FAIRSFAIR
Fostering Fair Data Practices in Europe

How to be FAIR with your data:
a teaching and training handbook for higher education institutions

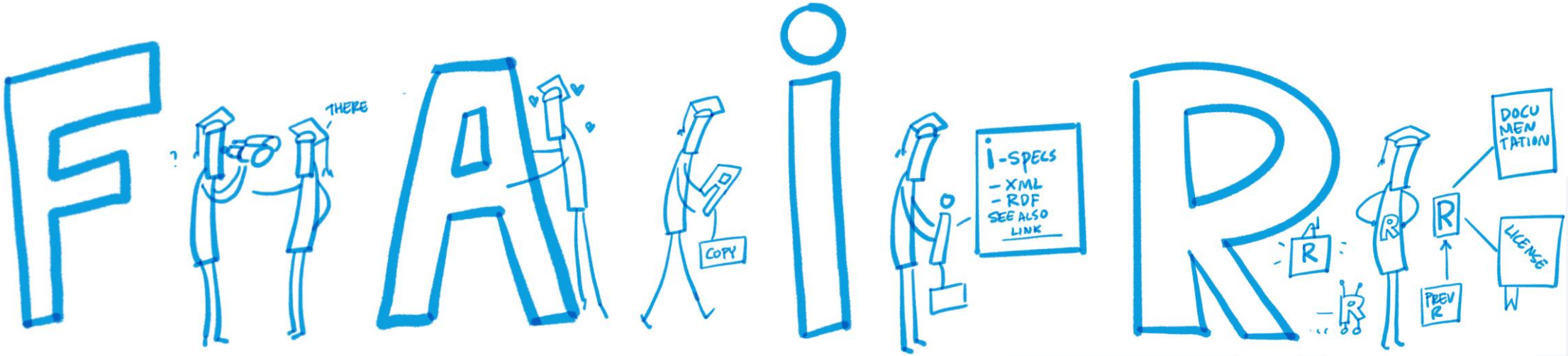
Claudia Engelhardt, Göttingen State and University Library

Supporting the uptake of FAIR data teaching and training: tools and reflections from the FAIRsFAIR project
15 February 2022, online

FAIRsFAIR "Fostering FAIR Data Practices in Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 Grant agreement 831558

Content

- Creation of the handbook
- Components
- Future



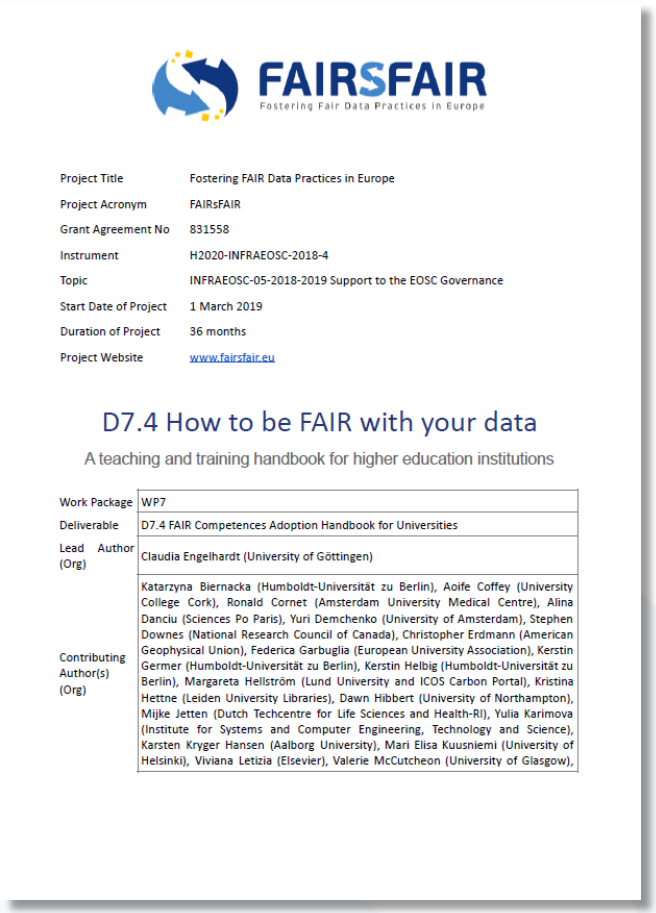
How to be FAIR with your research data – A teaching and training handbook for HEIs

