



Open Archives Initiative

Frequently Asked Questions (FAQ)

Protocol Version 2.0 of 2002-06-14

Document Version 2002/06/10T11:00:00Z

<http://www.openarchives.org/documents/FAQ.html>

Editors

[Carl Lagoze](#) (OAI Executive; [Cornell University - Computer Science](#))

[Herbert Van de Sompel](#) (OAI Executive; [Los Alamos National Laboratory - Research Library](#))

[Michael Nelson](#) ([Old Dominion University - Computer Science](#))

[Simeon Warner](#) ([Cornell University - Computer Science](#))

Contents

[What is the purpose of this document?](#)

[What is the mission of the Open Archives Initiative?](#)

[Who manages the Open Archives Initiative?](#)

[Who supports the Open Archives Initiative?](#)

[How can I participate in the Open Archives Initiative?](#)

[How can I find out more about the Open Archives Initiative?](#)

[What is a data provider?](#)

[What is a service provider?](#)

[What do you mean by an "Archive"?](#)

[What do you mean by "Open"?](#)

[Is the Open Archives Initiative only about E-Prints?](#)

[Is the Open Archives Initiative only concerned with metadata?](#)

[Can I use the OAI-PMH to harvest content?](#)

[How long will it take me to implement the protocol?](#)

[What is the relationship between the Open Archives Metadata Harvesting Protocol and other protocols such as Z39.50?](#)

[What advantage is there for me in participating in the OAI?](#)

[How many participants are there in the OAI?](#)

[What is the advantage of registering my repository with the OAI?](#)

[Is there a way of affiliating my repository with other OAI-PMH conformant repositories?](#)

[What is the advantage of using OAI identifiers for metadata records?](#)

[What is the difference between the versions of the OAI-PMH?](#)

[Is version 2.0 of the OAI-PMH compatible with earlier versions?](#)

[Will there be continued support for version 1.1 of the OAI-PMH?](#)

[Why does the protocol mandate a common metadata format \(and why is that common format Dublin Core\)?](#)

[What if I want to expose metadata in other formats than Dublin Core?](#)

[How do I let others know that my repository supports the OAI-PMH?](#)

[What about intellectual property issues?](#)

[What if I have more questions?](#)

What is the purpose of this document?

This FAQ will hopefully answer the most common questions about the Open Archives Initiative, its mission, and the interoperability protocol that it develops and promotes. All the information included here is also included in other documents available from the main Open Archives Initiative web site at <http://www.openarchives.org>. This is a dynamic document - We expect this FAQ to evolve with increasing experience with the OAI interoperability framework. Check the online copy of this page for latest updates.

Implementers note: This document is not a replacement for the [Open Archives Initiative Protocol for Metadata Harvesting specification](#) (referred to as OAI-PMH for the remainder of this document), which you should read before working with the protocol.

[Top](#)

What is the mission of the Open Archives Initiative?

The Open Archives Initiative develops and promotes interoperability standards that aim to facilitate the efficient dissemination of content. The Open Archives Initiative has its roots in an effort to enhance access to e-print archives as a means of increasing the availability of scholarly communication. Continued support of this work remains a cornerstone of the Open Archives program. The fundamental technological framework and standards that are developing to support this work are, however, independent of the both the type of content offered and the economic mechanisms surrounding that content, and promise to have much broader relevance in opening up access to a range of digital materials. As a result, the Open Archives Initiative is currently an organization and an effort explicitly in transition, and is committed to exploring and enabling this new and broader range of applications. As we gain greater knowledge of the scope of applicability of the underlying technology and standards being developed, and begin to understand the structure and culture of the various adopter communities, we expect that we will have to make continued evolutionary changes to both the mission and organization of the Open Archives Initiative.

[Top](#)

Who manages the Open Archives Initiative?

Policy decisions about the Open Archives Initiative are made by a [Steering Committee](#). The interoperability infrastructure was developed by a technical committee, which continues to advise on the infrastructure as experience with it develops. Herbert Van de Sompel and Carl Lagoze are responsible for coordination of OAI activities, which are centered at Cornell University. The mail address for OAI correspondence is openarchives@openarchives.org.

[Top](#)

Who supports the Open Archives Initiative?

Support for Open Archives Initiative activities comes from the [Digital Library Federation](#), [Coalition for Networked Information](#), and National Science Foundation Grant No. IIS-9817416 ([Project Prism](#)).

[Top](#)

How can I participate in the Open Archives Initiative?

