Zaragoza



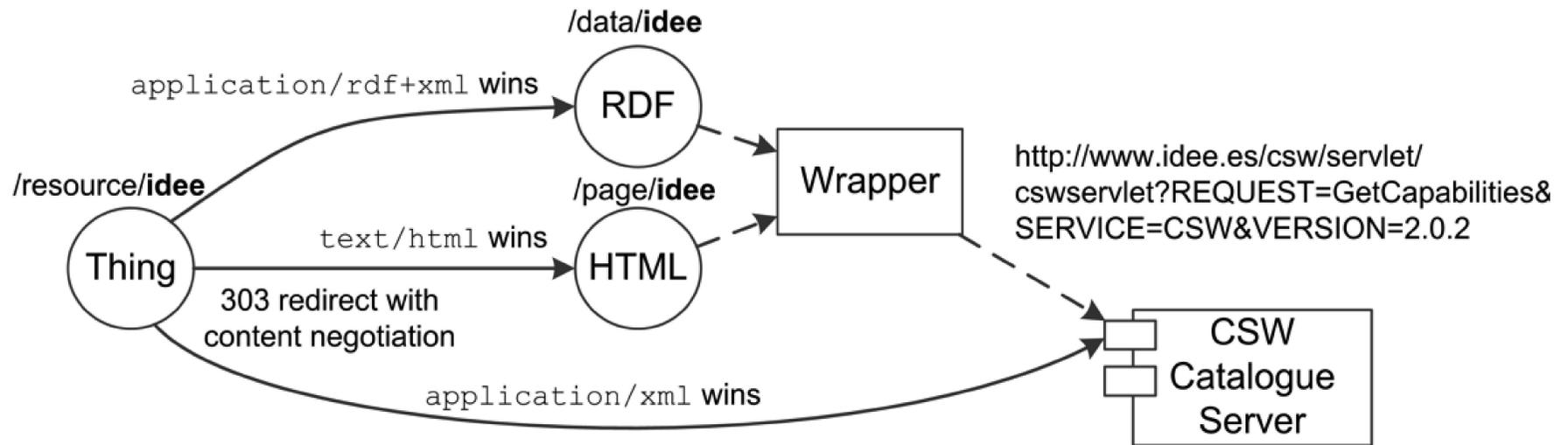
Laboratorio de Sistemas de
Información Avanzados (IAAA)
<http://iaaa.unizar.es>

CSW vs LD: Conceptual differences

- Record oriented
- RCP: Web resolves endpoints
- Closed world
- UML/XML Schema semantics
- Resource oriented
- ROA: Web resolves resources
- Open world
- Formal semantics



CSW2LD : Core idea



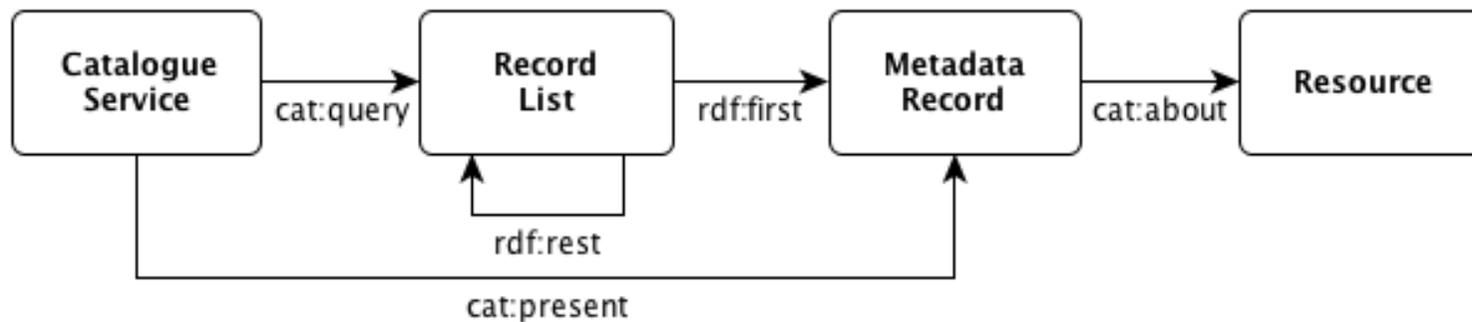
CSW2LD : Concepts

- Catalogue service (grounded in GetCapabilities)
 - Template /resource/{service}
 - The IDEE Catalogue:
/resource/idee
- Record list (grounded in GetRecords)
 - Template /resource/{service}/records{?keys+}
 - The list of records between 50 and 64 in the IDEE Catalogue:
/resource/idee/records?start=50&max=15
- Metadata record (grounded in GetRecordById)
 - Template /resource/{service}/record/{id}
 - The record identified as 100000_full1038_es in the IDEE Catalogue:
/resource/idee/record/100000_full1038_es

