

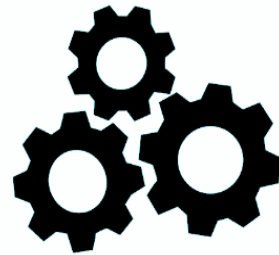


Example of FAIRification of data from AnaEE Denmark

Klaus Steenberg Larsen
Associate Professor, AnaEE Denmark Coordinator
Department of Geosciences and Natural Resource Management
University of Copenhagen

Understanding FAIRification - What is FAIR data?

Findable **A**ccessible **I**nteroperable **R**eusable



FAIR \neq Open but Open \implies FAIR
As open as possible, as closed as necessary

AnaEE has a decentralized data strategy =
you are in charge of your data



Upholding the FAIR Principles

To be Findable:

- F1. (meta)data are assigned a globally unique and persistent identifier
- F2. data are described with rich metadata (defined by R1 below)
- F3. metadata clearly and explicitly include the identifier of the data it describes
- F4. (meta)data are registered or indexed in a searchable resource

To be Accessible:

- A1. (meta)data are retrievable by their identifier using a standardized communications protocol
 - A1.1 the protocol is open, free, and universally implementable
 - A1.2 the protocol allows for an authentication and authorization procedure, where necessary
- A2. metadata are accessible, even when the data are no longer available

To be Interoperable:

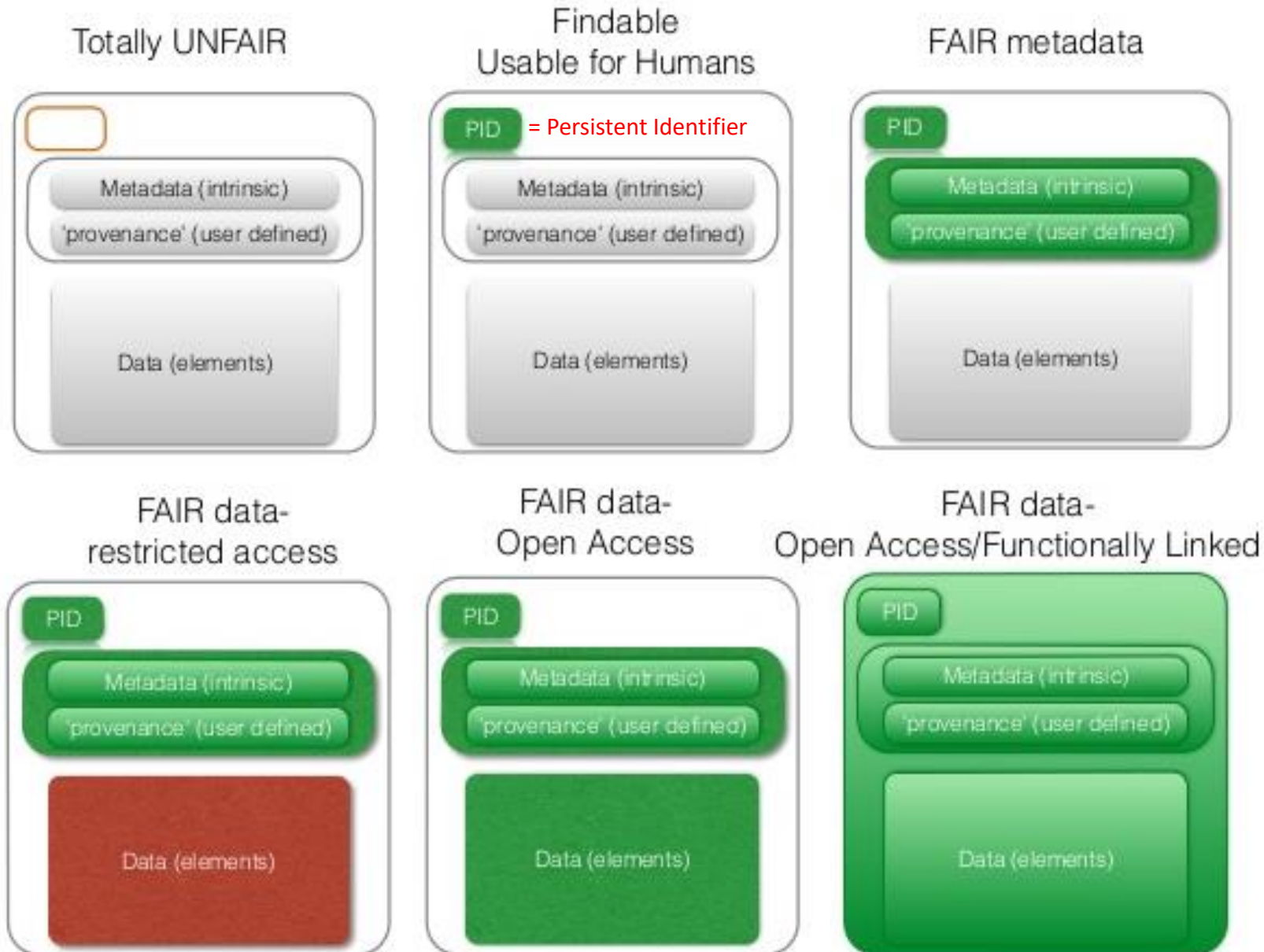
- I1. (meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.
- I2. (meta)data use vocabularies that follow FAIR principles
- I3. (meta)data include qualified references to other (meta)data

