OpenAIRE
Belgium

Open Research Data

Open research data and data management for Horizon 2020 projects

Emilie Hermans
Project Assistant OpenAIRE, UGent

Challenge

Challenge


CREATING DATA
PROCESSING DATA
ANALYSING DATA
PRESERVING DATA
GIVING ACCESS TO DATA
RE-USING DATA
PUBLICATIONS

DATA MANAGEMENT AND OPEN DATA

Prevents data loss

Credit & longer shelf life [1]

Maximize usefulness

Increases transparency

Fosters creativity

Promote integrity

Citizens science

1. e.g. Piwowar HA, Vision TJ. (2013) Data reuse and the open data citation advantage. PeerJ 1:e175 https://doi.org/10.7717/peerj.175, Piwowar HA, Day RS, Fridsma DB (2007) Sharing Detailed Research Data Is Associated with Increased Citation Rate. PLoS ONE 2(3): e308. doi:10.1371/journal.pone.0000308

EC: The Open Research Data Pilot

Flexible ORD pilot:

From limited to default in 2017

- Foster Open Science
- Avoid duplication of research and loss of resources

Data Management Planning

Open Access to research data (or partially opt-out)

Which areas are participating?

Projects started in 2014-2016

Limited ORD Pilot

- Limited ORD pilot: some areas: Check Article 29.3
- DMP/Dataset
- Possibility to opt-in or opt-out

From 2017

Extended ORD Pilot

- Participating is **default** option for **all** projects
- 1 DMP/Project
- Possibility to opt-out

Costs eligible (Article 6.2.D.3 of the Model Grant Agreement)

(PARTIALLY) OPTING-OUT

Reasons e.g.

- Exploitation of results
- Confidentiality
- Protection of personal data
- Would jeopardize the main aim of the action
- No data generated
- Any other legitimate reason

- Projects can opt out at any stage:
  - Complete opt-out via project amendment
  - Complete or partially opt-out:

    describe issues in project DMP

As open as possible as closed as necessary

Requirements of the Data Pilot

1. Data Management Plan (DMP)

2. Deposit data in data repository

3. Provide information to validate results

4. Open up data

What is a DMP?

Handling of data during and after project

Living document: update

Reflects on curation, preservation, sustainability and security

What parts will be open and how?