Purpose:

- provide guidance and practical support with integrating the FAIR principles and related content into curricula and teaching

Published as a project deliverable/pdf on Zenodo in December 2021

- https://doi.org/10.5281/zenodo.5665492



FAIRSFAIR
Fostering Fair Data Practices in Europe

Project Title	Fostering FAIR Data Practices in Europe
Project Acronym	FAIRsFAIR
Grant Agreement No	831558
Instrument	H2020-INFRAEOSC-2018-4
Topic	INFRAEOSC-05-2018-2019 Support to the EOSC Governance
Start Date of Project	1 March 2019
Duration of Project	36 months
Project Website	www.fairsfair.eu

D7.4 How to be FAIR with your data

A teaching and training handbook for higher education institutions

Work Package	WP7
Deliverable	D7.4 FAIR Competences Adoption Handbook for Universities
Lead Author (Org)	Claudia Engelhardt (University of Göttingen)
Contributing Author(s) (Org)	Katarzyna Biernacka (Humboldt-Universität zu Berlin), Aoife Coffey (University College Cork), Ronald Cornet (Amsterdam University Medical Centre), Alina Danciu (Sciences Po Paris), Yuri Demchenko (University of Amsterdam), Stephen Downes (National Research Council of Canada), Christopher Erdmann (American Geophysical Union), Federica Garbuglia (European University Association), Kerstin Germer (Humboldt-Universität zu Berlin), Kerstin Helbig (Humboldt-Universität zu Berlin), Margareta Hellström (Lund University and ICOS Carbon Portal), Kristina Hettne (Leiden University Libraries), Dawn Hibbert (University of Northampton), Mijke Jetten (Dutch Techcentre for Life Sciences and Health-RI), Yulia Karimova (Institute for Systems and Computer Engineering, Technology and Science), Karsten Kryger Hansen (Aalborg University), Mari Elisa Kuusniemi (University of Helsinki), Viviana Letizia (Elsevier), Valerie McCutcheon (University of Glasgow),

Collaborative creation involving the community

- Six 3-hour book sprint sessions in June 2021
- 40 collaborators (incl. FsF people)
- Experts with diverse disciplinary backgrounds
- Subsequent editorial process led by Editorial Team

