LibraryCloud OAI-PMH Data Provider

OAI-PMH (https://www.openarchives.org/OAI/openarchivesprotocol.html) provides an application-independent interoperability framework based on metadata harvesting.

The LibraryCloud OAI-PMH Data Provider provides access to metadata about pre-defined sets of items in the Harvard Library collections. OAI "sets" are pre-defined and named as "collections" in the LibraryCloud collection database.

Metadata is provided in the MODS schema (with some LibraryCloud extensions) and Dublin Core.

Base URI

The service's base address is: https://api.lib.harvard.edu/oai

ListSets

The ListSets verb will return a list of all harvestable OAI sets defined as LibraryCloud collections.

https://api.lib.harvard.edu/oai/?verb=ListSets

ListRecords

The ListRecords verb will return full metadata records for a given LibraryCloud collection. Results can be returned as MODS or DC.

https://api.lib.harvard.edu/oai/?verb=ListRecords&metadataPrefix=mods&set=scores
https://api.lib.harvard.edu/oai/?verb=ListRecords&metadataPrefix=oai_dc&set=scores

ListIdentifiers

The ListIdentifiers verb will return the identifiers of records in a given LibraryCloud collection.

https://api.lib.harvard.edu/oai/?verb=ListIdentifiers&metadataPrefix=mods&set=scores

Identify

The Identify verb will return information about Harvard's LibraryCloud repository.

https://api.lib.harvard.edu/oai/?verb=Identify

GetRecord

The GetRecord verb will return an individual metadata record from the LibraryCloud repository. Required arguments specify the identifier of the item to be requested and the format of the metadata that should be included in the record.

https://api.lib.harvard.edu/oai/?verb=GetRecord&metadataPrefix=mods&identifier=000796856 (using Aleph ID)
https://api.lib.harvard.edu/oai/?verb=GetRecord&metadataPrefix=oai_dc&identifier=990007968560203941 (using Alma ID)

ListMetadataFormats

The ListMetadataFormats verb will return the metadata formats available from the LibraryCloud repository. An optional argument restricts the request to the formats available for a specific item.

https://api.lib.harvard.edu/oai/?verb=ListMetadataFormats
https://api.lib.harvard.edu/oai/?verb=ListMetadataFormats&identifier=990007968560203941

Resumption tokens

The OAI requests for ListSets, ListRecords, and ListIdentifiers retrieve 10 results at a time, so the repeated use of a resumptionToken in the request URL is needed to retrieve the full amount of data.

For a ListRecords request in the Scarlet collection -- https://api.lib.harvard.edu/oai/?verb=ListRecords&metadataPrefix=mods&set=scarlet -- this is how to use the resumption token to retrieve additional data beyond the first 10 items:

2. Copy the resumptionToken value from the bottom of the list of 10 results, for example:
   
   <resumptionToken completeListSize="445">10:0001-01-01:9999-12-31:scarlet:mods</resumptionToken>
The first part of the token value indicates the number at which the next list will start – “10” in the example above.

3. Remove the “metadataPrefix=x&set=y” parameters from the original request url and replace with the resumptionToken as an argument, like this:


4. Re-issue the request to get the next set of 10 items.