The OAI invites anyone to participate in the interoperability framework that is defined in the [Open Archives Metadata Harvesting Protocol](#). Participation has two dimensions:

- *Data Providers* Data Providers administer systems that support the OAI-PMH as a means of exposing metadata. A data provider may choose to register via the [OAI data provider registration page](#), and thereby [publicize](#) the fact that they have adopted the OAI-PMH.
- *Service Providers* use metadata harvested via the OAI-PMH as a basis for building value-added services. A service provider may choose to register by following the instructions available at the [service provider registration page](#). As a result of this registration, service providers [publicize](#) their existence.

Participants may also subscribe to either of the open mailing lists:

1. OAI-general - This is a general discussion group for people interested in the OAI. Subscribe to the list at [mailman/listinfo/oai-general](#). This mailing list is archived and can be searched at [pipermail/oai-general/](#).
2. OAI-implementers - This is a technical discussion group for sharing information on implementation experience. Subscribe to the list at [mailman/listinfo/oai-implementers](#). This mailing list is archived and can be searched at [mailman/listinfo/oai-implementers](#).

[Top](#)

How can I find out more about the Open Archives Initiative?

Information about the Open Archives Initiative is available in a number of web-accessible documents:

- OAI web site (<http://www.openarchives.org>) - This is the best source for up-to-date information and news about the OAI, providing links to all documents in the remainder of this list.
- OAI-PMH specification (<http://www.openarchives.org/OAI/openarchivesprotocol.html>) - The specification of the metadata harvesting protocol.
- OAI-PMH implementation guidelines (<http://www.openarchives.org/OAI/2.0/guidelines.htm>) - A set of documents supplemental to the OAI-PMH specification that describe best practices and community-specific extensibility features.
- OAI-PMH migration document (<http://www.openarchives.org/OAI/2.0/migration.htm>) - A document supplemental to the OAI-PMH specification that describes differences between the version 2.0 and 1.1 specification.
- OAI FAQ (<http://www.openarchives.org/documents/FAQ.html>) - This document: consult the web version for the latest revision.

- oai-general archives (<http://www.openarchives.org/pipermail/oai-general/>) - The archives of the mail list for discussion of non-technical OAI issues and announcements.
- oai-implementers archives (<http://www.openarchives.org/pipermail/oai-implementers/>) - The archives of the mail list for discussing technical OAI issues.

What is a data provider?

A data provider maintains one or more repositories (web servers) that support the OAI-PMH as a means of exposing metadata.

[Top](#)

What is a service provider?

A service provider issues OAI-PMH requests to data providers and uses the metadata as a basis for building value-added services.

[Top](#)

What do you mean by an "Archive"?

The term "archive" in the name *Open Archives Initiative* reflects the origins of the OAI – in the E-Prints community where the term archive is generally accepted as a synonym for repository of scholarly papers. Members of the archiving profession have justifiably noted the strict definition of an "archive" within their domain; with connotations of preservation of long-term value, statutory authorization and institutional policy. The OAI uses the term "archive" in a broader sense: as a repository for stored information. Language and terms are never unambiguous and uncontroversial and the OAI respectfully requests the indulgence of the professional archiving community with this broader use of "archive".

[Top](#)

What do you mean by "Open"?

Our intention is "open" from the architectural perspective – defining and promoting machine interfaces that facilitate the availability of content from a variety of providers. Openness does not mean "free" or "unlimited" access to the information repositories that conform to the OAI-PMH. Such terms are often used too casually and ignore the fact that monetary cost is not the only type of restriction on use of information – any advocate of "free" information recognize that it is eminently reasonable to restrict denial of service attacks or defamatory misuse of information.

[Top](#)

Is the Open Archives Initiative only about E-Prints?

The roots of the OAI lie in the E-Print community, which promotes and maintains web-accessible archives of scholarly papers as a means of increasing access to scholarly

research. Initial work in the OAI was motivated by a desire to develop interoperability frameworks for federating E-Print archives. It soon became evident, however, that the concepts in the OAI interoperability framework - exposing multiple forms of metadata through a harvesting protocol - had applications beyond the E-Print community. Therefore, the OAI has adopted a mission statement with broader application: opening up access to a range of digital materials. The participants in the OAI have an ongoing interest in publishing alternatives of interest to a variety of stakeholders - E-Print providers, publishers, authors - and view the OAI as a forum for discussions and experimentation with those alternatives.

[Top](#)

Is the Open Archives Initiative only concerned with metadata?

