Policy needs to go hand in hand with practice
The learning and listening approach to data management

Cruz, Maria; Dintzner, Nicolas; Dunning, Alastair; Kuil, Annemiek van der; Plomp, Esther; Teperek, Marta; Türkyilmaz-van der Velden, Yasemin; Versteeg, Anke

DOI
10.5334/dsj-2019-045

Publication date
2019

Document Version
Final published version

Published in
Data Science Journal

Citation (APA)

Important note
To cite this publication, please use the final published version (if applicable). Please check the document version above.
PRACTICE PAPER

Policy Needs to Go Hand in Hand with Practice: The Learning and Listening Approach to Data Management

Maria Cruz1, Nicolas Dintzner2, Alastair Dunning2, Annemiek van der Kuil3, Esther Plomp2, Marta Teperek2, Yasemin Turkyilmaz-van der Velden2 and Anke Versteeg2

1 Vrije Universiteit Amsterdam, NL
2 Delft University of Technology, NL
3 Utrecht University, NL
Corresponding author: Marta Teperek (M.Teperek@tudelft.nl)

In this paper, we explain our strategy for developing research data management policies at TU Delft. Policies can be important drivers for research institutions in the implementation of good data management practices. As Rans and Jones note (Rans and Jones 2013), “Policies provide clarity of purpose and may help in the framing of roles, responsibilities and requisite actions. They also legitimise making the case for investment”. However, policy development often tends to place the researchers in a passive position, while they are the ones managing research data on a daily basis. Therefore, at TU Delft, we have taken an alternative approach: a policy needs to go hand in hand with practice. The policy development was initiated by the Research Data Services at TU Delft Library, but as the process continued, other stakeholders, such as legal and IT departments, got involved. Finally, the faculty-based Data Stewards have played a key role in leading the consultations with the research community that led to the development of the faculty-specific policies. This allows for disciplinary differences to be reflected in the policies and to create a closer connection between policies and day-to-day research practice. Our primary intention was to keep researchers and research practices at the centre of our strategy for data management. We did not want to introduce and mandate requirements before adequate infrastructure and professional support were available to our research community and before our researchers were themselves willing to discuss formalisation of data management practices.

This paper describes the key steps taken and the most important decisions made during the development of RDM policies at TU Delft.

Keywords: data stewardship; TU Delft; policy; policy development; code; research data management; RDM; open science; policy implementation; data champions; data repository; data archive

Introduction

In recent years, Research Data Management (RDM) has become an increasingly important part of the research process. Funders at the national and international level (NWO n.d.; NSF n.d.; European Commission 2016) require Data Management Plans (DMPs) and stipulate the open sharing of datasets resulting from publicly-funded projects. Publishers are increasingly asking for the data underlying research papers to be made publicly available (Byrne 2017; Naughton and Kernohan 2016; Hrynaszkiewicz et al. 2017; Nature Research n.d.; PLOS ONE n.d.; OAD 2018; COS n.d.). Research organisations, such as CERN, the European Organization for Nuclear Research, are also increasingly opening up their data and promoting transparent research practices (Chen et al. 2019). Governments across the world are encouraging the accessibility of datasets (Gunawan and Amalia 2016; van Belle 2018). Moreover, professional RDM practices increase the efficiency, transparency and reproducibility of research, and boost the visibility and social impact of the
research outputs (Piwowar and Vision 2013; Vandewalle 2012; Lowndes et al. 2017; Pryor, Jones, & Whyte 2013). Recognising the importance of good practices in RDM, research institutions have started introducing policies on research data management.

In the Netherlands, RDM policies are in place at most universities (LCRDM n.d.). Introducing policies early in the process of creating institutional RDM support can help align provisions offered by multiple service providers, and also to make institutional business cases for RDM (Rans and Jones 2013). Such institutional policies often outline responsibilities regarding data management for both the academic and support staff, as they are concerned not only with research practices but also with the support services necessary for good data management, such as Information and Communication Technologies (ICT) services, legal support, ethical committees, etc. However, if researchers are already subjected to additional funders and publishers requirements surrounding research data, yet do not receive the support needed to fulfill them, a policy first approach can be problematic. Researchers might be unable to put these policies into practice and policies might as a result turn into aspirational documents, out of touch with daily research practices. In addition, the experience gained through implementation of policies on open access to journal articles suggests that policies in general tend to be rather unpopular among researchers. Contrary to their intentions, these policies can be perceived as “box-ticking” exercises, rather than tools facilitating cultural change (Johnson and Chiarelli 2017).

