

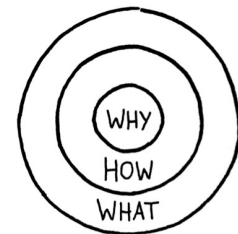
CEG Research data management policy

2020-04-30

Definitions/abbreviations

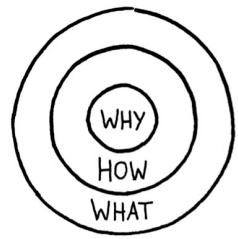
- Data:
 - Observations: field, lab
 - Simulations: input and results
 - (Research) software: software code, analysis scripts,
- FAIR: Findable, Accessible, Interoperable, Reusable
- DMP: Data management plan

Why



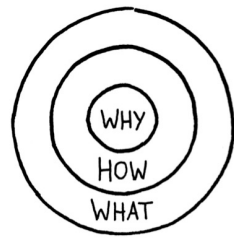
- Data has always been at the core of science. Without data there can not be science. Proper data and code management is the *foundation* for good science.
- Providing FAIR data is the *default*. It makes our work transparent and creates opportunities for ourselves and others to (re-)use our data.
- Proper data management, data re-use and providing FAIR data is (should become) *rewarding*

How



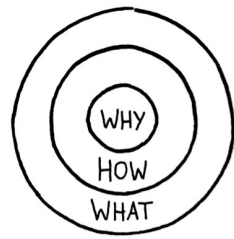
- All staff (faculty, Postdocs and PhD students) are aware of and given the opportunity to follow a training on data management. (Training push & -pull)
- All TU Delft researchers will deposit data FAIR and – preferably – open accessible in a sealed datacentre. This is outcome of proper data management.
- All staff is supported during their work to execute data management by e.g. support in writing a DMP, budgeting, storage, archiving and citing, etc..
- Data management is a matter of leadership:
 - On individual level influencing recruitment, performance, R&D and promotion
 - On organisational level creating awareness on choices and setting targets

What



- FAIR data can be cited, re-used or combined in meta studies by others
- Data management is recognized as integrated part of research practice
- Reproducibility of research results is ensured
- Re-use of available data is encouraged

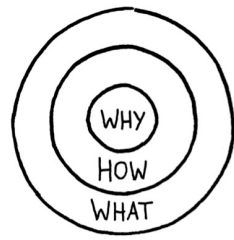
How (detailed) *training on data management*



- Training exists of
 - Initial training: awareness, principles, process within TU Delft
 - Training on the job (data steward support and peer-feedback)
 - Dedicated training: licenses, privacy sensitive data, programming skills, geographic data
- We distinguish two types of scientists within TU Delft
 - Temporary staff (Postdocs and PhD students): *awareness, practical skills*
 - Faculty: *awareness, practical skills, judgement, vision*
- Data management training
 - Temporary staff: GS (collectively) and specialists such as DS and Library (e.g. DMP, ICT-possibilities)
 - Faculty: DS, Library, privacy team, ...

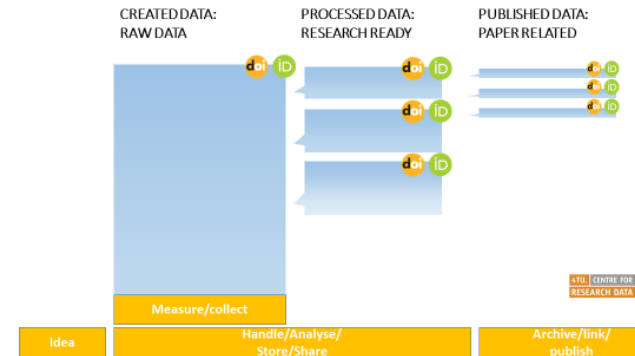
Promotion of available training via FIC, OCW, ...

How (detailed) *deposit data FAIR*

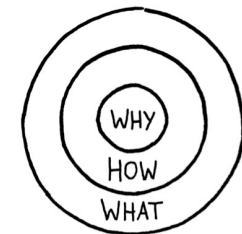


FAIR data needs to be integral part of the project lifecycle

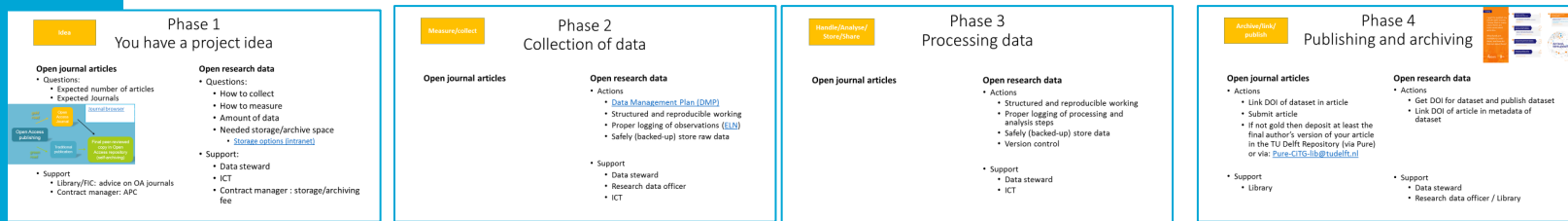
- Preparation:
 - Proposal: data related costs are estimated and budgeted and, if partners are involved, the discussion on the FAIR ambitions is started
 - DMP: further detail FAIR ambitions, associated costs and all project members/partners commit themselves to the DMP
 - Infrastructure: infrastructure is arranged before starting the actual data collection; using version control is considered
 - Roles and responsibilities: explicitly assign someone as the project data manager
- Execution:
 - PI takes responsibility to follow DMP, with help of data manager
 - Reproducibility is ensured
 - Keep track of documentation and metadata throughout
- Finishing
 - Relation between article findings/conclusions and sup_j is non-ambiguous
 - Metadata is carefully reviewed



