

**‘Research Data Management’**

by Lora Armstrong and Kees den Heijer

Department of Geoscience & Engineering/Faculty of Civil Engineering and Geosciences

**Entrance Requirements:** MSc level in science or technology; computer with internet connection and microphone required

**Course date:** 17-21 August, 2020

**Time:** 13:00-16:30 each day

**Place:** Online workshop

**GS credits:** 2

**Spaces available:** 25

**Entrance requirements:** This course is aimed at PhD students who are interested in gaining an overview of good data management practices and learning how to apply them to their own research. There are no prerequisites for this course, but a computer with an internet connection and microphone is needed to participate in the sessions. The course will be taught online.

**Scope:** The aim of this course is to provide an in-depth overview of research data management (RDM) topics, tools and best-practices that will help during different stages of a research project. We will discuss what is required by the university and funding agencies in terms of RDM, what resources are available (including resources specifically for geoscientists), and how to go about implementing them. The course will be a mix of lectures, discussions, and hands on activities where participants will apply the concepts to their own project and data. By the end of the course, participants will have a data management plan for their project and the background knowledge needed to carry out each step of this plan.

The following topics will be addressed, with focus on topics that are of most interest to the participants:

- Research data management and what the benefits are
- RDM requirements from funding agencies, universities and journals
- FAIR data
- Where to store your data
- How to describe your data using metadata
- Data workflows, organization, structure
- Databases
- Electronic lab notebooks
- Data repositories and publications for sharing data
- How to write a data management plan

- Managing research code: Version control basics
- Managing research code: Collaboration & continuous integration
- Helpful resources

### **Learning Objectives:**

At the end of the course, students will be able to:

- Write a data management plan for a research project
- Explain what constitutes research data
- Explain why research data management is needed and what benefits it brings to a project and the wider scientific community
- Link specific data management tasks to different parts of the research lifecycle
- Explain what constitutes FAIR data and how to improve the FAIRness of a dataset
- Identify the data policies and requirements that apply to a research project (e.g. from funders, universities, journals)
- Design a file organization system for a project based using best-practices for organization, file naming, and documentation
- Understand what metadata is and how it is used to describe data
- Understand what metadata standards are and how they are used
- Identify metadata standards applicable to a particular research project
- Identify what metadata should be recorded for a research project
- Understand how Electronic Lab Notebooks can be incorporated into a project for documenting and sharing lab notes
- Understand best practices for backing up and securely storing data
- Explain what the main types of databases are and understand how a database can be incorporated into a research project
- Understand how version control systems such as Git can be used to manage research code in a project with one or more contributors
- Explain the different ways to share research data (e.g. different types of data repositories, data publications)
- Identify a repository for sharing data from a particular research project
- Identify preferred file types for repositories

**Assessment:** There will not be an assessment, but participation in the in-class activities is required.