Meteorologisk institutt

DOIs implementation at MET

Lara Ferrighi, FOU-FD

25.03.2020

Overview

- What are persistent identifiers and DOI?
- Why is a DOI important for Data?
- How to get a DOI at MET?
- How is DOI implemented at MET?

What is a Persistent Identifier?

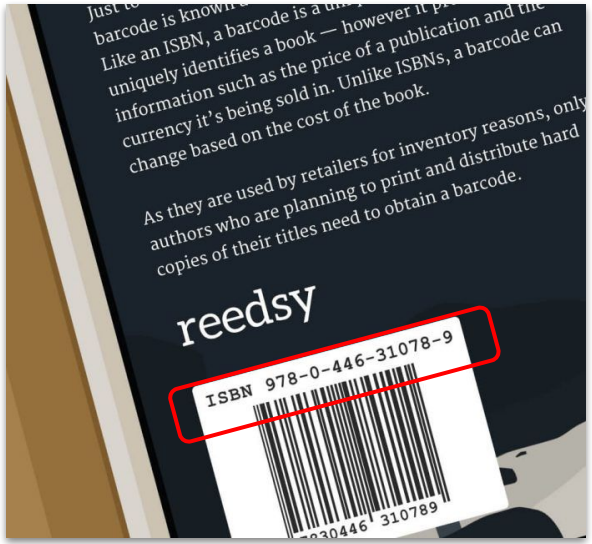
DOI is a unique and persistent identifier.

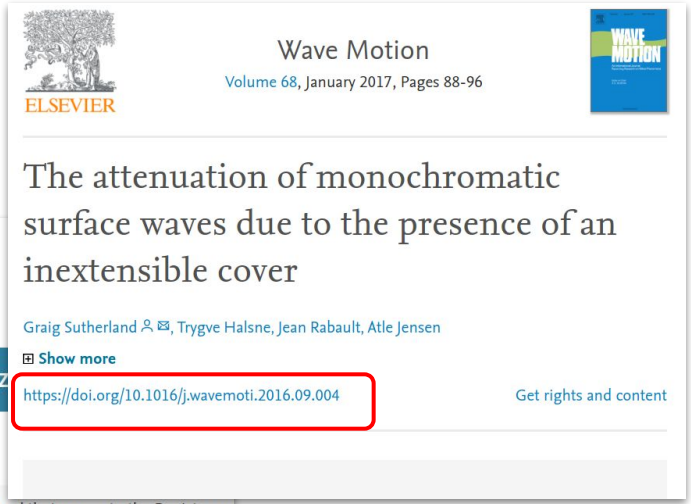
A **persistent identifier** (PI or PID) is

- a *long-lasting* reference to a document, file, web page, or other object.
- usually refers to digital objects that are accessible over the Internet.
- typically, such an identifier is not only **persistent** but **actionable**: you can plug it into a web browser and be taken to the identified source.

What is a Persistent Identifier?

- International Standard Book Number (ISBN)
- ORCID iDs (for researchers)
- Digital Object Identifiers (DOIs)
- ….





What is a Persistent Identifier?

- International Standard Book Number (ISBN)
- ORCID iDs (for researchers)
- Digital Object Identifiers (DOIs)
- ….

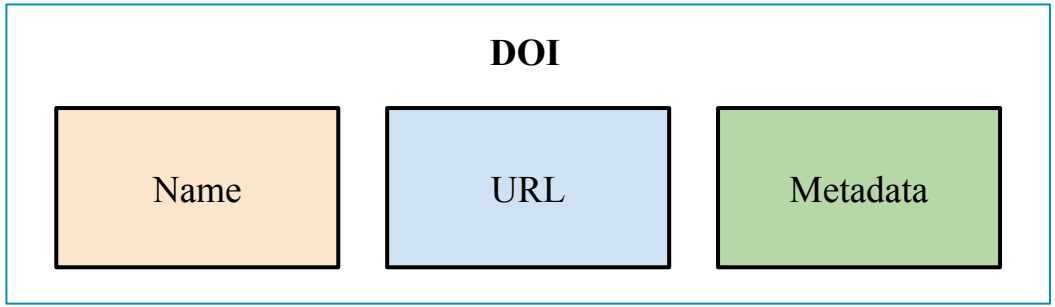
Meteorologisk institutt

DataCite
FIND, ACCESS, AND REUSE DATA

Our registration agency

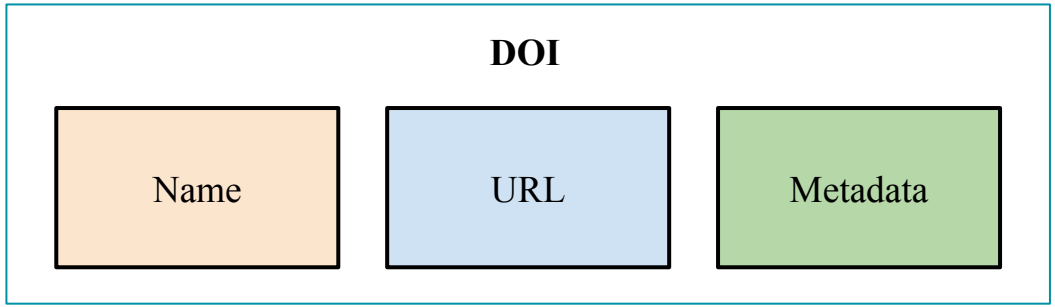
What is a DOI?

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 a DOI?

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.



Name
Prefix/Suffix

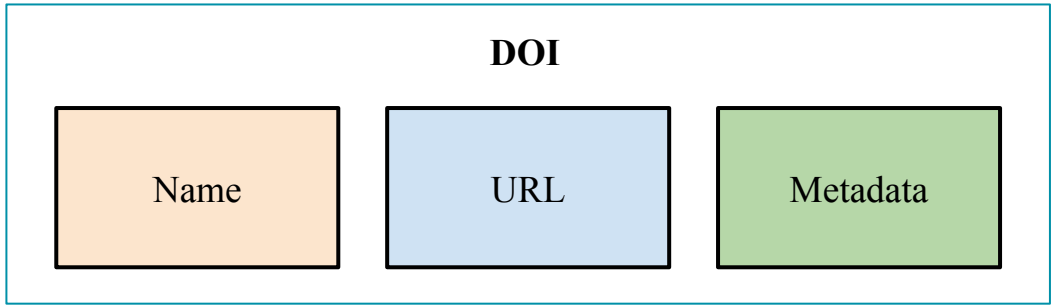
doi.org/**10.21343**/z9n1-qw63

Assigned by Datacite to METNO

UNIQUENESS

What is a DOI?