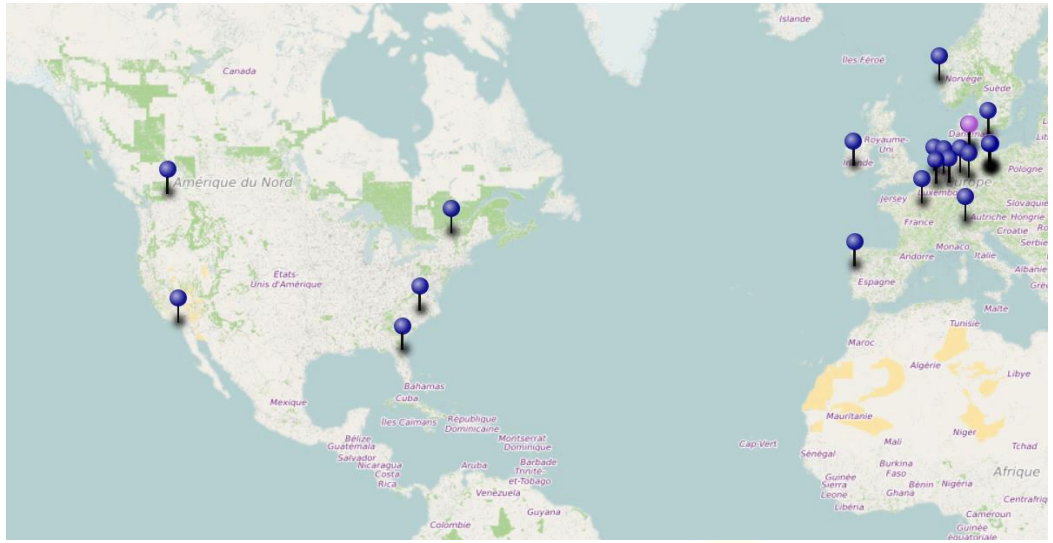


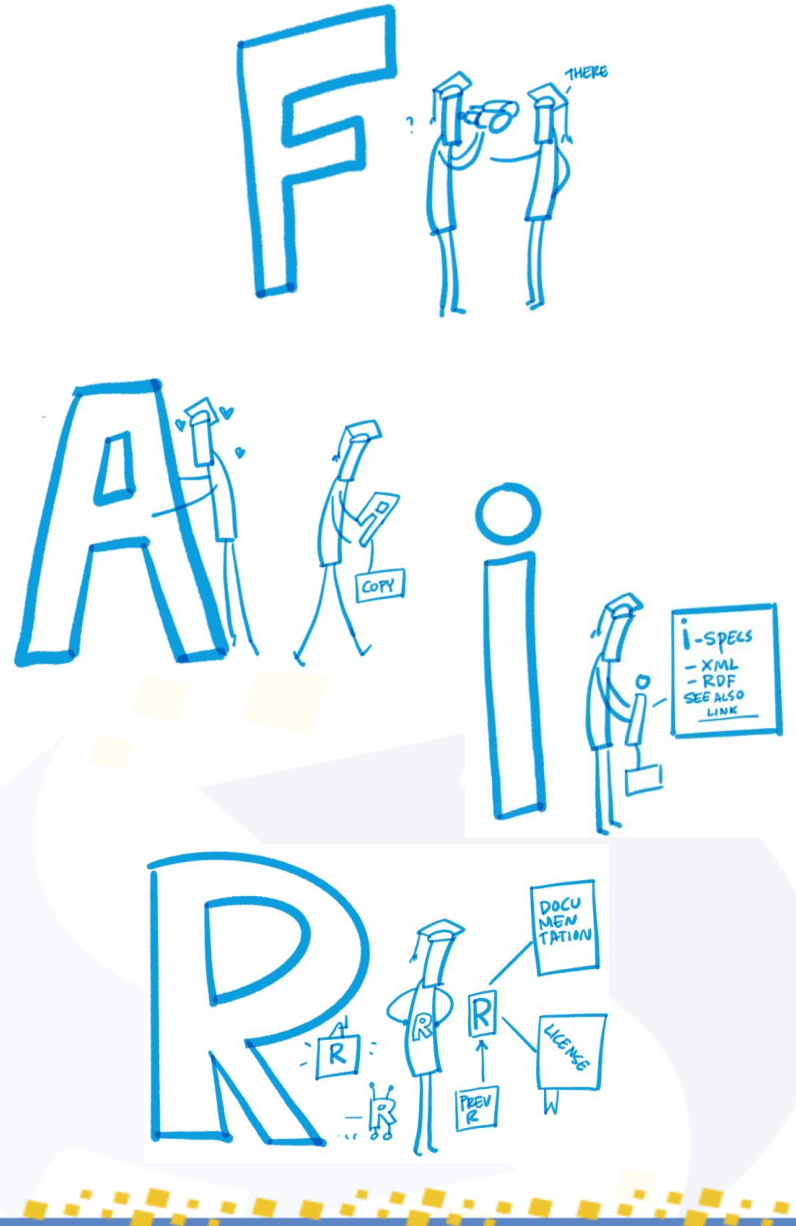
Image from OpenStreetMap

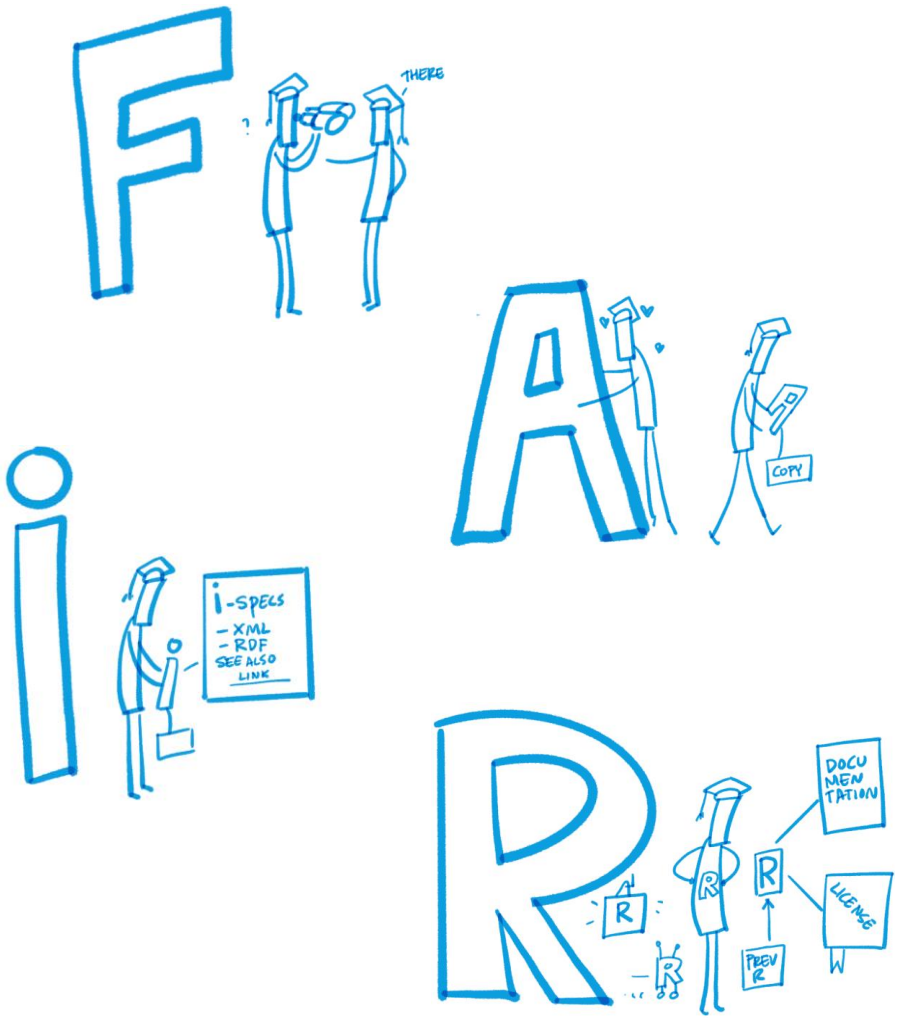
- Public consultation in August & September 2021
- Revised draft presented at workshop in October 2021 (https://tinyurl.com/543vf4hu)
- Project deliverable published in December 2021
- GitBook end of February, print (and OA PDF) edition in spring 2022

Primary target audience(s)

HEI staff members who

- Create and teach lessons and courses,
  e.g. lecturers, professors, trainers
- Design, adapt and implement curricula,
  e.g. doctoral programme managers, deans
- Train and consult PhD students and early career researchers ,
  e.g. support staff, lecturers, trainers
- Implement FAIR in institutional strategies, policies, administrative workflows etc.
  e.g. vice rectors/presidents, offices of research



