

# CONTENTS

<i>List of Figures and Table</i>	viii
<i>Data Stories</i>	x
<i>About the Authors</i>	xi
<i>Acknowledgements</i>	xii
<i>Overview of the Book</i>	xv
<i>Introduction</i>	xvii
<b>Part I Data in Society</b>	<b>1</b>
1 Data in Society	3
1.1 Introduction: Who cares about data?	3
1.2 Datafication and its components	6
1.3 Data, ethics and knowledge production	10
1.4 Conclusion: The impact of Datafication	13
<b>Part II Data Creation</b>	<b>15</b>
2 Big Data in Context	21
2.1 Introduction: The rise of Big Data	21
2.2 The Big Data mythology: Data transforms society	25
2.3 A historical perspective: Society transforms data	28
2.4 Conclusion: Data do not speak for themselves	32
3 Characteristics of Data	35
3.1 Introduction: Data do not stay still	35
3.2 Data are not neutral	40
3.3 Data are context-dependent	46
3.4 Conclusion: Characteristics of data	48
4 Data, Evidence and Knowledge	51
4.1 Introduction: The representational and the relational views on data	52
4.2 What is evidence? The path from data to knowledge	55
4.3 Examples of data within the knowledge production cycle	57
4.4 Contrasting the representational and relational perspectives	59
4.5 Conclusion: Data science in a relational perspective	62



<b>Part III</b>	<b>Data Circulation</b>	<b>65</b>
5	Putting Data to Work	73
5.1	Introduction: The complexity of putting data to work	74
5.2	The challenge of ‘messy’ data	75
5.3	Infrastructures	77
5.4	Conventions and metadata	81
5.5	Models	84
5.6	Visualisations: Forms, tools and interfaces	87
5.7	Curation	91
5.8	Conclusion: Forms of data work	93
6	New Data Skills	94
6.1	Introduction: Data expertise	94
6.2	What is data science?	95
6.3	Data science skills	101
6.4	Bringing skills together	107
6.5	Conclusion: Becoming a data scientist today	113
7	Governance of Data Journeys	116
7.1	Introduction: What is data governance?	117
7.2	Data as private commodities: Closed data	119
7.3	Data as public goods: Open data	122
7.4	A hard case: The journeys of health-related data	125
7.5	Shifting focus to usable data: The FAIR principles	129
7.6	International data journeys and the problem of data inequities	131
7.7	Conclusion: Governance is not a silver bullet	134
<b>Part IV</b>	<b>Data Value, Innovation and Responsibility</b>	<b>135</b>
8	Data as a Source of Value	143
8.1	Introduction: What makes data valuable?	144
8.2	Assumptions about the value of data	145
8.3	Data and innovation	147
8.4	The data economy	149
8.5	Who benefits from the value of data?	152
8.6	Allocating value, responsibility and profit	154
8.7	How does AI add value to data?	156
8.8	The value of prediction	159
8.9	The value of metrics	160
8.10	Conclusion: Making data valuable	161



9	Data Justice and Ethics	163
9.1	Introduction: From data value to data ethics	164
9.2	Which data are ethically sensitive?	165
9.3	Data justice: Implementing fairness	168
9.4	Ethics for data work: General frameworks	172
9.5	Ethics in data work: Assessing technical decisions	175
9.6	Responsibilities of data workers	178
9.7	Conclusion: From analysis to action, from rules to power	182
10	Responsible Use of Data as Evidence	184
10.1	Introduction: Data matters	185
10.2	What is evidence-based decision making?	187
10.3	Ensuring responsible use of data	191
10.4	Legal frameworks and formal regulation	192
10.5	Codes of conduct	194
10.6	Computational metrics and design	195
10.7	Organisational and cultural interventions	197
10.8	Institutional Review Boards	198
10.9	Social participation and slow science	199
10.10	Conclusion: Responsibility, monitoring and trust	201
<b>Part V</b>	<b>Conclusion: Data and the Knowledge We Need</b>	<b>203</b>
11	Towards Good Data Science	211
11.1	Lesson 1: 'Data' is a relational category	212
11.2	Lesson 2: Infrastructures and data stewardship are essential to extract knowledge from Big Data	212
11.3	Lesson 3: Data workers must use data sources with discernment and be aware of the risks of discrimination and inequality connected to data	213
11.4	Lesson 4: Ethics, security and social responsibility are a fundamental part of data work	215
11.5	Lesson 5: Responsible data work requires social dialogue, community engagement and contributions to data literacy	216
	<i>Glossary</i>	218
	<i>References</i>	221
	<i>Index</i>	240



# DATA STORIES

Data Story 1: Big Data on Consumer Habits	16
Data Story 2: Remote Sensing for Conservation Research	17
Data Story 3: Geolocation: It's a GIS World	66
Data Story 4: Tracking Tuberculosis Using Phone Data	68
Data Story 5: Uber Drivers	136
Data Story 6: Dating Apps	140
Data Story 7: Emergency Data Science and the Use of COVID-19 Contact Tracing Apps	204
Data Story 8: Combining Data to Advance the Sustainable Development Goals	207