

1	Einführung	1
1.1	Graphen und Netzwerke	1
1.2	Netzwerkwissenschaft als Wissenschaftsdisziplin	6
1.3	Zusammenfassung	7
1.4	Übungen	7
	Quellen	8
2	Umgang mit Netzwerkdaten	11
2.1	Eingrenzung der Untersuchungspopulation (Knoten)	12
2.2	Eingrenzung der Relationen (Kanten)	12
2.3	Eingrenzung von Attributen	13
2.4	Modellierung von Netzwerkdaten	14
2.4.1	Property Graph Model (PGM)	15
2.4.2	Beispiel: PGM	16
2.5	Erhebung von Netzwerkdaten	19
2.5.1	Auswahl des Erhebungsverfahrens	19
2.5.2	Merkmalsträger, Attribute und Analyseebenen	19
2.5.3	Dateiformate	20
2.5.4	Technologien	20
2.6	Netzwerkanalyse	23
2.6.1	Analyseperspektiven und -ebenen	23
2.6.2	Analyseverfahren	24
2.7	Aspekte der Visualisierung von Netzwerken	25
2.7.1	Layout	25
2.7.2	Visualisieren von Attributen	27
2.7.3	Färbungen	28

2.7.4	Gruppierungen	28
2.8	Beispiel: Netzwerk des Musikprojekts „Desert Sessions“	29
2.9	Zusammenfassung	37
2.10	Übungen	40
	Quellen	40
3	Grundlagen der Graphentheorie	43
3.1	Netzwerkrepräsentationen	44
3.1.1	Adjazenzmatrix	45
3.1.2	Inzidenzmatrix	46
3.2	Typen	46
3.2.1	Multigraphen	46
3.2.2	Gewichtete Netzwerke	47
3.2.3	Hypergraphen	47
3.2.4	Bipartite Graphen	48
3.2.5	Bäume	50
3.2.6	Planare Graphen	51
3.3	Einfache Kennzahlen	51
3.3.1	Grad	51
3.3.2	Durchschnittlicher Grad	53
3.3.3	Gradverteilung	54
3.3.4	Dichte	55
3.3.5	Distanzen, Pfade und Wege	58
3.3.6	Kürzester Pfad	60
3.3.7	Diameter	61
3.3.8	Durchschnittliche Pfadlänge	62
3.4	Komponenten	62
3.5	Suchprozesse	64
3.5.1	Breitensuche (BFS)	64
3.5.2	Tiefensuche (DFS)	67
3.6	Zusammenfassung	68
3.7	Übungen	69
	Quellen	70
4	Knoten: Position und Zentralität	73
4.1	Zentralitätsmaße	74
4.1.1	Gradzentralität	74
4.1.2	Nähezentralität	75
4.1.3	Zwischenzentralität	76
4.1.4	Clustering-Koeffizient	79
4.1.5	Eigenvektorzentralität	80
4.1.6	Katz-Zentralität	82
4.2	Zusammenfassung	83

Inhaltsverzeichnis		XV
4.3	Übungen	84
	Quellen	84
5	Knotengruppen	85
5.1	Dyade	86
5.1.1	Dyadenkonstellationen und -zensus	86
5.1.2	Reziprozität in Dyaden	87
5.2	Triaden	88
5.2.1	Triadenkonstellationen und -zensus	88
5.2.2	Transitivität	92
5.3	Ego-zentrierte Netzwerke	95
5.4	Gemeinschaften	97
5.4.1	Similarität und Äquivalenz	98
5.4.2	Algorithmische Aspekte	99
5.4.2.1	Partitionierung vs. Detektion	99
5.4.2.2	Agglomeration vs. Division	99
5.4.2.3	Bisektion	100
5.4.2.4	Güte und Aussagekraft von Zerlegungen: Modularität	101
5.4.3	Gemeinschaftsdetektion (reguläre Äquivalenz)	103
5.4.3.1	Signierte Graphen und strukturelle Balance	103
5.4.3.2	CONCOR	104
5.4.4	Gemeinschaftsdetektion (strukturelle Äquivalenz)	105
5.4.4.1	Cliquen und k-Konzepte	105
5.4.4.2	Hierarchisches Clustering	107
5.4.4.3	Label Propagation (LP)	107
5.4.5	Beispiel: LP	108
5.4.5.1	R	108
5.4.5.2	Julia	109
5.4.5.3	Python	111
5.4.5.4	Vergleich der Ergebnisse	112
5.5	Zusammenfassung	112
5.6	Übungen	113
	Quellen	115
6	Netzwerkmodelle	117
6.1	Wahrscheinlichkeitsverteilungen	118
6.1.1	Binomialverteilung	121
6.1.2	Poisson-Verteilung	122
6.1.3	Exponentialverteilung	123
6.1.4	Potenzgesetz	124
6.2	Strukturmodelle: Verdrahtung	126
6.2.1	Zufallsgraphen	126

