



## Technical documentation of OpenEdition OAI-PMH repository

The OpenEdition OAI-PMH repository contains the metadata of the documents available on [OpenEdition Books](#), [Revue.org](#), [Hypotheses](#), and [Calenda](#).

The repository is available online at <http://oai.openedition.org>.

It uses the version 2.0 of the OAI-PMH protocol. The specifications of this protocol are available online at: <http://www.openarchives.org/OAI/openarchivesprotocol.html>.

### Contents

|                                    |   |
|------------------------------------|---|
| Features .....                     | 2 |
| Available formats .....            | 2 |
| Available metadata .....           | 2 |
| Selective harvesting.....          | 3 |
| Selective harvesting by set.....   | 3 |
| Selective harvesting by date ..... | 4 |
| ResumptionToken.....               | 4 |
| Deleted records .....              | 5 |
| Sample requests .....              | 5 |



# Features

## Available formats

Records are available to all in the following metadata formats:

- [Dublin Core \(oai\\_dc\)](#)
- [Qualified Dublin Core \(qdc\)](#)
- [Metadata Encoding and Transmission Standard \(METS\)](#)

The METS format is available only for journal issues and books. It is not available for academic blogs, Calenda events, and electronic journals publishing only streamed articles.

Furthermore, OpenEdition's partners have access to the [TEI](#) (Text encoding initiative) of Revues.org and OpenEdition Books documents. It allows retrieving and indexing the full text of the documents. If you want to get an access to the OpenEdition's TEI, please contact us at [contact@openedition.org](mailto:contact@openedition.org).

## Available metadata

All metadata formats contains basic information about the documents, when applicable: document's title, type, abstract, authors, publication date, publisher name, language, keywords, rights, DOI, ISBN, URL...

METS and Qualified Dublin Core contain some additional metadata.

Qualified Dublin Core contains amongst other things: metadata about the document's extent, availability date, and journal's ISSN.

METS contains metadata about the structure of books and journals issues. It can be useful to retrieve metadata of articles published in a specific issue.

A look-over at the sample requests in the following sections can give a good idea of all available metadata.

## Selective harvesting

The repository allows [selective harvesting](#), by set and by date. Two simple sample queries:

- [http://oai.openedition.org/?verb=ListRecords&metadataPrefix=oai\\_dc&set=journals:chs](http://oai.openedition.org/?verb=ListRecords&metadataPrefix=oai_dc&set=journals:chs) for retrieving a list of records from the journal *Crime, History & Societies*
- [http://oai.openedition.org/?verb=ListRecords&metadataPrefix=oai\\_dc&from=2013-03-13](http://oai.openedition.org/?verb=ListRecords&metadataPrefix=oai_dc&from=2013-03-13) for retrieving a list of records added or updated since 13 march 2013.

## Selective harvesting by set

The repository allows selective harvesting with the *set* parameter. The repository is divided in two kinds of set: platform sets, and publication sets.

The platform sets are:

- A *journals* set, containing all records from Revues.org.
- A *books* set, containing all records from OpenEdition Books.
- A *blogs* set, containing all records from Hypotheses.
- An *events* set, for all records from Calenda.

As for the publication sets, they allow retrieving records from specific publication within a platform (except Calenda). The name of the publication sets follows the pattern <platform set ID>:<publication ID>:

| Set name pattern              | Example      |
|-------------------------------|--------------|
| journals:<journal website ID> | journals:chs |
| books:<publisher website ID>  | books:obp    |
| blogs:<academic blog ID>      | blogs:cpa    |

For instance:

- For retrieving the document metadata from the *Publications de l'École française de Rome* only, you will have to query the “books:efr” set:  
[http://oai.openedition.org/?verb=ListRecords&metadataPrefix=oai\\_dc&set=books:efr](http://oai.openedition.org/?verb=ListRecords&metadataPrefix=oai_dc&set=books:efr)
- For retrieving metadata from all Revues.org's documents, you will have to query the “journals” set:  
[http://oai.openedition.org/?verb=ListRecords&metadataPrefix=oai\\_dc&set=journals](http://oai.openedition.org/?verb=ListRecords&metadataPrefix=oai_dc&set=journals)

The comprehensive list of the sets can be retrieved with the query <http://oai.openedition.org/?verb=ListSets>.

## Selective harvesting by date

The OpenEdition repository allows harvesting by date, that is to say the harvesting of records added or updated before or after the date specified in the query.

The parameters to use are *from* or *until*. The date can have the *dd-mm-yyyy* format or the *dd-mm-yyyyThh:mm:ssZ* format.

For instance:

[http://oai.openedition.org/?verb=ListRecords&metadataPrefix=oai\\_dc&from=2013-03-13T16:47:48Z](http://oai.openedition.org/?verb=ListRecords&metadataPrefix=oai_dc&from=2013-03-13T16:47:48Z) will retrieve a list of records added or updated since the 13 march 2013 at 4.47 pm.

## ResumptionToken

The repository uses the resumptionToken system. Therefore, it is not possible to retrieve all documents with a single request.

For instance, for retrieving the metadata of all documents from the journal *Revista Crítica de Ciências Sociais*, it will be necessary to make a request [http://oai.openedition.org/?verb=ListRecords&set=journals:rccs&metadataPrefix=oai\\_dc](http://oai.openedition.org/?verb=ListRecords&set=journals:rccs&metadataPrefix=oai_dc). It will return a first list of 10 documents.

For retrieving the next 10 documents, it is necessary to pass the content of the <resumptionToken> tag as an argument of a new URL request: [http://oai.openedition.org/?verb=ListRecords&resumptionToken=metadataPrefix%253Doai\\_dc%2526set%253Djournals%25253Arccs%2526until%253D2013-10-01T08%25253A35%25253A54Z%2526cursor%253D10](http://oai.openedition.org/?verb=ListRecords&resumptionToken=metadataPrefix%253Doai_dc%2526set%253Djournals%25253Arccs%2526until%253D2013-10-01T08%25253A35%25253A54Z%2526cursor%253D10).

The OAI-PMH documentation available at <http://www.openarchives.org/OAI/openarchivesprotocol.html> gives a more detailed insight of the resumptionToken parameter.

## Deleted records

The repository does not provide information about deleted records.

## Sample requests

- Retrieving general information about the repository:  
<http://oai.openedition.org/?verb=Identify>
- Listing available sets: <http://oai.openedition.org/?verb=ListSets>
- Retrieving all documents metadata from all platforms, in Qualified Dublin Core format:  
<http://oai.openedition.org/?verb=ListRecords&metadataPrefix=qdc>
- Retrieving all metadata from Calenda added since 1 January 2013, in Qualified Dublin Core format:  
<http://oai.openedition.org/?verb=ListRecords&set=events:calendaorg&from=2013-01-01&metadataPrefix=qdc>
- Retrieving all documents of *Journal of TEI*, in METS format:  
<http://oai.openedition.org/?verb=ListRecords&set=journals:jtei&metadataPrefix=METS>
- Listing the documents' identifiers of the *Publications de l'École française de Rome*: <http://oai.openedition.org/?verb=ListIdentifiers&set=books:efr>
- Listing the available metadata formats for a post in the academic blog *The Recipes Project*.

<http://oai.openedition.org/?verb=ListMetadataFormats&identifier=oai:hypotheses.org:recipes/2294>