To be Reusable:

- R1. meta(data) are richly described with a plurality of accurate and relevant attributes
 - R1.1. (meta)data are released with a clear and accessible data usage license
 - R1.2. (meta)data are associated with detailed provenance
 - R1.3. (meta)data meet domain-relevant community standards



Gradually Increasing the Levels of FAIRness



Why make FAIR data?

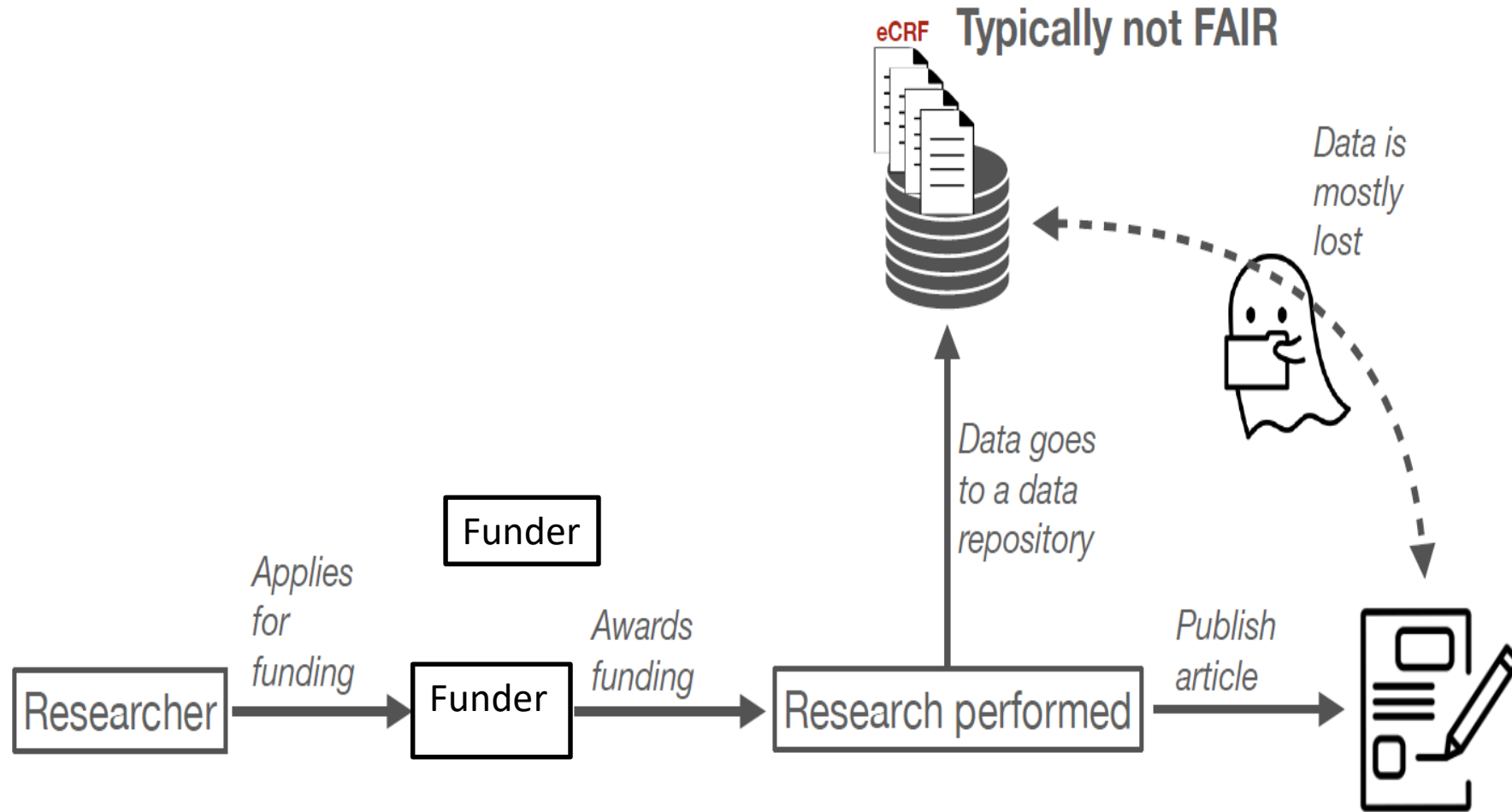
- Obligation to funders
 - + National ministries and funding bodies
 - + European commission
- AnaEE criteria (ENVRIFAIR will set the recommendations)

