



Procedure for minting DOIs to data

Ref.id.: DIG.NMD.FORV-04

Prosedyre

Side 1 av 3

Purpose

The purpose of this procedure is to describe how NMD assigns DOIs to data.

Target audience and responsibilities

The service to assign DOIs to data is delivered as part of NMDC, but managed by NMD.

Persons who wish to utilize NMDC to assign DOIs to their research data.

Employees at NMD who are responsible for:

- assigning DOIs for data
- curating metadata associated with the data
- aggregation of the data

Definitions

NMD stands for "Norsk Marint Datasenter" (eng: Norwegian Marine Data Centre) and is a section under IMR.

NMDC stands for "Norwegian Marine Data Centre" (not to be confused with NMD) and is a national infrastructure for marine data, which is a catalog of marine datasets and their associated metadata across all Norwegian marine regions. NMDC is a consortium of 16 national partners with IMR as the coordinator.

IMR stands for the Institute of Marine Research.

A data manager is an employee at NMD who is responsible for curating mandatory metadata associated with the data and may be involved in converting the data to the published format.

DOI stands for Digital Object Identifier and is a unique text string associated to an object.

A dataset is an organized collection of data. A dataset can consist of one or more files.

In Work refers to datasets currently undergoing production or data that is continuously being collected or updated, i.e. ongoing time-series data.

Completed refers to datasets in which no updates or further data collection will be made.

A Landing Page is a homepage for the data. The data's DOI points to its Landing Page.

A description of the data including mandatory metadata and information on how to obtain the data is found on the Landing Page. Each DOI must have a Landing Page.

If a dataset is changed the updated dataset must be considered as a new dataset.

Creative Commons is an international organization who provides free licenses used for making data available. The license used by IMR is Creative Commons CC BY 4.0.

Link: <https://creativecommons.org/licenses/by/4.0/>

NLOD is a Norwegian license for public data and is an license agreement for public companies who makes public data available. IMR is also using the NLOD license, which is very similar to CC BY 4.0. Link: <https://data.norge.no/nlod/no>



To assign the metadata associated with the data the metadata standard Global Change Master Directory (GCMD)/Directory Interchange Format (DIF) is being used.

To mint a DOI is to create and assign a DOI to a dataset.

Description

For the purposes of assigning DOIs data is regarded as either Completed or In Work. The format of the DOIs will be slightly different for each type.

Completed data

For a completed dataset the DOI has the following format: **10.21335/NMDC-Number**

A dataset cannot be changed once it has been given a DOI. If a dataset with a DOI is changed, a new DOI may need to be assigned for the dataset. Whether or not a new DOI is needed will depend on the impact of the changes. If the changes have a low impact it is regarded as Minor changes and will not result in a new DOI. If the changes have a big impact it is regarded as Major changes and will result in a new DOI. See the description of Minor and Major Changes below.

In Work data

We use DOI versioning when a dataset is regarded as In Work. For DOI versioning, the DOI has the following format: **10.21335/NMDC-Number.V**, where V is the version number. If the version number is omitted it resolves to the most current version. The versioning is linear and starts with 1, then 2, 3, etc. If a dataset, which is regarded as In Work, is changed, a new version number may need to be assigned for the dataset. Whether or not a new version is needed will depend on the impact of the changes. If the changes have a low impact it is regarded as Minor changes and will not result in a new version. If the changes have a big impact it is regarded as Major changes and will result in a new version. See the description of Minor and Major Changes below.

If a dataset is regarded as In Work all its files should have a DOI.

If a dataset is regarded as In Work and consists of only one file the above description can also be used.

Minor and Major changes

Minor changes have a low impact and will not result in the creation of a new DOI, whereas major changes have a big impact and will result in the creation of a new DOI (with a new version number for a dataset regarded as In Work).

Minor changes are: correction of spelling in metadata, minor changes in documentation, publication reference added.

Major changes are: new variable added, fixing errors in data, significant changes in documentation, adding new fields, adding additional files to the dataset, processing the data to give a new result.

Why is data regarded as either Completed or In Work?

At IMR we have ongoing data and we have decided to use versioning to be able to assign DOIs to this data type. We could also have used versioning for completed data, but once you add a version number to the DOI you also imply that several versions of the dataset exist or will exist.



Creation of DOIs

1. All necessary metadata associated to the dataset including its files and the dataset itself are collected. The data owner may have to be contacted to obtain the dataset and associated metadata.
2. The DOI will be minted to the dataset and points to a Landing Page with the associated metadata and information on how to obtain the data.

Data from IMR follows the license Creative Commons Attribution 4.0 International (CC BY 4.0) and the license agreement NLOD. If IMR is not the data owner the data owner will decide under which license the data is available.

Kryssreferanser

Eksterne referanser