

The OAI2LOD Server

Exposing OAI-PMH Metadata as Linked Data

Motivation

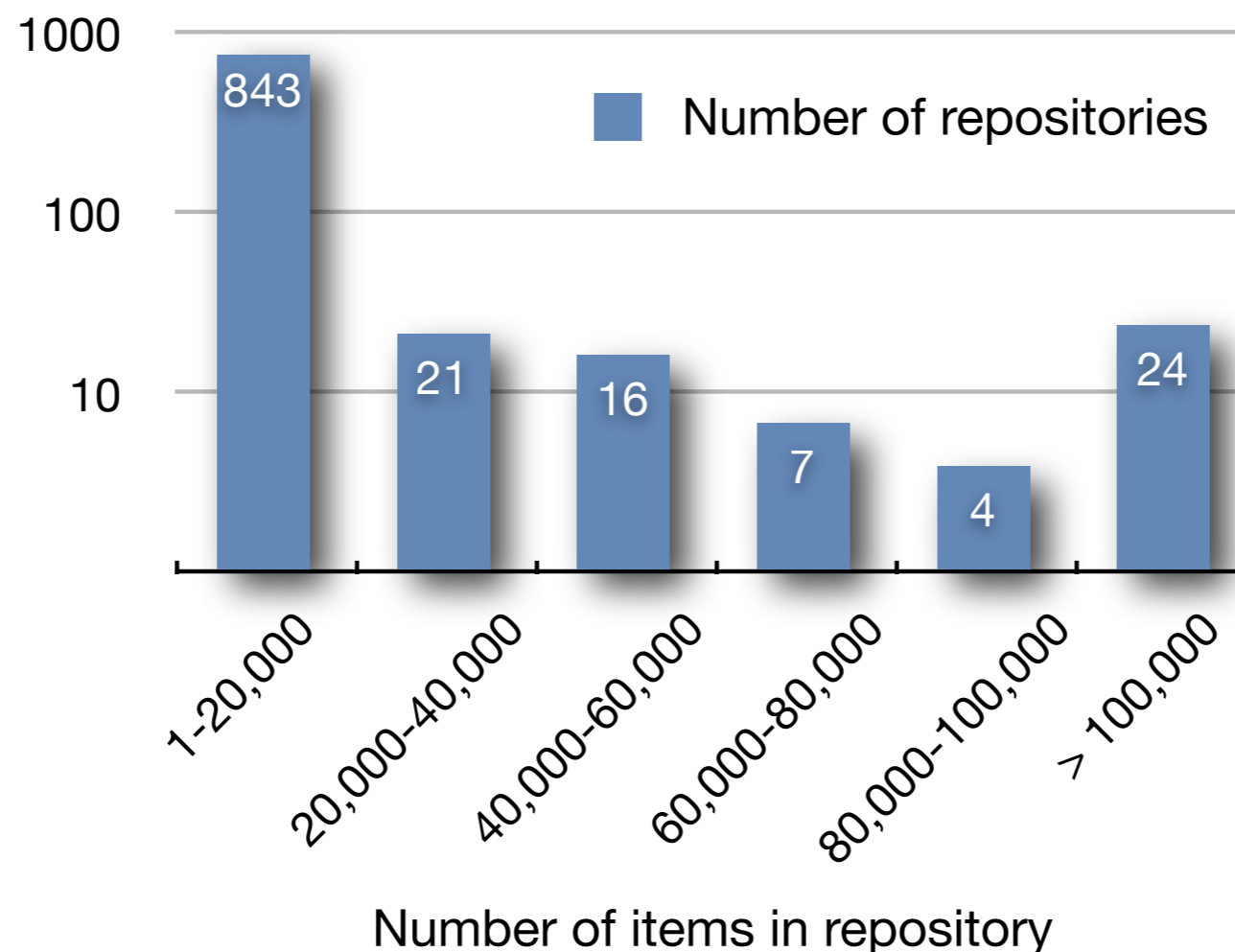
- more than 1700 institutions worldwide expose metadata via the

Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH)

