TIDY, SAFE AND SHARABLE DATA

Maggie Hellström
ICOS Carbon Portal

ICOS Cities webinar, May 11, 2022

Let’s start with some poll questions!

Feel free to introduce yourself in the chat, adding a sentence of what aspects of research data management interests you!

ICOS INTEGRATED CARBON OBSERVATION SYSTEM

Surfing the wave…

Image:unknown

- In 2021, approximately 80 zettabytes of data were created – that's almost 10^{22} bytes
- Access to "new" data gives fantastic opportunities for science
  - Ask – and answer – new scientific questions!
  - Applications of AI and machine learning
  - Address "grand challenges" like the Covid crisis
- Data is also expected to be a driver for the "global digital economy"
- Much resources are being spent on data infrastructure, e.g. for the EOSC

ICOS INTEGRATED CARBON OBSERVATION SYSTEM

Your data is everyone's data?!

Image:imgflip.com

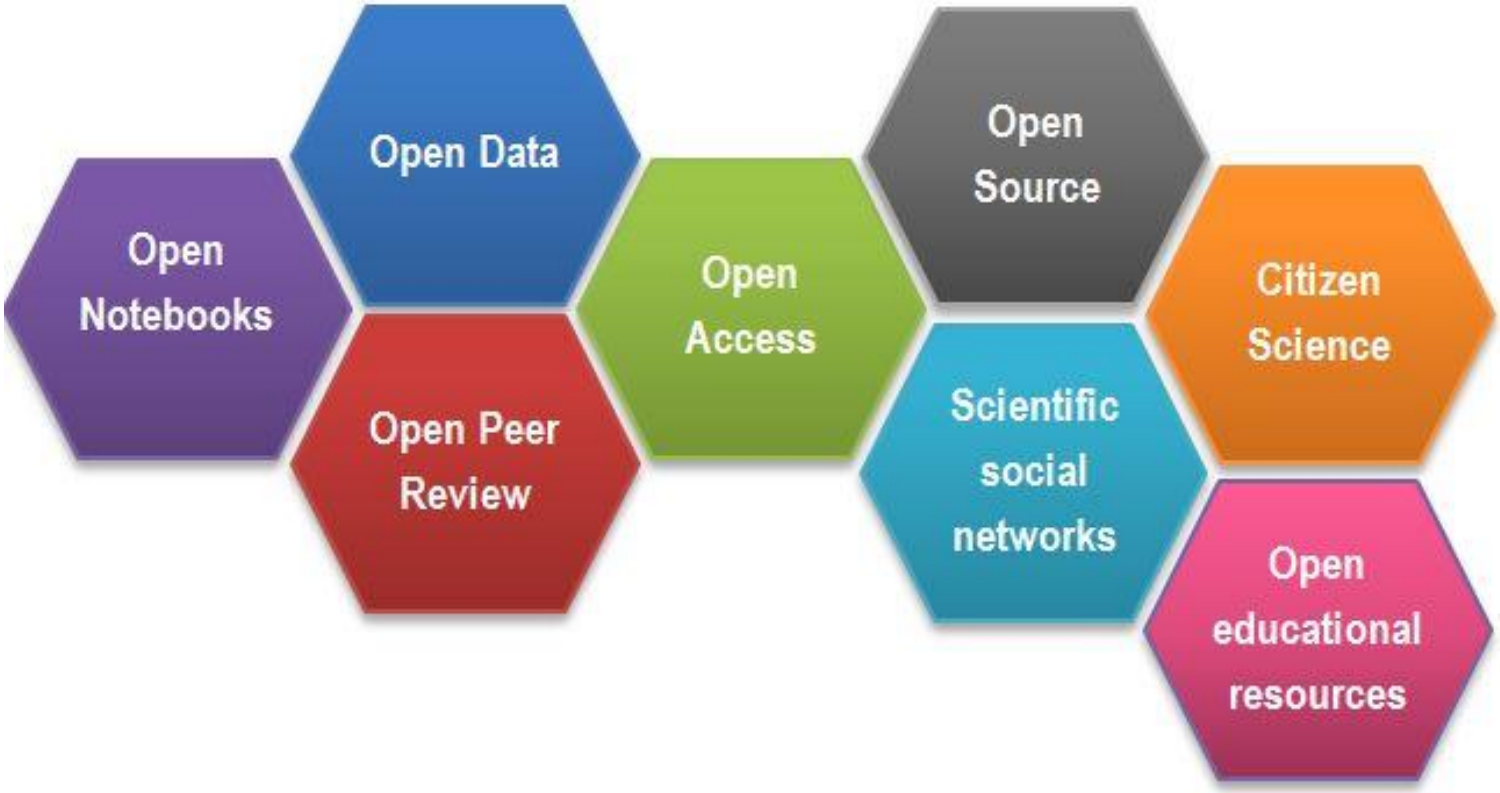
Pressure from funders and stakeholders:

Public-funded data should/must be

- easy to find,
- well documented,
- follow standards,
- simple to reuse,
- free from unnecessary restrictions & conditions
- "as open as possible, as closed as necessary"

ICOS INTEGRATED CARBON OBSERVATION SYSTEM

Open Science: the new paradigm



https://www.fosteropenscience.eu/content/what-open-science-introduction

Sharing is caring...

