

Challenges, Tensions, Possibilities for Research Data at TU Delft

Alastair Dunning
Head, 4TU.Research Data
Marta Teperek
Data Stewardship Coordinator

Open Data Meeting, TU Delft, October 2017



Some background

TU Delft Research Data Support

Helps researchers at Delft on Data Management
Plans, training, metadata, publishing etc.

Developing Research Data Policy Framework

Implementing [Data Stewardship Project](#) to embed
stewards at all 8 faculties

Done research on [FAIRness of 37 data repositories](#) in
the Netherlands

4TU Centre for Research Data

[Trusted Data Repository for technical sciences](#)

Nearly 7,500 datasets with Digital Object Identifier
(DOI), since its start in 2008

Offers at least 15 years long-term curation and
access for research data

[Open to upload data](#) for researchers from 4TU, ie
includes Delft (up to 100GB for free; get in touch for
higher costs)

Uploading data takes 10 minutes

- Archive funded via Delft, Twente, Eindhoven
- All 4TU researchers have 100GB per year free; others can also use archive
- Uploading takes 10 minutes ... we help with the metadata
- But *only* 7,500 datasets published ... how many datasets might there be at TU Delft in total?
- (Unsatisfied? Become a data champion and contribute to new archive/policy)

Uploading data takes 10 minutes

Long-term curation needs metadata

- DOI is part of way of ensuring long-term sustainability of the data
- All datasets copied at Delft, backed up in Leiden
- Long-term curation - *preferred formats* can be *shifted* over time; for other formats we promise *bit-preservation* only
- But good metadata and documentation is essential if you want your data to be easily understood reused in future
- Research undertaken on FAIR principles

Uploading Datasets, Publishing DOIs

Long-term curation needs metadata

Privacy: More common than you think

Many TU Delft researchers combine make use of human subjects. Two Examples

- Stefan vd Spek - Wifi Campus data
- Depression Study Videos at EWI

- Data needs to be anonymised (are the tools for this good enough)
- Consent forms need to be archived and give right permissions (for archiving as well as re-use)
- Even then is the archive secure enough?
- Policies and workflows across the university need to be harmonised
- [Forthcoming](#) EU General Data Protection Regulation

Uploading Datasets, Publishing DOIs

Long-term curation

Privacy: More common than you think

Commercial Pressures

- H2020 and NWO require data sharing as pre-requisite of funding.
- But commercial companies often have reverse attitude
- Expectation is that data will only be shared between company and research group
- Data often shared on “gentleman’s agreements”; or very restrictive licences that are determined by the company
- (Often) Data will not be published; metadata will not be published

Uploading Datasets, Publishing DOIs

Long-term curation

Privacy: More common than you think

Commercial Pressures

Open Science

Requests to make data Open (or FAIR) embedded at High Political Levels

- [G7 science ministers](#)
- [G20 governments have made statements on open science and FAIR principles](#)
- The Dutch and German governments have also made [a joint statement on earlier this year](#) (and are working with the French government on expanding this)
- [Rutte III Government mentions open science](#)

[Embedded at TU Delft](#)

But definition of Research Data is still ... open

- [FAIR principles](#) embedded in H2020
- Getting from FAIR principles to community standards
- Archiving data often means archiving software

- Are we trying to make science *reproducible*? Is that always possible? Are we trying to make science *transparent*? Is that enough to build trust?

Uploading Datasets, Publishing DOIs

Long-term curation

Privacy: More common than you think

Commercial Pressures

Open Science

Green Village

[Green Village](#), a 'living lab' at TU Delft is representative of these legal, commercial and open tensions



Uploading Datasets, Publishing DOIs

Long-term curation

Privacy Issues

Commercial Pressures

Open Science

Green Village

Black Box Society

Pressure on public and private institutions to share their data, yet at the same time protect the

Adam Greenfield - Radical Technologies

Cathy O'Neil - Weapons of Math Destruction

Frank Pasquale- Black Box Society

Government, tech firms, financial institutions, private corporations, police forces are all shown to be misusing data, and hiding behind 'black boxes' (ie closed data and algorithms)

What is the role of a university in the face of these tensions?

Universities must operate according to different values than the above

Can we create FAIRER data? Findable, Accessible, Interoperable, Re-usable, Ethically Responsible (via Deb Verhoeven)

Uploading take 10 minutes

Long-term curation

Privacy Issues

Commercial Pressures

Open Science

Green Village Example

Black Box Society

Embedding Research Data Practices

Research Data Policy Framework for TU Delft

Framework is [Open for Comment](#)

Forwarded to CvB in December
Faculties then develop their own policies
based on the framework

And to take this further forward ... Data
Stewardship → Marta