The current OAI technical infrastructure, which is specified in the [Open Archives Initiative Protocol for Metadata Harvesting](#), defines a mechanism for data providers to expose their metadata. There is nothing in the OAI mission that restricts the work of the OAI to metadata alone. However, we are guided by the goal to define a low-barrier and widely applicable framework for cross-repository interoperability and believe that exposing metadata is plausible route to such a goal. We may, in the future, explore and define other mechanisms for interoperability.

[Top](#)

Can I use the OAI-PMH to harvest content?

The [Open Archives Initiative Protocol for Metadata Harvesting](#) defines a mechanism for harvesting XML-formatted metadata from repositories. The protocol does not provide a mechanism for harvesting data (content) that is not encoded in XML. The protocol also does not mandate the means of association between that metadata and related content. Since many clients may want to access the content associated with harvested metadata, data providers may deem it appropriate to define a link in the metadata to the content. The mandatory Dublin Core format provides the *identifier* element that can be used for this purpose.

[Top](#)

How long will it take me to implement the protocol?

The OAI-PMH has been designed with easy implementation in mind. Therefore, the generic task of configuring a web server to handle OAI-PMH requests and parsing out the arguments should involve less than a day of work for someone experienced with setting up Web servers and writing CGI scripts. We expect that over time a number of generic front-ends for handling protocol requests will be available either at the [OAI web site](#) or through the [implementers mail list](#).

Implementing the protocol, however, involves more than simply parsing the protocol requests. Responding to protocol requests also involves accessing or extracting your metadata. Giving a time estimate for this is difficult since the nature of the task is entirely idiosyncratic. If your data is well-organized, already has metadata, and has

established mechanisms for extracting or deriving metadata, we assume that this task should not be onerous.

[Top](#)

What is the relationship between the Open Archives Metadata Harvesting Protocol and other protocols such as Z39.50?

The OAI technical framework is intentionally simple with the intent of providing a low barrier for participants. Protocols such as Z39.50 have more complete functionality; for example, they deal with session management and results sets and allow the specification of predicates that filter the records returned. However, this functionality comes at an increase in difficulty of implementation and cost. The OAI technical framework is not intended to replace other approaches but to provide an easy-to-implement and easy-to-deploy alternative for different constituencies or different purposes than those addressed by existing interoperability solutions. Continued use of the OAI-PMH will prove whether such low-barrier interoperability is realistic and functional.

[Top](#)

What advantage is there for me in participating in the OAI?

We can't promise any immediate benefit from adopting the OAI-PMH. The motivations for adoption depend on the type of participation.

- *Data Providers* adopt the OAI-PMH for a variety of reasons including membership in a community that collectively has decided to use the OAI-PMH as a means of increased exposure of their content, interest in the functionality offered by specific service providers, or general interest in the notion of exposing metadata as a means of knowledge sharing.
- *Service Providers* adopt the OAI-PMH due to their interest and ability in combining and adding value to the metadata available from a set of data providers.

[Top](#)

How many participants are there in the OAI?

The available lists of registered data providers (<http://www.openarchives.org/Register/BrowseSites.pl>) and registered service providers (<http://www.openarchives.org/service/listproviders.html>) provide one indication of the number of participants. However, since registration in both cases is optional, the actual number of adopters of the OAI-PMH is unknown.

[Top](#)

What is the advantage of registering my repository with the OAI?

Data providers who support the OAI-PMH may choose to register through the registration page at http://www.openarchives.org/pmh/register_data_provider. The registration data base serves as a publicly accessible list of OAI-PMH conformant repositories, making it easy for service providers to discover sites from which metadata can be harvested. The

registry database contains all the information available through the OAI *Identify* request including:

- repositoryName: a human readable name for the repository, in the example "The University of Spa E-print System";
- baseURL: the BASE-URL for making protocol requests to the repository;
- protocolVersion: the version of the OAI-PMH supported by the repository;
- adminEmail: the e-mail address of the administrator of the repository.
- additional protocol information and repository-specific description packages.

NOTE: data providers who register with the OAI agree to make this information public. There are no provisions for restricting access to the registry database.

[Top](#)

Is there a way to affiliate my repository with other OAI-PMH conformant repositories?

Data providers may be aware of other OAI-PMH conformant repositories; one example is a data provider that is a member of a group or community of affiliated data providers. In such cases, the [friends container](#) in the [Identify](#) response may be used by repositories to list confederate repositories. This provides an automatic way for harvesters to discover other repositories that may be of interest. Widespread use of this container would provide a decentralized mechanism by which harvesters can discover other repositories.

