

## Teil I Grundlagen

|          |                                      |    |
|----------|--------------------------------------|----|
| <b>1</b> | <b>Einleitung und Überblick</b>      | 3  |
| 1.1      | Was sind Computational Methods?.     | 6  |
| 1.2      | Der Werkzeugkoffer                   | 13 |
| 1.3      | Mit Fehlern umgehen                  | 21 |
| 1.4      | Überblick über das Buch.             | 23 |
|          | Literatur                            | 27 |
| <b>2</b> | <b>Datenquellen</b>                  | 31 |
| 2.1      | Webseiten                            | 34 |
| 2.2      | Application Programming Interfaces   | 37 |
| 2.3      | Datenbanken und Datensätze           | 42 |
|          | Literatur                            | 48 |
| <b>3</b> | <b>Datenformate</b>                  | 53 |
| 3.1      | Datentypen                           | 55 |
| 3.2      | Textformate (MD)                     | 58 |
| 3.3      | Tabellenformate (CSV)                | 61 |
| 3.4      | Auszeichnungssprachen (HTML und XML) | 62 |
| 3.5      | Objektdatenformate (JSON)            | 67 |
| 3.6      | Datenbanken (SQL und NoSQL)          | 68 |
| 3.7      | Datenmodelle (RDF)                   | 72 |
|          | Literatur                            | 79 |

|                                             |     |
|---------------------------------------------|-----|
| <b>4 Datenextraktion</b> .....              | 83  |
| 4.1 Selektionsverfahren und -sprachen ..... | 85  |
| 4.2 Transformationsverfahren .....          | 112 |
| Literatur .....                             | 128 |

## Teil II Programmierkonzepte

|                                                                    |     |
|--------------------------------------------------------------------|-----|
| <b>5 Programmierung</b> .....                                      | 133 |
| 5.1 Einführung in die Datenanalyse mit R .....                     | 136 |
| 5.2 Einführung in die Datenanalyse mit Python .....                | 174 |
| Literatur .....                                                    | 202 |
| <b>6 Scaling up. Daten und Skripte organisieren</b> .....          | 207 |
| 6.1 Versionsverwaltung .....                                       | 209 |
| 6.2 Entwicklungsumgebungen .....                                   | 214 |
| 6.3 Laufzeitumgebungen und virtuelle Boxen .....                   | 220 |
| 6.4 Parallelisierung mit Cores, Clustern und Cloud Computing. .... | 235 |
| Literatur .....                                                    | 249 |

## Teil III Anwendungsfelder

|                                                                |     |
|----------------------------------------------------------------|-----|
| <b>7 Automatisierte Datenerhebung</b> .....                    | 255 |
| 7.1 Webscraping .....                                          | 257 |
| 7.2 Application Programming Interfaces (APIs) .....            | 282 |
| Literatur .....                                                | 301 |
| <b>8 Maschinelles Lernen</b> .....                             | 305 |
| 8.1 Überwachte Lernverfahren .....                             | 308 |
| 8.2 Unüberwachte Lernverfahren .....                           | 336 |
| Literatur .....                                                | 350 |
| <b>9 Textanalyse</b> .....                                     | 355 |
| 9.1 Wortfrequenzanalysen .....                                 | 357 |
| 9.2 Diktionärbasierte Inhaltsanalyse .....                     | 373 |
| 9.3 Syntax und Semantik .....                                  | 380 |
| Literatur .....                                                | 386 |
| <b>10 Netzwerkanalyse</b> .....                                | 389 |
| 10.1 Grundlegende Konzepte der Netzwerkanalyse .....           | 392 |
| 10.2 Erhebung, Analyse und Visualisierung von Netzwerken ..... | 402 |
| Literatur .....                                                | 419 |

---

|           |                             |     |
|-----------|-----------------------------|-----|
| <b>11</b> | <b>Simulationsverfahren</b> | 423 |
| 11.1      | Eine Welt entsteht          | 428 |
| 11.2      | Die Welt entwickelt sich    | 429 |
| 11.3      | Analyse der Ergebnisse      | 431 |
|           | Literatur                   | 433 |
| <b>12</b> | <b>Zusammenfassung</b>      | 435 |
|           | Literatur                   | 438 |
|           | <b>Literatur</b>            | 439 |