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.

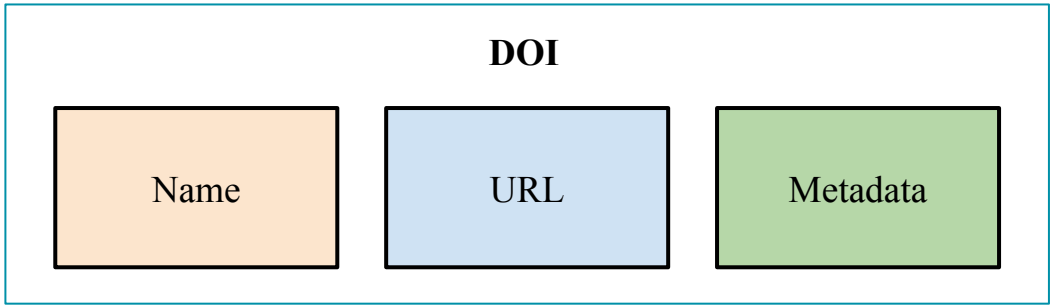


URL
landing page displaying metadata and how to access the object.

https://adc.met.no/datasets/10.21343/z9n1-qw63

What is a DOI?

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.



Metadata

Properties chosen for the accurate and consistent identification of a resource

Metadata required from the DOI provider

Mandatory	Recommended	Optional
Identifier	Subject	Language
Creator	Contributor	Alternate ID
Title	Date	Size
Publisher	Related identifier	Format
Publication year	Description	Version
Resource Type	GeoLocation	Rights

Metadata required from the DOI provider

Mandatory	Recor	ional
Identifier	S	uage
Creator	Cor	ate ID
Title		ize
Publisher	Relate	mat
Publication year	Des	sion
Resource Type	Geo	hts

- Audiovisual
- Collection
- DataPaper
- Dataset
- Event
- Image
- InteractiveResource
- Model
- PhysicalObject
- Service
- Software
- Sound
- Text
- Workflow
- Other

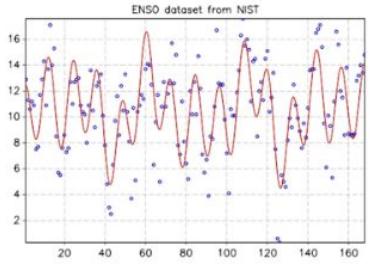
Metadata required from the DOI provider

Mandatory	Recommended	Optional
Identifier	Subject	Language
Creator	Contributor	Alternate ID
Title	Date	Size
Publisher	Related identifier	Format
Publication year	Description	Version
Resource Type	GeoLocation	Rights

This is only a subset of what we internally require!

Using a DOI - made simple

1. Take a dataset



2. Describe it

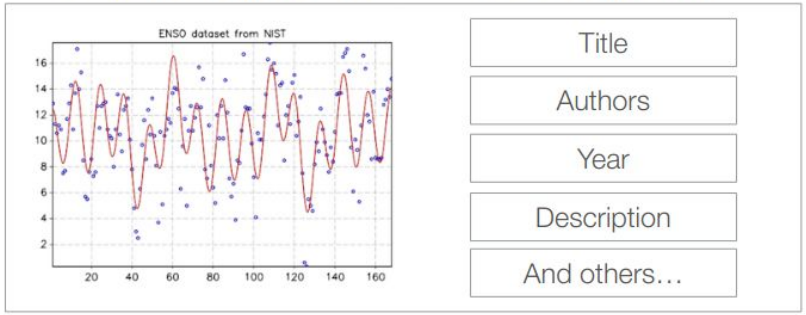
- Title
- Authors
- Year
- Description
- And others…

DataCite
FIND, ACCESS, AND REUSE DATA

utt

Using a DOI - made simple

1. Take a dataset



2. Describe it

- Title
- Authors
- Year
- Description
- And others…

3. Assign a DOI

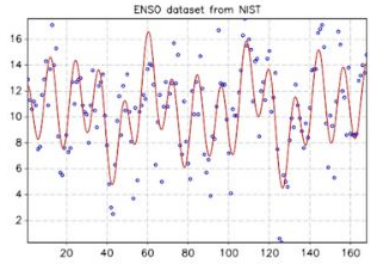
10.1234/exampledata

DataCite
FIND, ACCESS, AND REUSE DATA

utt

Using a DOI - made simple

1. Take a dataset



2. Describe it

- Title
- Authors
- Year
- Description
- And others…

3. Assign a DOI

10.1234/exampledata

4. Reuse and reference!

ATLAS Collaboration, "Data from Figure 7 from: Measurements of Higgs boson production and couplings in diboson final states with the ATLAS detector at the LHC: $H \to \gamma\gamma$,"
http://doi.org/10.7484/INSPIREHEP.DATA.A78C.HK44

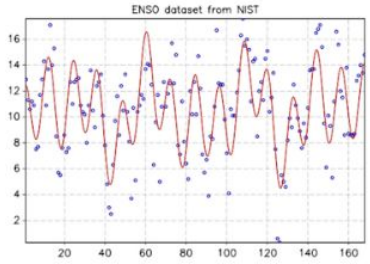
☑ Unique ☑ Persistent

DataCite
FIND, ACCESS, AND REUSE DATA

utt

Using a DOI - made simple

1. Take a dataset



2. Describe it

- Title
- Authors
- Year
- Description
- And others…

3. Assign a DOI

10.1234/exampledata

4. Reuse and reference!

ATLAS Collaboration, "Data from Figure 7 from: Measurements of Higgs boson production and couplings in diboson final states with the ATLAS detector at the LHC: $H \to \gamma\gamma$,"
http://doi.org/10.7484/INSPIREHEP.DATA.A78C.HK44