...but not always so easy in practice!

[link to video](#)

Data Sharing and Management Snafu in 3 Short Acts (Higher Quality)

ICOS INTEGRATED CARBON OBSERVATION SYSTEM

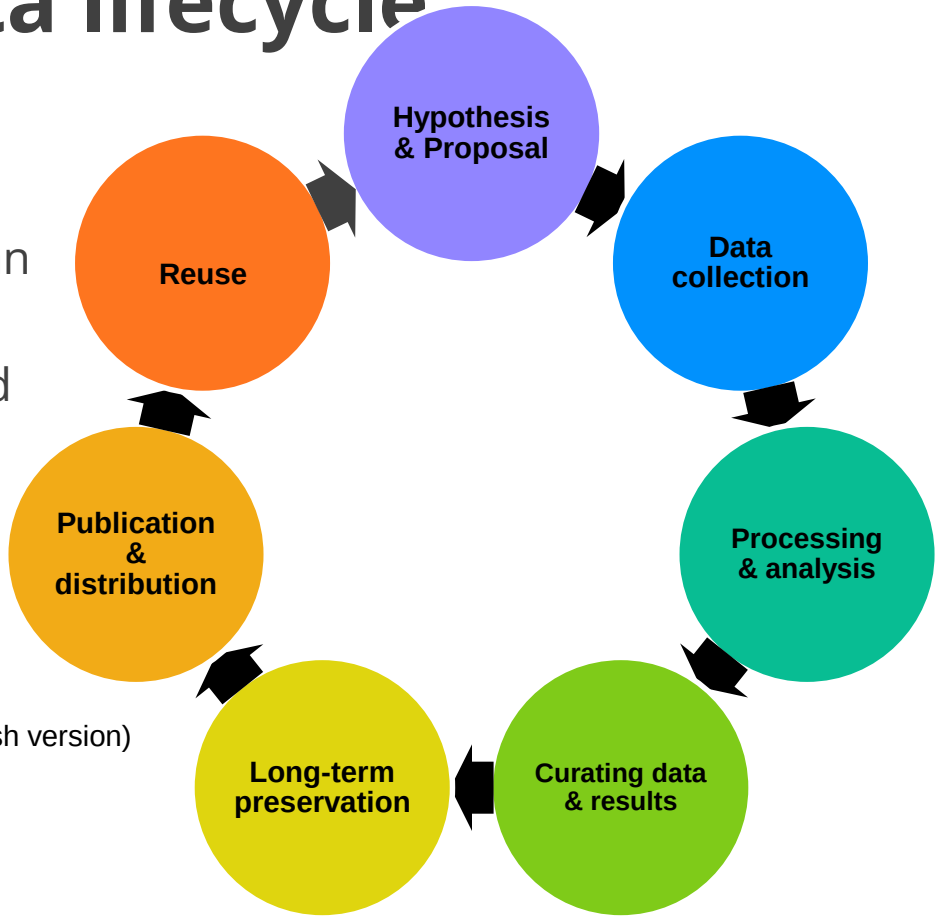
How to cope (and improve)

- Take a deep breath (and put the kettle on)
- Find out more about RDM, FAIR, OS, …
- Talk to your colleagues & analyse current practices
  - documentation of the research process?
  - data storage during a project?
  - archiving afterwards?
  - publishing data with papers?
- Use a data management plan (DMP) tool
- Ask for help, if needed!

ICOS INTEGRATED CARBON OBSERVATION SYSTEM

The research data lifecycle

- All steps not always included
- Starting point – and order - can vary
- Some iteration to be expected
- More spiral than circle?



Avoin tiede (2014): Open Science Handbook (English version)