[Top](#)

What is the advantage of using OAI identifiers for metadata records?

Every metadata record harvested by the OAI-PMH has an identifier that corresponds to the item from which the metadata was derived. The only restriction for conformance is that this identifier should be a URI that is unique within the respective repository. Data providers may choose, however, to adopt an identifier scheme whereby their identifiers are globally unique within the *oai* naming scheme. This choice is made at the time of registration and instructions for making this choice are available at the registration page at http://www.openarchives.org/pmh/register_data_provider. The specification of the oai identifier format is available in the [implementations guidelines document](#). The advantage for repositories of adopting this naming convention is that record identifiers will be resolvable via future oai name resolution services.

[Top](#)

What is the difference between the versions of the OAI-PMH?

The initial version of the OAI-PMH, 1.0, was released in January 2001. A minor update version, 1.1, was released in July 2001 to conform to changes in the XML schema specification. The current version of the OAI-PMH, 2.0, was released in June 2002.

Technical changes between the different versions are described in respective migration documents. The migration document for version 2.0 is available at

<http://www.openarchives.org/OAI/2.0/migration.htm>. An executive summary of the technical changes in version 2.0 is available at <http://www.openarchives.org/news/oaiv2press020614.html>.

In addition to technical changes, OAI-PMH version 2.0 also marks the end of the initial experimentation period for the protocol. This initial 18 month period was meant to test both the functional aspects of the protocol and the utility of the general notion of metadata harvesting. Substantial use of the OAI-PMH during that period has both validated metadata harvesting as a tool for information federation and provided the foundation for the technical changes and enhancements in OAI-PMH version 2.0.

While any claim of "stability" for a technical infrastructure would be disingenuous, we are committed to version 2.0 as a *production release*, meaning that future changes, if any, will be made with strict attention to backward compatibility issues.

[Top](#)

Is version 2.0 of OAI-PMH compatible with earlier versions?

OAI-PMH version 2.0 is *not* backward compatible. Results from the version 1.x experimentation period indicated that a number of fundamental changes were necessary, and maintaining backward compatibility would have resulted in a compromised production release. We are committed to strongly consider backward compatibility issues in future releases.

[Top](#)

Will there be continued support for version 1.1 of the OAI-PMH?

Our goal is to phase out version 1.1 of the OAI-PMH by the end of 2002. Certainly individual service and data providers may continue to support the earlier version, but the data and service provider registries will only contain and accept sites that are OAI-PMH version 2.0 compliant.

[Top](#)

Why does the protocol mandate a common metadata format (and why is that common format Dublin Core)?

Mapping among multiple metadata formats would place a considerable burden on service providers, who harvest the metadata and use it to build higher level services. While there is research work on creating services such as common search interfaces across heterogeneous metadata formats, a less burdensome and ultimately more deployable solution is to require repositories to map to a simple and common metadata format. The fifteen elements Dublin Core has over the past several years evolved as a de facto standard for simple cross-discipline metadata and is thus the appropriate choice for a common metadata set. Cooperation between the OAI and the Dublin Core Metadata Initiative has led to a common xml schema for unqualified dublin core that is available at <http://dublincore.org/schemas/xmls/simpledc20020312.xsd>.

[Top](#)

What if I want to expose metadata in other formats than Dublin Core?

The metadata harvesting protocol supports the notion of multiple metadata sets, allowing communities to expose metadata in formats that are specific to their applications and domains. The technical framework places no limitations on the nature of such parallel sets, other than that the metadata records be structured as XML data, which have a corresponding XML schema for validation.

[Top](#)

How do I let others know that my repository supports the OAI-PMH?

A data provider may choose to register via the [OAI registration page](#), and thereby publicize the fact that they have adopted the OAI-PMH.

[Top](#)

What about intellectual property issues?

The OAI does not define or prescribe any rights management scheme. Issues of access restriction and management of intellectual property in exposed metadata are the responsibility of the data providers that adopt the protocol. We expect that some repositories that adopt the protocol will permit more or less unrestricted access to their metadata, while others may use various methods to restrict access (e.g., by IP address of origin).

[Top](#)

What if I have more questions?

General inquiries can be sent to openarchives@openarchives.org, which is monitored by the OAI coordinators. Suggestions for additions to the FAQ are always welcome.

[Top](#)