This briefing paper introduces the concept of the Open Archives Initiative-Protocol for Metadata Harvesting (OAI-PMH), its uses and importance in the field of digital repositories.

What is it & why is it important?

The Open Archives Initiative (OAI) was founded with the specific aim of developing and promoting interoperability standards which facilitate open access and allow the dissemination of content in digital repositories. In essence, OAI has been an attempt to create a "low-barrier interoperability framework" for access to services that contain digital content.

The OAI framework is based on a client/server architecture using two groups of services: the data provider (for example; repositories) and the service provider (for example; a metadata harvester such as OAIster). The data provider services are generally embedded into the repository or open archive software and provide a mechanism to allow open access to the metadata held within the repository. Service providers make use of the OAI-PMH interfaces provided by the data providers to harvest the metadata of a central location. Using this aggregated metadata, the service provider can then offer value added services such as search and dissemination tools.

The complete process of a service provider gathering metadata from a number of distributed repositories in to a combined data store is known as harvesting.

The OAI Request & Response

OAI-PMH requests have two parts to their structure: the base-url which is the path to the handler for the OAI-PMH requests, and the keyword arguments, a set of key-value pairs which specify the type and details of the request. OAI-PMH provides support for six request types known as verbs. The response by the data provider to an OAI request is formatted as an HTTP reply encoded in XML. When harvesting, the service provider requests metadata records from the data provider using one of the six verbs: Identify, ListMetadataFormats, ListSets, ListIdentifiers, ListRecords and GetRecord.

For OAI compliance, a repository must expose and allow dissemination of unqualified Dublin Core. The Dublin Core (DC) metadata schema is defined as the minimum standard to allow interoperability between repositories. This schema contains 15 elements defined by the Dublin Core Metadata Initiative. In addition, other metadata schemas can be exposed via OAI-PMH, providing the metadata schema can be encoded as XML and a valid XML schema has been defined for this metadata.

How does OAI-PMH work?

The OAI framework is based on a client/server architecture using two groups of services: the data provider (for example; repositories) and the service provider (for example; a metadata harvester such as OAIster). The OAI Registration

While a repository can be harvested simply by providing an OAI-PMH interface, registration provides a useful means of promoting the visibility of your repository to service providers for harvesting. The Open Archives Initiative provides a service allowing your repository to be registered as a data provider in the OAI registry.

The complete process of a service provider gathering metadata from a number of distributed repositories in to a combined data store is known as harvesting.

The OAI Request & Response

OAI-PMH requests have two parts to their structure: the base-url which is the path to the handler for the OAI-PMH requests, and the keyword arguments, a set of key-value pairs which specify the type and details of the request. OAI-PMH provides support for six request types known as verbs. The response by the data provider to an OAI request is formatted as an HTTP reply encoded in XML. When harvesting, the service provider requests metadata records from the data provider using one of the six verbs: Identify, ListMetadataFormats, ListSets, ListIdentifiers, ListRecords and GetRecord.

For OAI compliance, a repository must expose and allow dissemination of unqualified Dublin Core. The Dublin Core (DC) metadata schema is defined as the minimum standard to allow interoperability between repositories. This schema contains 15 elements defined by the Dublin Core Metadata Initiative. In addition, other metadata schemas can be exposed via OAI-PMH, providing the metadata schema can be encoded as XML and a valid XML schema has been defined for this metadata.

How does OAI-PMH work?

The OAI framework is based on a client/server architecture using two groups of services: the data provider (for example; repositories) and the service provider (for example; a metadata harvester such as OAIster). The data provider services are generally embedded into the repository or open archive software and provide a mechanism to allow open access to the metadata held within the repository. Service providers make use of the OAI-PMH interfaces provided by the data providers to harvest the metadata to a central location. Using this aggregated metadata, the service provider can then offer value added services such as search and dissemination tools.

OAI Registration

While a repository can be harvested simply by providing an OAI-PMH interface, registration provides a useful means of promoting the visibility of your repository to service providers for harvesting. The Open Archives Initiative provides a service allowing your repository to be registered as a data provider in the OAI registry.
The registry is a publicly accessible list of all OAI conformant repositories which allows easy discovery of data providers by service providers. When registering your repository, the OAI service will perform conformance testing to ensure your repository complies to OAI-PMH. If validation is successful, your repository will be added into the registry.

