Plan ahead: practical tools to make your data and software more FAIR

NWO BioPhysics 2022 -
October 10 2022

https://tinyurl.com/BioPhysics2022FAIR

https://doi.org/10.5281/zenodo.7133499

Lourens Veen

Research Software Engineer
Netherlands eScience Center
l.veen@esciencecenter.nl

Kim Ferguson

Research Data Management Specialist
DANS
kim.ferguson@dans.knaw.nl

Mateusz Kuzak

Community Manager
Netherlands eScience Center
m.kuzak@esciencecenter.nl

Jaro Camphuijsen

Research Software Engineer
Netherlands eScience Center
j.camphuijsen@esciencecenter.nl

Esther Plomp

Data Steward
TU Delft, Faculty of Applied Sciences
e.plomp@tudelft.nl

1

Introduction

Findable, Accessible, Interoperable and Reusable (**FAIR**), Data/Software Management Plans (**DMP**)
Esther

Pitches + Break out sessions!

Learn more: The Turing Way

DMP

- Tool to structure the management of your data/software
  - Increase quality of documentation
  - Prevents loss of data
  - No unfindable files through thanks to file naming conventions
- How can I make my data/software FAIR?
- Data/software sharing
  - Validation
  - Increased impact (Citation and collaborations)


Findable
Accessible
Interoperable
Reusable

Findable

- Deposit your data in a data repository with metadata and a persistent identifier

F

Findable

- Deposit your data in a **data repository** with metadata and a persistent identifier

F

an online archive that curates research datasets and provides long-term access

- Finalised datasets
- ~10-15 years

Repositories

ENA European Nucleotide Archive

DANS

DataverseNL

DRYAD

4TU.Centre for Research Data

European Genome-phenome Archive

PCDDB

figshare

zenodo

UniProt

IDR Image Data Resource

[Recommended Repositories](#) (nature)
[Registry of Research Data Repositories](#)
[Fairsharing.org](#)

Findable

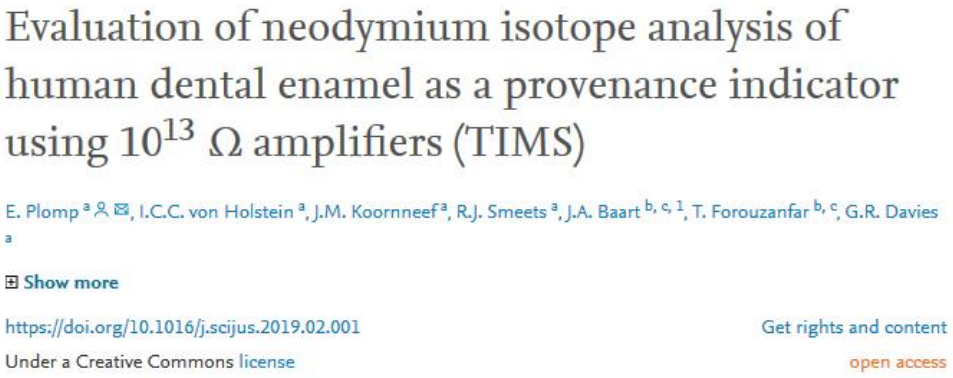
- Discipline common metadata standards
  - FAIRsharing.org
  - Research Data Alliance metadata directory
  - Digital Curation Center

- Deposit your data in a data repository with **metadata** and a persistent identifier

Metadata = information about data

- Contextual information
- Title, author, keywords
- When? For what purpose?
- Size? Standards?

F



Evaluation of neodymium isotope analysis of human dental enamel as a provenance indicator using 10^{13} Ω amplifiers (TIMS)

E. Plomp [a], I.C.C. von Holstein [a], J.M. Koornneef [a], R.J. Smeets [a], J.A. Baart [b, c, 1], T. Forouzanfar [b, c], G.R. Davies [a]

Show more

https://doi.org/10.1016/j.scijus.2019.02.001

Get rights and content

Under a Creative Commons license

open access

- Deposit your data in a data repository with metadata and a **persistent identifier**

A persistent identifier is a long-lasting reference to a file, web page, or other object