☑ Unique ☑ Persistent

5. Enjoy the benefits

Findability	Track citations
Reusability	Measure impact

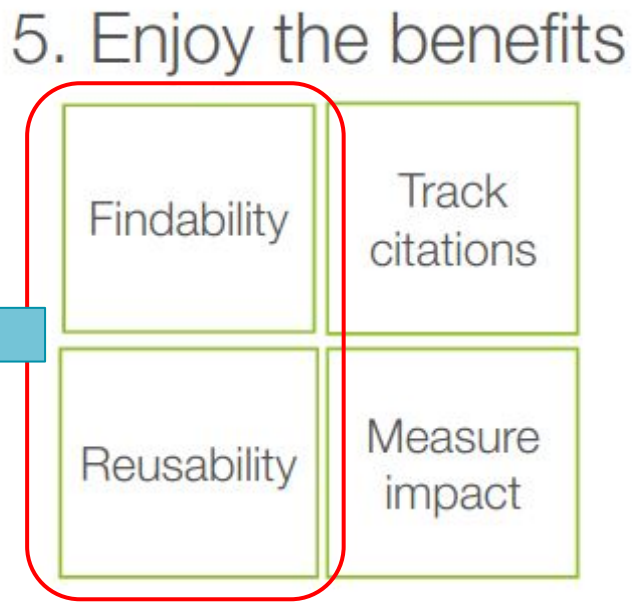
DataCite
FIND, ACCESS, AND REUSE DATA

utt

Why is DOI important?

[The FAIR Guiding Principles for scientific data management and stewardship]

https://www.go-fair.org/fair-principles/



Why is DOI important?

The DATA can be found

- Anywhere
- anytime
- with a click on a link
- NO separate search

GET VISIBLE!

DataCite
FIND, ACCESS, AND REUSE DATA

Meteorologisk institutt

Why is DOI important?

Incorporated into a **citation**, the DOI becomes a guaranteed location for the item cited because it will always resolve to the right web address (URL)

GET CITED!

DataCite
FIND, ACCESS, AND REUSE DATA

Meteorologisk institutt

Why is DOI important?

With a DOI you can easily track who and where your data has been cited.

Think also about social media :-)

GET CREDITS!

DataCite
FIND, ACCESS, AND REUSE DATA

Meteorologisk institutt

Why is DOI important?

With a DOI you can easily show the value of your data within the community and to your funding agencies

BE VALUABLE!

DataCite
FIND, ACCESS, AND REUSE DATA

How to get a DOI?

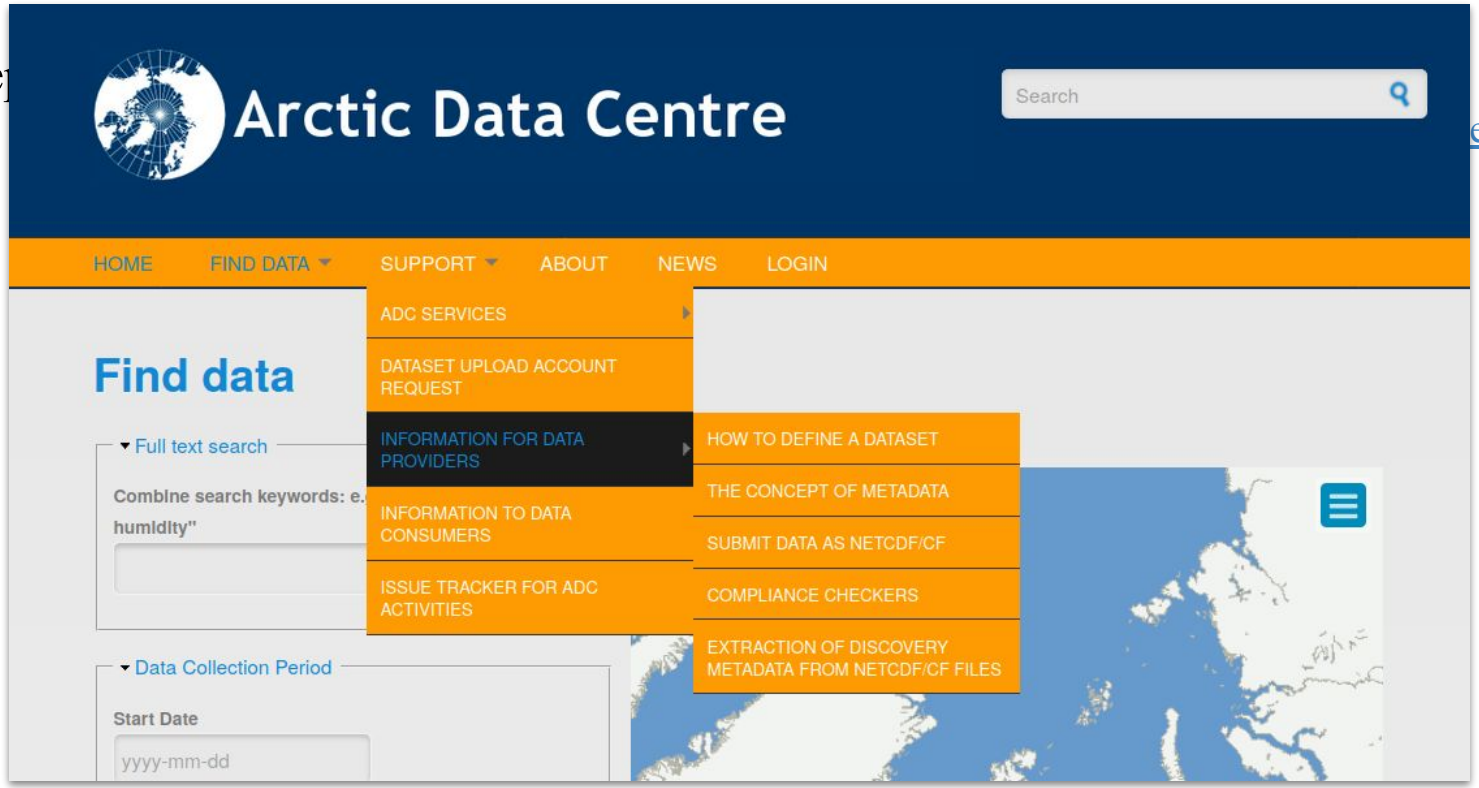
"All of this is great, but I am starting to guess it is quite a lot of work, isn't it?"

