#dsip2018

DataDay
tudelft2018

Towards cultural change in data management - data stewardship in practice

Thursday 24 May, 09:30 – 18:00

09:30	**Coffee, registration**				
10:00	**Welcome remarks:** Alastair Dunning, TU Delft Library and 4TU.Centre for Research Data				
10:10	**Keynote: Danny Kingsley from the University of Cambridge**				
	The 'end of the expert': why science needs to be above criticism				
10:55	**Marta Teperek from TU Delft**				
	Vision for Data Stewardship at TU Delft				
11:05	**Coffee break**				
11:30	**Case studies – Data Stewardship at TU Delft**				
12:30	**Lunch break + signing up for interactive sessions**				
13:15	**Data Stewardship case studies from other institutions**				
13:15	**Martine Pronk from Utrecht University**				
	Maximizing flexibility and cooperation: *Governance of Data stewardship at Utrecht University*				
13:30	**Inge Slouwerhof from Radboud University**				
	Institutional data stewardship changes the RDM landscape: *experiences from the Radboud University*				
13:45	**Joakim Philipson from Stockholm University**				
	Research Data Management at Stockholm University by 'piecemeal social engineering'				
14:15	**Parallel interactive sessions**				
	Carrot or stick, my role as a Data Steward in Research data management	Data Management and Academic Integrity	Software reproducibility - how to put it into practice?	Why is this a good Data Management Plan?	Tagging privacy-sensitive data according to the new European privacy legislation: GDPR DataTags – a prototype
	Kubus	Yellow Brickroad	Prism	Wit Licht	Green Room
15:15	**Reports from interactive sessions *(main room)***				
15:40	**Coffee break**				
16:00	**Keynote: Kim Huijpen from the VSNU**				
	Giving scientists more of the recognition they deserve				
16:30	**Keynote: Ingeborg Verheul from LCRDM**				
	Data Stewardship? Meet your peers!				
16:50	**Closing remarks:** Wilma van Wezenbeek, TU Delft Library Director				
17:00	**Drinks reception**				
18:00	**Conference close**				

Interactive sessions

Carrot or stick, my role as a Data Steward in Research data management

- **Session organisers:** Paulien Adamse, Jen Banach, Martijn Staats
- **Room:** Kubus

In an interactive workshop, we would like to gather ideas about the role of Data Stewards and creative strategies for stimulating (carrot) and/or enforcing (stick) proper data management.

At RIKILT Wageningen University & Research, data is generated which, next to the original purpose of the analyses, could also be used for direct or indirect (food) concerns. To enable finding those datasets and combining them with data from within or outside RIKILT, it is necessary to make the datasets FAIR. We have prepared guidelines for writing a Data Management Plan and for storing data sets and meta-data (contact person, type of data, software used, etc.) in a practical and insightful way.

We are planning to train data stewards to support their colleagues in achieving FAIR data at RIKILT and to ensure that the guidelines are applied consistently. But, what should be the role of data stewards? And how can they ensure that everyone participates? Pitch your creative solutions, inspire others and compete for the winning prize.

Data Management and Academic Integrity

- **Session organiser:** Lotte Melenhorst
- **Room:** Yellow Brickroad

What is the relationship between data management and academic integrity? And how can data management policies stimulate researchers to uphold the highest standards of academic integrity? These are the type of questions that will be addressed during this interactive session. Currently, a committee is reassessing the TU Delft integrity policy. The committee secretary, dr. Lotte Melenhorst, invites you to reflect on the relationship between data management and academic integrity and share best practices.

Dr. Lotte Melenhorst is policy advisor Strategic Development at TU Delft and currently secretary of the Committee Reassessment Integrity Policy.

Software reproducibility - how to put it into practice?

- **Session organisers:** Maria Cruz, Shalini Kurapati, Yasemin Türkyilmaz-van der Velden
- **Room:** Prism

There is a reproducibility crisis in science. In some fields, over half of published studies fail reproducibility tests; in a survey of 1576 scientists conducted by Nature in 2016, most respondents agreed that there was a crisis and over 70% said they had tried and failed to reproduce another group's experiments. Given the ubiquitousness of software in many areas of contemporary scientific research, it could be argued that there can't be reproducibility in science without reproducible software.

In a recent Comment in Water Resources Research, in response to "Most computational hydrology is not reproducible, so is it really science?", Hut, van de Giesen & Drost (2017) argue that documenting and archiving code and data is not enough to guarantee the reproducibility of computational results. They suggest the use of software containers and that researchers work more closely with research software engineers to learn best practices in software design. This advice is presented in the context of hydrology, but it should apply more generally.

Inspired by the article and its advice, this workshop will explore the topic of software reproducibility, how some of the advice could be put in practice, and what role could libraries and data stewards play in this regard.

Interactive sessions

Tagging privacy-sensitive data according to the new European privacy legislation: GDPR DataTags – a prototype

- **Session organiser**: Ilona von Stein
- **Room**: Wit Licht

Many researchers do not know the details of privacy legislation and its consequences for the data they collect, deposit and/or share in a data repository. The managers and curators of such repositories typically have only superficial insight in the data that are being deposited. Research funders requiring the deposition and sharing of research data cannot know to what degree legal impediments hinder their requirements, and hence whether objections by researchers against such requirements are justified or not.

In May 2018 the General Data Protection Regulation (GDPR) will go into force throughout the EU. This set of rules was designed to harmonise data privacy laws across Europe, to protect and empower all EU citizens' data privacy. The GDPR will have significant impact on the preservation and sharing of research data sets containing information on persons, especially in the social and life sciences (social surveys, interviews, medical data, etc.).

This workshop presents the prototype of a concrete tool to evaluate whether datasets can be shared, and under which conditions, using the notion of DataTags earlier described by Latanya Sweeney and Mercè Crosas. The approach has been adapted to comply with the GDPR requirements. The DataTags tool presented in this workshop essentially is an online questionnaire tool, which assists data depositors and curators to assess whether a dataset contains privacy-sensitive data in terms of the GDPR. The tool recommends a security level for the data and suggests restrictions for data archiving and sharing, and conditions or under which data sharing is permitted. The future GDPR DataTags questionnaire tool should ultimately be usable for any repository in a country where this EU legislation is enforced.

Why is this a good Data Management Plan?

- **Session organisers:** Ellen Leenarts and Marjan Grootveld
- **Room:** Green Room

Data management plans are still relatively new and so is the practice of evaluating them. In this session we will share experiences in reviewing (draft) DMPs. In the introduction we will mention some developments and tools. The main part of the session, however, is interactive. We will look at sample texts taken from public DMPs and discuss what is good about them and why - or what should be improved. It will be interesting to see if there is consensus in the room…

We welcome everyone who has reviewed DMPs or "previewed" them, for instance by feeding back on draft DMPs or co-authoring them. By way of managing expectations: yes, there are differences between research funders' DMP requirements, as well as between research domains. And no, we won't have time to go into all these differences. However, we do expect that the participants will benefit from the discussion and the examples.

Session organisers Ellen Leenarts and Marjan Grootveld (DANS) are involved in RDM and DMP support in international projects. Ellen also coordinated the CESSDA Expert Tour Guide on data management, whereas Marjan has reviewed DMPs for the European Commission.

Bouwcampus venue map