The OAI will also periodically test your repository for conformance. If the analysis fails, your repository will be removed, and a notification email sent to the administrator detailing the reason for removal. This ensures the integrity of the OAI registry and will assist in increasing the visibility of your repository to other service providers, direct registration with these services is also possible to guarantee your repository is harvested by them. The main service providers which require additional registration are Intute Search 6, OAIster 2 and OpenDOAR 7.

Registering With Other Services

Although registering with the Open Archives Initiative will assist in increasing the visibility of your repository to other service providers, direct registration with these services is also possible to guarantee your repository is harvested by them. The main service providers which require additional registration are Intute Search 6, OAIster 2 and OpenDOAR 7.

References & further information:

1Open Archives Initiative http://www.openarchives.org/
The Open Archives Initiative develops and promotes interoperability standards that aim to facilitate the efficient dissemination of content.

2OAIster http://www.oaister.org/
OAIster is a search service for Open Access repositories, providing access to an extensive range of academic content.

3Open Archives Forum http://www.oaforum.org/
The Open Archives Forum provided a Europe-based focus for dissemination of information about European activity related to open archives and, in particular, to the Open Archives Initiative.

4Dublin Core Metadata Initiative http://dublincore.org/
The Dublin Core Metadata Initiative is an open organisation engaged in the development of interoperable online metadata standards.

5OAI Registration
http://www.openarchives.org/data/registerasprovider.html
Information on registering with the OAI as an OAI-PMH conformant data provider.

6Intute Search http://irs.ukoln.ac.uk/
Intute repository search provides a service to search for various types of scholarly literature across many UK academic repositories.

7OpenDOAR http://www.opendoar.org/
OpenDOAR provides a list of academic open access repositories, alongside information and tools that help support both repository administrators and service providers.

Repositories Support Project http://www.rsp.ac.uk/
The Repositories Support Project (RSP) aims to co-ordinate and deliver good practice and practical advice to HEIs to enable the implementation, management and development of digital institutional repositories.

Open Archives Initiative - Repository Explorer http://re.cs.uct.ac.za/
This site presents an interface to interactively test archives for compliance with the OAI-PMH.

The registry is a publicly accessible list of all OAI conformant repositories which allows easy discovery of data providers by service providers. When registering your repository, the OAI service will perform conformance testing to ensure your repository complies to OAI-PMH. If validation is successful, your repository will be added into the registry.

The OAI will also periodically test your repository for conformance. If the analysis fails, your repository will be removed, and a notification email sent to the administrator detailing the reason for removal. This ensures the integrity of the OAI registry and will assist in increasing the visibility of your repository to other service providers, direct registration with these services is also possible to guarantee your repository is harvested by them. The main service providers which require additional registration are Intute Search 6, OAIster 2 and OpenDOAR 7.

Registering With Other Services

Although registering with the Open Archives Initiative will assist in increasing the visibility of your repository to other service providers, direct registration with these services is also possible to guarantee your repository is harvested by them. The main service providers which require additional registration are Intute Search 6, OAIster 2 and OpenDOAR 7.

References & further information:

1Open Archives Initiative http://www.openarchives.org/
The Open Archives Initiative develops and promotes interoperability standards that aim to facilitate the efficient dissemination of content.

2OAIster http://www.oaister.org/
OAIster is a search service for Open Access repositories, providing access to an extensive range of academic content.

3Open Archives Forum http://www.oaforum.org/
The Open Archives Forum provided a Europe-based focus for dissemination of information about European activity related to open archives and, in particular, to the Open Archives Initiative.

4Dublin Core Metadata Initiative http://dublincore.org/
The Dublin Core Metadata Initiative is an open organisation engaged in the development of interoperable online metadata standards.

5OAI Registration
http://www.openarchives.org/data/registerasprovider.html
Information on registering with the OAI as an OAI-PMH conformant data provider.

6Intute Search http://irs.ukoln.ac.uk/
Intute repository search provides a service to search for various types of scholarly literature across many UK academic repositories.

7OpenDOAR http://www.opendoar.org/
OpenDOAR provides a list of academic open access repositories, alongside information and tools that help support both repository administrators and service providers.

Repositories Support Project http://www.rsp.ac.uk/
The Repositories Support Project (RSP) aims to co-ordinate and deliver good practice and practical advice to HEIs to enable the implementation, management and development of digital institutional repositories.

Open Archives Initiative - Repository Explorer http://re.cs.uct.ac.za/
This site presents an interface to interactively test archives for compliance with the OAI-PMH.