

Data (and code) Roles of the Future at TU Delft

15 May 2019 by Alastair Dunning and Marta Teperek

	Data Steward	Data Manager	Research Software Engineer ¹	Data Scientist ²	Data Champion
Type of position	Dedicated staff member	Dedicated staff member	Dedicated staff member	Researcher, with expertise in data analysis and programming	Researcher, with voluntary data management tasks
Affiliation	Faculty	Research Group	Faculty/Research Group	Research Group/Non-university	Research group
Funding source	Faculty budget	Research group budget (project funding)	Faculty budget / Research group budget (project funding)	Research group budget (project funding)	Research group budget (project funding)
Expertise	Good understanding of Faculty's research; understanding of research data	Understanding of research specific to the research group	Background in software development	Data analysis, programming	Understanding of the importance of good data management practices
Education	PhD or equivalent	PhD, Msc or BSc	BSc or higher	MSc or higher	Doctoral Candidate or higher
Skills & knowledge³	Data management Disciplinary workflows and metadata standards Oral and written communication skills Negotiation skills Project management	Data management Disciplinary workflows and metadata standards Analytical thinking Project coordination Written communication skills	Programming and data modelling Expertise using version control environments Software management Data management Disciplinary workflows and metadata standards	Analytical thinking Programming Data modelling Oral and written communication skills	Data management Disciplinary workflows and metadata standards Data sharing
Tasks	Analyse faculty needs for data management; Create policies and workflows; Identify skill gaps and ensure appropriate training provisions; Identify support service gaps and design strategies to fill them; Advise faculty on internal and external developments in data management.	Analyse data management needs of a research group; Propose, implement and monitor workflows and tools to improve data management practice; Oversee collection, description, cleaning, merging, licensing, sharing and publishing of data.	Develops code for applications / interfaces / data processing / web services etc.; Champion software sustainability practices, oversee and ensure maintenance and documentation of code produced by group members Design and maintain databases for specific use cases;	Find, clean, merge, analyse and interpret datasets; Ensure intellectual consistency of data across various projects; Model, visualise, present and communicate data insights and findings to various audiences, and recommend applications.	Advise other researchers in the local setting Advocate for research data management and sharing in the local setting Speak to researchers about own experience with data management Participate in Data Champions meetings
Liaison	Faculty researchers Cross-faculty stakeholders University-wide support services Data Stewards from other faculties National and international stakeholders	Research group members Faculty Data Steward Research Software Engineer Faculty-level support service providers Third party providers of specific tools for data management	Research group members Data Manager Data Scientist Faculty Data Steward Programming communities	Research group members Data Manager Research Software Engineer Project leaders Other stakeholders, including third parties, as appropriate	Faculty Data Steward Other Data Champions Research group members

¹ While Research Software Engineers are not dedicated to work on data as such, software and data are inseparable in most research projects, and therefore Research Software Engineers are also included in this table

² While in the past data scientists were dedicated team members with expertise on data analysis and programming skills, nowadays data science skills are more and more becoming the essential skills for every modern researcher. Hence, data scientist can be simply perceived as a scientist in the 21st century.

³ Skills and knowledge, as well as tasks, are discussed in the context of data (and code) management needs