

CONTENTS

About the author

Acknowledgements

Chapter 1 The data problem

- 1.1 Why is everyone talking about data management? 2
- 1.2 What is data management? 4
- 1.3 Why should you do data management? 8

Chapter 2 The data lifecycle

- 2.1 The data lifecycle 9
- 2.2 The data roadmap 11
- 2.3 Where to start with data management 14
- 2.4 Chapter summary 15

Chapter 3 Planning for data management

- 3.1 How to plan for data management 16
- 3.2 Creating a data management plan 19
- 3.3 Data policies 24
- 3.4 Case studies 31
- 3.5 Chapter summary 34

Chapter 4 Documentation

- 4.1 Research notes and lab notebooks 36
- 4.2 Methods 44
- 4.3 Other useful documentation formats 46
- 4.4 Metadata 50
- 4.5 Standards 57
- 4.6 Chapter summary 61

Chapter 5 Organization

- 5.1 File organization 62
- 5.2 Naming conventions 69

5.3 Documenting your conventions	73
5.4 Databases	75
5.5 Chapter summary	79
Chapter 6 Improving data analysis	80
6.1 Raw versus analyzed data	80
6.2 Preparing data for analysis	82
6.3 Managing your research code	88
6.4 Chapter summary	93
Chapter 7 Managing sensitive data	94
7.1 Types of sensitive data	94
7.2 Keeping data secure	98
7.3 Anonymizing data	108
7.4 Chapter summary	115
Chapter 8 Storage and backups	116
8.1 Storage	116
8.2 Backups	122
8.3 Case studies	125
8.4 Chapter summary	125
Chapter 9 Long-term storage and preservation	127
9.1 What to retain and how long to retain it	127
9.2 Preparing your data for the long term	132
9.3 Outsourcing data preservation	137
9.4 Chapter summary	139
Chapter 10 Sharing data	140
10.1 Data and intellectual property	140
10.2 Local data sharing and reuse	145
10.3 Collaborations	146
10.4 Public data sharing	147
10.5 Getting credit for shared data	160
10.6 Chapter summary	162
Chapter 11 Data reuse and restarting the data lifecycle	163
11.1 Finding and reusing data	163
11.2 Citing data	167
11.3 Restarting the data lifecycle	169
References	171
Index	186