Therefore, at Delft University of Technology (TU Delft) in the Netherlands, we have decided to try a different approach: policy development that goes hand in hand with practice. By combining a top-down, policy-driven approach, with a bottom-up approach, researchers are more inclined to invest their time and effort in RDM and sharing their findings with others (Teperek, Higman & Kingsley 2018). Therefore, a systematic approach to engaging with stakeholders across the university was devised (Figure 1, and also section “Policy development”). This approach not only raised awareness about data management but also allowed the various stakeholders across the university to experiment with new ideas and approaches related to management and sharing of research data (see below). In addition, these discussions helped to build trust among the different partners and to gain a better understanding of the needs of the different stakeholder groups, which were then reflected in the policy. As a result, an extended period of informal consultations helped build an engaged community of interested people, who were willing to do some initial experimentation without having to be tied to policy obligations. Such engagement with researchers also proved fruitful in the formulation of RDM policies at the Dutch Universities of Wageningen (van Zeeland and Ringersma 2017), Leiden (Verhaar et al. 2017) and Erasmus Rotterdam (Domingus 2017).

**Technical and human infrastructure at TU Delft**

Our ‘policy hand-in-hand with practice’ approach would not have been possible without the technical infrastructure at TU Delft, adequate training and research support staff provision, which are the foundations for good research data management.

![Figure 1: Stakeholders involved in the policy development at TU Delft.](image-url)
**Technical infrastructure**

TU Delft has the advantage that significant pieces of technical infrastructure were in place far in advance of policy development. Infrastructure for data storage during the active phase of the research lifecycle is provided by the central ICT department. Researchers are also supported when it comes to long-term data archiving. The 4TU.Centre for Research Data (4TU.ResearchData 2019) was launched in 2010 (Figure 2) as part of the 4TU federation of technical universities in the Netherlands (4TU.Federation 2019). TU Delft Library took the lead in establishing the archive. The archive is hosted at TU Delft and many staff members of TU Delft Research Data Services team work also for the 4TU.Centre for Research Data. The archive is a certified, trusted repository for long-term preservation and sharing of research data (CoreTrustSeal 2016).

4TU.Research Data also worked as a test bed for experimenting with issues germane to the later policy development, such as establishing guidelines for large file sets and reflecting disciplinary differences. An early use case supported by the archive was the publication of several terabytes of data collected from a high resolution meteorological International Research Centre for Telecommunications and Radar Drizzle Radar (IDRA) (Otto et al. 2010; Figueras i Ventura and Russchenberg 2008). The sheer volume of the data presented challenges for their effective sharing and long-term preservation. However, given that the observations consisted of unique meteorological records, it did not seem appropriate to archive only processed (and thus smaller) parts of the dataset. Instead, it was deemed necessary to store all the collected data. To facilitate sharing and accessibility of these large datasets, they were made available via the archivist’s Open-source Project for a Network Data Access Protocol (OPeNDAP) server (OPeNDAP 2019; Cruz and Gramsbergen 2018). The use of the OPeNDAP server enables users to query the data online and access just the parts that are of interest to them. The IDRA dataset case study helped us develop the necessary infrastructure to ensure that when the policy was introduced, appropriate solutions to allow researchers comply with the policy requirements, were available.

Another interesting case was the DINED data (Molenbroek 2018). DINED is a large collection of anthropometric data which have been gathered for over 40 years, and which are widely used in various design projects to help designers make better products. The DINED research group was awarded the Dutch Data Prize in 2014 in recognition of the important value of their data collection (NWO 2014). This case study demonstrated the value of data sharing and helped engage with the research community, which we believed was essential for discussions around policy development.

Helping researchers who were interested in depositing their data in the 4TU.Research Data archive, allowed us to identify academics interested in data management and sharing within TU Delft faculties, who

---

1 DINED is the name of the whole project and of the dataset. The origin of this name comes from DIN + NED. DIN stands for the Deutsches Institut für Normung, and NED is for Nederlands. DINED was inspired by a German standard for anthropometrics, and this is a Dutch extension of it.
would act as early advocates and keen reviewers of the policy documentation. The team from the DINED project was based in the Faculty of Industrial Design, which was otherwise relatively unengaged in the process of data sharing. Further investigation within the faculty revealed myriad design processes and research methodologies at work in the faculty; this eclecticism would need to be reflected in the policy (Love 2019a, 2019b).