Content of H2020 DMP

Template: EC guidelines on FAIR Data Management

Data summary

FAIR Data principles

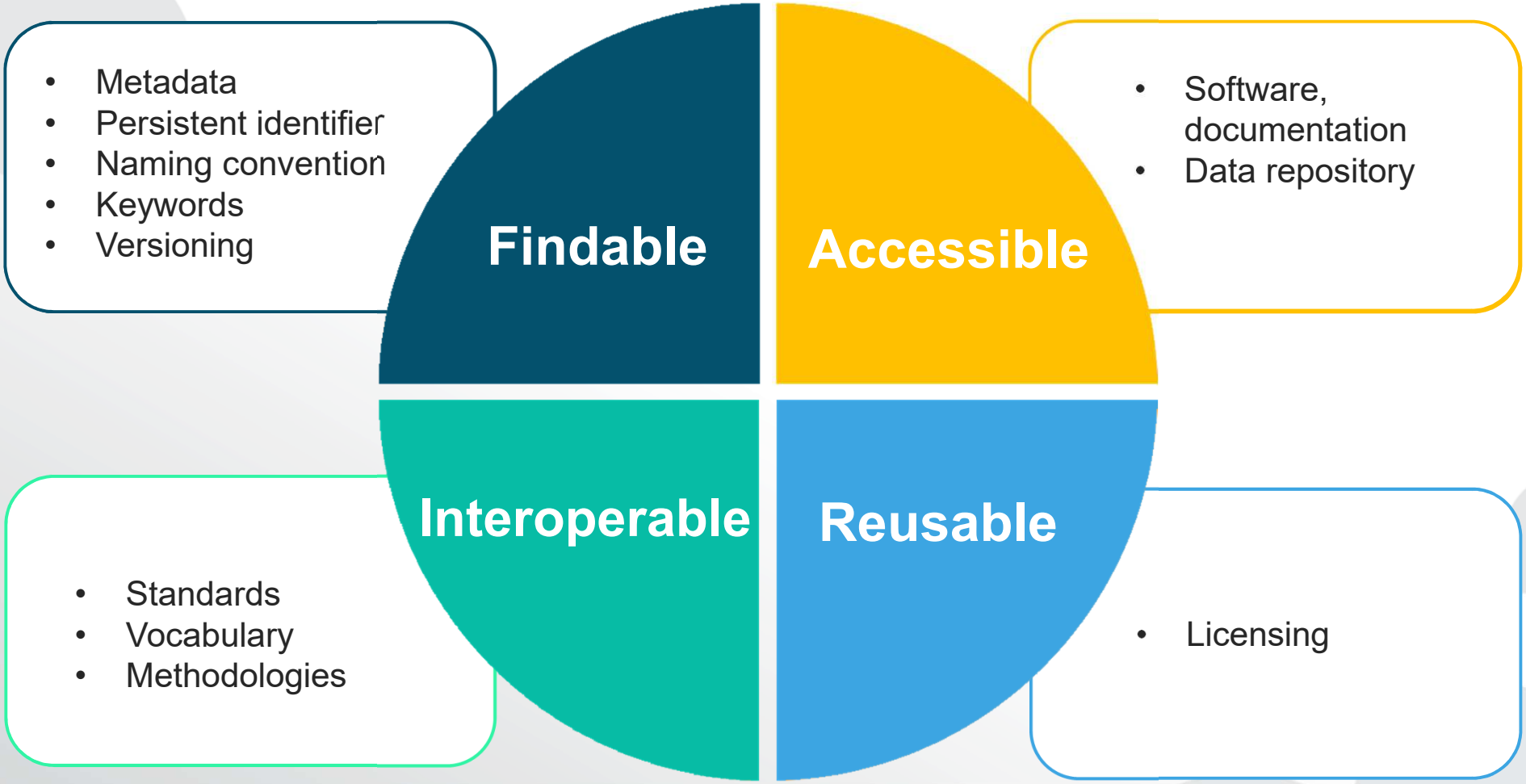
Resources

Data security

Ethical aspects

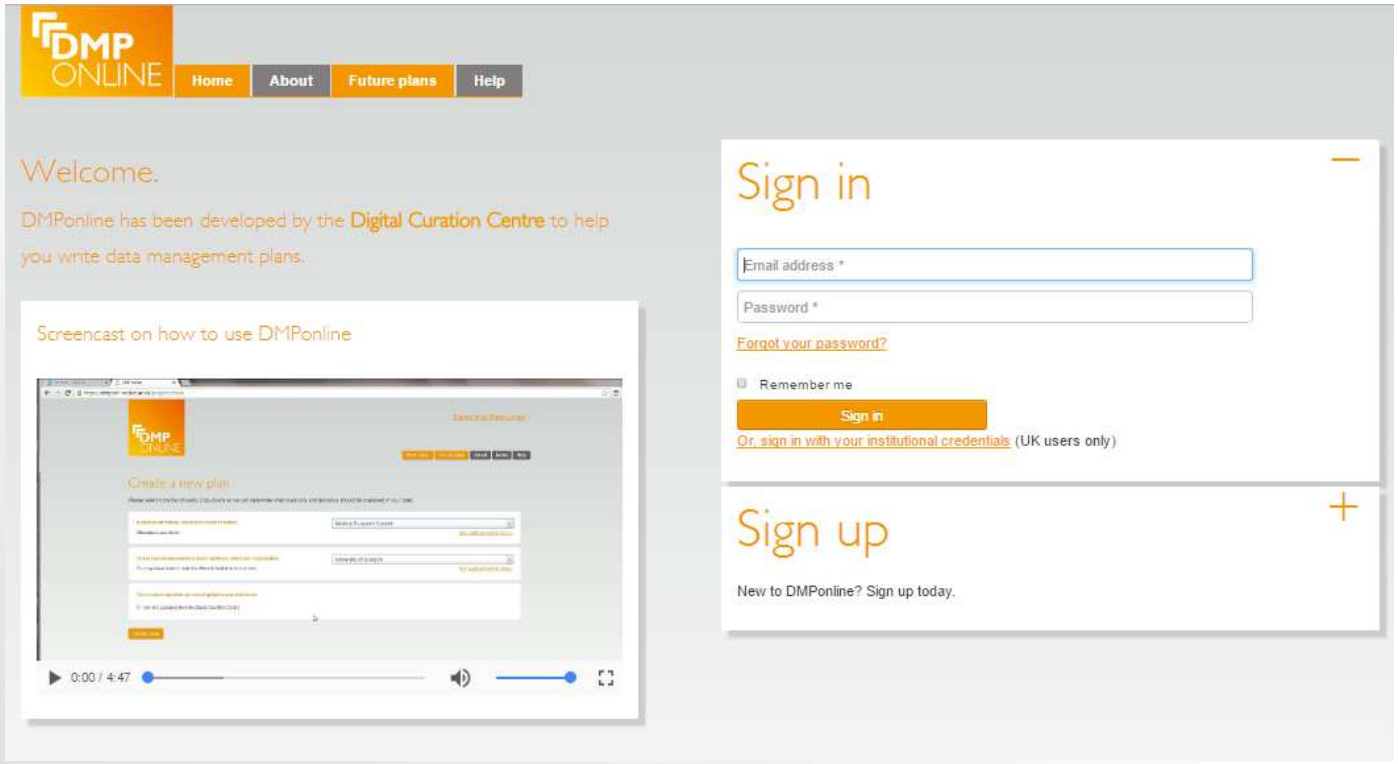
http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf

