Data Archiving and Networked Services

DANS

FAIR Aware

a tool to aid you in your open data adventure

Kim Ferguson (RDM Specialist)
w/ materials from: Maaike Verburg (RDM Specialist)
Cees Hof (Data Station Manager Life, Health and Medical Sciences)

October 10, 2022
NWO BioPhysics Workshop, “Plan ahead: practical tools to make your data and software more FAIR”

Who am I?

Kim Ferguson

→ Research Data Management Specialist at DANS-KNAW

DANS - Data Archiving & Networked Services

Database: EASY (Data Stations soon)

→ PhD in insect genetics, postdoc in metabarcoding

→ Open Science advocate

depositing data sets,

publishing preprints,

open access publishing, and

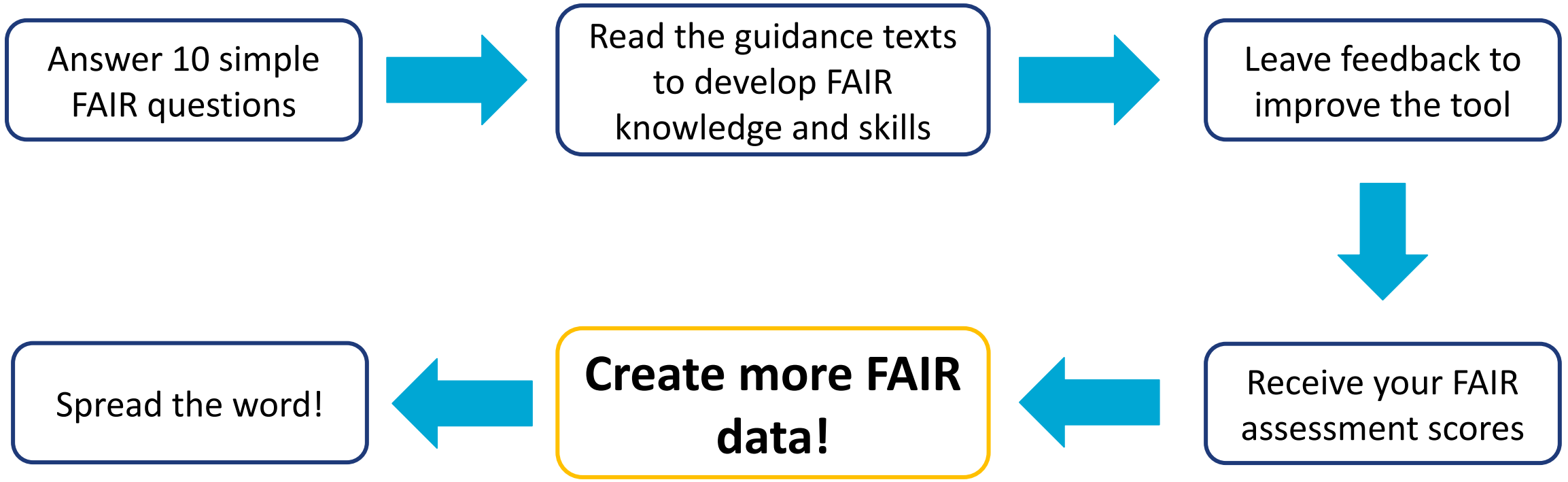
sharing posters & presentations

Plan for today

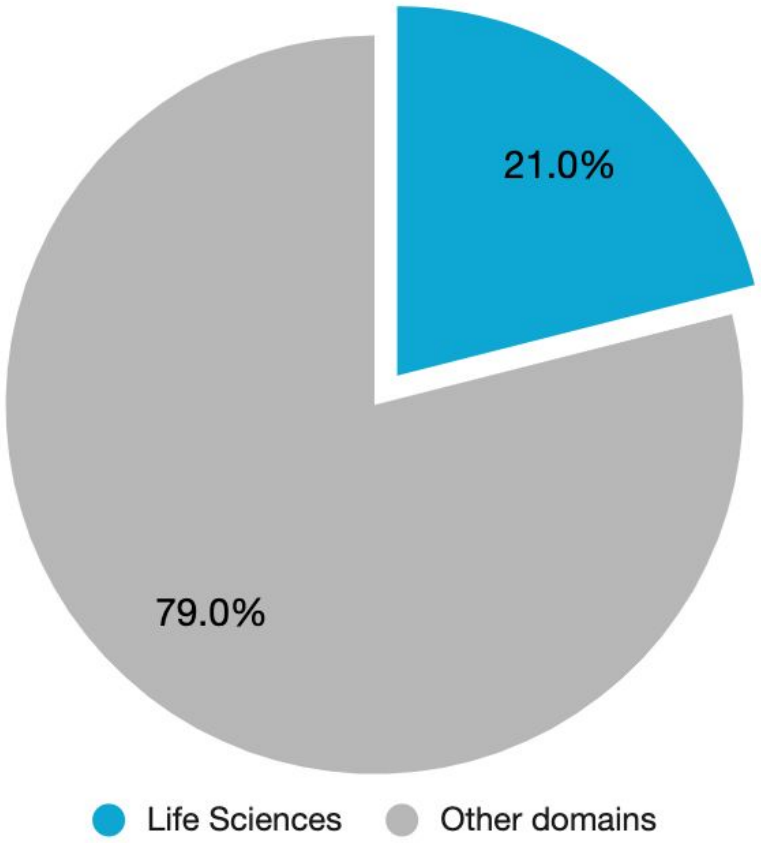
Time	Element
3 min	Introduction
15 min	FAIR-Aware and your data
5 min	Data repositories (re3Data)
4 min	Wrap-up & questions