OR

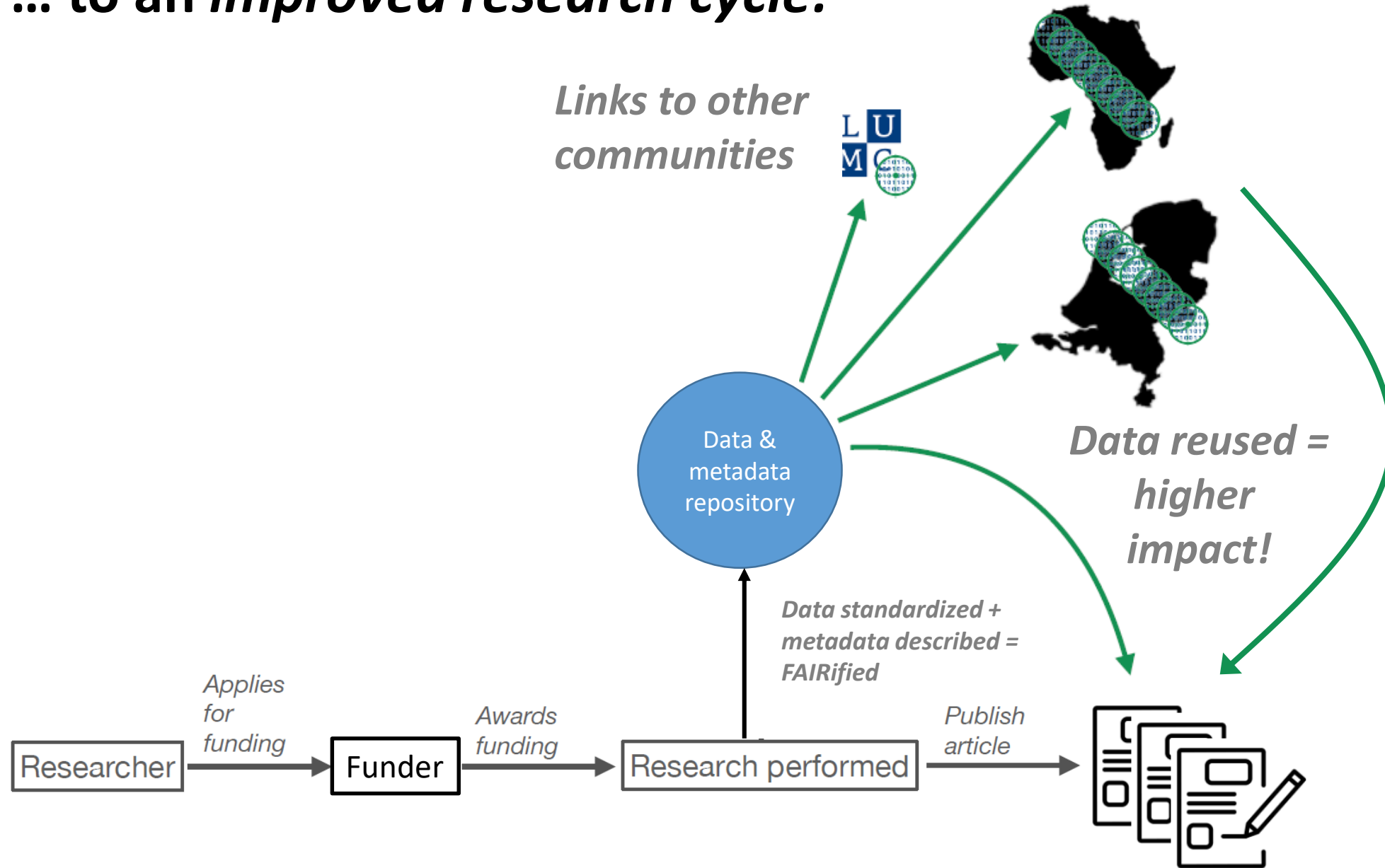
Better and more efficient science!!!