FAIR data principles

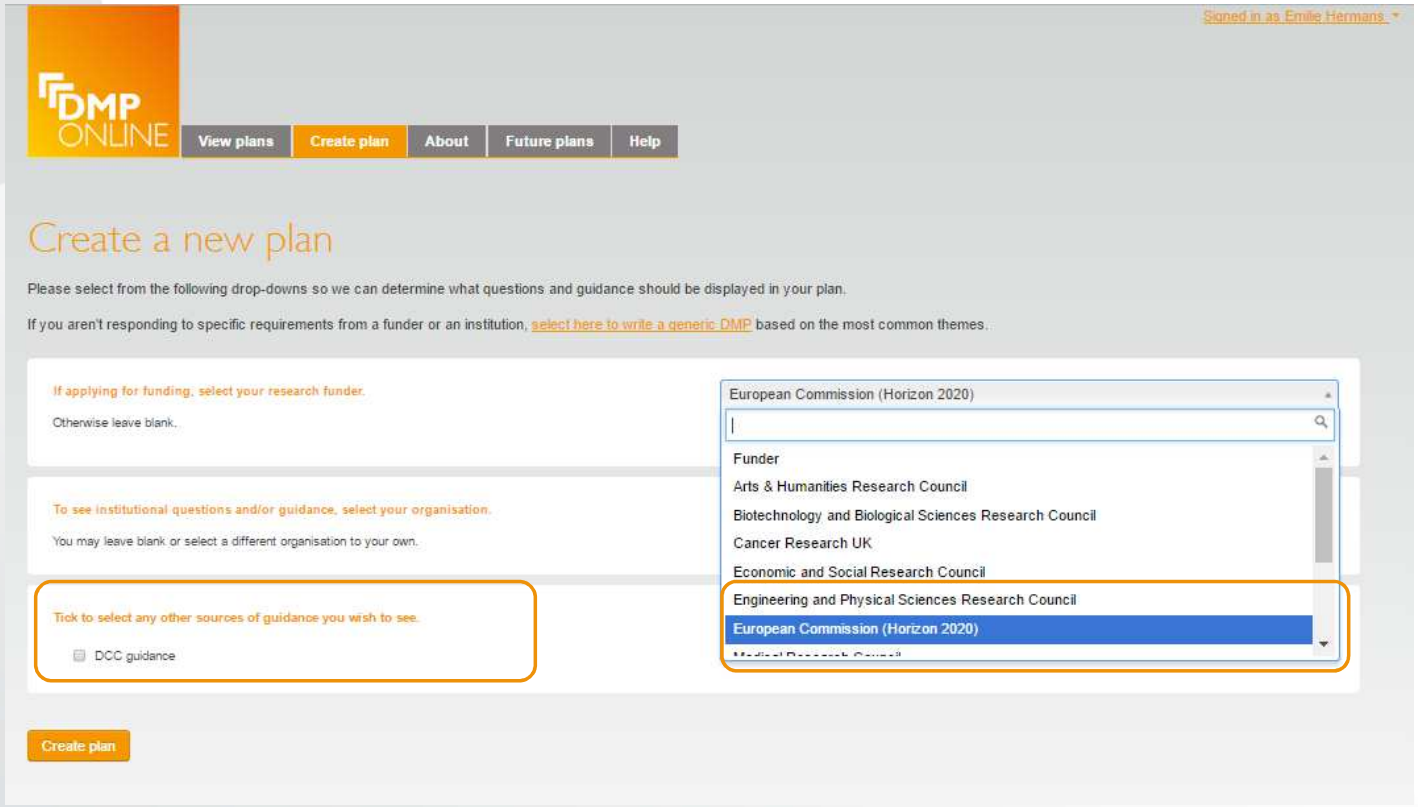


How to write a DMP

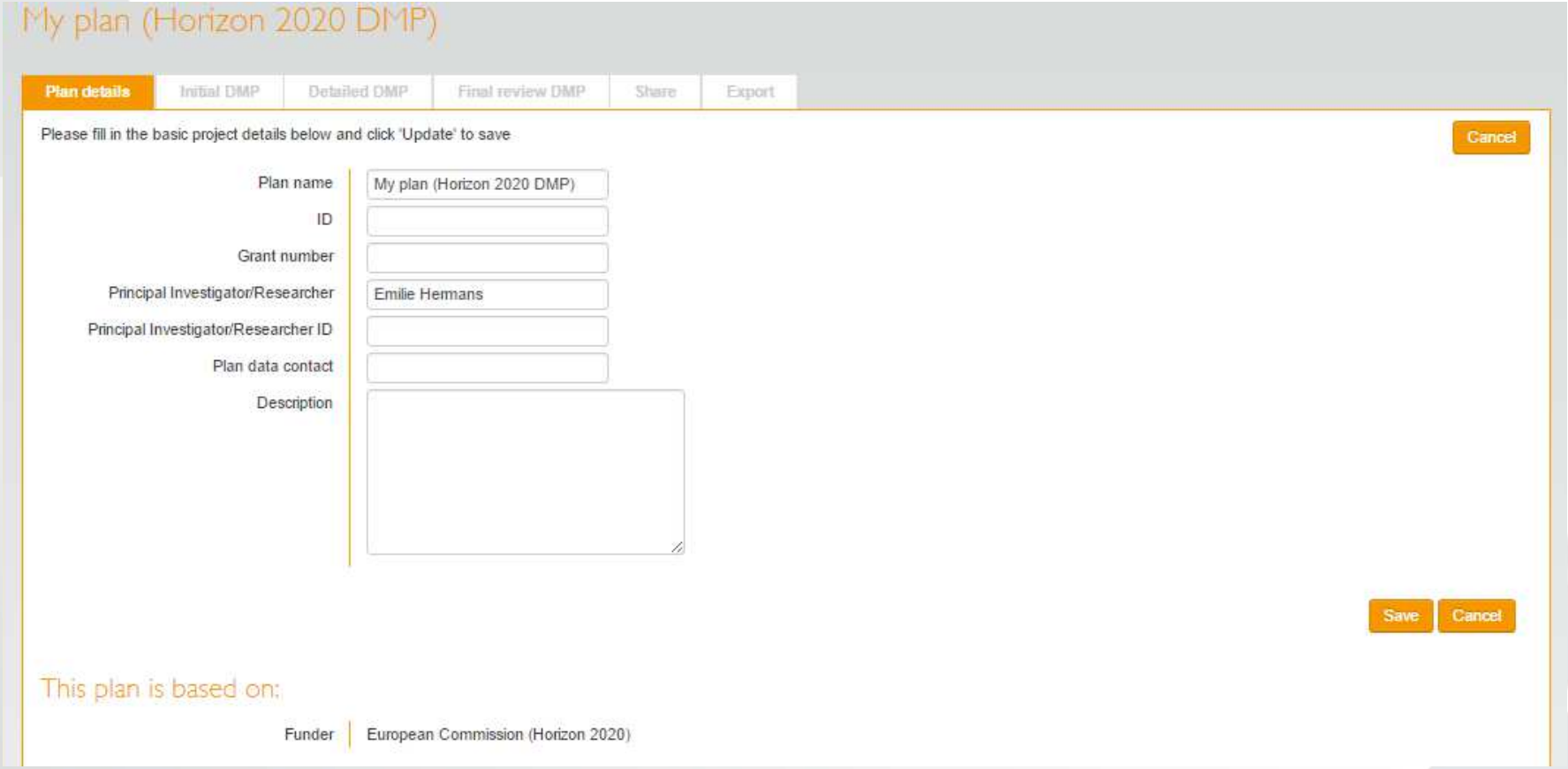
dmponline.dcc.ac.uk



Create and confirm plan

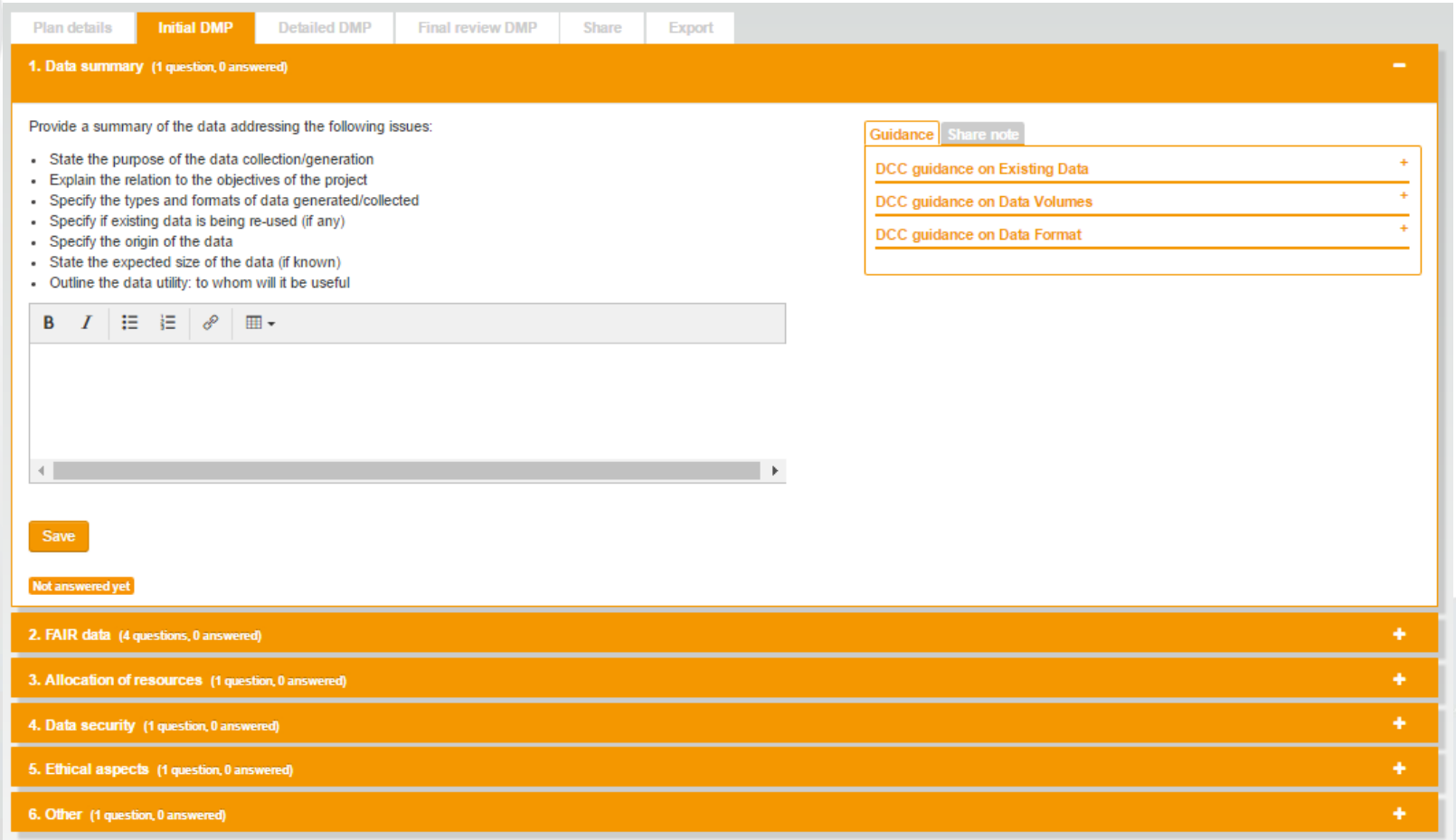


Plan Details

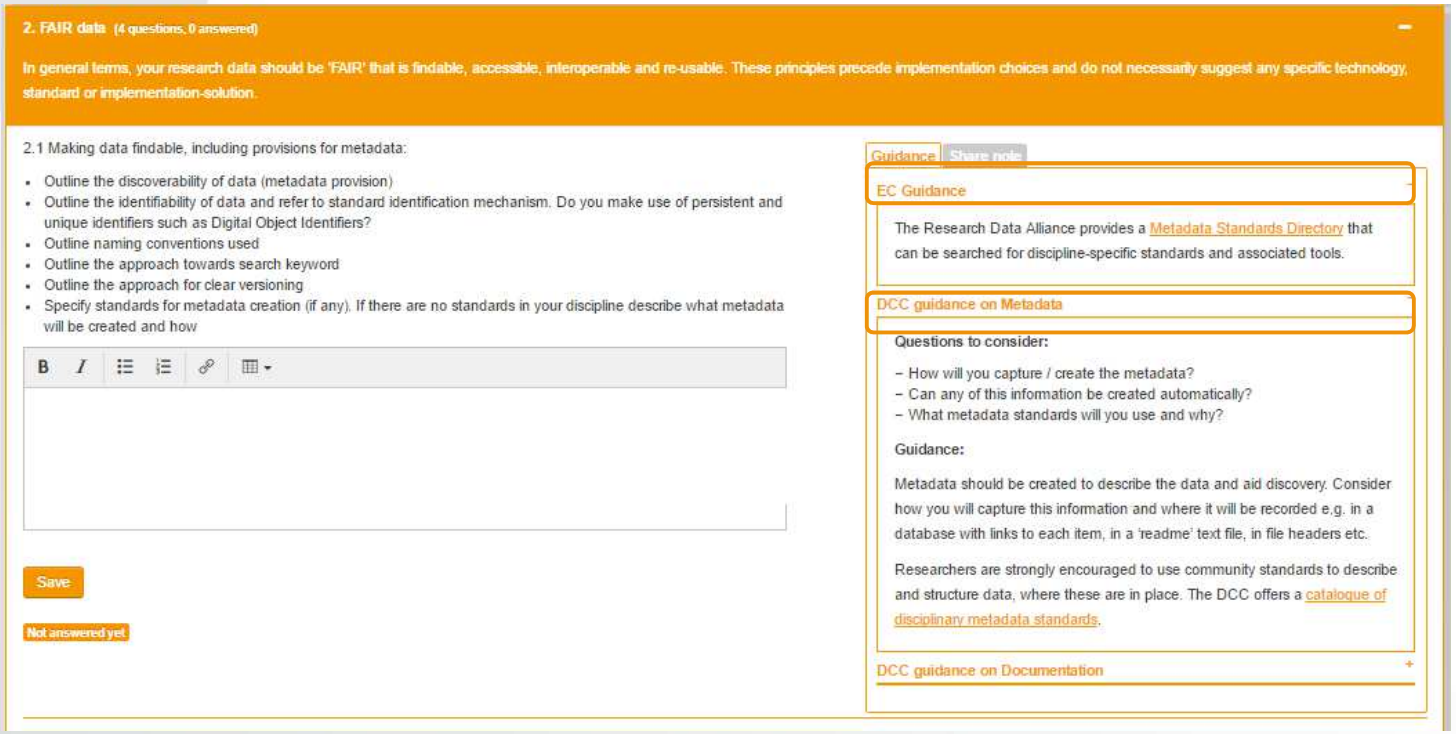


Versions

Share with partners



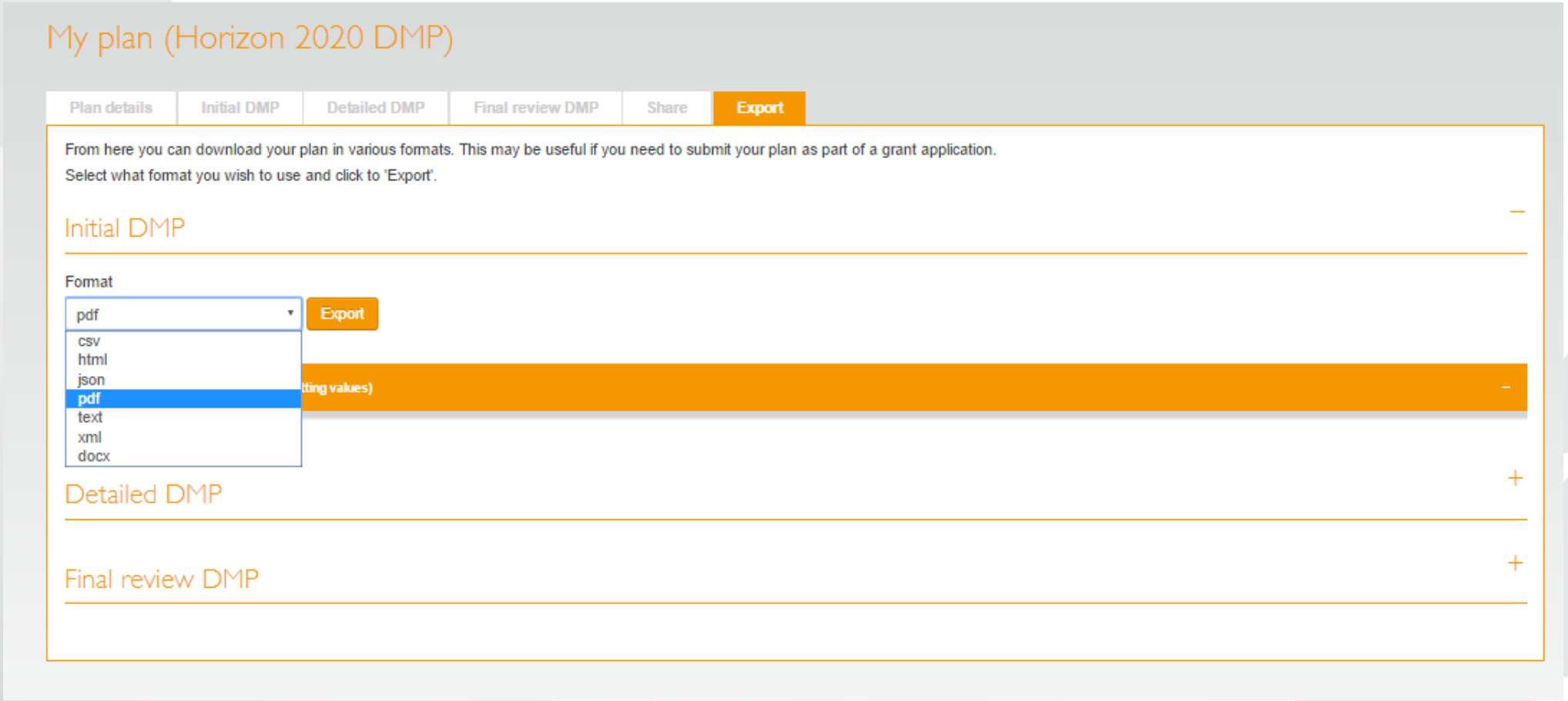
Guidance



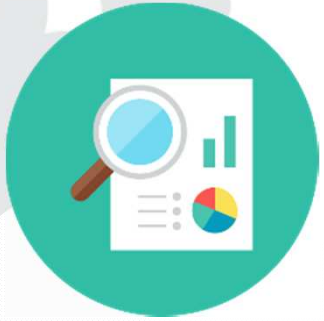
Guidance based on guidelines EC

Guidance and links from DCC

Export to various formats



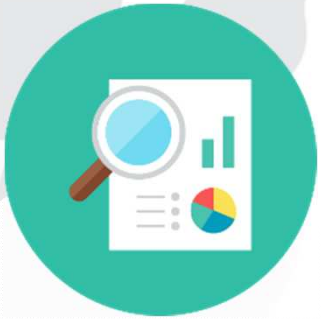
RECOMMENDATIONS



Data summary

Data collection

- Origin: generated, collected, reused
  - Does similar data exists? What about reintegration or reuse?
  - Re-use? Provide the source and check IPR
- Types e.g. digital/non-digital data, qualitative/quantitative, audio files, surveys, databases, field notes…..



Data summary

Data files formats

Use data formats that are:

- Open standard possible?
- In an easily re-usable format
- Commonly used by research community

Use consequent naming convention

Structured organizing of files

Examples of preferred format choices:	
Text	.odt, .txt, .xml, .html, .rtf
Tabular Data	.csv (comma separated values), .xml, .rdf, .SPSS portable
Images	.tif, .jpeg2000, .png, .svg,
Structured data	.xml, .rdf

Any standard used in your field

Accessible

Documenting data

- Make your data understandable: project level (context) and data level (e.g. codebooks, protocol)
- Be clear what methods you use
  - E.g. lab notebook, end-to-end code/scripts for statistics
  - Software can help: R, MatLab, Python…

Findable

Create searchable data

Using metadata

- Data about data
- Consists of set of attributes
- Machine readable
- Helps prevent inappropriate use
- Use metadata standards of your domain