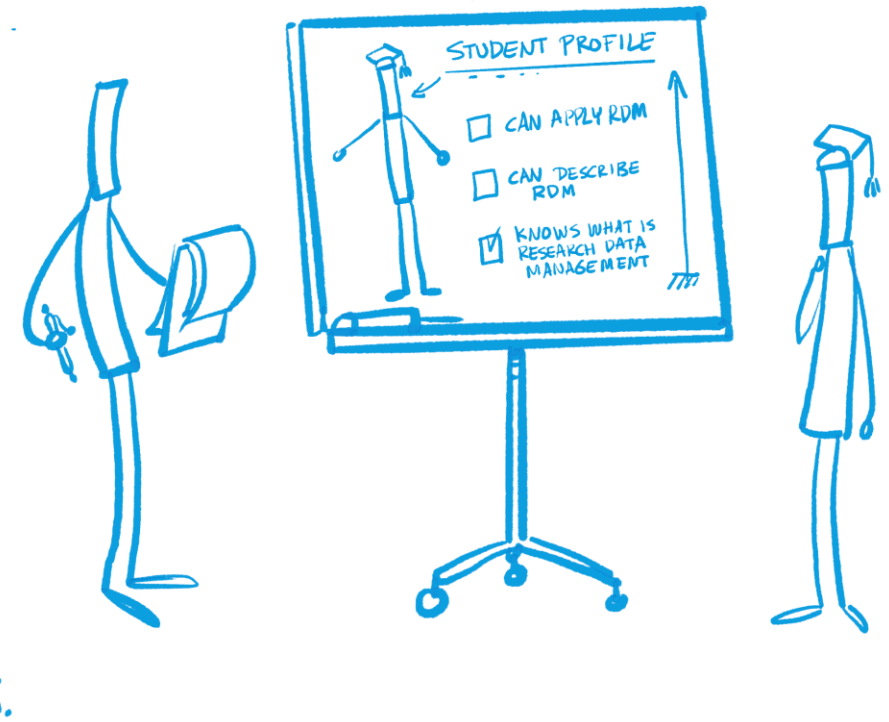


Contents – overview

FAIR Skills & Competences

- FAIR competence profiles (bachelor, master and doctoral level)
- Corresponding learning outcomes (bachelor, master and doctoral level)

- Based on "FAIR Competence Framework for Higher Education“ (https://doi.org/10.5281/zenodo.5361917) & experience and expertise of authors



Competence profiles

Table 2: Competence profiles for the bachelor, master and doctoral level

Topic	**Bachelor** (required level)	**Master** (required level)	**PhD** (required level)	**Entry-level content?**
General principles and concepts in data management – overview	basic	intermediate	advanced	yes
Overview of data types, data type registries and data formats	basic	basic	intermediate	yes
Metadata, metadata formats, standards and registries	basic	intermediate	advanced	yes
Open Research, Open Access, Open Data	basic	intermediate	advanced	yes
Metadata management, registries and publication	basic	basic	intermediate	no
Persistent Identifiers (PID), Open Researcher and Contributor ID (ORCID), Research Organization Registry (ROR)	basic	basic	intermediate	yes
FAIR (Findable, Accessible, Interoperable, Reusable) principles in	basic	basic	intermediate	yes