How to get a DOI?

- Prepare your dataset
  - Remember that only well documented datasets can be reused!
    - Document your dataset (add good metadata to it! Some guidelines: https://adc.met.no/node/4)

How to get a DOI?

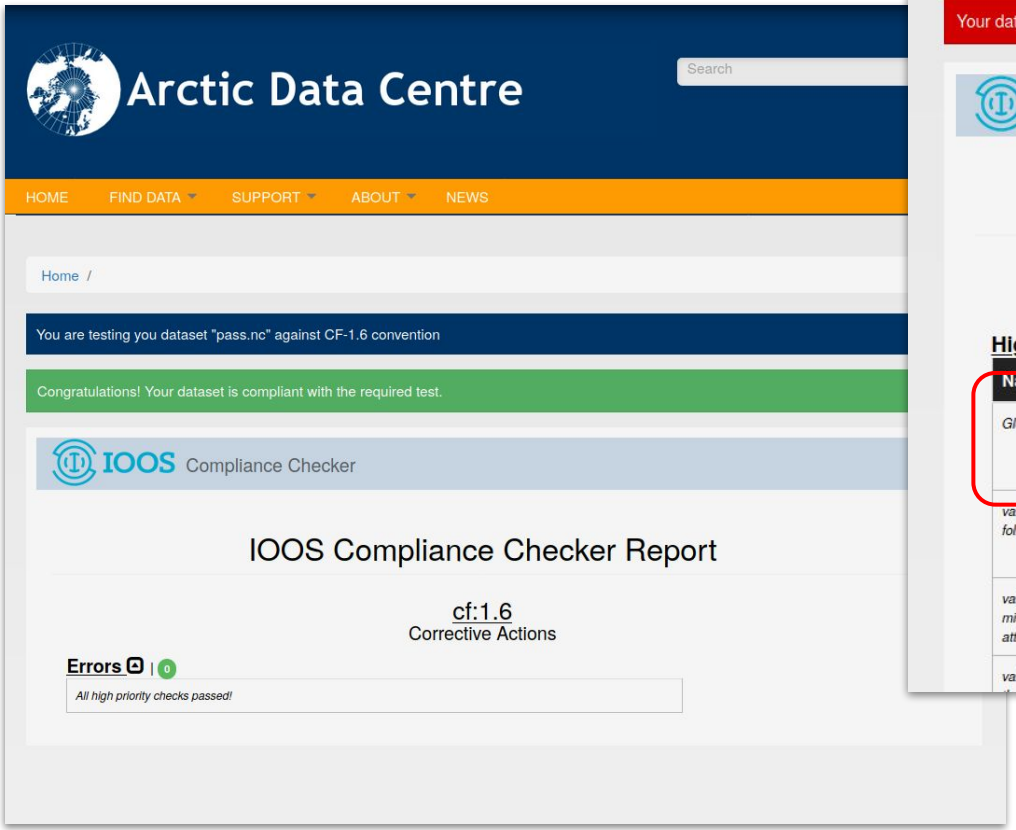
- Prep
  - e/4)

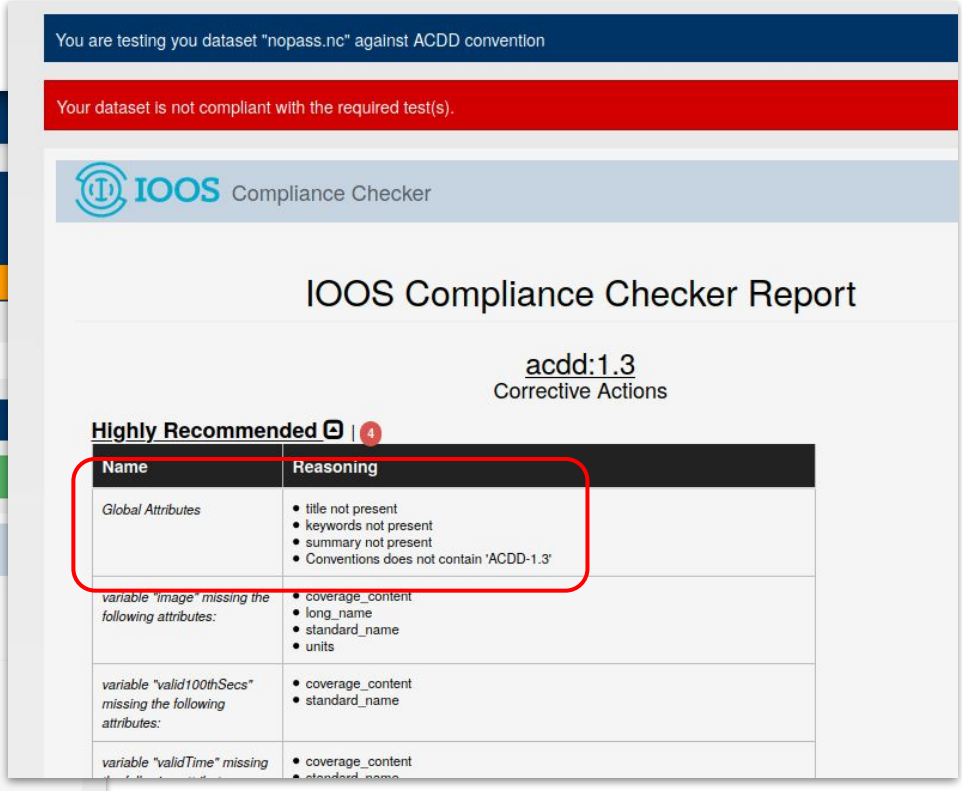


How to get a DOI?

- Prepare your dataset
  - Remember that only well documented datasets can be reused!
    - Document your dataset (add good metadata to it! Some guidelines: https://adc.met.no/node/4)
  - Remember that only understandable datasets can be reused!
    - Follow standards (https://adc-test.met.no/dataset_validation/form)

How to get a DOI?





How to get a DOI?