ORCID

F

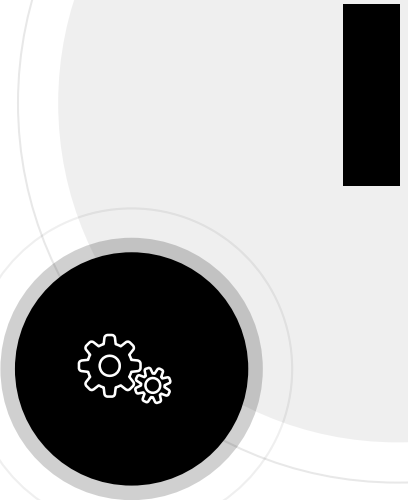
Accessible

- Consider what will be shared
- Obtain participant consent and perform risk management
- Determine access control
- Share your metadata

A

Interoperable

- Use open/common format
- Consistent vocabulary
- Discipline specific metadata standards



Reusable

- Apply a licence to specify how others can re-use your data/code
- Documentation

R

https://data.4tu.nl/info/en/use/publish-cite/upload-your-data-in-our-data-repository/licencing

Reusable

- Apply a **licence**
- Documentation

Data:

Creative Commons (Overview)

Creative Commons License Chooser

Software:

Choose a License

tl;dr Legal

R

Learn more:
[The Turing Way]

≠

Standard

Open

Intrinsic quality

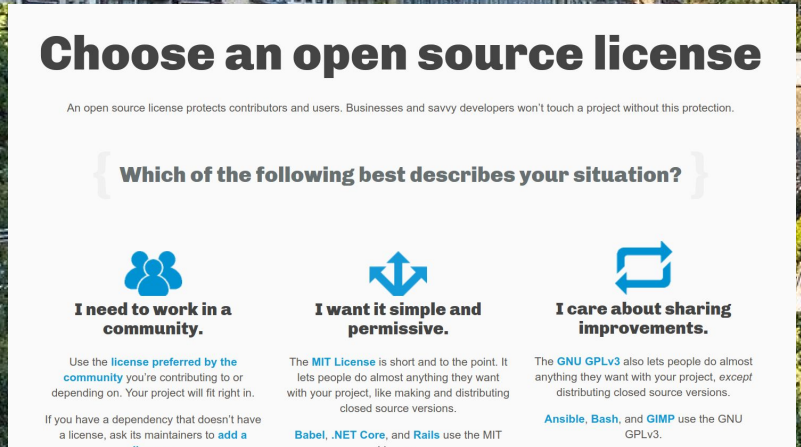
2

Pitches

1. Software licensing - Lourens Veen
2. Software citation - Jaro Camphuijsen
3. FAIR-Aware tool - Kim Ferguson

Software Licensing

tl;drLegal



Choose an open source license

An open source license protects contributors and users. Businesses and savvy developers won't touch a project without this protection.

Which of the following best describes your situation?

I need to work in a community.

Use the **license preferred by the community** you're contributing to or depending on. Your project will fit right in.

If you have a dependency that doesn't have a license, ask its maintainers to **add a**

I want it simple and permissive.

The **MIT License** is short and to the point. It lets people do almost anything they want with your project, like making and distributing closed source versions.

Babel, **.NET Core**, and **Rails** use the MIT

I care about sharing improvements.

The **GNU GPLv3** also lets people do almost anything they want with your project, *except* distributing closed source versions.

Ansible, **Bash**, and **GIMP** use the GNU GPLv3.

Software citation

How to supply **citation metadata** and how to get a DOI for your software

1. **Why** you should make your software citable
2. Using the **CITATION.cff** file
3. Automatically archiving software releases using **GitHub** and **Zenodo**