6.2.2	Kleine-Welt-Modell	130
6.2.3	Skalenfreie Netze	132
6.2.4	Konfigurationsmodell	133
6.2.5	Beispiel: Konfigurationsmodell „GitHub Social Network“	136
6.2.6	Exponential Random Graph Models (ERGMs)	138
6.2.7	Beispiel: ERGM (Freundschaftsnetzwerk „Volksschule“)	141
6.3	Dynamische Aspekte von Netzwerken	148
6.3.1	Veränderung	149
6.3.2	Modelle mit Fokus auf Kantenveränderung	152
6.3.2.1	Bevorzugte Verbindungen	152
6.3.2.2	Homophilie	154
6.3.3	Modelle mit Fokus auf Knotenveränderung	155
6.3.3.1	k -Schwellenwert	155
6.3.3.2	Epidemiologische Modelle	156
6.3.3.3	Verzweigung	162
6.4	Zusammenfassung	164
6.5	Übungen	165
	Quellen	166
7	Spezielles Kapitel: Vorgehensmodelle	171
7.1	Analyse bestehender Vorgehensmodelle	171
7.2	Graph Data Science Workflow (GDSW)	173
7.2.1	Exploration: Erkundung, Vorbereitung und Planung	173
7.2.2	Datenbeschaffung	175
7.2.3	Datenpräparation	175
7.2.4	Datenexploration	176
7.2.5	Modellierung/Hypothesentest	177
7.2.6	Interpretation	177
7.2.7	Publikation	177
7.2.8	Operationalisierung	178
7.3	Phasenübergreifende Kompetenzen und Aktivitäten	178
7.4	Zusammenfassung	178
	Quellen	179
8	Spezielles Kapitel: Aspekte der Ethik	181
8.1	Funktionen und Ebenen der Ethik in der Informationstechnologie (IT)	182
8.2	Ethisches Handeln: eine didaktische Perspektive	184
8.2.1	Ethische Grundsätze, Leitlinien und Kodizes	185
8.2.2	Methodische Normen verantwortungsvoller Forschung	188
8.3	Ethik im wissenschaftlichen Arbeitsprozess	188
8.3.1	Verzerrung und Varianz	189
8.3.2	Vermeidung von Diskriminierung	190

8.4	Zusammenfassung	191
	Quellen	193
9	Spezielles Kapitel: Datenschutz	197
9.1	Gesetzlicher Rahmen in Deutschland	198
9.2	Zweck des Datenschutzes	198
9.3	Begriffe im Umfeld des Datenschutzes	199
9.3.1	Personenbezogene Daten einer betroffenen Person	200
9.3.2	Rechtssubjekte im Umfeld betroffener Personen	201
9.3.3	Pseudonymisierung	201
9.4	Grundsätze im Umgang mit personenbezogenen Daten	202
9.4.1	Rechtmäßigkeit der Verarbeitung	202
9.4.2	Zweckbindung	203
9.4.3	Datenminimierung	203
9.4.4	Richtigkeit	203
9.4.5	Speicherbegrenzung	204
9.4.6	Integrität und Vertraulichkeit	204
9.5	Verbot mit Erlaubnisvorbehalt	204
9.6	Privacy Preserving Data Mining (PPDM)	205
9.6.1	Data Mining: Begriff, Prozessschritte und Rollen	205
9.6.2	PPDM-Framework	208
9.6.3	PPDM: Techniken, Methoden und Algorithmen	209
9.7	Zusammenfassung	210
	Quellen	212
10	Schlussbetrachtung	215
10.1	Zusammenfassung	215
10.2	Limitationen und Ausblick	221
A	Lösungsvorschläge: Kap. 1	225
B	Lösungsvorschläge: Kap. 2	227
C	Lösungsvorschläge: Kap. 3	229
D	Lösungsvorschläge: Kap. 4	233
E	Lösungsvorschläge: Kap. 5	235
F	Lösungsvorschläge: Kap. 6	247
	Stichwortverzeichnis	255