- Prepare your dataset
  - Remember that only well documented datasets can be reused!
    - Document your dataset (add good metadata to it! Some guidelines: https://adc.met.no/node/4)
  - Remember that only understandable datasets can be reused!
    - Follow standards (https://adc-test.met.no/dataset_validation/form)
  - Place your dataset on thredds.met.no
    - Structure your data (e.g. consider granularity, separate stations, specify featureType)

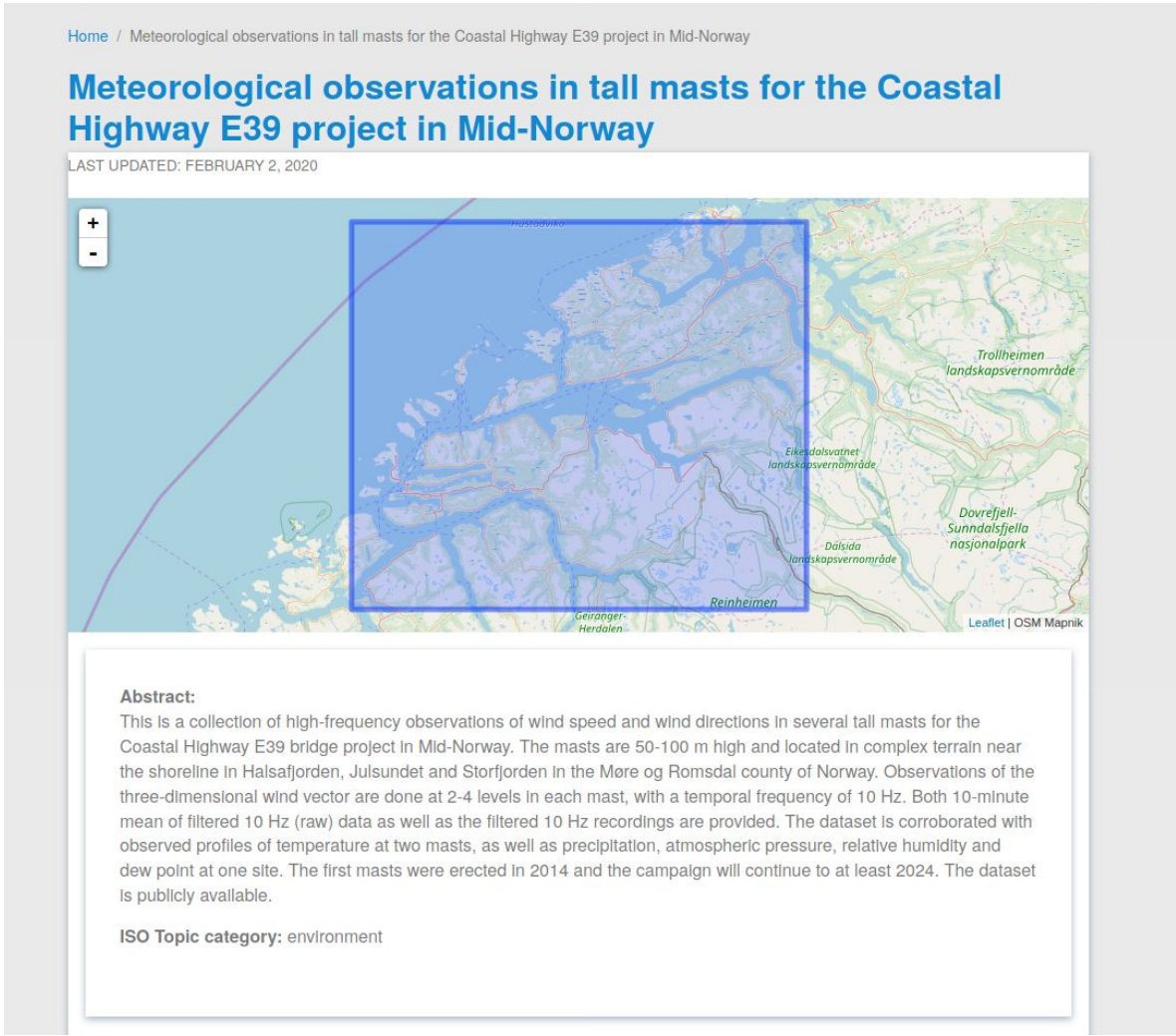
How to get a DOI?

- Prepare your dataset
  - Remember that only well documented datasets can be reused!
    - Document your dataset (add good metadata to it! Some guidelines: https://adc.met.no/node/4)
  - Remember that only understandable datasets can be reused!
    - Follow standards (https://adc-test.met.no/dataset_validation/form)
  - Place your dataset on thredds.met.no
    - Structure your data (e.g. consider granularity, separate stations, specify featureType)
- Send me an email
  - Have patience