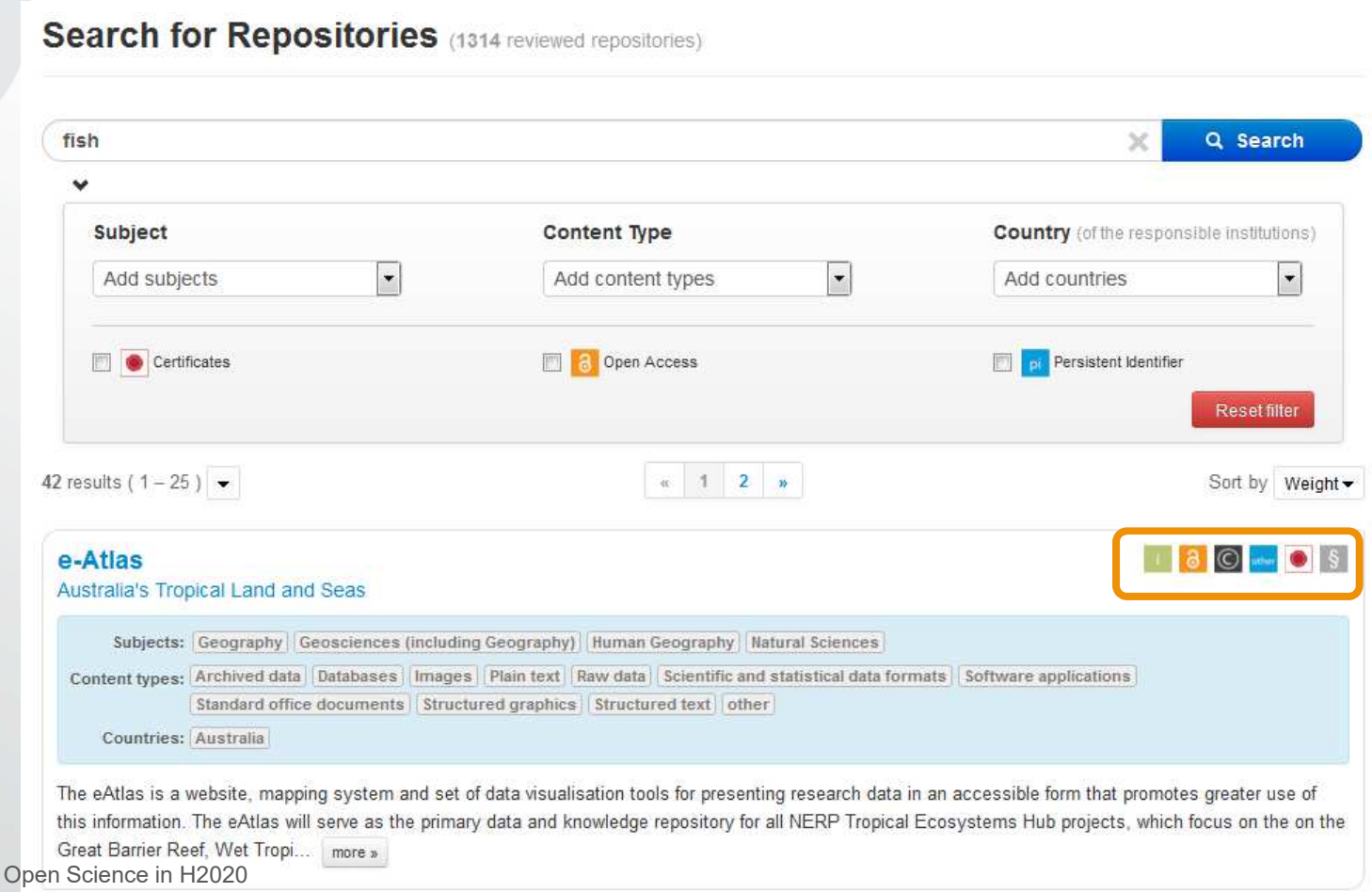
Accessible

Where to deposit data?

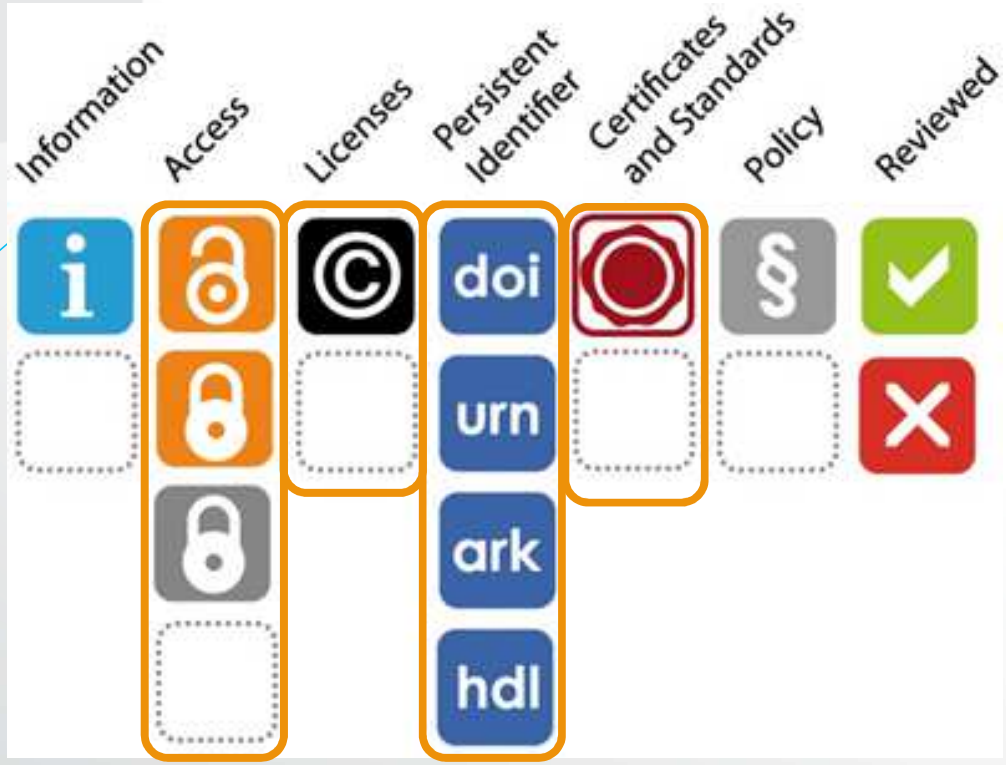
Research data repository

- Matches data needs
- Disciplinary/Institutional data repository
- Directory of data repositories: www.Re3data.org
- Zenodo

Re3data.org


Search for Repositories (1314 reviewed repositories)
fish
Search
Subject
Add subjects
Content Type
Add content types
Country (of the responsible institutions)
Add countries
Certificates
Open Access
Persistent Identifier
Reset filter
42 results (1 – 25)
« 1 2 »
Sort by Weight
e-Atlas
Australia's Tropical Land and Seas
Subjects: Geography Geosciences (including Geography) Human Geography Natural Sciences
Content types: Archived data Databases Images Plain text Raw data Scientific and statistical data formats Software applications Standard office documents Structured graphics Structured text other
Countries: Australia
The eAtlas is a website, mapping system and set of data visualisation tools for presenting research data in an accessible form that promotes greater use of this information. The eAtlas will serve as the primary data and knowledge repository for all NERP Tropical Ecosystems Hub projects, which focus on the on the Great Barrier Reef, Wet Tropi... more »

Re3data



- Access
- Licenses
- Persistent identifier

→ Trustworthy Digital repository

What to deposit?

Everything needed to validate results presented in scientific publications

DATA

- Validate results
- Selection

METADATA

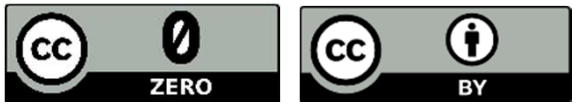
DOCUMENTATION

- Tools, software….
- Read_me file?

OPEN DATA?