FAIR Aware

Your first step towards your FAIR data(set)



Results so far



Over **750** users with a **76%** usability rating!

"The FAIR principles were not familiar to me at the time when I was trying to get published. The FAIR-Aware tool is informative and easy to use as a point of reference for my further publication career"

Researcher from the life science / medicine domain

So what does it do?

FAIR Aware

Your first step towards your FAIR data(set)

→ Let's take a look at FAIR Aware

Note 1: The tool works best when you're about to deposit a data set, but can be useful to think about what you need to consider or when you're writing a data management plan

Note 2: A lot of the elements can be taken care of by a trustworthy repository, which we will look at shortly after

Data repositories: good for finding data, good for storing data

What is archiving? What about publishing a dataset?

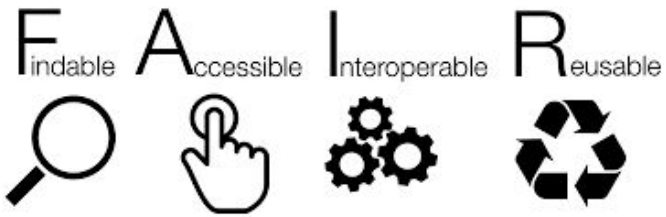
Internal shared drives are useful for colleagues, but aren’t the same as publishing a dataset in a repository.

Repositories & FAIR

A good repository:

- Provides a Persistent Identifier (PID) to your dataset = **Findable**
- Enables the inclusion of rich metadata = **Reusable**
- Allows for human and machine findable and readable metadata = **Findable + Accessible**
- Uses Knowledge Organisation Systems (KOS), such as vocabularies, to improve the quality of metadata = **Reusable**
- Provides long term and easy access to your data = **Accessible + Reusable**
- …….

A good repository means FAIR data



Where to deposit data, **or find data for re-use**

Depositing sequence data:

- **ENA** <> GenBank <> DDBJ (share sequences, trustworthy, all use MIBBI)
- Organism domain or species/family-specific databases? **In addition**, never only

Depositing everything else, or finding (non-sequence) data for re-use?

- Domain-specific repositories likely have more applicable metadata or keywords
- Some institutions or universities have their own repositories, making it easier for you to deposit
- National repositories like DANS-EASY are also good options, but you need to bring in your supervisor

So what else is out there? Browse-by-subject in **Re3Data** is a great starting point.