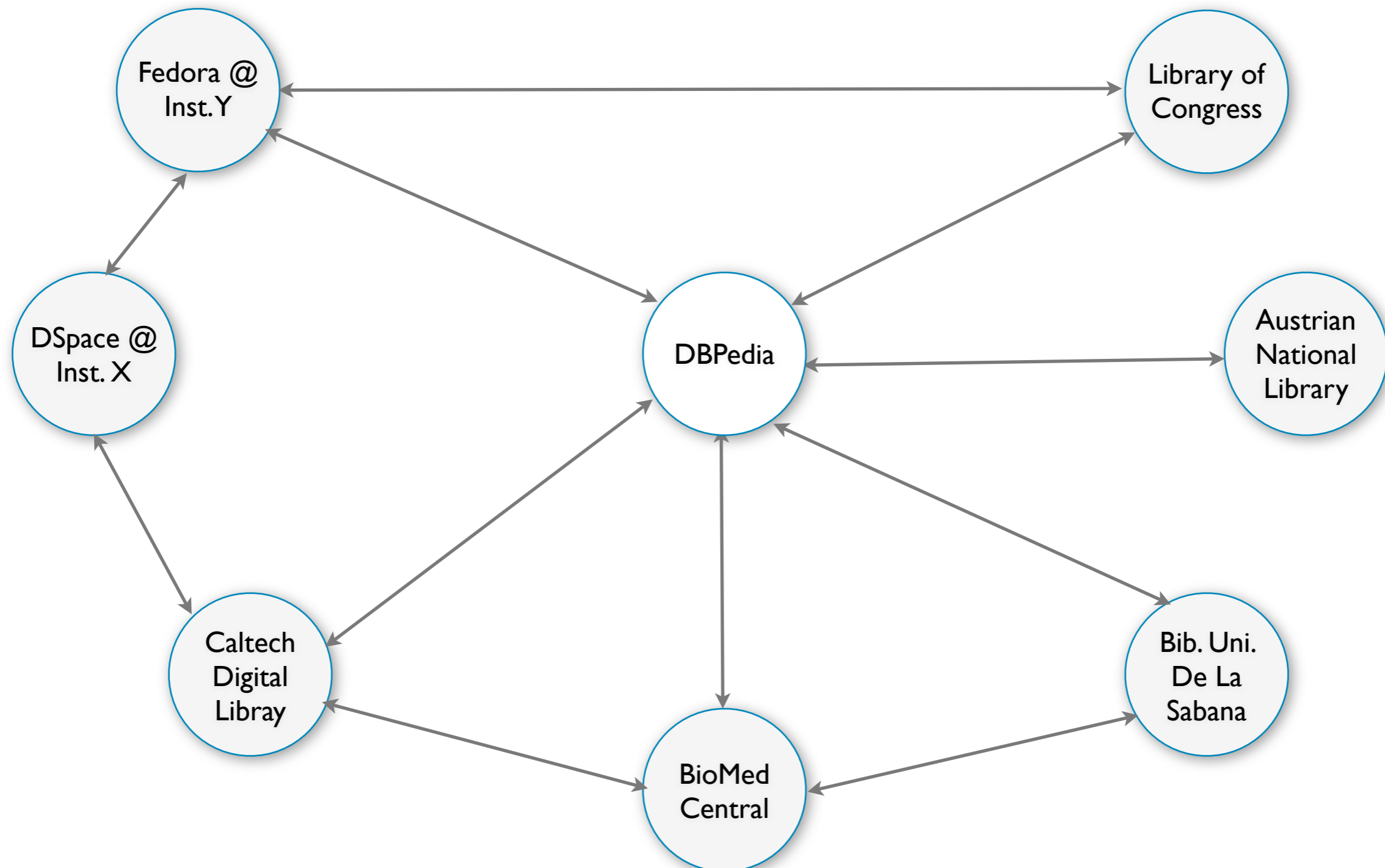
using open standards like URI, HTTP, XML

