

DataCite

DataCite MDS API Application

The DataCite Metadata Store (MDS) API[1] is going to be used for Hub dataset Digital Identifier Object (DOI) and corresponding metadata creation and update.

1. Introduction

The DataCite Metadata Store (MDS) API is the default API for registering DataCite DOIs and associated metadata. Below shows the requirements for using the MDS API.

(1) The API requires authentication. PURR is member of DataCite and has credentials for production PURR. To develop and test the DOI creation application program on PURR, we obtain test account and password from DataCite support.

(2) DataCite assigns the DOI prefix 10.4231 to PURR. The prefix for Sandbox is 10.5072.

(3) DOI request needs to transfer through HTTPS.

(4) HTTP basic authentication is performed in the DataCite Metadata Store API usage.

2. DOI register

The DOI registration takes two steps, first is to register DOI metadata, and second is to register the DOI name and dataset URL.

2.1 Registering DOI metadata

The DOI metadata for dataset includes title, author, dates, description, and so on. All the metadata that PURR collects for a dataset shows in table 1. Metadata element

| | Explanation |
|--|-------------|
|--|-------------|

DATA CITE

| | |
|---|---|
| Identifier | Identifier type is doi, and the value of identifier is the doi. |
| Creators/creator/creatorName nameIdentifier | Author(s) of the dataset nameIdentifierScheme is ORCID? and value is the ORCID. |
| titles/title | Name of the dataset. |
| publisher | Holder of the dataset |
| publicationYear | The year when the dataset is published |
| Contributors/contributor/contributorName dates | Project leader name There are two metadata elements, one is when dataType is Valid, the value is the date when the dataset is submitted on PURR; Another is when the datatype is Accepted, the value is the date when the dataset is approved for publish on PURR. |
| language | It is set to en by default |
| resourceType | resourceTypeGeneral is set to the resourceType designated by PURR. The value is the resourceTypeTitle. |
| RelatedIdentifiers/relatedIdentifier | relatedIdentifierType is DOI, relationType is IsNewVersionOf, and the value is the relatedDoi, which is DOI of the resource that reference this dataset. |
| version | Publication version |
| rights | The license that user selects in the dataset submission workflow on PURR |
| Descriptions/description | DescriptionType is Abstract, and the value is the abstract of the dataset |

The metadata elements in table 1 are saved in xml format as showing below.

```
<resource xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xmlns="http://datacite.org/schema/kernel-4"
xsi:schemaLocation="http://datacite.org/schema/kernel-4
http://schema.datacite.org/meta/kernel-4.1/metadata.xsd">
```

```
<identifier identifierType="DOI">doi</identifier>
```

```
<creators>
```

```
<creator>
```

```
<creatorName>name 1</creatorName>
```

DATA CITE

</creator>

<creator>

<creatorName>name 2</creatorName>

</creator>

<creator>

<creatorName>name 3</creatorName>

</creator>

<creator>

<creatorName>name 4</creatorName>

<nameIdentifier nameIdentifierScheme="ORCID">orcid goes here</nameIdentifier>

</creator>

</creators>

<titles>

<title>title</title>

</titles>

<publisher>Purdue University Research Repository</publisher>

<publicationYear>year</publicationYear>

<contributors>

<contributor contributorType="ProjectLeader">

<contributorName>name</contributorName>

</contributor>

</contributors>

<dates>

DATA CITE

```
<date dateType="Valid">date</date>
<date dateType="Accepted">date</date>
</dates>
<language>en</language>
<resourceType resourceTypeGeneral="Dataset">Datasets</resourceType>
<version>1.0</version>
<rights>Attribution 3.0 Unported</rights>
<descriptions>
<description descriptionType="Abstract">
Description goes here
</description>
</descriptions>
</resource>
```

All the metadata elements conform to the DataCite metadata schema [2] standards.

The MDS API can generate random DOI name when prefix is included in the metadata registration. The identifier in the metadata xml can be left empty in this case.

The curl package can be used to submit the metadata to MDS service, see example below.

```
curl -H "Content-Type:application/xml;charset=UTF-8" -X PUT -i --user username:password -d
@10.5072.xml https://mds.test.datacite.org/metadata/10.5072
```

The response to the call above is OK (10.5072/XXXX-XXXX). We need take the DOI out from there and use for further process.

The identifier value in the metadata xml on DataCite will be replaced by the DOI name that is created in this step.

The DOI is set to Draft state [3] after the metadata are registered.

1.2.2 Registering the URL

The second step is to register the DOI name and the URL that refers to the dataset. Below is an example of the curl command. Note that DOI name and URL to the dataset is included.

```
curl -H "Content-Type:text/plain;charset=UTF-8" -X PUT --user username:password -d
'$doi=10.5072/0000-03VCnurl=http://example.org/publications/1'
https://mds.test.datacite.org/doi/10.5072/0000-03VC
```

The DOI transits from Draft state to Findable state after the DOI name and URL is registered.

Below table includes the endpoint and other useful information that need to use in the DOI registration in production and sandbox environment. The sand box endpoint, prefix, username and password need to be configured on the admin interface for dev PURR and any other local development instances. The production endpoint, username, and password need to be configured on production PURR.

| Term | Value |
|-------------------------|---------------------------------|
| PURR DOI prefix | 10.4231 |
| MDS production endpoint | https://mds.datacite.org |
| MDS sandbox endpoint | https://mds.test.datacite.org |
| Sandbox DOI prefix | 10.5072 |
| Sandbox username | apply through DataCite Support. |
| Sandbox password | apply through DataCite Support. |

2. Update DOI Metadata Records

Update the DOI metadata is the same as storing a new version of the DOI metadata. According to the MDS API document, we need to issue a post of the metadata xml to DataCite. Below is an example on what the curl command should look like for such operation.

```
$ curl -H "Content-Type: application/xml;charset=UTF-8" -X POST -i --user username:password
-d @10.5072/JQX3-61AT.xml https://mds.test.datacite.org/metadata
```

Reference list

1. <https://support.datacite.org/docs/mds-api-guide>
2. <http://schema.datacite.org/>

3. <https://support.datacite.org/docs/doi-states>

Switching from EZID to DataCite

1. Navigate to **/administrator** and login to the backend of the Hub
2. Click on **Components** and then select **Publications**
3. Click on **Options** and then click on **DOI Configuration**
 1. Note: You will only be able to access the Options button if you are a super user.
4. Locate *Enable DataCite or EZID DOI Service* and select from the drop-down the appropriate option
 1. Note: When the "DataCite" is selected, the DataCite DOI service is activated and only DataCite related options on the configuration form are going to show up. When "EZID" is selected, EZID DOI service is activated and only EZID related options on the configuration form are going to show up. When "None" is selected, all options related to DOI service will be hidden. When users publish a dataset in this case, a message saying that "No DOI service is selected and please contact site administrator" is popping up.
5. Click **Save & Close**