**Training and people support**

We also reasoned that before policies can be implemented, the research community needs to be adequately supported, so that the requirements can be adhered to in practice. We therefore focused on addressing two gaps: in disciplinary support for data management, and in adequate training. In response to research community requests for a more dedicated, subject-specific support, the Data Stewardship programme was launched in 2017 (Figure 1). The programme aims to develop mature, subject-specific practices for working with research data across the whole campus (Teperrek et al. 2018). To achieve this, a dedicated Data Steward with disciplinary knowledge (at least a PhD degree in a relevant research area) was appointed at every faculty at TU Delft. Data Stewards serve as the first contact points for any research-data related questions. Thanks to this ‘first line support’ approach, the Data Stewards are very closely connected with the research communities in their faculties. They meet regularly with researchers for individual consultations and proactively discuss with researchers what their needs and attitudes towards data management are. These meetings and discussions have been instrumental in the development of faculty data policies (see Policy development section).

In addition, to ensure that training needs are appropriately addressed before the policies are implemented, one of the initial activities of the Data Stewards was to investigate the specific training needs of the research community. We discuss here training in the area of research software management and development as an illustrative example of our approach to training development in general. A TU Delft-wide survey on data management needs (van Dijck 2018) revealed that 32% of researchers were interested in training on using version control software and 18% specifically in software carpentry workshops.

To address this growing need for research software management support and version control tools, the 4TU. Centre for Research Data became a Gold Member of The Carpentries (The Carpentries n.d.). The membership entitles 4TU.Centre for Research Data to organise carpentry workshops for its research community and also to train local carpentry instructors (Kurapati and Teperrek 2018). In addition, the Data Stewards launched monthly coding consultations (Dintzner, den Heijer & Teperrek 2019): “Coding Lunch & Data Crunch”. During these open doors sessions, researchers are welcome to drop by and receive help not only with coding problems but also other RDM related issues, such as data storage and archiving, working with confidential data, preparing Data Processing Agreements and informed consent forms, and other data-related issues.

Altogether, through the provision of dedicated disciplinary support for researchers and flexible training, we were able to deepen our understanding of the needs of our research community. This not only allowed us to further strengthen the connections with our researchers and make them more data and software-savvy but also helped us better shape the content of the future policy. For example, the scope of the central policy was extended from data only, to data and code, as result from early interactions between data stewards and their research communities. But most importantly, adequate support was introduced in parallel to policy discussions – to ensure that researchers are able to implement the policy in their day to day practices.

**Policy development**

In preparation for policy development, a series of qualitative interviews, quantitative surveys and consultations with researchers were conducted, which led to a better understanding of the barriers to data management and sharing. While surveys (done in 2015 and in 2017/2018) allowed us to gather feedback from larger groups of researchers (221 and 680, respectively) (Versteeg, van der Kuil & Dunning 2016; TU Delft Data Stewards et al. 2018; Andrews Mancilla 2019), individual consultations about the policy provided more in-depth input. One to one interviews were held with researchers at different stages of their career (from PhD students to professors). Faculty Executive Secretaries identified several researchers at every faculty, who would be the most representative for their research communities. Representatives from support services (e.g. legal, HR, ICT) were also consulted, as well as senior administrative staff. While the approach was systematic and a template to guide the discussions was used (Data Stewards 2018), the conversations were deliberately informal, to allow participants to freely share their opinions. We asked about people’s daily
practices and experience with data sharing, but also about problems and needs regarding data management practices (Teperek, 2018; Cruz and Gramsbergen 2018; Cruz and Dunning 2018).

After every interview, the interviewer(s) prepared reports, which were shared within the Research Data Services team, who carefully analysed feedback looking for synergies and recurring concerns and recommendations in order to make policy proposals adequate to the given stakeholder groups. For example, when thinking about the roles and responsibilities of PhD students and their supervisors, considerations were given to the voices of PhD students, as well as those of their supervisors (Dunning 2019). In addition, any requirements were carefully considered alongside possible obstacles and problems, and solutions which needed to be put in place before any formal requirements could be introduced (Figure 3).

One framework policy to rule them all... and multiple disciplinary policies

While discussing the future policy, one key requirement articulated by our research community was that diverse disciplinary practices and behaviours had to be reflected in the policy. Although TU Delft specialises in the technical sciences and engineering, research conducted across the campus is diverse. Consequently, data workflows used by researchers studying social behaviours of people in cities differ from those used by researchers collecting and analysing real-time weather data or working on 3D printing projects.