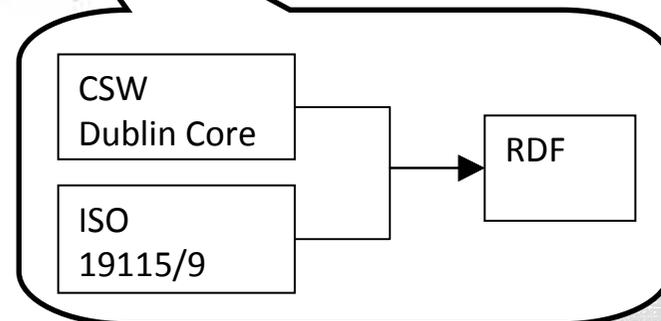
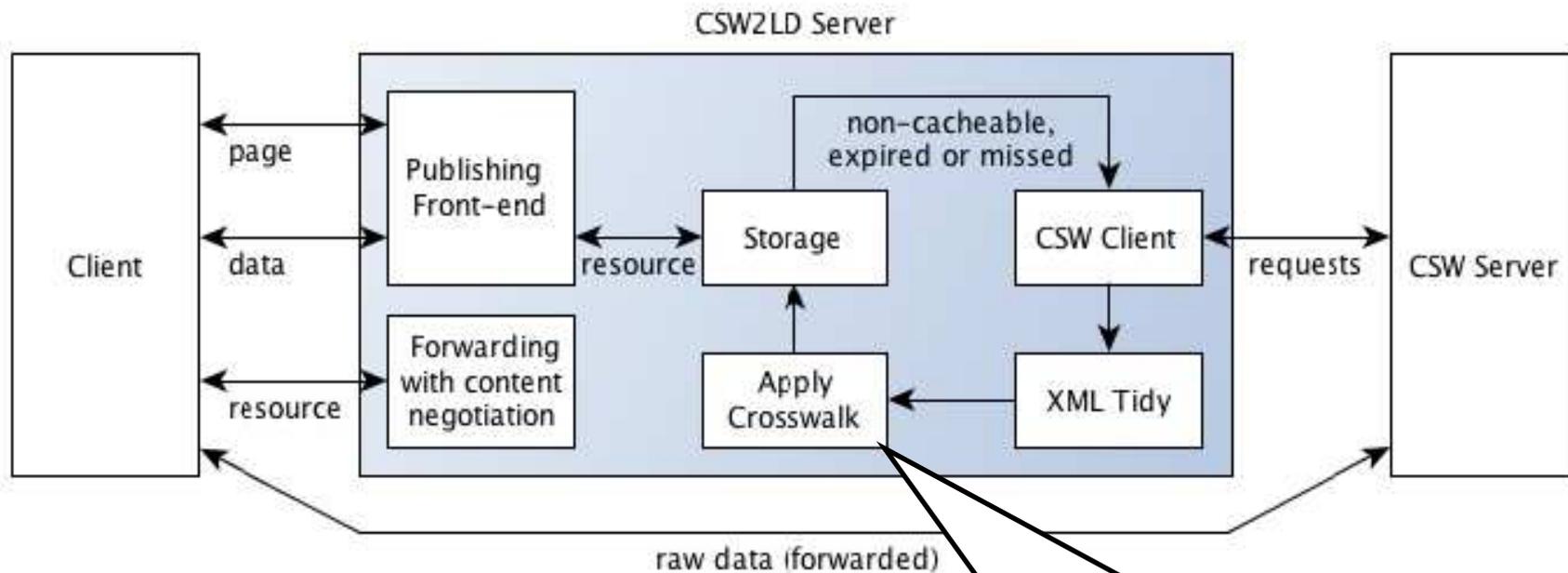


CSW2LD : Vocabularies

- Dublin Core : content
 - Resource description
- CSW2LD vocabulary : navigation
 - Catalogue Service (GetCapabilities)
 - Record List (GetRecords)
 - Metadata Record (GetRecordById)



CSW2LD : Prototype architecture



Conclusions

- Standard services are an opportunity for the LD community
 - Well documented and supported
 - Reusable wrapper
 - Should the LD community propose LD bindings?
- Need a shared approach for paging Web APIs
 - Vocabularies
 - Best practices
- Tools as CSW2LD need additional features for production
 - Obvious: Personalized UI, User profiles, Annotation, Edition
 - Not so obvious: Integration with existing workflows and applications