Better and more efficient science – moving from a *typical research cycle*:



... to an *improved research cycle*:



Links to other communities



Data reused = higher impact!

Data standardized + metadata described = FAIRified

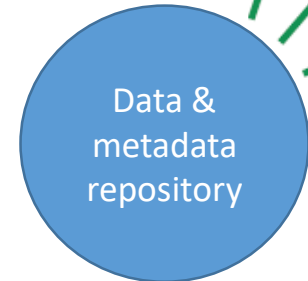


... to an *improved* research cycle:

Goal: Reuse and INCREASE scientific usage of AnaEE Denmark generated data by creating FAIR data!

Requirements: generation of standard data and metadata templates, change in culture (clarify benefits for individual researchers)

Links to other communities



Data reused = higher impact!

Data standardized + metadata described = FAIRified



AnaEE Denmark FAIR data projects

- Co-funded by DeiC (Danish e-infrastructure cooperation)
- Project 1: 10 MM in total in 2019
- Project 2: 5 x 2-3 days workshops in June-October 2020 with DeiC + GOFAIR. No MM but funded workshops

**Strategy: Lowest hanging fruits first:
Do what you can – skip what you can't**

Main outcomes

- A lot of learning!
- 3 “FAIRified” meteorological datasets
- A FAIRification roadmap for AnaEE Denmark
- First machine-readable metadata near-ready
- FAIRified GHG data near-ready



The steps:

- Find a persistent data repository (ERDA at UCPH)
- Create templates for data and meta data
- Get a PID (in our case DOI – ERDA can provide)
- Create machine-readable metadata

Link to data:

anaee.dk/access/



Machine-readable metadata: Openly available tools:

BioPortal – a catalogue of ontologies from 1000 different communities

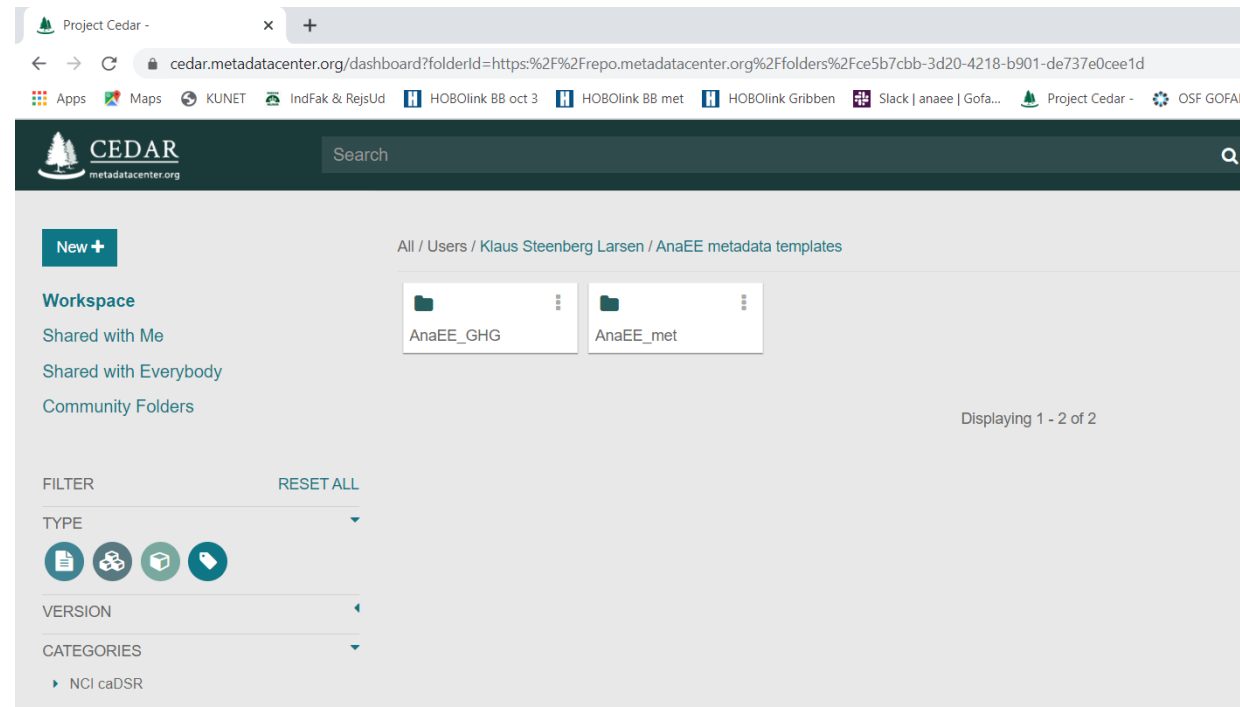
<https://bioportal.bioontology.org/>

CEDAR:

cedar.metadatacenter.org

CEDAR can help you
map your data to ontologies

= create machine-readable
metadata code –
without doing code!