Policies also need to be closely aligned with community practices if they are to be relevant. As noted by Prof. Bas Teusink, the Scientific Director of the Amsterdam Institute for Molecules, Medicines and Systems (AIMMS) at VU University in Amsterdam: “There’s a huge gap between policy and implementation for people doing the daily work” and there is also a need to “translate general policies and general infrastructure into daily practical solutions that fit (...) local needs” (Cruz 2018). Policies that address these local needs facilitate the development of guidelines for best practises in data management and help to provide clarity and address researchers’ concerns around research data management (Jahnke, Asher & Keralis 2012). To address these local needs, while also ensuring that there is an overarching vision for research data management at TU Delft, the approach chosen was to have a central framework policy on data management – the TU Delft Research Data Framework Policy (Dunning 2018) – and to allow every faculty to develop their own policy, which will reflect disciplinary norms and guidance, under the umbrella of the Framework Policy. Similar approaches with one central, overarching policy, and multiple disciplinary policies were used at the University of Twente (University of Twente 2018) and the University of Utrecht (Utrecht University 2016).

Furthermore, departments and individual research groups, assisted by the Data Stewards, are encouraged to come up with their own guidance, aiming to translate the faculty policy into their day-to-day practice.

---

Figure 3: Timeline for the policy development. Initial consultations started in 2015, which resulted in the initial draft of the central TU Delft Research Data Framework Policy, approved on the 26th of June 2018. The Faculty policies have been in development since June 2018, based on the central Framework.
Key requirements of the framework policy

The framework policy defines the roles and responsibilities of TU Delft-wide stakeholders, for example, ICT services, TU Delft Executive Board, the Library. In addition, a template (Dunning 2018) is provided to guide the faculties through their policy development efforts. To ensure that all TU Delft stakeholders share the same vision, and that there is consistency between the faculty policies, the framework policy defines some immutable and common key roles and responsibilities of the various faculty stakeholders. We arrived at this list of common requirements after analysing the interviews conducted with the representatives of the different groups, which helped us identify common requirements (see policy development section). These key requirements include:

- The responsibility of principal investigators for good data management practices within their research groups;
- The responsibility of PhD supervisors for good data management practices of their PhD students;
- The requirement on every researcher to appropriately document and share research data, code and any other materials needed to reproduce research findings in a research data repository in accordance with the FAIR principles (Wilkinson et al. 2016);
- The requirement for PhD students to deposit research data and code underlying their PhD theses in a data repository.

Bottom-up activities

To ensure wide policy adoption across the university, we believed it was essential that the top-down, policy-driven approaches happen in parallel to bottom-up, community-driven efforts (for rationale see Introduction). In September 2018, inspired by a successful initiative from the University of Cambridge (Higman, Teperek & Kingsley 2018), TU Delft launched its Data Champions programme (TU Delft Library 2018). Data Champions are researchers who volunteer to advocate good data management and sharing among their local communities. In December 2018, TU Delft has seen the kick off meeting of its Data Champions community (Plomp 2019), which, as of August 2019, already has over 45 members (TU Delft Library 2019).

Having a community of Data Champions has already assisted policy development at TU Delft. Two of the Data Champions from the Faculty of Applied Sciences, Anton Akhmerov and Gary Steele, took the lead and developed a dedicated policy on sharing data supporting publications for their department (Quantum Nanoscience) (Akhmerov and Steele, 2019). This has been a source of inspiration not only for other research groups at TU Delft (Plomp 2019) but also for senior management, who decided to use the departmental policy as a baseline for the policy of the entire faculty. In addition, at the management team meeting of the Faculty of Applied Sciences, the dean expressed his support of the Data Champions programme and suggested that the faculty should have at least one Data Champion at every department to help create group-specific data management guidance.

Reflections and next steps

All eight TU Delft faculties are anticipated to develop draft policies, organise consultations and publish the final versions of the policies by the end of 2019 (Figure 3). At every faculty, Data Stewards facilitate policy development and Data Champions are involved in policy consultations; however, the strategies for policy development differ at each faculty. For example, at the Faculty of Applied Sciences and the Faculty of Mechanical, Maritime and Materials Engineering, the policy draft is discussed during meetings with the management teams and researchers of every single department. At the Faculty of Technology, Policy and Management, the development of the policy is done through two channels simultaneously. First, the Data Steward, with the help of a member of the faculty management team, organises focus group meetings with researchers and support staff to gather their suggestions and comments. Secondly, the draft of the policy is made available online for comments through the faculty newsletter, encouraging feedback from all faculty members.