Motivation

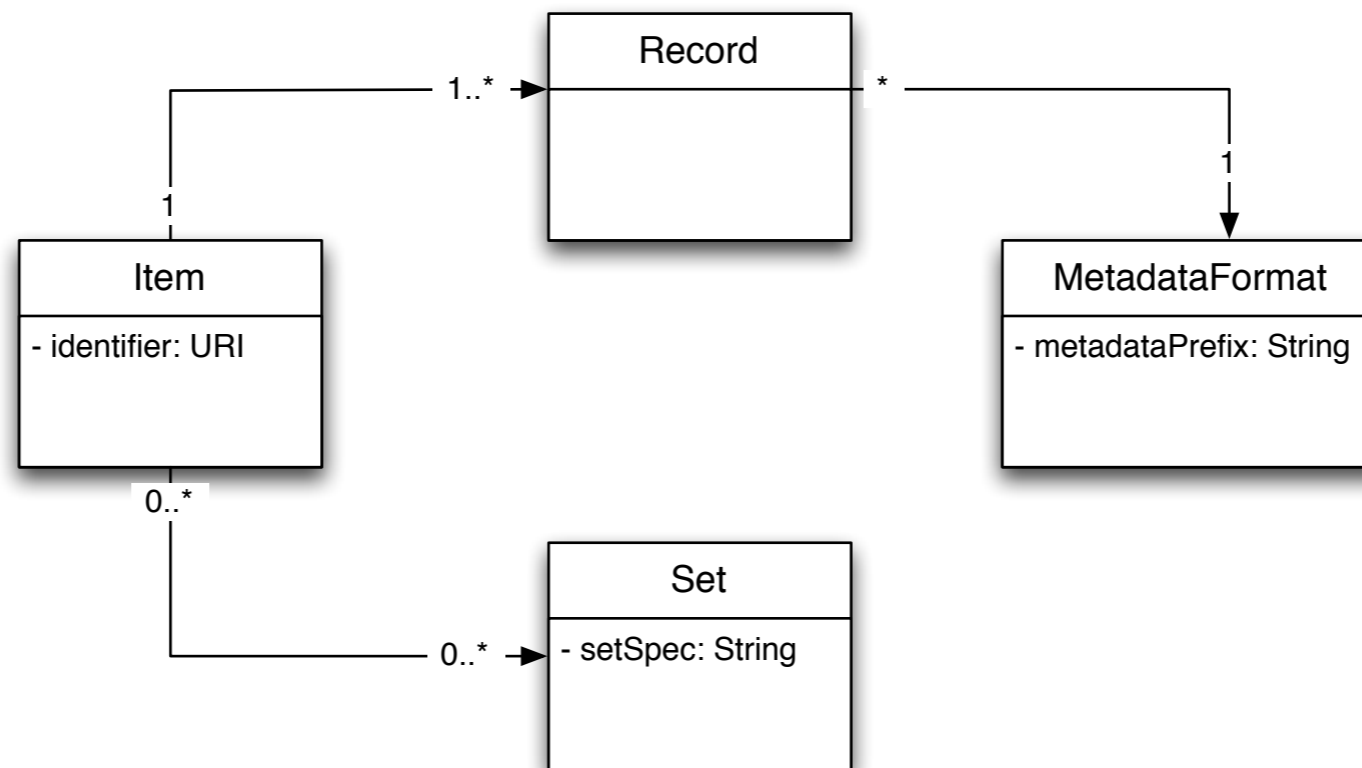
- from 900 investigated OAI-PMH repositories
- avg. number of Items/Data Provider: **~14,000**



Goal



OAI-PMH at a glance



OAI-PMH at a glance

sample request:

```
http://memory.loc.gov/cgi-bin/oai2_0?  
verb=GetRecord&  
identifier=oai:lcoal.loc.gov:loc.gdc/gcfr.0018_0163&  
metadataPrefix=oai_dc
```

OAI-PMH at a glance

sample response:

```
<OAI-PMH...>
```

```
...
```

```
<header>
```

```
  <identifier>oai:lcoal.loc.gov:loc.gdc/gcfr.0018_0163</identifier>
```

```
  <setSpec>ascfrbib</setSpec>
```

```
</header>
```

```
<metadata>
```

```
  <dc:title>Don Christopher Columbus to his friend, Don Louis de Santangel</dc:title>
```

```
  <dc:creator>Columbus, Christopher</dc:creator>
```

```
...
```

```
</metadata>
```

```
</GetRecord>
```

