the OAI-PMH

Requests

Replies

service provider
data provider

harvester

repository

6 Requests

Replies
Core concepts in OAI-PMH

• low-barrier interoperability
• data-provider & service-provider model
• metadata harvesting model
  OAI-PMH
  HTTP based
  Reply • XML Schema
  • Self contained
• shared metadata format and parallel, community-specific metadata formats
  Dublin Core
• acceptable use
  Community specific, oai-rights
OAI-PMH data model

- Resource
- Item
- Records

Entry point to all records pertaining to the resource

OAI-PMH identifier

OAI-PMH sets

Dublin Core metadata

MARCXML metadata

Datestamp

MetadataPrefix
OAI-PMH harvesting tools

service provider

Supporting protocol requests:
- Identify
- ListMetadataFormats
- ListSets

data provider

repository
OAI-PMH harvesting tools

service provider

Supporting protocol requests:
• Identify
• ListMetadataFormats
• ListSets

Harvesting protocol requests:
• ListRecords
• ListIdentifiers
• GetRecord

data provider

server provider

repoistory

harvester

harvester

repository
OAI-PMH harvesting tools

service provider

Datestamp
Identifier Set
Records

data provider

harvester

repository
harvesting requests

**service provider**

* from=a
* until=b
* set=klm

**data provider**

ListRecords * metadataPrefix=dc

ListRecords / Time / Request

REPEAT

• Identifier
• Datestamp
• Metadata

/REPEAT
OAI-PMH identifiers

• Not (necessarily) identifier of the resource
• Each item **must** have a **globally unique** identifier

identifiers **must follow rules for valid URIs**

• Example:
  - oai:<archiveId>:<recordId>
  - oai:etd.vt.edu:etd-1234567890

• Each identifier **must resolve to a single item and always to the same item**
  - Can't reuse OAI item identifiers
OAI-PMH datestamps

- Needed for every OAI record to support incremental harvesting
- Must be updated when addition or modification or deletion made in order to ensure changes are correctly propagated to harvesters
  - Also for dynamically generated metadata formats
- Different from dates within the metadata – OAI datestamp is used only for harvesting
- Can be either YYYY-MM-DD or YYYY-MM-DDThh:mm:ssZ (must be GMT timezone)