Policy publication by the end of 2019 is just the beginning of the process. In parallel to developing policies, faculties will also have to produce implementation strategies with realistic timelines. This will consist not only of a communication plan but also of a clear approach to address any gaps in training and infrastructure provisions (both technical and human). Finally, faculties will establish the means to monitor, evaluate and review the policy implementation.
Researchers practising good data management need to be rewarded

Finally, to ensure that research data management becomes an integral part of good research practices, and policies are not perceived merely as bureaucratic requirement, monitoring and evaluation need to be coupled with changes to academic rewards. If researchers are to consider data management as an activity worth the investment of additional time and effort, these activities need to become part of hiring criteria and performance evaluation.

Because researchers are mobile and typically work at multiple research institutions throughout their careers, it is important that these issues are addressed at higher levels than individual faculties and institutions. In line with this, the Dutch National Plan on Open Science has a commitment to review academic rewards systems (van Wezenbeek et al. 2017). A similar intention has been articulated by funding bodies as a result of a joint workshop at the LIBER (Ligue des Bibliothèques Européennes de Recherche) 2018 conference (Ball et al. 2018). However, while it is necessary that changes to academic rewards systems are implemented at national and international levels, reaching a consensus between multi-stakeholder groups at varying degrees of readiness can be time-consuming. Therefore, local initiatives are important. For example, the University Medical Center Utrecht is no longer using the number of papers as the sole indicator of academic performance (Benedictus, Miedema & Ferguson 2016). Instead, researchers are asked to submit narrative explanations about the impact of their research.

At TU Delft, the issues around academic rewards were discussed during consultations for faculty policy development. While solutions to these issues might not be included in the final policies, awareness is raised and discussions are triggered. For example, the Faculty of Civil Engineering and Geosciences at TU Delft wishes to include the commitment to open science as one of its hiring criteria, and also to ensure that researchers are asked about their data management practices during their progress review meetings. Both commitments are to be included in the draft faculty policy (which will be subsequently consulted with the faculty members). The Faculty of Applied Sciences and the Faculty of Mechanical, Maritime and Materials Engineering are also planning to stress the importance of rewarding the faculty Data Champions for their local advocacy efforts.

Conclusion

In this paper, we argue that to successfully implement an institutional policy on data management, policy needs to go hand in hand with practice. Otherwise there is a risk that the policy will become a purely aspirational document, with no reflection in daily practices. At TU Delft we strived to avoid that risk and put a lot of effort to ensure that we had a holistic approach towards the implementation of good research data management practices across the whole campus before the policies were introduced. We believed it was essential that all three elements – infrastructure, culture and policy – were in place and were aligned. We therefore invested in robust infrastructure for data management and sharing, in disciplinary support for data management and appropriate training. In parallel, we started changing culture through systematic discussions with individual researchers, as well as members of support staff. We also invested in bottom-up community efforts, such as the Data Champions initiative, which proved instrumental to get the research community engaged in policy development. At TU Delft we have reached a stage where the basic data management requirements of our research community were addressed, so that mature policies for data management could have been introduced. However, good data management is a moving target, and therefore both the services, as well as the policies, need to be frequently reviewed and adapted to ensure that they remain relevant. We are therefore planning to review our policies on a regular basis and we hope that with the presence of the Data Stewards, as well as the Data Champions, we can ensure that our policies will continue to evolve in close connection with the research community practices and needs.

Competing Interests
The authors have no competing interests to declare.

Author Contributions
- Maria Cruz: Conceptualization (supporting), Writing – original draft, Writing – review & editing
- Nicolas Dintzner: Conceptualization (supporting), Investigation (equal), Writing – original draft
- Alastair Dunning: Conceptualization (equal), Funding acquisition (lead), Project administration (Supporting), Supervision (equal), Writing – review & editing
- Annemiek van der Kuil: Conceptualization (equal), Investigation (Supportingl), Funding acquisition
References


Cruz, MJ and Dunning, A. 2018. Research Data Management within the 4TU Research Centres. DOI: https://doi.org/10.31219/osf.io/sgfhw


Molenbroek, JFM. 2018. DINED – anthropometric database. 4TU.Centre for Research Data. DOI: https://doi.org/10.4121/uuid:199467d8-5c40-4a1f-a2f2-f2040db26270


Teperek, M. 2018. Views on Data Stewardship – report of preliminary findings at the Faculty of Technology, Policy and Management (TPM) at TU Delft. Open Science Framework. DOI: https://doi.org/10.31219/osf.io/8ce5v


Versteeg, A, van der Kuil, A and Dunning, AC. 2016. Motivations for Sharing Research Data at Delft University of Technology. TU Delft. DOI: https://doi.org/10.4121/uid:d20e37bc-3f74-4c64-b88af089ffae2f2