adc.met.no

Landing Page Creator

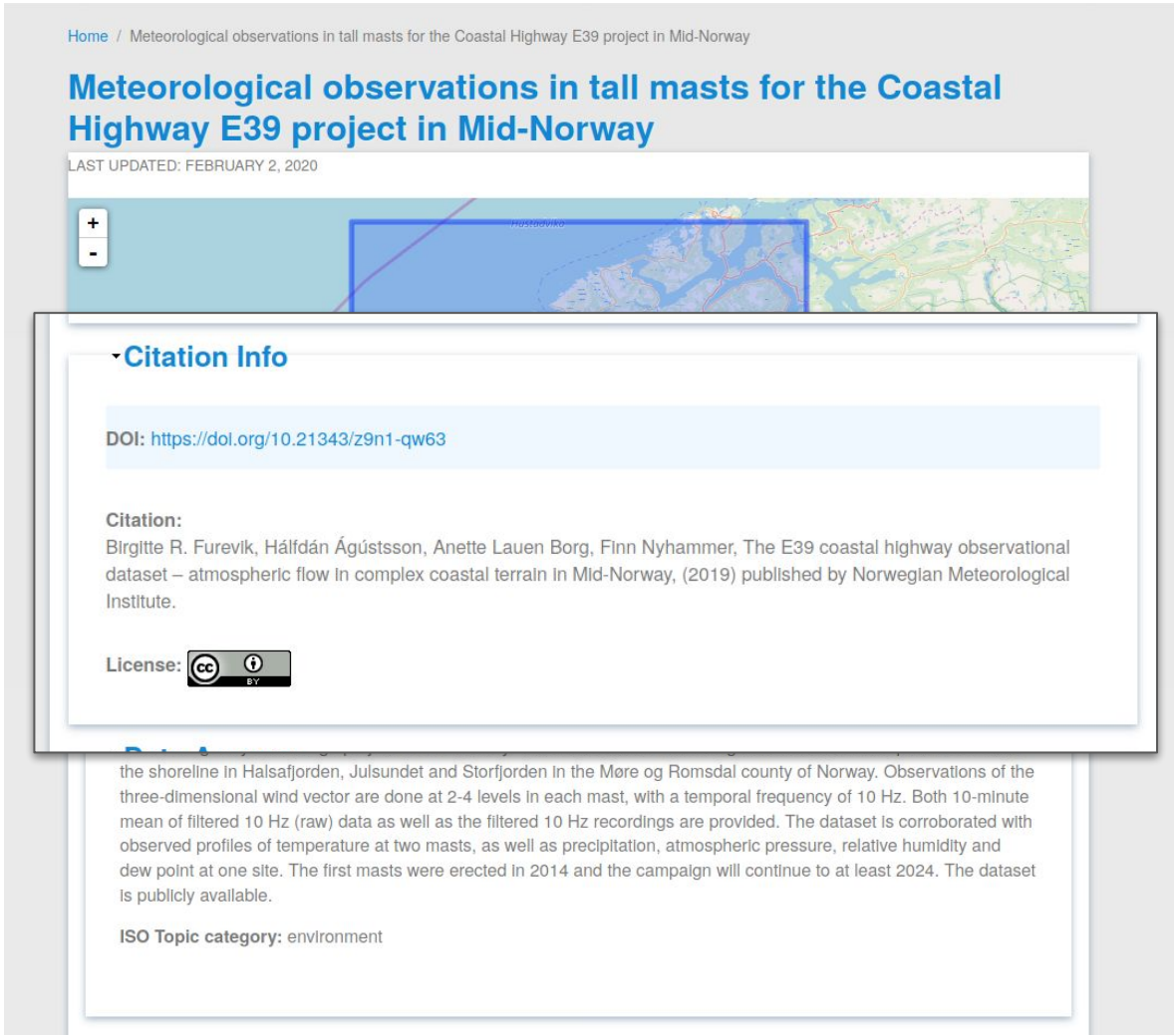
1) enable drupal module (M)
2) create content type on adc (M)
3) configure datacite account (M)



adc.met.no

Landing Page Creator

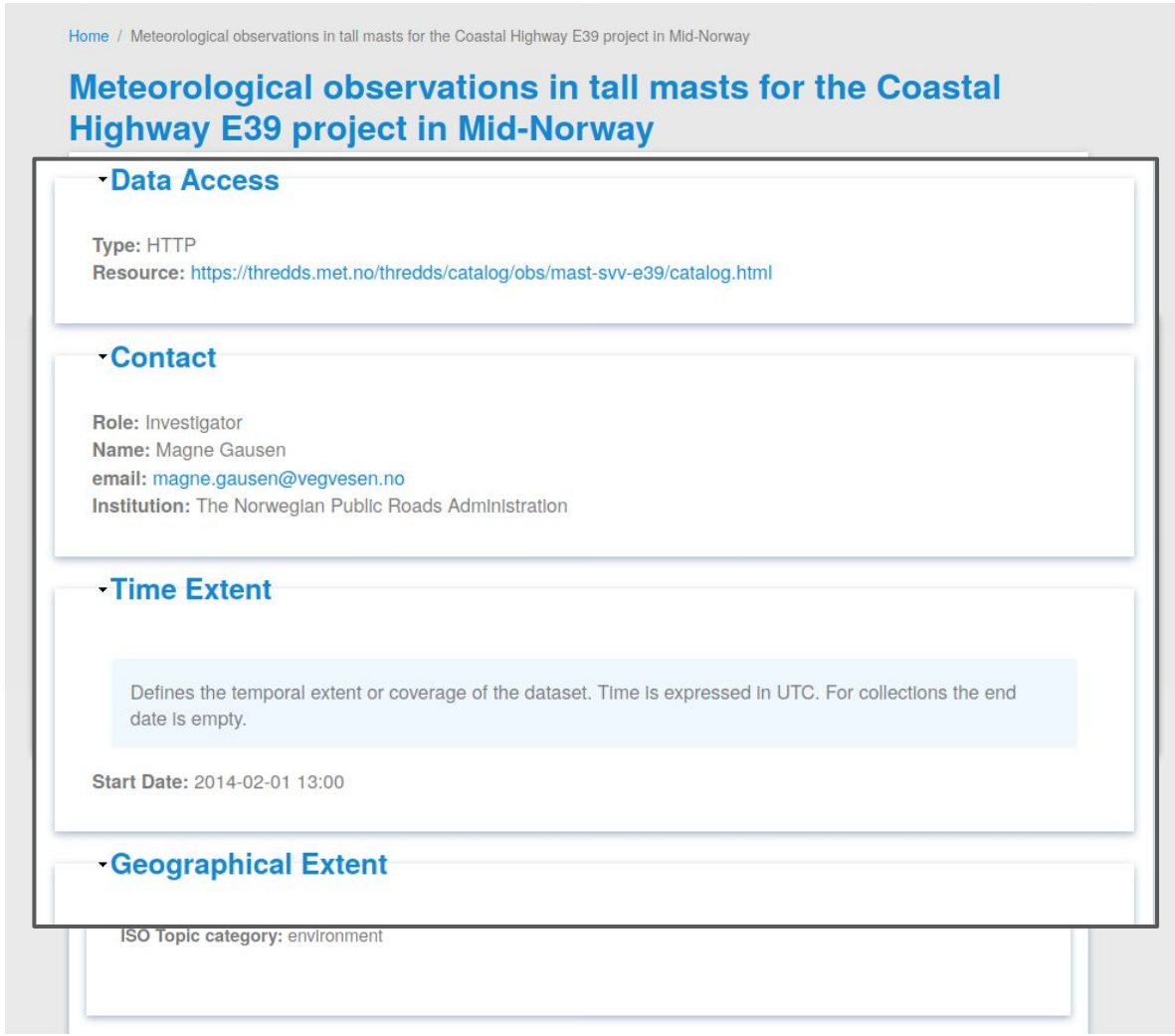
1) enable drupal module (M)
2) create content type on adc (M)
3) configure datacite account (M)



adc.met.no

Landing Page Creator

1) enable drupal module (M)
2) create content type on adc (M)
3) configure datacite account (M)



