**OAI-PMH Data Provider**

**Architecture Document**

**Prepared by : Irwan Krisna**

**Australian National University**

1. **Background:**

As part of the ANDS metadata Store project, institutions are required to provide information about their researchers to Research Data Australia (RDA) system. One way to identify each researcher is to use a unique Identifier provided by the National Library of Australia (NLA). The NLA together with ANDS have developed an infrastructure (called Trove) to facilitate the generation of Researchers IDs.

For that purpose, the ANU Metadata Team has done some feasibility studies on several OAI-PMH software packages to support Party Records contributions to the NLA. The team has finally chosen and used an open source software which was built by software developers from University of Adelaide. It was developed to meet the ANDS Data Commons Project requirements (Source code link: <http://code.google.com/p/oai-pmh-2/>).

1. **The System Architecture:**
2. **System Main Model:**



**Note**: The (dotted) arrow refers to an implementation to a certain model.

As seen in the above diagram, the technologies used to implement the *OAI-PMH provider* are the following :

1. *PHP 5.4.5* : Scripting Language
2. *Apache* : web server
3. *MySQL 5.0.95* : Database

Application Web address: <http://dc7-dev2.anu.edu.au/oai/>

1. **State Chart Diagram model for “XML OAI – PMH Requests and Responses” :** 

**State Chart Diagram Description:**

* The requests are submitted by data harvester (e,g: the National Library of Australia).
* OAI-PMH [requests](http://www.openarchives.org/OAI/openarchivesprotocol.html#ProtocolMessages) are expressed as [HTTP](http://www.ietf.org/rfc/rfc2616.txt) requests.They **must** be submitted using either the HTTP GET or POST methods.
* A record is returned in an XML-encoded byte stream in response to an OAI-PMH request for metadata from an item.
* There are Six types of OAI – PMH requests*: Identify, List Metadataformat, List Sets, Get Record, List Identifiers* and *List Records* as described in the next section.

1. **The Six OAI-PMH Protocol Requests:**
2. **Identify**:This verb is used to retrieve information about a repository. Some of the information returned is required as part of the OAI-PMH. Repositories **may** also employ the Identify verb to return additional descriptive information.
3. **List MetaData Format**:This verb is used to retrieve the metadata formats available from a repository. An optional argument restricts the request to the formats available for a specific item.
4. **List Sets:** This verb is used to retrieve the set structure of a repository, useful for [selective harvesting](http://www.openarchives.org/OAI/openarchivesprotocol.html#SelectiveHarvestingandSets).
5. **Get Record:**This verb is used to retrieve an individual metadata record from a repository. Required arguments specify the identifier of the item from which the record is requested and the format of the metadata that should be included in the record. Depending on the level at which a repository tracks [deletions](http://www.openarchives.org/OAI/openarchivesprotocol.html#deletion), a header with a "deleted" value for the status attribute **may** be returned, in case the metadata format specified by the metadataPrefix is no longer available from the repository or from the specified item.
6. **List Identifiers**: This verb is an abbreviated form of [ListRecords](http://www.openarchives.org/OAI/openarchivesprotocol.html#ListRecords), retrieving only [headers](http://www.openarchives.org/OAI/openarchivesprotocol.html#header) rather than [records](http://www.openarchives.org/OAI/openarchivesprotocol.html#Record).
7. **List Records:**This verb is used to harvest records from a repository. Optional arguments permit [selective harvesting](http://www.openarchives.org/OAI/openarchivesprotocol.html#Datestamp) of [records](http://www.openarchives.org/OAI/openarchivesprotocol.html#header) based on [set](http://www.openarchives.org/OAI/openarchivesprotocol.html#Set) membership and/or datestamp. Depending on the repository's support for [deletions](http://www.openarchives.org/OAI/openarchivesprotocol.html#deletion), a returned [header](http://www.openarchives.org/OAI/openarchivesprotocol.html#header)**may** have a status attribute of "deleted" if a record matching the arguments specified in the request has been deleted. No metadata will be present for records with deleted status.
8. **XML Schema (RIF –CS):**

This section describes briefly about the RIF-CS ( Registry Interchange Format, Collections and Services) that is being developed and used as a data interchange format for supporting the submissions of collections metadata to the ANDS collections Registry. These below are its data elements:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **Party registry object (Required)** |  |  |  |  |
|  | **Label** | **Schema element name** | **Required/Recommended** | **Requirement status** |  |
|  | registry objects | [<registryObjects>](http://www.ands.org.au/guides/cpguide/cpgrecstructure.html) | Required | All requirements specified March 2010, unless otherwise noted (RIF-CS 1.0.1) |  |
|  | registry object | [<registryObject>](http://www.ands.org.au/guides/cpguide/cpgrecstructure.html) | Required |  |  |
|  | originating source | [<originatingSource>](http://www.ands.org.au/guides/cpguide/cpgorigsource.html) | Required |  |  |
|  | Group | [@group](http://www.ands.org.au/guides/cpguide/cpggroup.html) | Required |  |  |
|  | key | [<key>](http://www.ands.org.au/guides/cpguide/cpgkey.html) | Required |  |  |
|  | class type (party) | [<party>](http://www.ands.org.au/guides/cpguide/cpgparty.html) | Required |  |  |
|  | party type | [<party type>](http://www.ands.org.au/guides/cpguide/cpgparty.html) | Required |  |  |
|  | name | [<name>](http://www.ands.org.au/guides/cpguide/cpgname.html) | Required |  |  |
|  | identifier | [<identifier>](http://www.ands.org.au/guides/cpguide/cpgidentifiers.html) | Recommended |  |  |
|  | location/address | [<location>](http://www.ands.org.au/guides/cpguide/cpglocationintro.html) | Recommended |  |  |
|  | related object(activity) | [<rel tedObject>](http://www.ands.org.au/guides/cpguide/cpgrelatedobject.html) | Recommended |  |  |
|  | related object(collection) | [<relatedObject>](http://www.ands.org.au/guides/cpguide/cpgrelatedobject.html) | Required |  |  |
|  | description | [<description>](http://www.ands.org.au/guides/cpguide/cpgdescription.html) | Recommended |  |  |
|  | subject | [<subject>](http://www.ands.org.au/guides/cpguide/cpgsubject.html) | Recommended |  |  |
|  | existence dates | [<existenceDates>](http://www.ands.org.au/guides/cpguide/cpgexistencedates.html) | Recommended | Added Dec 2011 (RIF-CS v1.3.0) |  |
|  |  |  |  |  |  |

**d. Database Design :**



**References:**

* <http://services.ands.org.au/documentation/rifcs/1.3/guidelines/rif-cs.html>
* [**http://www.ands.org.au/guides/cpguide/cpgparty.html**](http://www.ands.org.au/guides/cpguide/cpgparty.html)**)**