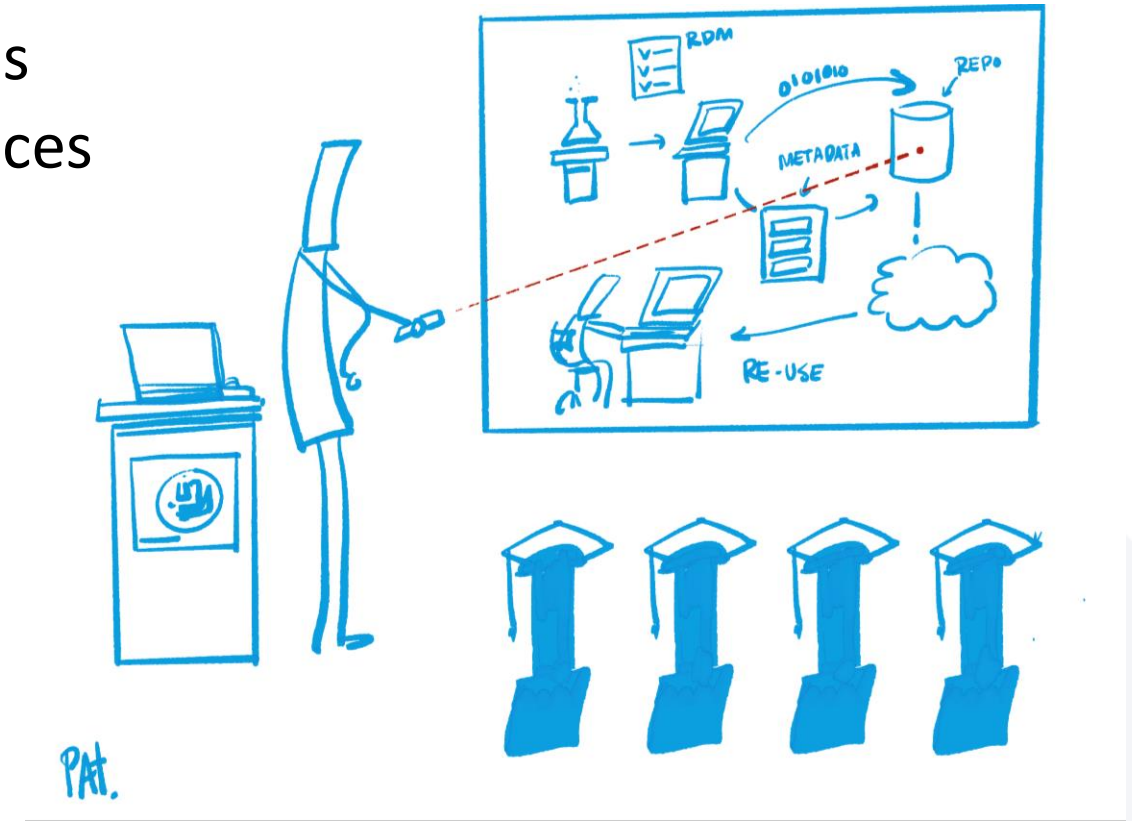
Learning outcomes

Table 5: Entry-level content including learning outcomes – doctoral level

Topic	**Required level**	**Learning outcomes** [b]=basic, [i]=intermediate, [a]=advanced]
General principles and concepts in data management – overview	advanced	- [b] Can define Research Data Management (RDM) and can describe its relevance and benefits. - [i] Can describe RDM measures to be taken (including explaining why) at different stages of the research process. - [a] Can practically apply theoretical knowledge about proper RDM measures to be taken at different stages to their own research process/project.
Overview of data types, data type registries and data formats	inter-mediate	- [b] Can describe what types of data exist (Knowledge). - [b] Can explain what data type registries are (Knowledge). - [b] Can identify data formats (Knowledge). - [i] Can determine proper data types for a resource (Analyse). - [i] Can use a data type registry (Apply). - [i] Can use proper data formats to express resources (Apply).
Metadata, metadata formats, standards and registries	advanced	- [b] Can describe types of metadata. - [b] Can recognise metadata formats.

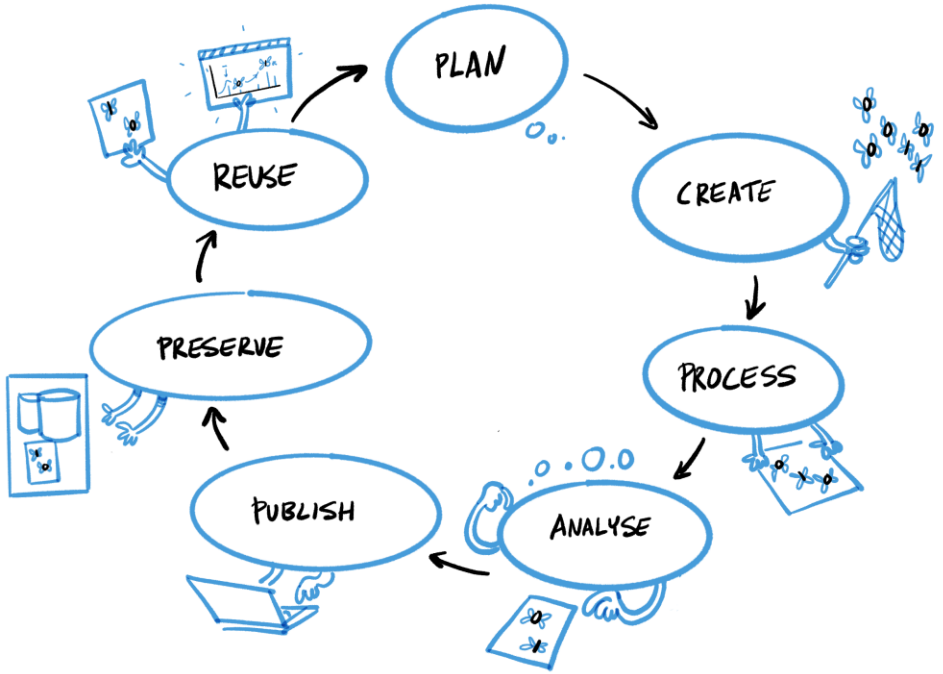
Teaching and training designs

- Phases of course design
  - Select/Identify Learning Outcomes
  - Select/Develop Learning Experiences
  - Select content
  - Identify/develop assessments
  - Evaluate course effectiveness