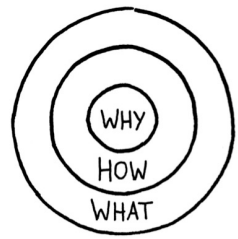
How (detailed) *supported during their work*



- As underlying process we use the lifecycle of a project (see ppt Open science roadshow)
- Initiating a project: budgeting, writing a DMP (DMPonline)
- During a project: storage, version control, workflow
- After a project: archiving and linking to publications, etc..
- Defining process, responsibilities, roles and support.



How (detailed) *matter of leadership*



- Data management is a collective responsibility to prioritize and therefore a matter of leadership;
- On individual level:
 - We recruit scientists who are active with open science and/or proper data management,
 - We judge performance on data management among others during R&D, VLC and promotion
- On organisational level
 - We create awareness on deliberate choices (opting out)
 - We set relevant targets and
 - We translate organisational targets to individual goals

Lifecycle of a project



Training

DMP
Budgeting
IT-possibilities
Journals

Version control
Storage

Reporting
Archiving
Link to articles

Support on the job of staff
by Data Steward, ICT, Contract man., etc. and
PEERS

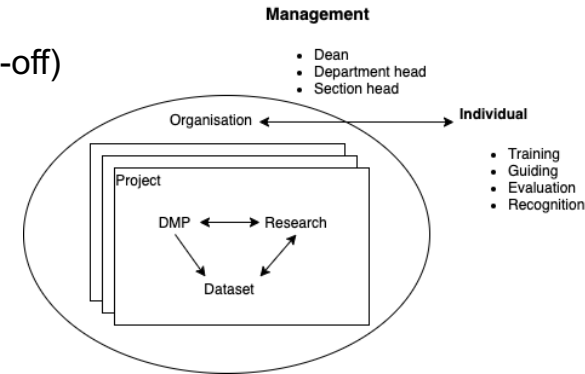
Career path of a scientist

Recruitment, performance (R&D, VLC, promotion)

Reporting/rewarding

Reporting takes place on several levels

- Individual to give feedback on performing on data management (VLC, R&D)
 - Insight in number of links between publications and datasets,
 - (Registration of opting out)
- Project level to ensure compliance to funder and TU Delft policy
 - End report to funder
 - Links between project, publications and datasets
- Department level (Preview & Q-meeting)
 - Performing on content level: output FAIR data, degree of data management
 - Performing on personnel level: insight in degree of data management
- Faculty level
 - Outreach (e.g. stories of science, valorisation, spin-off)



The DEAN is responsible for:

- Ensuring that Data Steward is embedded within faculty
- Ensuring that all research staff are compliant with of this policy

PRINCIPAL INVESTIGATORS are responsible for:

- Ensuring that every research project starts with a data management plan, which needs to be regularly updated and adhered to by all project members.

PHD SUPERVISORS are responsible for:

- Ensuring that PhD students attend relevant training on data management.
- Ensuring that their PhD students make all data and code underlying their completed PhD theses appropriately documented and accessible for at least 10 years from the end of the research project, in accordance with the FAIR principles (Findable, Accessible, Interoperable and Reusable), unless there are valid reasons which make research data unsuitable for sharing.
- Supporting their PhD students in preparation of a written data management plan for managing research outputs within the first 12 months of the PhD study.

All RESEARCHERS are responsible for:

- Ensuring that research data, code and any other materials needed to reproduce research findings are appropriately documented and shared in a research data repository in accordance with the FAIR principles (Findable, Accessible, Interoperable and Reusable) for at least 10 years from the end of the research project, unless there are valid reasons not to do so.
- Should data not be made available in a repository, ensure that the relevant metadata is published in a suitable repository and any research publications resulting from the project have a statement explaining what additional datasets/materials exists; why access is restricted; who can use the data and under what circumstances

PHD STUDENTS are responsible for:

- Ensuring that all data and code underlying completed PhD theses are appropriately documented and accessible for at least 10 years from the end of the research project, in accordance with the FAIR principles (Findable, Accessible, Interoperable and Reusable), unless there are valid reasons which make research data unsuitable for sharing
- Developing a written data management plan for managing research outputs within the first 12 months of the PhD study, and attend the relevant training.

Contact / more information

- Data Steward
 - Kees den Heijer
 - datasteward-citg@tudelft.nl
- Executive secretary
 - Stijn van Boxmeer