M = manual
A = automated

dataset.nc

Extract metadata (M/A)

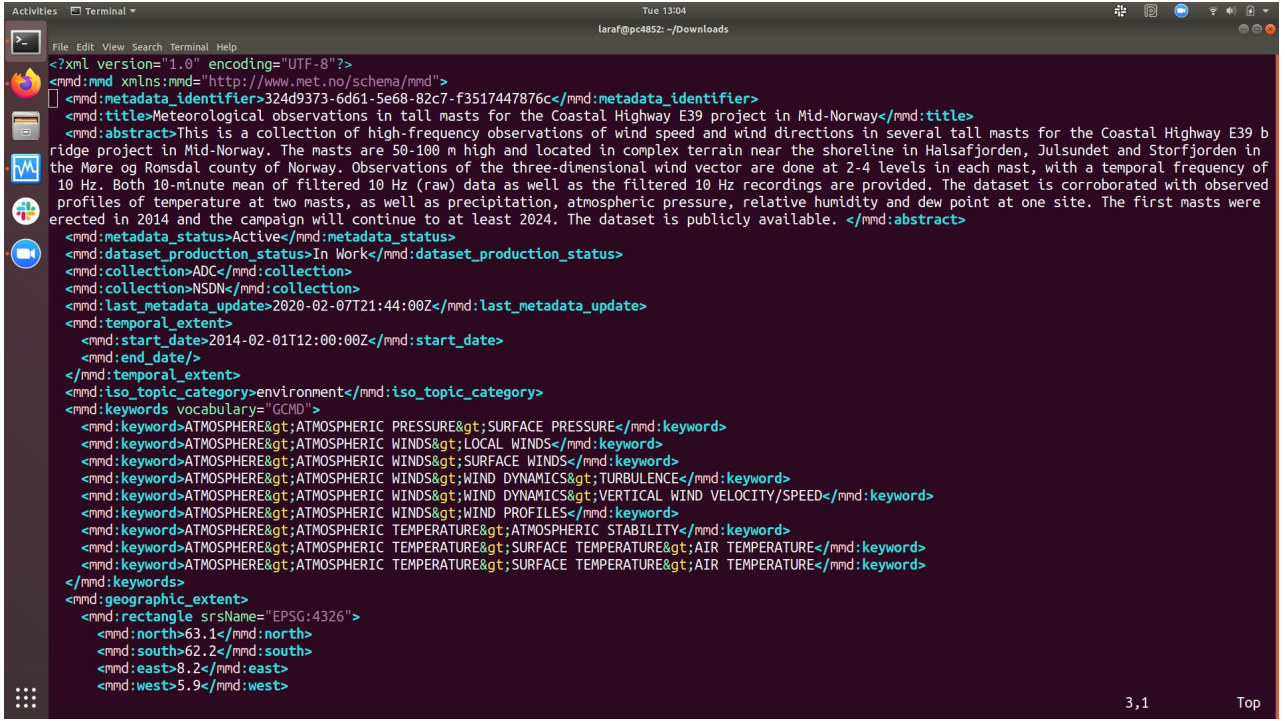
mmd.xml

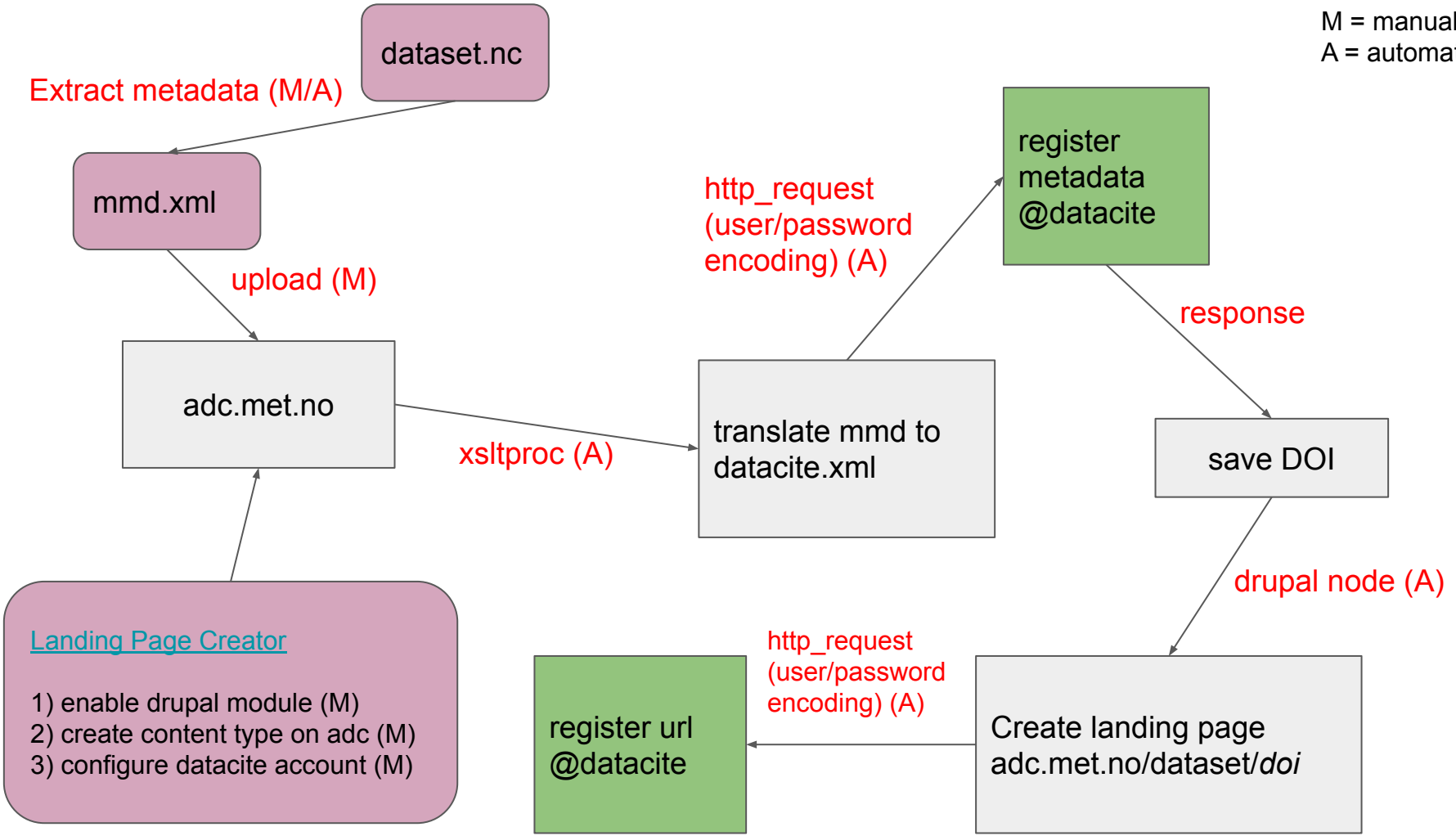
upload (M)

adc.met.no

[Landing Page Creator]