Lesson plans

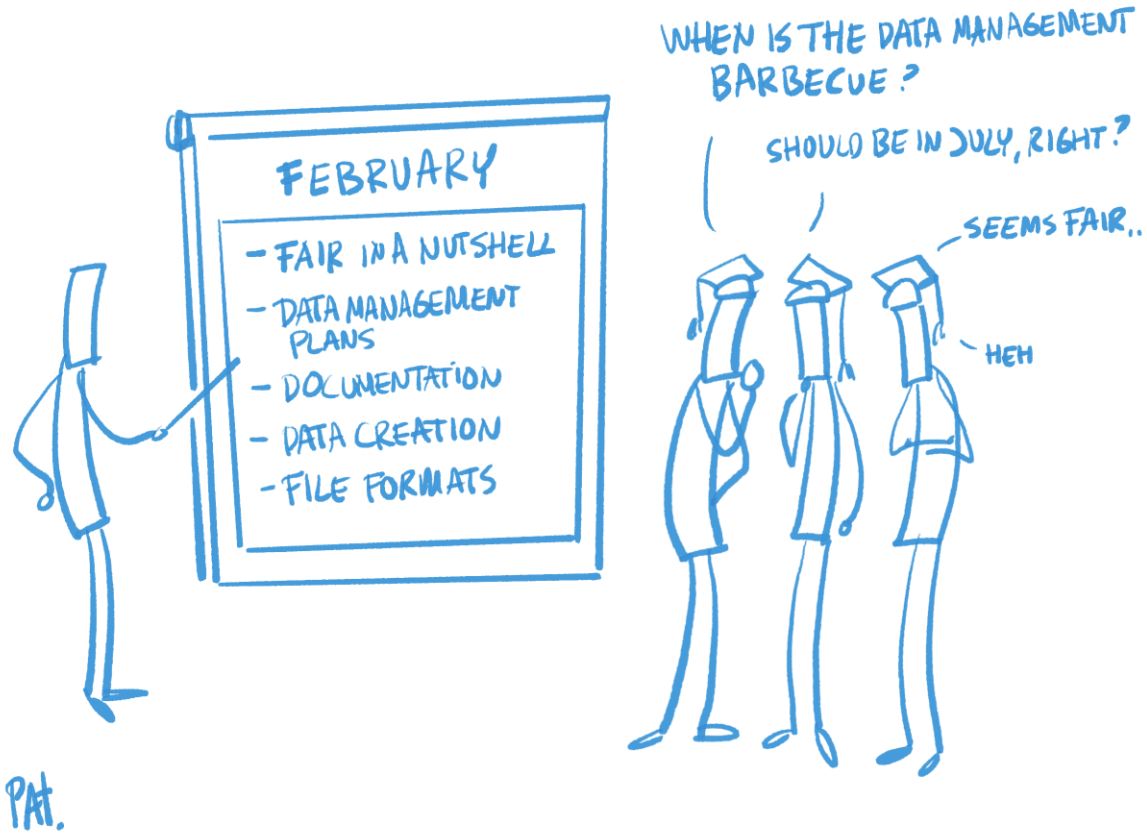
- FAIR in a nutshell
- DMPs
- Documentation
- Data Creation
- File Formats
- Data Standardisation and Ontologies
- PIDs
- Licences, Copyright, IPR
- Data reuse
- Repositories
- Sensitive data & ethical aspects
- Data Access



- FAIR Software/citable code
- RDM – overview & best practices
- Data Management and Governance in Industry and Research

Lesson plan structure

- FAIR elements concerned
- Primary audience(s)
- Learning outcomes
- Summary of tasks/actions
- Materials/equipment
- References
- Take-home task(s)



Lesson plan 3: Documentation

FAIR element(s):

Re-usable
The ultimate goal of FAIR is to optimise the reuse of data. To achieve this, metadata and data should be well-described so that they can be replicated and/or combined in different settings.