Reusable

Open data

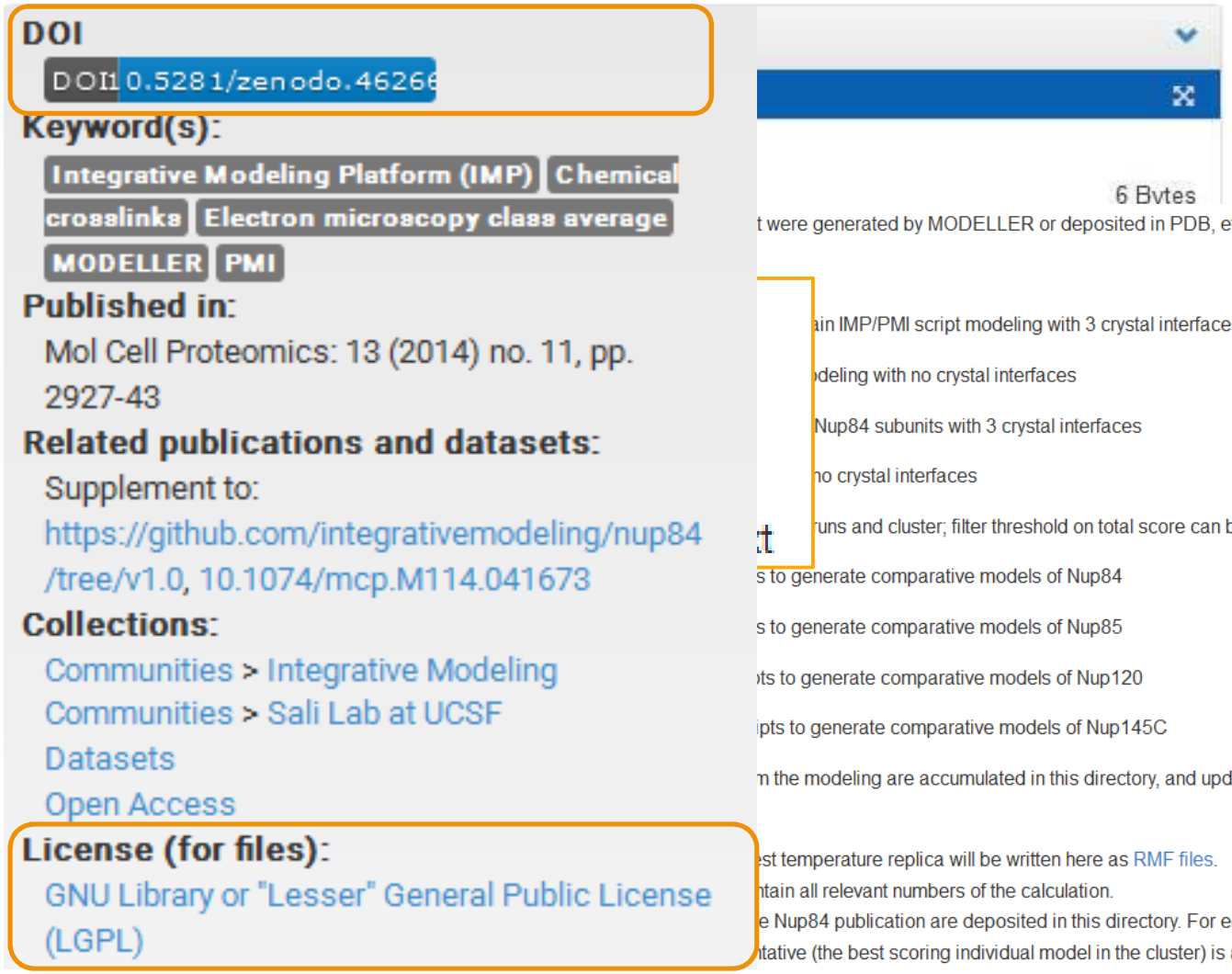
- Keep it simple
- Apply an open license:
  - e.g. creative commons: 

  - Data repositories can provide licenses
    - Re3data.org

The terms of use and licenses of the data are provided by the research data repository.

Example

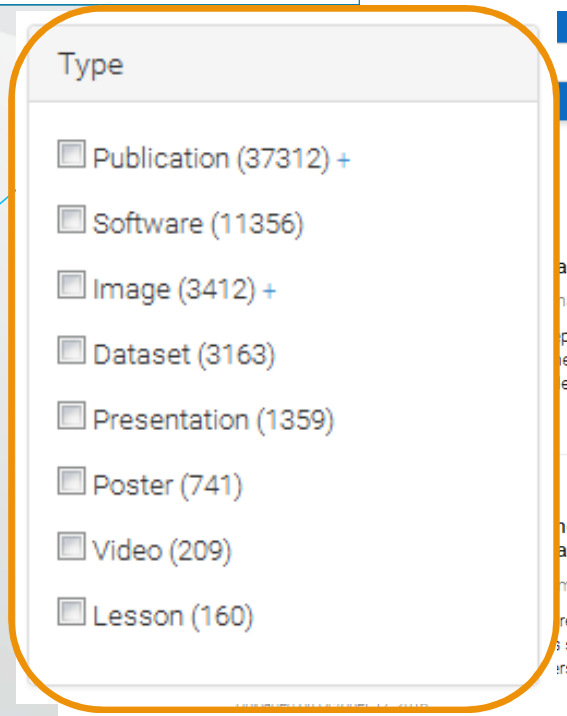
- Understandable for humans
- Machine readable metadata
- Tools
- Open Data
- Open license