The screenshot shows the CEDAR metadata center dashboard. The browser address bar displays the URL: cedar.metadatacenter.org/dashboard?folderId=https%2F%2Frepo.metadatacenter.org%2Ffolders%2Ffce5b7cbb-3d20-4218-b901-de737e0cee1d. The dashboard header includes the CEDAR logo and a search bar. The main content area shows a workspace titled "All / Users / Klaus Steenberg Larsen / AnaEE metadata templates". There are two folders visible: "AnaEE_GHG" and "AnaEE_met". A filter sidebar on the left includes options for "TYPE", "VERSION", and "CATEGORIES", with "NCI caDSR" selected under categories. The text "Displaying 1 - 2 of 2" is visible at the bottom right of the folder list.



FAIR data with machine-readable metadata

Project Cedar - Template Design x Public Archive: ad280f6ee8ad896 x +

/archives/ad280f6ee8ad89643c7e1f8d2909eb51/published-archive.html

IndFak & RejsUd HOBOLink BB oct 3 HOBOLink BB met HOBOLink Gribben Slack | anaee | Gofa... Project Cedar - OSF GOFAIR

Format: ASCII (csv), comma separated. Data variable names and data files were created following a structure similar to the ICOS project (<https://www.icos-ri.eu/>), with some additions when needed. The ICOS meta-data standards are described at <https://meta.icos-cp.eu/resources/cpmeta/atcMeteoTimeSer>. Table-driven variable codes (GRIB codes) are provided for each data variable according to WMO guidelines of variable codes (<https://public.wmo.int/en>). For more information on variable codes of WMO, see https://www.wmo.int/pages/prog/www/WMOcodes/WMO306_vI2/LatestVERSION/LatestVERSION.html.

Keywords: CLIMAITE, AnaEE Denmark, Brandbjerg, Denmark, climate change experiment, meteorological data, INCREASE EU project.

Data policy: This dataset is licensed under a Creative Commons Attribution 4.0 international license (<https://creativecommons.org/licenses/by/4.0/>).

Acknowledgements: The CLIMAITE experiment was originally made possible by a large donation from the Villum Foundation.

For other AnaEE Denmark datasets, see http://www.erda.dk/vgrid/AnaEE_DK/.

Archive DOI Data

Archive DOI

<https://doi.org/10.17894/ucph.e58a99c2-da7b-444a-b1c0-11f00e70041c>

DataCite Entry

Complete DOI meta data and citation info is available in the [DataCite DOI Collection](#).

DOI Details

Show/hide all DOI meta data registered with DataCite.

Archive Files

1 to 2 of 2 rows 25 files per page

Name	Date	Size
AnaEE_DK/Brandbjerg/BRANDBJERG 20051001 20131231 HOURLY.csv	2019-11-18 12:22:23	12347812
AnaEE_DK/Brandbjerg/BRANDBJERG REF METADATA 20051001 20131231 HOURLY.csv	2019-11-18 12:22:23	6586

LINK TO MACHINE-READABLE METADATA HERE!



A FAIRification roadmap for AnaEE Denmark - 2020

Content:

1. Strategic Rational for a FAIRification roadmap
2. Understanding FAIRification - What is FAIR data?
3. Automated Machines Access to FAIR data
4. Upholding the FAIR Principles
5. Gradually Increasing the Levels of FAIRness
6. A necessary move from *typical research cycle*
7. ... to *improved research cycle*
8. Reassessing Organization, Technology, Politics, Economy
9. Coordinating the AnaEE FAIRification roadmap
10. Aiming for community harmonization and specialization
11. Aiming for an Improved Division of Labor
12. Roadmap & shorter-term FAIRification Workplans
13. Roadmap & Technical Breakdown and To-do's
14. Local and National Data Stewardship Support
15. Formulating FAIRification Success Criteria
16. Summary of AnaEE Denmark FAIRification roadmap