R1. (Meta)data are richly described with a plurality of accurate and relevant attributes

R1.2. (Meta)data are associated with detailed provenance

Primary audience(s): UG, Masters, PhD

Learning outcomes:

- Can explain the purpose (benefits) of the documentation, and its relation to FAIRness
- Can identify different types of data documentation, and which are suitable to a specific discipline/domain
- Can use existing documentation
- Can modify existing documentation
- Can identify considerations and strategies for documentation

Summary of Tasks / Actions:

1) Introduce concept of documenting research data
    a) Outline that a key aspect of data reusability is that it is easily interpreted by people outside of the study, and that this can be achieved by proper documentation
2) Link to relevant section/question of DMP tool used in your country/region (The examples used below are from the Canadian DMP Assistant, https://assistant.portagenetwork.ca/).
    a) What documentation will be needed for the data to be read and interpreted correctly in the future?
        (1) Project-level
        (2) File-level
        (3) Item-level
        (4) Any other contextual information necessary for others to interpret
    b) How will you make sure the documentation is created or captured consistently throughout the project?
        (1) Clear articulation of how this will be done and by whom
        (2) Standardised process for accurate, consistent, and complete documentation
3) Depending on discipline/domain of the group, introduce relevant documentation formats
    a) Readme file
    b) Data dictionary
    c) Codebook
    d) Commented code
    e) Lab/field notebook (including Jupyter Notebooks, R markdown, electronic lab notebooks, etc.)
        i) If introducing multiple formats, outline similarities/differences and use cases
        ii) For each format that is showcased, articulate considerations and other important aspects by using exemplars and other material from the "References" section
4) Conduct an exercise in which learners complete one or more of the documentation formats, based on course/project work that is relevant to learners. Blank templates can be found/created using material from the "References" section. Review and discuss challenges, as well as strategies to mitigate challenges.

References

READMEs
- Guide to writing "readme" style metadata
- README template

Data Dictionaries
- How to Make a Data Dictionary
- Data Dictionary Template
- Community defined models and formats in FAIRsharing

Codebooks
- Codebook Cookbook
- Sample Questionnaire with Coding

Commented Code
- Coding and Comment Style

Lab/field Notebook
- Examples of notebook pages and entries
- Guide for Taking Field Notes
- Electronic Lab Notebooks
- Jupyter
- R Markdown

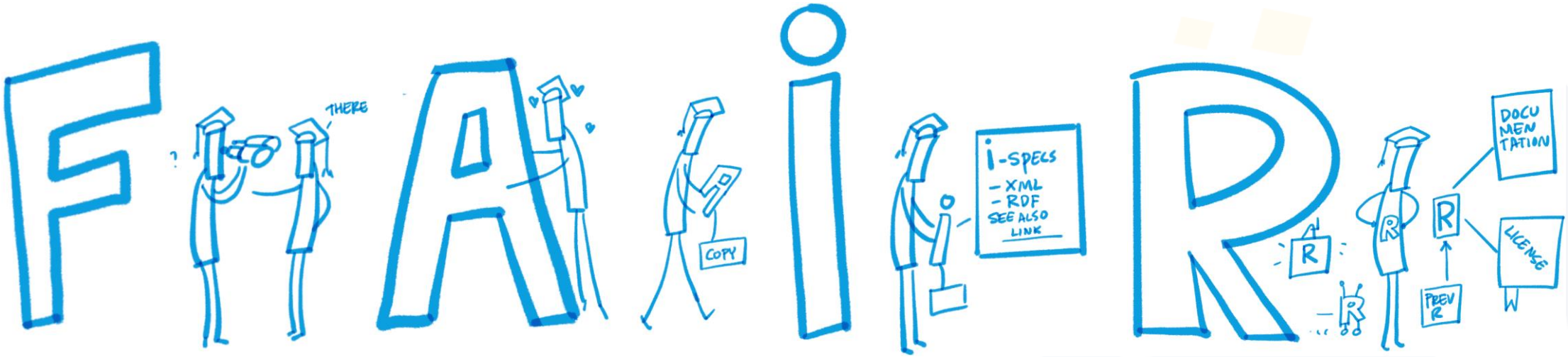
Exercises
- LEGO® Metadata for Reproducibility game pack - Enlighten: Publications

Implementing FAIR

- Getting to FAIR institutional policies
- Data management planning
- Data processing and documentation
- Support infrastructure
- Data publication
- Data reuse

Future

- Print and Open Access PDF version with Göttingen University Press
- GitBook version hosted by University of Minho
- Project ending in February 2022, but discussions with community initiatives regarding potential future revisions/updates



FAIRSFAIR
Fostering Fair Data Practices in Europe

Thanks – any questions?

Contact:
claudia.engelhardt@sub.uni-goettingen.de