ICOS INTEGRATED CARBON OBSERVATION SYSTEM

The FAIR principles

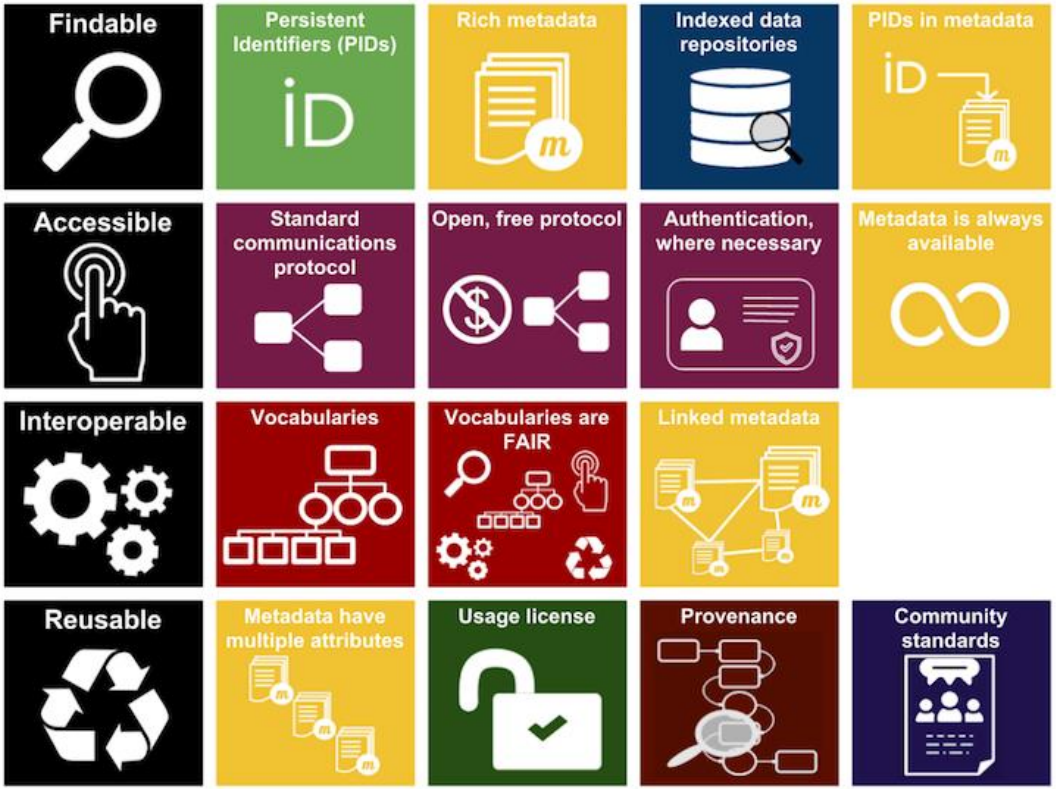
- Findable, Accessible, Interoperable, Reusable
- a set of guiding principles that
  - encourage comprehensive & sustainable data management
  - support & facilitate computer-based data processing
- relevant for “everyone”: data producers, data stewards and data users
- FAIR is not equal to Open
- FAIR is no guarantee for high quality or “fitness for use”

https://force11.org/info/the-fair-data-principles/

ICOS INTEGRATED CARBON OBSERVATION SYSTEM

Components of FAIR

Read more at and
https://www.go-fair.org/



ICOS INTEGRATED CARBON OBSERVATION SYSTEM

What's in it for me?

Making your data "FAIR enough" gives you better control of what happens to your data, and

- helps make your data sustainable, also for you & your co-workers!
- ensures your data can be found – and used - by others
- facilitates collection & storage of metadata
- guarantees data can be cited, giving you credit
- facilitates extraction of data usage statistics
- simplifies reporting to funders & streamlines cost estimates
- makes data management plans easier to write

ICOS INTEGRATED CARBON OBSERVATION SYSTEM

Making your data FAIR(er)

- Make a **plan** for the data **before you start** a project!
- Design a **logical folder structure**, and keep to it!
- Collect detailed descriptive information (= **metadata**) throughout
- **Document all steps** of the data processing & analysis
- Use **standards** and **formats common** to your discipline
- Keep **at least 2 separate backups** of data, metadata and code

ICOS INTEGRATED CARBON OBSERVATION SYSTEM

Making your data FAIR(er), cont'd

- Store the data in a **trusted & sustainable** repository or data center
- See to that the data gets **persistent identifiers** (DOIs or similar)
- Apply a suitable **usage license**
- Provide end users with information on "**intended use**"
- Make the data **"as open as possible, as closed as necessary"**
- Ensure that **metadata remain available** even if the data cannot be accessed any more
- **Test the FAIRness** of your data using some online tool

ICOS INTEGRATED CARBON OBSERVATION SYSTEM

Some useful links

- DMPonline data management plan tool: https://dmponline.dcc.ac.uk/
- ORCiD identifiers for people: https://orcid.org/
- Re3data catalogue of data repositories: https://www.re3data.org/
- EUDAT license selector tool: https://ufal.github.io/public-license-selector/
- DataCite's metadata schema: https://schema.datacite.org/
- RDA metadata standards catalog: https://rdamsc.bath.ac.uk/
- The FAIR principles: https://www.go-fair.org/
- FAIRness assessment tools: https://satifyd.dans.knaw.nl/ (for humans) & https://www.f-uji.net/ (for machines)
- ICOS Carbon Portal: https://icos-ri.eu/

ICOS INTEGRATED CARBON OBSERVATION SYSTEM

Let’s wrap up with more poll questions!

If you want, please add your “take home message(s)” from today’s webinar in the chat!

ICOS INTEGRATED CARBON OBSERVATION SYSTEM

Thanks for listening!

You’re welcome to get in touch with Maggie with your data management questions via margareta.hellstrom@nateko.lu.se

ICOS INTEGRATED CARBON OBSERVATION SYSTEM