A convenient tool/site: re3data.org

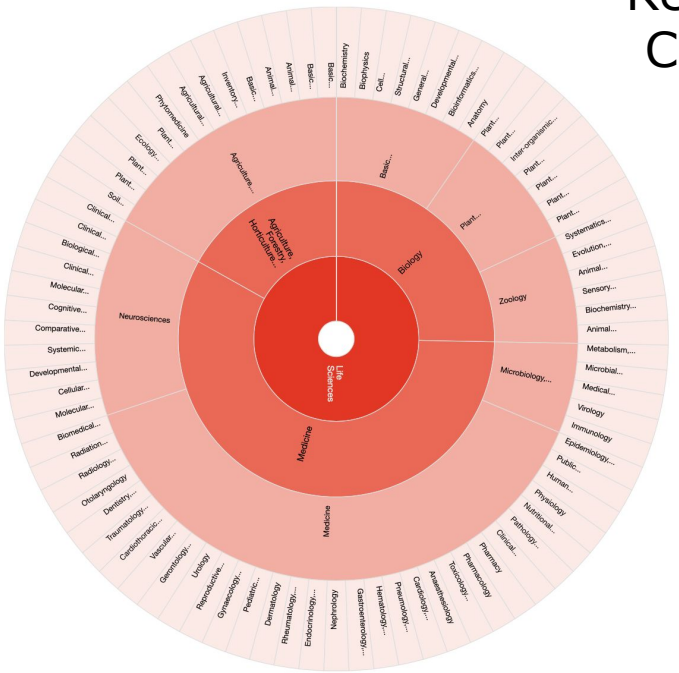
https://www.re3data.org

Over 2600 international repositories for a wide spread of domains

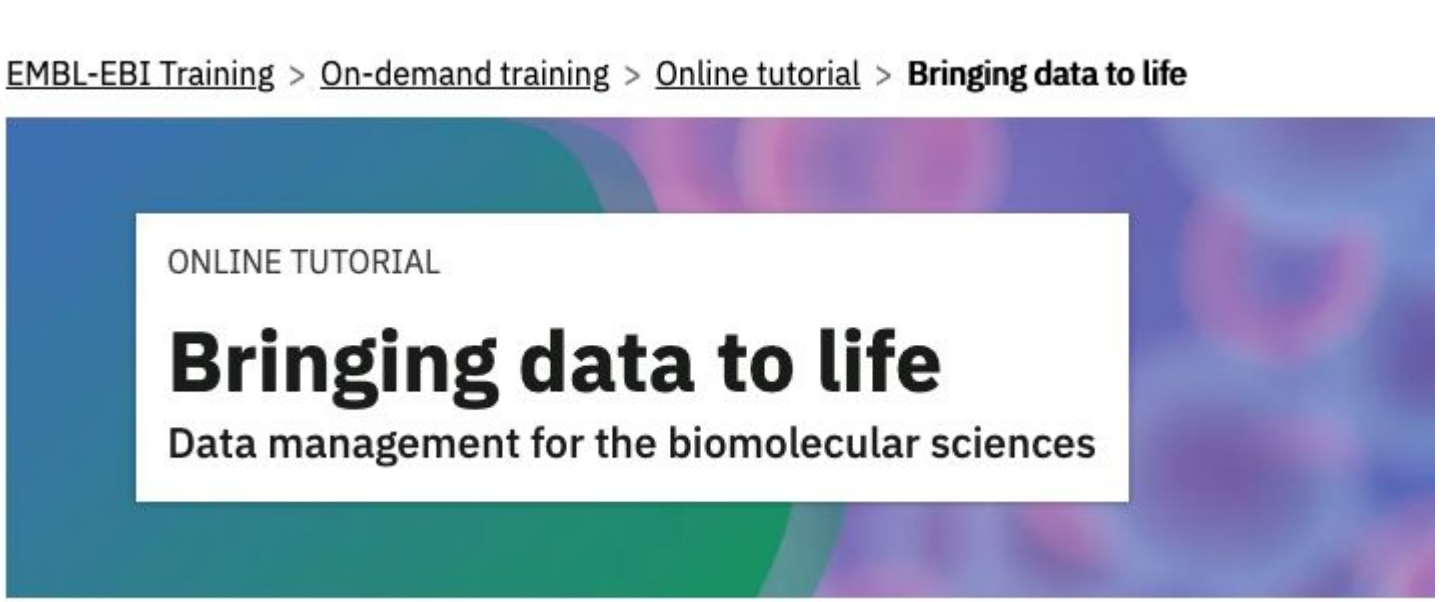
Plentiful filters to find appropriate repositories

All Categories

Life Science Repositories Categories



But how much can we cover in 25 minutes?



This online tutorial from EMBL-EBI is self-paced and can provide a great structure for learning data management, but you don't necessarily need it.
.: recommended for those who want to know more.

References/Further Reading

- FAIR Aware: https://fairaware.dans.knaw.nl/
- Re3Data: https://www.re3data.org/
- All emojis designed by OpenMoji – the open-source emoji and icon project. License: CC BY-SA 4.0

Additional resources

- BioSB - Netherlands Bioinformatics and Systems Biology research school: https://www.dtls.nl/biosb/
- Additional courses/workshops:
  - EMBL-EBI Course on Research Data Management in Bioinformatics: https://www.ebi.ac.uk/training/online/courses/bringing-data-to-life-data-management/
  - DTL page for training: https://www.dtls.nl/training-and-education/
  - eScience Center for training: https://www.esciencecenter.nl/digital-skills/

Thanks for your attention!

kim.ferguson@dans.knaw.nl

Twitter: @DANS_knaw_nwo

dans.knaw.nl/en

FAIR Aware slides: Maaike Verburg
Re3Data slides: Cees Hof