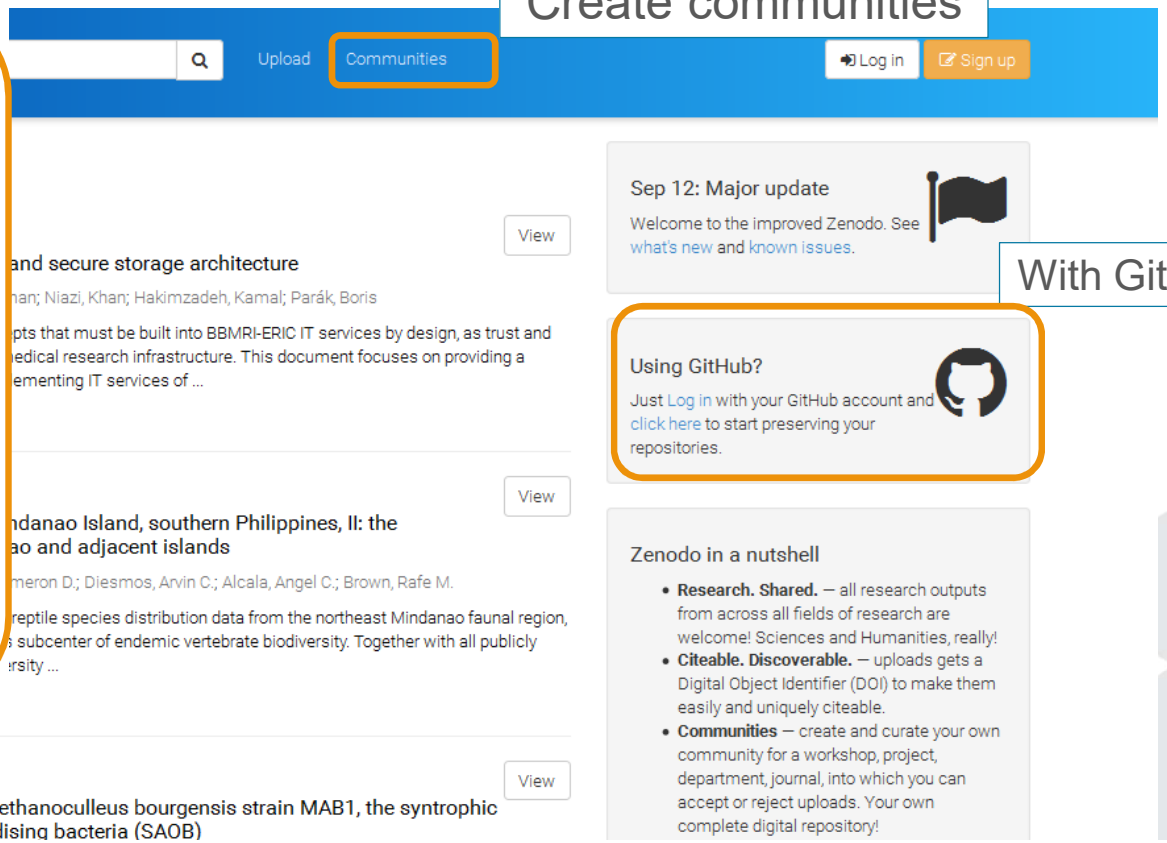
Zenodo

upload → describe → publish

For all content types



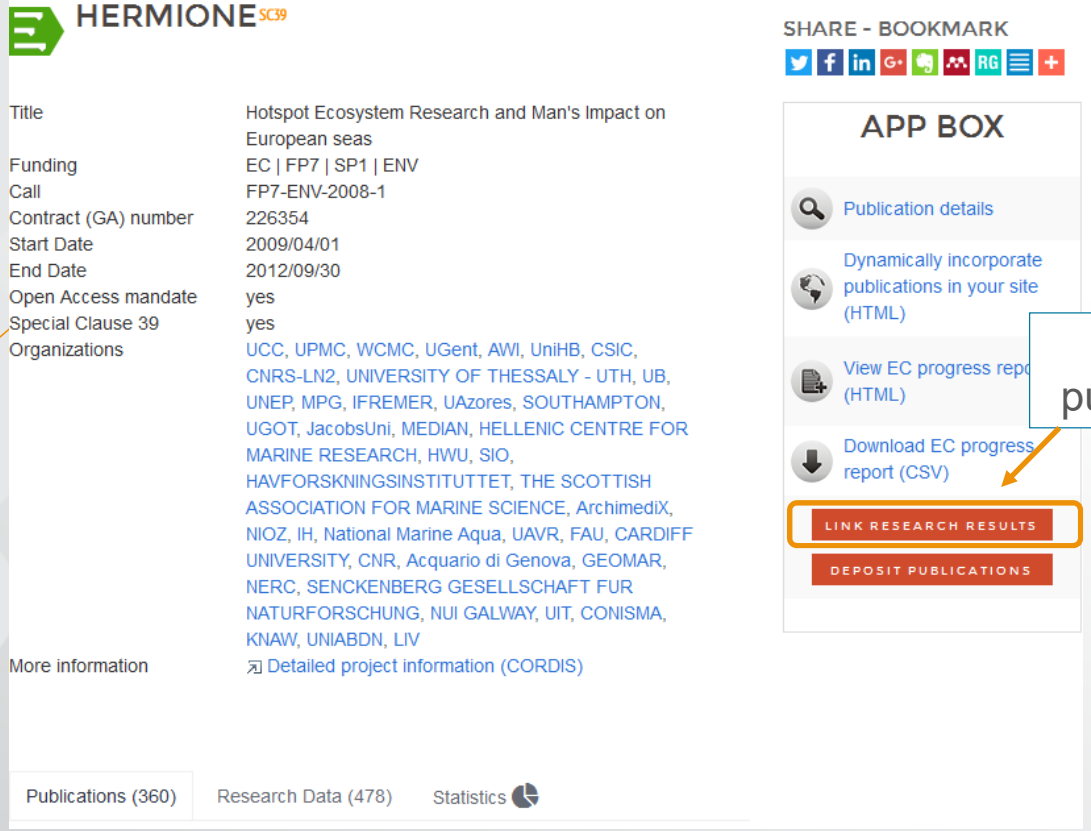
Create communities



With GitHub Integration

OpenAIRE

www.openaire.eu/search



Link your data to publications or project

OpenAIRE

Training and support material

www.openaire.eu/opendatapilot

Information on:

- Open research data pilot
- Creating a data management plan
- Selecting a data repository
- Dealing with personal data

Support material:

Briefing papers, factsheets, webinars, workshops , FAQs, helpdesk

STEP 1

WRITE A DMP
dmponline.dcc.ac.uk

Update at

- 6 months
- Periodic evaluation
- Final review

STEP 2

FIND REPOSITORY

Matches data needs

Data Repositories

- discipline/institutional
- www.re3data.org
- Zenodo

STEP 3

DEPOSIT DATA

(Open) Data

Metadata

Other tools

- Standard File Formats
- Standards metadata schema
- (Open) Licences

SUPPORT

Supporting infrastructure and information

Open Research Data Pilot

- EC guidelines
- OpenAIRE.eu
- dcc.ac.uk

Questions!

www.openaire.eu

@openaire_eu

Facebook.com/groups/openaire

https://www.linkedin.com/groups/OpenAIRE3893548

Emilie.Hermans@UGent.be

info@openaccess.be