Engaging Researchers with Data Management
The Cookbook

CONNIE CLARE, MARIA CRUZ, ELLI PAPADOPOULOU, JAMES SAVAGE,
MARTA TEPEREK, YAN WANG, IZA WITKOWSKA, AND JOANNE YEOMANS

Cover image: Photo by Johannes Groll on Unsplash, https://unsplash.com/photos/mrIaqKh9050

Cover design: Anna Gatti.

I. Introduction

Good Research Data Management (RDM) is a key component of research integrity and reproducible research, and its value is increasingly emphasised by funding bodies, governments, and research institutions. However, discussions about data management and sharing are often limited to librarians, data professionals, and researchers who are already passionate about data stewardship and open science. In order to implement good RDM practice throughout research communities a cultural shift is necessary, and effective engagement with researchers, who are the main data producers and re-users, is essential for this shift to happen.

What Is this Book About?

This book contains 24 RDM case studies, each describing an innovative activity used by a research institution to engage with its researchers about research data. These case studies, collected from research institutions worldwide, illustrate the diversity of feasible initiatives that could be implemented in other institutional settings.

The aim of this book is to inspire and inform those responsible for RDM using activities that have already been implemented and reflected upon elsewhere, and to help drive overall cultural change towards better data management. Our focus is not on what constitutes good research data management, but rather how it can be effectively communicated to the research community.

Who Is this Book For?

This book has been written for anyone interested in RDM, or good research practice more generally. It will be particularly useful to those interested in how to effectively engage with researchers about research data management. This might include librarians, data managers, data stewards, archivists, members of ICT (Information and Communication Technology) departments, colleagues from legal and financial support, faculty management, senior executives at institutions, funders, policymakers, publishers, members of the commercial sector, and researchers at any career stage who want to change practices among their peers. In short, if you have read this far, then this book is for you.

Why Read this Book?

We hope that reading this book will:

- inspire you to implement new activities to engage with researchers about research data;
- help you find the activities most suitable for your institutional setting (according to size, research profile, resources available for data management, target audience, etc.);
- inform you about the ease of implementing each case, identifying the specific challenges associated with them and possible tips to overcome these;
- give you a general overview of what other institutions around the world do to engage their researchers with research data;
- provide you with tangible suggestions for actions that you could present to senior management at your institution;
- stimulate collaboration. We hope that reading our case studies and learning about the initiatives adopted by contributing institutions will lead to new connections and cooperation.

How Should I Read the Book?

However you want! We designed this book with a diverse audience in mind, and while some might be keen to read everything from beginning

to end, others may decide to focus on a selection of the most relevant chapters or case studies. All of our case studies have been carefully selected for your interest, however, you might find our 'How to Use This Cookbook' infographic helpful to navigate the cases of most interest to you. This book is analogous to a cookbook in the sense that it presents each individual case study in a similar format to that of a recipe. Each case study contains a list of 'key ingredients', that is, the institutional context and key elements required to successfully implement the initiative, as judged by the people directly involved. For example, information on the number of researchers involved, the target audience, the cost and ease of its implementation are presented in comprehensible, visual manner so that the reader can easily understand and compare case studies.

How Did this Book Come About?

There are many interesting initiatives utilised by research institutions all over the world to effectively engage with their research communities about research data. Typically, those interested in RDM support and engagement learn about these diverse activities at conferences, by going to a talk from someone who implemented such an initiative and/or discussing it in person. In addition, some RDM units and institutional libraries may have blogs that report their ongoing activities, and well-connected individuals may pick up useful information directly through their networks. . But is this really the most effective way to share good practice? What about those who cannot attend conferences, or those who are just starting with RDM and don't yet have established connections or know where to look online for more information? How do they get started?

With these concerns in mind, the authors, together with members of Research Data Alliance[1] (and the Libraries for Research Data Interest Group[2] in particular) decided to collect information from various institutions worldwide on how they engage researchers about managing their research data. The goal was to make this body of knowledge about good practice more readily available by collecting it into a book that

[1]   Research Data Alliance, https://www.rd-alliance.org/
[2]   Libraries for Research Data Interest Group, https://www.rd-alliance.org/groups/libraries-research-data.html

would be more discoverable and accessible to the wider community of research data supporters. Our goal was to make it as easy as possible for others to get started supporting good practice in RDM, and rather than reinventing the wheel, facilitate the adoption and adaptation of existing methods from similar institutional settings. We hope you find this book as interesting to read, as we found collecting the information and putting it all together.[3]

---

3   A blog post about the book sprint during which we wrote the *Cookbook* is available online: Connie Clare, 'Book Sprint Success: A Team Writing Exercise for the Win', 23 July 2019, https://www.rd-alliance.org/blogs/book-sprint-success-team-writing-exercise-win.html

II. Methodology

The aim of the project 'Research Engagement with Data Management' was to collect case studies from different organisations around the globe that focus on how to engage with the research community about research data management. By asking various questions about the models used and also about the organisational context, we created a useful resource for organisations that are looking to increase their engagement with their research communities.