eScience center

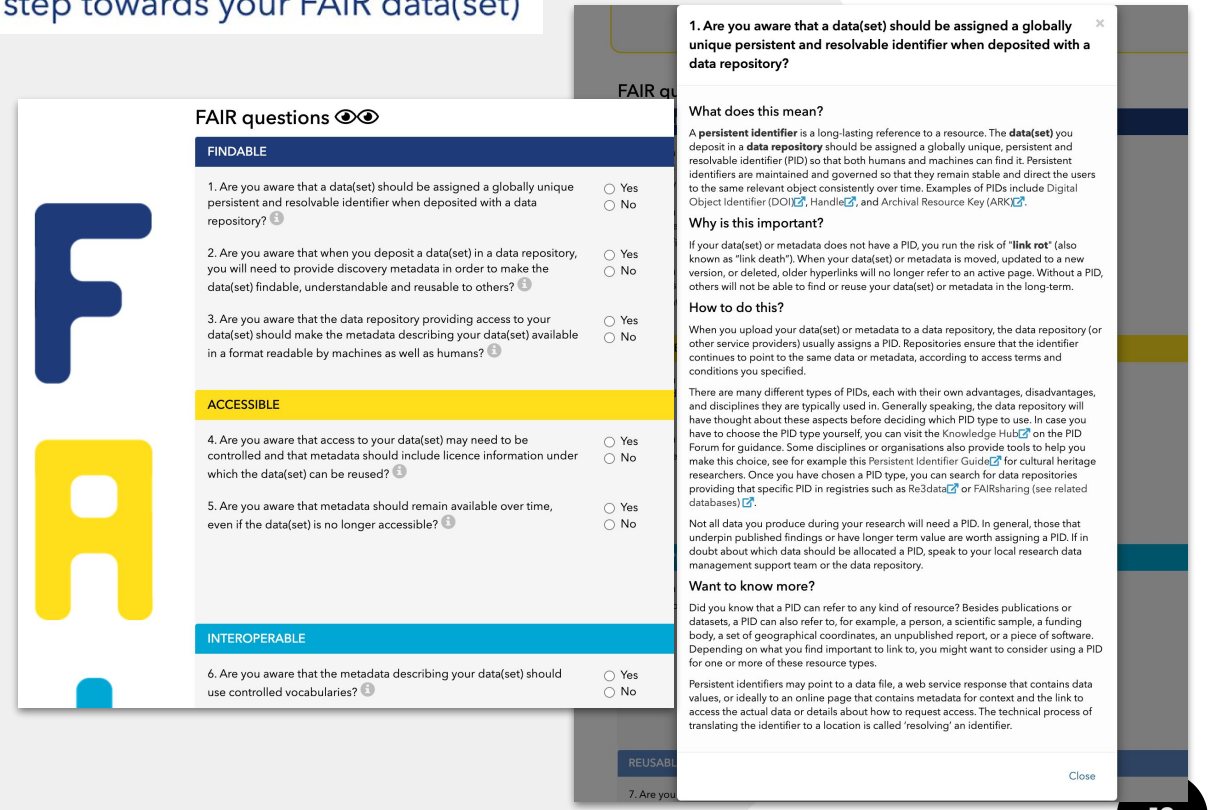
For those who can't attend:
`https://fair-software.nl`

FAIR Aware

Your first step towards your FAIR data(set)

- Your first step towards your FAIR data(set).
- What FAIR practices will help you create more FAIR data?
- 10 questions + tips and tricks
- **In this workshop:** working with biological data and elements to keep in mind

fairaware.dans.knaw.nl



3

Break Out Sessions

1. Software licensing - Lourens Veen
2. Software citation - Jaro Camphuijsen
3. FAIR-Aware tool - Kim Ferguson

4

Wrap up

See also the
LCRDM website

TU Delft

RDM support & Data Stewards & DCC
Esther (e.plomp@tudelft.nl), Yasemin (y.turkyilmaz-vandervelden@tudelft.nl)

Vrije Universiteit Amsterdam

Research Data Services
researchdataservices.ub@vu.nl

VU University Medical Center

Data Management
datamanagement@vumc.nl
(Locatie AMC: Rudy Scholte)

University Medical Center Utrecht

Data Steward
Irina van Dijk
(I.C.vanDijk-9@umcutrecht.nl)

Utrecht University

RDM support
info.rdm@uu.nl

Leiden University

Centre for Digital Scholarship
cds@library.leidenuniv.nl

Wageningen University & Research

Data Management Support
data@wur.nl

University of Amsterdam

RDM support
rdm-support@uva.nl

University of Groningen

RDM Support
Researchdata@rug.nl

Radboud University Nijmegen

RDM Support
rdmsupport@ubn.ru.nl

Radboud University Medical Center

Data Stewardship
datastewardship.im@radboudumc.nl

Thanks to:

- Presentation template by SlidesCarnival (Cymbeline)
- Naomi Chrispijn & Renée Calon