
Contents

Preface..... xiii

Introductionxv

Author..... xvii

1. Collecting and Managing Research Data..... 1

 Introduction..... 1

 Topics 1

 Data and Science 1

 Data Gone Awry 3

 Case 1: False Alarm..... 3

 Case 2: Case of the Missing Eights 4

 Case 3: Unnoticed Outliers 4

 Case 4: Data Leaks 5

 Case 5: Surprisingly Consistent Responses..... 5

 Case 6: What Data?..... 6

 Case 7: Different Data..... 6

 Data Management Quality System 7

 Data Management’s Minimum Standards..... 9

 Determinates of Data Management Rigor 10

 Ethics..... 10

 Regulations 10

 Institutional Policies 11

 Funder Requirements..... 12

 Plans for Data Use or Reuse 12

 Frameworks for Thinking about Managing Research Data..... 13

 Summary 16

 Exercises 16

 References 20

2. Defining Data and Information..... 23

 Introduction..... 23

 Topics 23

 Data about Things in the Real World..... 23

 Data as a Model of the Real World 26

 Defining Data 27

 Three Fundamental Kinds of Data..... 28

 Data Elements..... 30

 Stevens’s Scales of Measurement..... 31

 Back to the Definition of Data and Information..... 33

 Categories of Data..... 34

Summary 40

Exercises 41

References 44

3. Deciding, Defining, and Documenting Data to Be Collected..... 47

Introduction 47

Topics 47

Constructs and Concepts..... 48

Operational and Conceptual Definitions 49

Reliability and Validity 51

Data Elements Revisited—ISO 11179 52

Special Considerations for Continuous Data Elements..... 54

Specifics for Discrete Data Elements..... 55

Deciding What Data to Collect 57

Choosing the Right Data Element When Options Exist 59

Documenting Data Element Definitions 60

Summary 62

Resources..... 63

Exercises 63

References 64

4. Data Management Planning 67

Introduction 67

Topics 67

Data Sharing and Documentation Requirements..... 68

What Is in a DMP? 72

Extent of Data Management Planning..... 75

Data Flow and Workflow 77

Summary 78

Exercises 78

References 80

5. Fundamental Dynamic Aspects of Data 83

Introduction 83

Topics 83

Time Aspects of Research Design 84

Defining Data Accuracy..... 84

Key Data-Related Milestones 85

The Ideal Case 88

Data Growing (or Shrinking) Over Time 90

 Addition or Inactivation of Data Values 91

 Addition of Experimental or Observational Units..... 92

 Addition of Measures 93

 Addition of Time Points 93

Data Moving and Changing Over Time..... 95

Summary 98

Exercises 98

References 102

6. Data Observation and Recording..... 103

Introduction..... 103

Topics 103

Four Possibilities for Observation and Recording 104

Observation by Humans..... 104

Obtaining Data by Asking Those Who Directly Observed or Experienced 106

Recording by Humans 109

Automated Observation and Recording 113

Tasks Associated with Observation and Recording..... 117

Summary..... 118

Exercises 118

References 122

7. Good Data Recording Practices Applicable to Man and Machine... 125

Introduction..... 125

Topics 125

Necessity of Good Documentation 125

Attributable..... 126

Legible 128

Contemporaneous..... 128

Original 129

Accurate..... 130

An Urban Legend about Data Recording..... 131

Summary..... 132

Exercises 132

References 135

8. Getting Data into Electronic Format 137

Introduction..... 137

Topics 137

Getting Data into Electronic Format..... 138

Preprocessing 140

Scan-Based Methods 140

Bar Codes 141

OMR and OCR 143

Speech Recognition 146

Key Entry 147

Quality Assurance and Control for Data Entry 150

Timing of Data Entry	151
Direct Electronic Capture of Data	151
Summary	153
Exercises	154
References	156
9. Data Structures	157
Introduction.....	157
Topics	157
Data Collection Formats	157
Data Storage Structures	162
Data Integration	168
Data Exchange Structures and Concepts	170
Summary	174
Exercises	174
References	177
10. Data Processing.....	179
Introduction.....	179
Topics	179
Data Cleaning.....	180
Imputations.....	185
Standardization.....	185
Mapping	186
Formatting Transformations	189
Conversions and Calculations	189
Workflow for Transformations	189
Data Enhancement.....	190
Summary.....	192
Exercises	193
References	195
11. Designing and Documenting Data Flow and Workflow.....	197
Introduction.....	197
Topics	197
Importance of Diagrams.....	197
Diagrams as Models	199
Two Types of Diagrams	199
Data Flow and Context.....	200
DFD Symbols and Conventions.....	202
Workflow Diagrams	206
Workflow Diagram Symbols and Conventions.....	208
Summary	212
Exercises	212
References	217

- 12. Selecting Software for Collecting and Managing Data 219**
 - Introduction..... 219
 - Topics 219
 - Automation 220
 - Data Collection and Processing Tasks for Which Computers Are
Often Used 225
 - Types of Computer Systems Used for Data Processing
and Storage..... 227
 - Software Selection 229
 - Summary 235
 - Exercises 236
 - Reference 238

- 13. The Data Management Quality System..... 239**
 - Introduction..... 239
 - Topics 239
 - Quality Management Systems 240
 - Defining Quality 241
 - Frameworks for Managing Quality 243
 - A Framework for Quality Management of Research Data 246
 - Summary 253
 - Exercises 253
 - References 256

- 14. Calculating the Time and Cost for Data Collection and
Processing..... 259**
 - Introduction..... 259
 - Topics 259
 - The Hidden Value of Project Budgeting..... 260
 - Modeling Data Collection and Management for a Research
Project 261
 - The Project Budget..... 265
 - Cost Forecasting, Tracking, and Management..... 266
 - Summary 267
 - Exercises 268

- 15. Research Data Security 273**
 - Topics 273
 - Security as Protection against Loss of Confidentiality,
Integrity, or Availability..... 273
 - Steps Investigators Can Take to Prevent Loss 276
 - Risk-Based Categorizations of Data and Information Systems 279
 - Implications for Contractual Data Protections..... 282
 - Components of a Research Data Security Plan 283
 - Summary 284

Exercises 285

References 289

16. Data Ownership, Stewardship, and Sharing..... 291

 Introduction..... 291

 Topics 291

 Institutional Policies for Research Data..... 291

 Data Ownership and Stewardship..... 292

 Federal Requirements for Data Sharing 293

 Effective Data Sharing..... 294

 Special Considerations for Sharing Data about Human Subjects 294

 Special Considerations for Sharing Data Received from Others 296

 Summary 297

 Exercises 297

 References 298

17. Data Archival..... 299

 Introduction..... 299

 Topics 299

 Archiving and Reuse of Research Data..... 300

 What to Archive..... 302

 Formats for Archival 308

 Records Retention 308

 Discoverability of Archived Data and Research Records..... 308

 Summary 309

 Exercises 309

 References 311

Glossary 313

Index 325