In order to achieve this, we first designed and sent out a survey 'Researcher Engagement with Data Management: What Works?' to 60 funders, 80 scientific institutions, and 28 relevant mailing lists worldwide, as well as social media channels including blogs and Twitter. The survey was open from 18 January until 14 February 2019, and is available through Zenodo.[4]

Respondents were asked to think about which of their methods of researcher engagement would be of interest to other organisations. Each respondent was encouraged to mention as many initiatives used to engage researchers as they thought relevant, and to fill in the survey separately for each initiative. In addition, they were asked to characterise their research institution (number of researchers, number of PhD students, number of full-time employees providing data management support), as well as their engagement activity (target group, main drivers, activity cost, ease of implementation at a different institution) by responding to quantitative questions (see the survey template[5] for the details of questions and possible answers). For example, to estimate the

---

4   Iza Witkowska, 'The Survey Researcher Engagement with Data Management: What Works?' (22 July 2019), *Zenodo*, http://doi.org/10.5281/zenodo.3345305

5   Ibid.

cost of running the activity, respondents were asked to select one of five ranked options, ranging from 'inexpensive' to 'expensive'. Respondents were not asked to elaborate on their choices, or justify them.

We received 234 responses. Of these, 90 were complete and provided details describing the engagement activities, such as the activity objective, description, challenges and opportunities associated with the activity, etc. Responses that provided enough information to understand the activity were considered as valid responses and used for further selection of the most innovative activities.

The final selection of case studies was done by five volunteers. Each volunteer was asked to select 20 to 25 cases, which, based on their RDM knowledge and experience, looked innovative, inspiring and applicable to research institutions worldwide. There were no other criteria used for the selection process. To make sure that important engagement activities were not omitted from this study, the volunteers were also asked to suggest other innovative activities that they were aware of, but which were not submitted through the survey.

All cases selected by volunteers were used for the final selection. This selection was made based on the overlap between these five different lists. If a case study was listed on three or more of the five lists, it made it to the final list. In this way, 24 cases made it to this final list. Out of these, 17 were activities submitted via the survey and 7 were new activities. The list was discussed and approved, first by the five volunteers and afterwards by the entire project group.

In the next stage, we undertook hour-long interviews with 'respondents' of all shortlisted cases in order to collect missing information, quotes and photos. The interviews were recorded, transcribed and shared with the writing team.

To write the book, we organized a three-day 'book sprint' in The Hague, Netherlands. Six writers and two editors (one on-site and one working remotely), took part in the book sprint. Cases were grouped into eight themes, based on the main focus of the activity: policy, data management plans, training, events, community networks, dedicated consultants, interviews and data from senior researchers. All cases were divided between the writers, written up using the collected information, and then reviewed and edited with the help of the writers. By the end of the three days the first draft of the book was finished. After the book

sprint, the editing work continued, and final versions of each case study were sent to respondents for their approval during the following week. Subsequently, the book was publicly shared for consultation, and editing continued on an ongoing basis in response to community feedback.[6]

6   Draft book with full version and comment history available online: https://docs.google. com/document/d/1XnXJeOocmaz-xU0oTmMLpBXrcFTdHmBDQG8bHMq7_GY/

Table I. Graphical representation of the key ingredients of each case study, CC BY 4.0.

III. How to Use this Cookbook

This infographic (Table I, left) has been designed to help you to navigate case studies of interest, and to select those most suitable for implementation within your research institution. Just like a cookbook recipe, we provide a list of 'key ingredients' in a graphical format: the elements you'll need to successfully implement each initiative.

In addition, we also provide a table (Table II, below) with a quick overview of all case studies and recipes (ingredients) necessary to implement them. You can use this quick overview to navigate directly to cases which might be most relevant to the situation at your institution (for example, the amount of resources available to you to engage with researchers).

	1.1	1.2	2.1	2.2	2.3	3.1	3.2	3.3	3.4	4.1	4.2
No of researchers											
No of PhDs											
Target audience											
Main drivers											
Ease of implementation											
FTEs for RDM											
Costs materials											
Costs infrastructure											
Costs people											

Table II. Overview of all cases and their key ingredients, CC BY 4.0.

4.3	4.4	4.5	5.1	5.2	5.3	6.1	6.2	6.3	7.1.1	7.1.2	8.1	8.2

Table II. (continued from previous page).