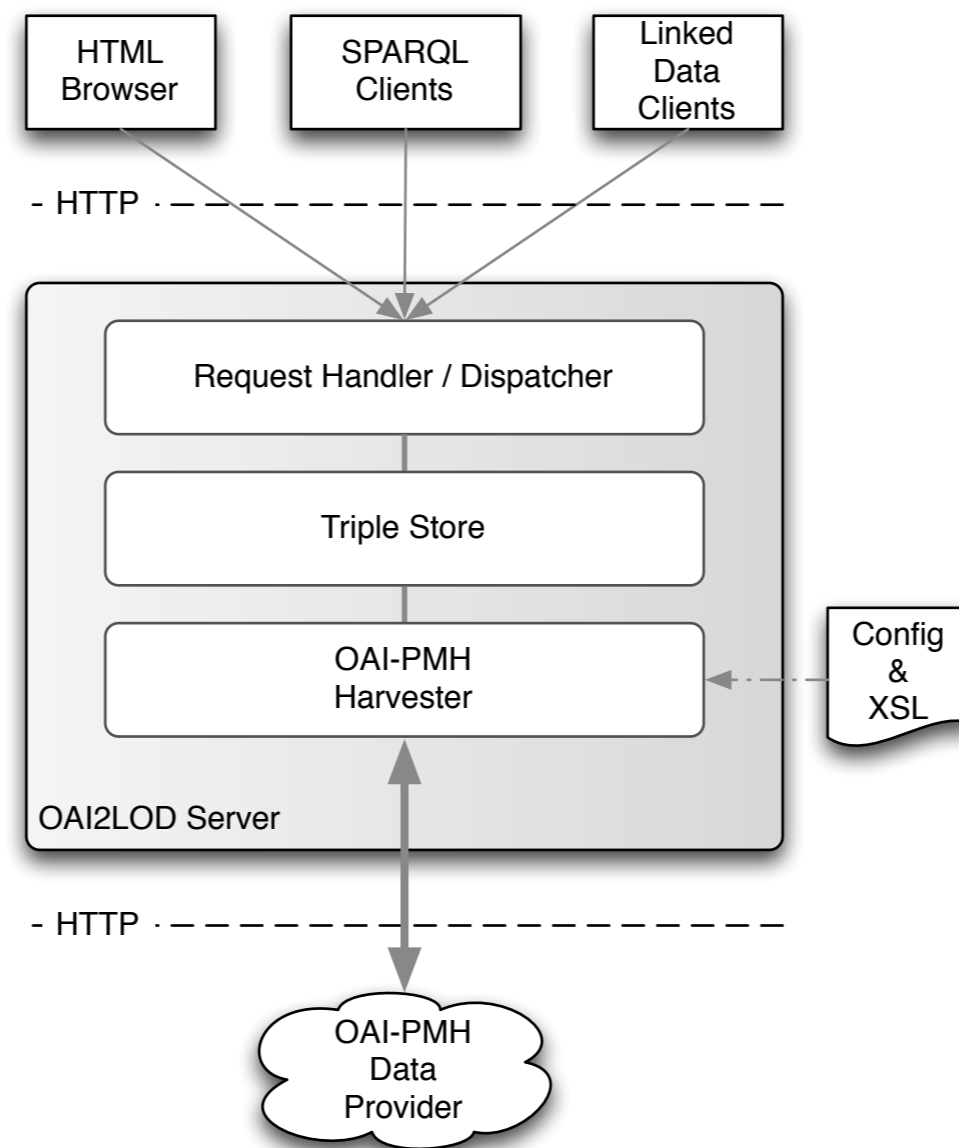
OAI-PMH at a glance

- **ListRecords**
 - batch retrieval of records
- **ListIdentifiers**
 - returns item identifiers
- **ListSets**
 - returns available sets
- **Identify**

The OAI2LOD Server

- makes OAI-PMH resources (items/sets) dereferencable via their URIs
- provides metadata access for humans and machines that “do not know” the OAI-PMH protocol
- exposes a SPARQL interface to these data
- links metadata with other LOD sources

The OAI2LOD Server



The OAI2LOD Server

- LOD Rule 1+2 - “Things should have (resolvable) URIs”
- Items: http://example.com/resources/item/oai:lcoal.loc.gov:loc.gdc/gcfr.0018_0163
- Sets: <http://example.com/resources/set/ascfribib>
- Vocabularies

The OAI2LOD Server

- LOD Rule 3: “Deliver useful information when URIs are dereferenced”
- Content negotiation based on HTTP Accept
 - RDF for machines
 - (X)HTML for humans

The OAI2LOD Server

- LOD Rule 4: “Metadata should contain links to other related resources”
 - link to any other OAI2LOD / LOD data sources
 - configurable linking property - e.g., [rdfs:seeAlso](#)
 - linking heuristics based on configurable string similarity metrics (e.g., Levensthein, SoundEx)

Outlook

- the number of OAI-PMH repositories will grow
- major initiatives push its adoption
 - e.g., “The European Library”
 - integrates 47 national libraries
 - provides access to approx. 150 M items

Outlook

- OAI-ORE (Object Reuse and Exchange)
- latest standardization effort for the “description and exchange of aggregations of Web resources”
- data model is based on RDF
- concepts have dereferencable URI identifier
- aggregations are the means to “link” resources

Further Infos

- OAI2LOD - Demos, Download & Instructions:

<http://www.mediaspaces.info/tools/oai2lod/>

- Contact

bernhard.haslhofer@univie.ac.at

- the end -

BACKUP

Motivation

- each OAI-PMH compliant repository
- **MUST** expose Dublin Core metadata
- **MAY** provide other MetadataFormats