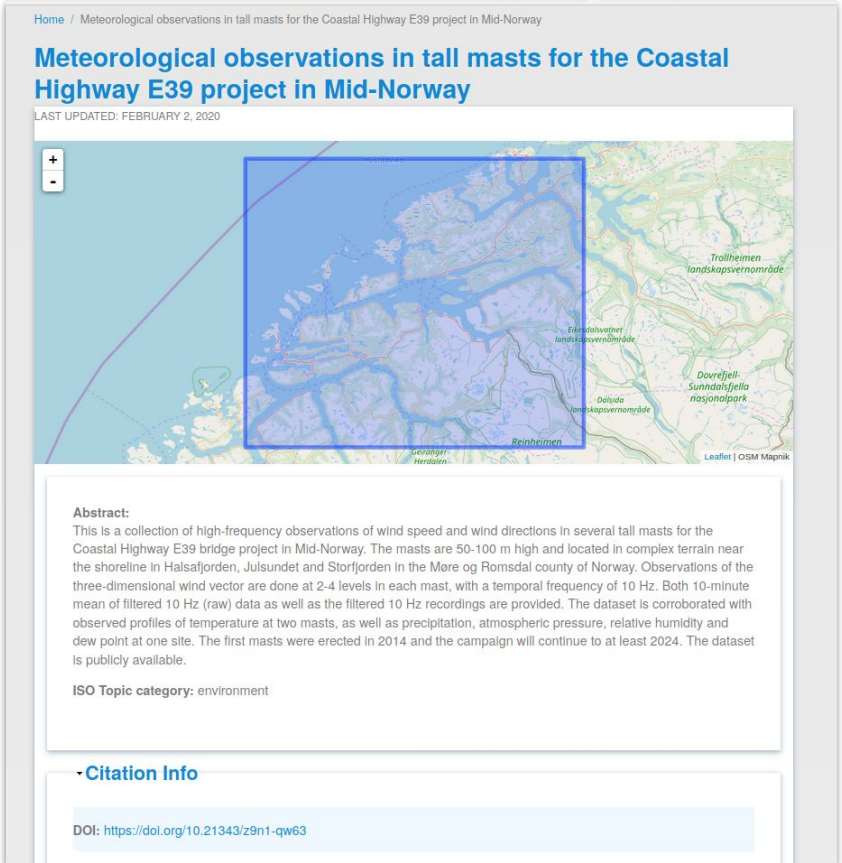
1) enable drupal module (M)
2) create content type on adc (M)
3) configure datacite account (M)



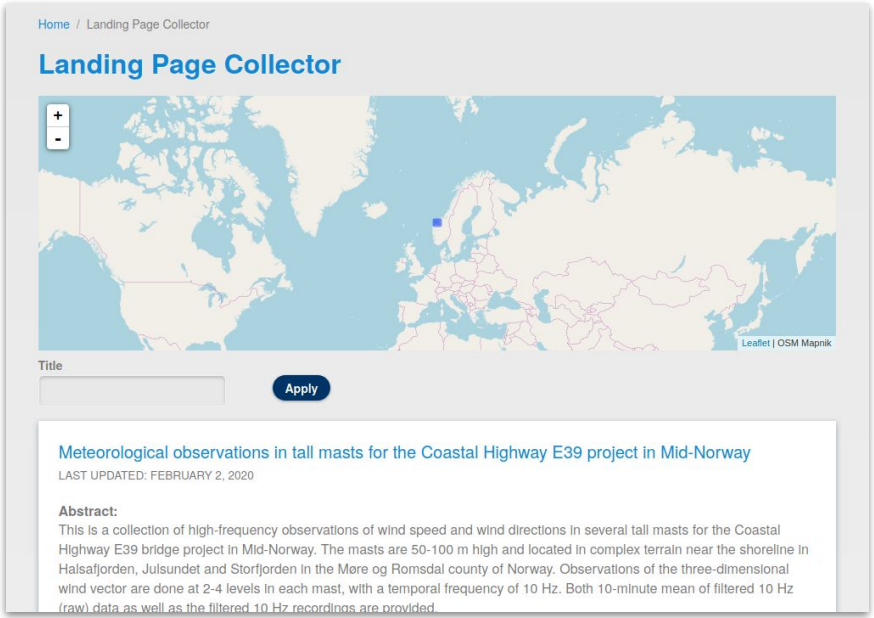

M = manual
A = automated
dataset.nc
Extract metadata (M/A)
mmd.xml
upload (M)
adc.met.no
xsltproc (A)
translate mmd to datacite.xml
http_request (user/password encoding) (A)
register metadata @datacite
response
save DOI
drupal node (A)
Create landing page adc.met.no/dataset/doi
http_request (user/password encoding) (A)
register url @datacite
Landing Page Creator
1) enable drupal module (M)
2) create content type on adc (M)
3) configure datacite account (M)

Celebrate your DOI!

https://doi.org/10.21343/z9n1-qw63



https://adc.met.no/landing-page-collector



Thank you!

Further reading:

FAIR data
https://www.go-fair.org/fair-principles/
https://septentrio.uit.no/index.php/SCS/issue/view/451

Conventions:
http://wiki.esipfed.org/index.php/Attribute_Convention_for_Data_Discovery_1-3
http://cfconventions.org/standard-names.html

DOI
https://datacite.org/
https://www.doi.org/faq.html