About DOIs

Digital Object Identifiers
DOI: Digital Object Identifier

• A DOI is a serial code used to **uniquely identify** content of various types of entities. The DOI system is particularly used for electronic documents such as journal articles or datasets.

• What is digital is the identifier, not the object!

<table>
<thead>
<tr>
<th>Name</th>
<th>DOI</th>
<th>URL</th>
<th>Metadata</th>
</tr>
</thead>
</table>

DataCite
FIND, ACCESS, AND REUSE DATA
Prefixes are assigned to different services. Each one of them manages the ‘suffix’ namespace freely.

Example: 10 . 1234 / data567
URL

- The DOI name points to a URL, and it can be repointed as many times as needed.
- This URL can be the object itself or a landing page displaying metadata and how to access the object.

DOI name ➔ http://page1.com
http://page2.com
http://page2.com/old/
A metadata schema is a list of core metadata properties chosen for the accurate and consistent identification of a resource.

<table>
<thead>
<tr>
<th>Mandatory</th>
<th>Recommended</th>
<th>Optional</th>
</tr>
</thead>
<tbody>
<tr>
<td>Identifier</td>
<td>Subject</td>
<td>Language</td>
</tr>
<tr>
<td>Creator</td>
<td>Contributor</td>
<td>Alternate ID</td>
</tr>
<tr>
<td>Title</td>
<td>Date</td>
<td>Size</td>
</tr>
<tr>
<td>Publisher</td>
<td>Related identifier</td>
<td>Format</td>
</tr>
<tr>
<td>Publication year</td>
<td>Description</td>
<td>Version</td>
</tr>
<tr>
<td>Resource Type</td>
<td>GeoLocation</td>
<td>Rights</td>
</tr>
</tbody>
</table>
Cool DOIs

- DataCite’s display guidelines: DOIs should be displayed as HTTPS URLs
  - https://doi.org/DOI name
- Designing your suffix:
  - Human readable
  - Avoid semantic information
  - Make them easy to generate
About the service
Service agreements

- DataCite serves its members
- Two types of members:
  - Allocating
  - Non-allocating
- Each member serves its affiliated data centers
- With different business models
Technical infrastructure

- DataCite serves all the data centers using a centralized infrastructure
- Same end-point for all of them
- Becoming a member does not require to set up any technical infrastructure
About the MDS

MetaData Store
The MetaData Store

• The DataCite Metadata Store is a service for data publishers to mint DOIs and register associated metadata.

• Two ways to use it:
  • Web interface - https://mds.datacite.org
  • API - https://mds.datacite.org/static/apidoc
Web interface

• “Single” or manual operations.

<table>
<thead>
<tr>
<th>Members</th>
<th>Datacenters</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Datacentre</strong></td>
<td><strong>Dataset</strong></td>
</tr>
<tr>
<td>Create new Datacentre</td>
<td>Register new Dataset</td>
</tr>
<tr>
<td>List all Datacentres</td>
<td>List all Datasets</td>
</tr>
<tr>
<td>Find by Symbol</td>
<td>Find by DOI</td>
</tr>
<tr>
<td>Find by Name</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th><strong>View</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>API documentation</td>
</tr>
</tbody>
</table>
Example: mint a DOI

- DOI latency: Be aware that it can take up to 24 hours until a DOI update is globally known. New DOIs should be resolvable after about 5 minutes.

  For testing purposes please only use our dedicated test prefix 10.5072

  DOI: [Input field]

  Name: [Input field]

  URL: [Input field]

  Metadata: [Input field]

  XML upload: Choose File

  Please select an XML file. It must reference a schema located under the following base URL: [http://schema.datacite.org/meta/](http://schema.datacite.org/meta/)

  XML: [Input field]

  SAVE
API

- RESTful API.
- HTTPS with basic authentication.
- Example, first upload the metadata, then mint the DOI:

  ```
  curl -u USER.NAME -H "Content-Type: application/xml" --data-binary @metadata.xml https://mds.datacite.org/metadata
  200 OK
  
  200 OK
  ```
Sandbox

• DataCite provides a test environment for developers.

• The endpoint is https://test.datacite.org and the resolver is http://dx.test.datacite.org

• A test prefix exists for all data centers 10.5072
Distributed System

- The Handle system (the technical infrastructure of DOIs) is distributed.

- Newly minted DOIs should be resolvable in less than 5 minutes (normally no more than 1 minute).

- Updates to DOIs (URL/metadata) can take up to 24h before being distributed through the whole network.
Services around DOIs

- Integrated search
- Service status
- re3data
- Statistics
- Citation formatter
- OAI-PMH
Questions?

support@datacite.org
References

• DOI Handbook: https://doi.org/10.1000/182

• DataCite Metadata Schema Documentation: https://doi.org/10.5438/0012

• DataCite Metadata XML Schema: https://doi.org/10.5438/0013